Stochastic Gradient Descent Optimization for Low Power Nano-CMOS Thermal Sensor Design

Oghenekarho Okobiah[1], Saraju P. Mohanty[2], Elias Kougianos[3], Oleg Garitselov[4], and Geng Zheng[5]
NanoSystem Design Laboratory (NSDL, http://nsdl.cse.unt.edu)
University of North Texas, Denton, TX 76207, USA.
E-mail ID: oo0032@unt.edu[1], saraju.mohanty@unt.edu[2], elias.kougianos@unt.edu[3],
omg0006@unt.edu[4], and gengzheng@my.unt.edu[5]

***Abstract*—The drive for ultra efficient and low-cost portable devices continues to push the need for low power circuit designs. The increasing transistor density and complexity of IC designs aggravates the task of producing efficient low power and low cost design. The short time to market (TTM) also increases this burden on designers, as optimal designs have to be produced in an ever decreasing amount of time. This paper presents an optimization design flow methodology that optimizes the power (accounting leakage) consumption of integrated circuits (ICs). The design flow incorporates a stochastic gradient descent (SGD) based algorithm and is implemented using a 45 nm thermal sensor circuit as case study. Power-efficient high-sensitive thermal sensors are important to reduce the burden on the systems or circuits that they are implanted to sense. Experiments are performed to apply the proposed design flow methodology on the thermal sensor with the power consumption as the design objective while keeping the temperature resolution as a constraint. Experiments on full-blown (RCLK) netlist of sense amplifier show a reduction in power consumption by 38%.**

***Keywords*-Optimization, Stochastic Gradient Descent, Nano-CMOS, Low Power, Design Flow, Thermal Sensor**

I. Introduction

The market desire for ultra efficient and low cost mobile devices continues to drive the needs for low power IC design. However with the increasing transistor density and complexity of System-on-Chip (SOC) designs, it is becoming more and more difficult for designers to produce optimal designs efficiently. The complexity of designs and the number of design parameters makes an exhaustive exploration of the design space to find optimal designs infeasible. To overcome this restriction, intelligent techniques and methods have been used during the design process to reduce the cost of producing optimal designs. Optimization techniques, a critical part of circuit design have been researched and are ubiquitous [1].

Common techniques include gate ordering, transistor sizing, clock gating, architecture analysis etc., and can be applied at different levels of abstraction including circuit, logic, behavioral and system level. Relatively recent techniques applicable to designs in the deep nanometer region also include the use of dual threshold voltage and thickness oxide [2]. Popular optimization algorithms have been combined with most techniques to improve the optimization process. However, with increasing design complexities, and conflicting design objectives, the research of optimization algorithms is still very important.

The current paper proposes a design optimization flow methodology incorporating a stochastic gradient descent based (SGD) algorithm. The SGD algorithm helps to improve the optimization time and also eliminates the problem of local optima. The design flow is presented using a 45 nm thermal sensor as case study circuit. The thermal sensor is a critical component of SoCs, even more so with the increasing power density of modern circuits. Over 50 % of integrated circuit (IC) failures can be attributed to thermal related issues [3]. The thermal sensors are needed for effective thermal management which helps to reduce power consumption and increase performance.

The **novel contribution** of this paper is the presentation of a design optimization flow model, which incorporates an SGD algorithm for the optimization of a design objective under constraints. The SGD algorithm has been modified to reiteratively have restart points to eliminate the local optimum problem. The proposed algorithm has been applied on a 45nm sensor design at the silicon level.

The rest of this paper is organized as follows. A brief review of selected related research is presented in Section II. In Section III, a description of the case study circuit is presented. The proposed design optimization flow methodology is presented in section IV. The experimental setup and results are presented in section V. In Section VI, conclusions and future research directions are discussed.

II. Related Research

Optimization algorithms are used extensively in IC design and have been well researched and documented. A few prominent algorithms include evolutionary algorithms, genetic algorithms, swarm algorithms, simulated annealing, tabu search and geometric programming [4], [5], [6], [7], [8]. Gradient descent algorithms have been used in many optimization problems. In particular, Stochastic Gradient Descent (SGD), has also been applied in optimization methods. SGD differs from ordinary gradient descent by estimating the gradient descent at random decision points instead of going down through the whole parameter set. In [9], a comparison of optimization methods including three variations of SGD has been presented. Its has the advantage of decreasing the search time by reducing the gradient computation time. In [10], a

form of gradient descent for multi-objective optimization is proposed.

Although the use of SGD in VLSI systems have been reported previously [11], [12], [13], it has not been applied as an optimization method for the circuit itself. In [11], [12], the SGD algorithm is implemented to optimize an adaptive optical system. In [13], SGD is used as a training optimization for implementation on VLSI neural networks.

The design of on-chip thermal sensors has been well researched including design for accurate temperature estimation and robust performance [14], [15], [16]. In [15], a class of thermal sensors based on Differential Ring Oscillators (DRO) is introduced. In [17], a low power thermal sensor has also been proposed. It employs a an oscillator based on RS register based structure. In [16], a statistical approach is taken to compensate for the effect of noise, process variations and V_{DD} fluctuations on the thermal sensor. In [18], [19], a proposed PTAT current source is proposed. The proposed circuit uses the ratio between the drain currents of two current source transistors operating in the subthreshold region which is PTAT for thermal sensing. An effort to reduce the effect of process variation and noise on thermal sensors, similar circuits have been proposed in [20]. The thermal sensor design used as a study in this paper is closely similar to the design presented in [14]. The thermal sensor used is implemented using the conventional ring oscillator topology in contrast to the current starved topology. The thermal sensor is also not operated in the subthreshold region which leads to a decrease in frequency with increasing temperature. We do not also include the frequency divider and multiplexer.

A summary of related sensors is shown in Table I. This provides a broad perspective of the state-of-the art.

III. Thermal Sensor Design for 45 nm

The thermal sensor design used to demonstrate the proposed design flow optimization is briefly discussed here. The 45 nm thermal sensor shown in Fig. 1 uses a ring oscillator as a major component for thermal sensing. The circuit also uses a combination of 10-bit binary counters and 10-bit registers for accurately expressing the output. The operational frequency of the ring oscillator is very sensitive to ambient temperature and thus the output frequency fluctuates in response to the effect from surrounding temperature.

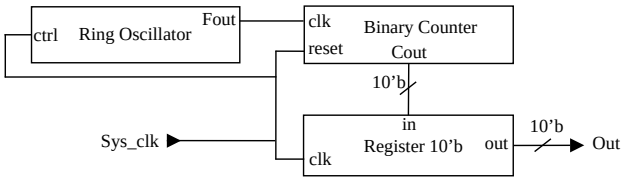

Fig. 1. Block Diagrammatic Representation of the Thermal Sensor System.

The ring oscillator is shown in Fig. 2. It consists of a cascade of an odd number of inverters that are connected in a loop leading to an unstable state which creates the oscillations. The ring oscillator shown in Fig. 2 has a total of 15 inverters, but the first inverter has been modified as a NAND gate and used to enable the ring oscillator operation.

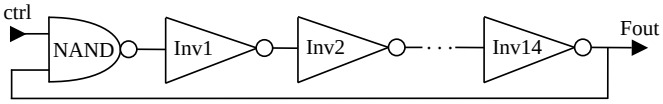

Fig. 2. Block Diagram of the Ring Oscillator.

The oscillation frequency of the ring oscillator is expressed using the following expression:

$$f_{osc} = \frac{1}{n(t_{pLH} + t_{pHL})}. \tag{1}$$

Where n is the number of stages used in the oscillator and t_{pLH} and t_{pHL} are the low-to-high and high-to-low propagation delays, respectively. The propagation delays can be expressed as follows [20]:

$$t_{pLH} = \frac{-2C_L V_{tp}}{\kappa_p (V_{dd} - Vtp)^2} + \frac{C_L}{\kappa_p (V_{dd} - Vtp)} \ln \frac{1.5V_{dd} + 2V_{tp}}{0.5V_{dd}}. \tag{2}$$

$$t_{pLH} = \frac{2C_L V_{tn}}{\kappa_n (V_{dd} - Vtn)^2} + \frac{C_L}{\kappa_p (V_{dd} - Vtn)} \ln \frac{1.5V_{dd} + 2V_{tn}}{0.5V_{dd}}. \tag{3}$$

Where C_L is the capacitive load and κ_n and κ_p are the transconductance values given as follows:

$$\kappa_{n/p} = \mu_n C_{ox} (\frac{W}{L})_{n/p}. \tag{4}$$

In equations 1 - 3, the threshold voltage V_t and mobilities are the factors most sensitive to temperature fluctuations. They are given in equations 5 and 6, respectively, as below [21]:

$$V_t(T) = V_t(T_0) + \alpha_{V_t}(T - T_0), \alpha_{V_t} = -0.5 - 3.0mV/^\circ K. \tag{5}$$

$$\mu(T) = \mu_0 \left(\frac{T}{T_0}\right)^{\alpha_\mu}, \alpha_\mu = -1.2 - 2.0. \tag{6}$$

An increase in temperature leads to an increase in the propagation delay which translates to a decrease in oscillating frequency.

The 10-bit binary counter is shown in Fig. 3 and consists of JK flip-flops, while the 10-bit register shown in Fig. 4 is used to store the value from the counter and is also implemented with JK flip-flops.

The thermal sensor shown in Fig. 1 was implemented using a 45 nm CMOS technology library. The thermal sensor design is characterized to sense temperatures between 0°C and 100°C. The *Sys_clk* signal is used to enable the thermal sensor. When the *Sys_clk* turns to logic zero, the ring oscillator is disabled, the counter is also reset and the register also stops saving the count, storing the last count value it had before the *Sys_clk* was set to logic "0". The binary counter is used to count the frequency difference between the ring oscillator output and the system clock. The count is stored in the 10-bit register and calibrated to measure the temperature change. The physical design of the thermal sensor is shown in Fig. 5.

TABLE I
COMPARISON OF RELATED THERMAL SENSORS

Sensor Design	Average Power	Temperature Resolution	Temperature Range	Process Technology	Area
[15]	25 μW	2°C	40 – 150°C	4 5nm	–
[17]	0.9 μW	1°C		180 nm	0.2 mm^2
[18]	3.8 μW	–	-40 – 125°C	180 nm	1000 μm^2
[19]	5.8 μW	–	-20 – 100°C	0.35 μm	–
[20]	0.09 μW	–	–	0.13 μm	0.0036 mm^2
[14]	0.95 μW	0.04°C	8 – 85°C	0.13 μm	0.04 mm^2
This paper	379.4 μW	9.42 MHz/°C	0 – 100°C	45 nm	1389.31 μm^2

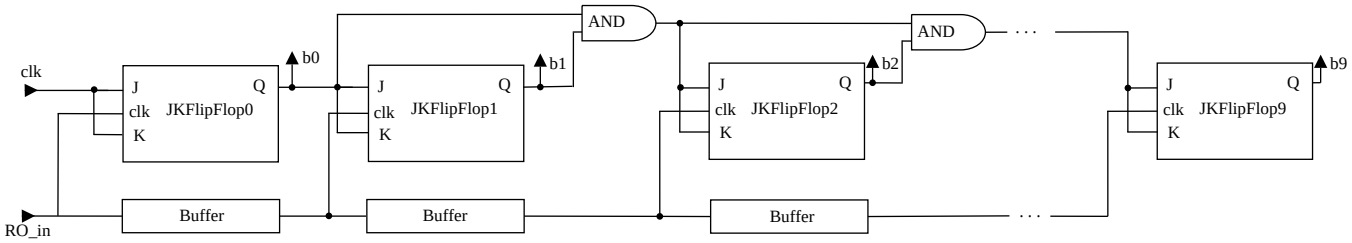

Fig. 3. Block diagram of the 10 bit binary counter.

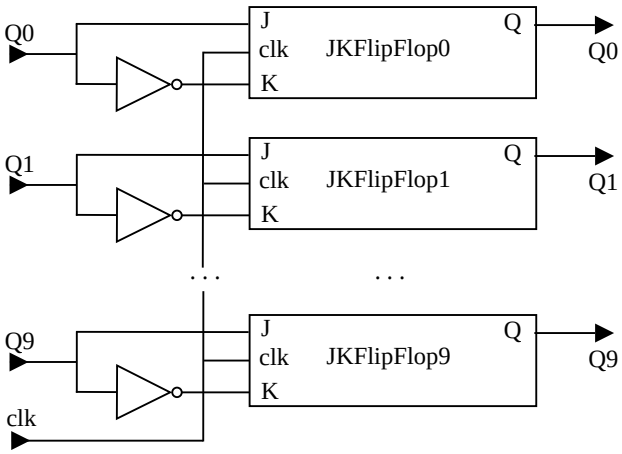

Fig. 4. Block diagram of the 10-bit register.

The performance and accuracy of the physical design is degraded when compared to the schematic design. This is expected due to parasitic effects from the layout. Table II shows a comparison between the schematic and physical design. Power consumption is increased by 29% while the sensitivity decreases by 44%. This circuit exhibits a linear dependence of oscillation frequency on junction temperature as shown in Fig. 6.

TABLE II
CHARACTERIZATION OF THE 45NM CMOS THERMAL SENSOR.

Design	Average Power (P_{TS})	Sensitivity (T_{TS})	Area (μm^2)
Schematic	293.1 μW	16.88 MHz/°C	-
Layout	379.4 μW	9.42 MHz/°C	1221.37
% Change	+29%	-44%	

As the temperature is increased, the frequency decreases.

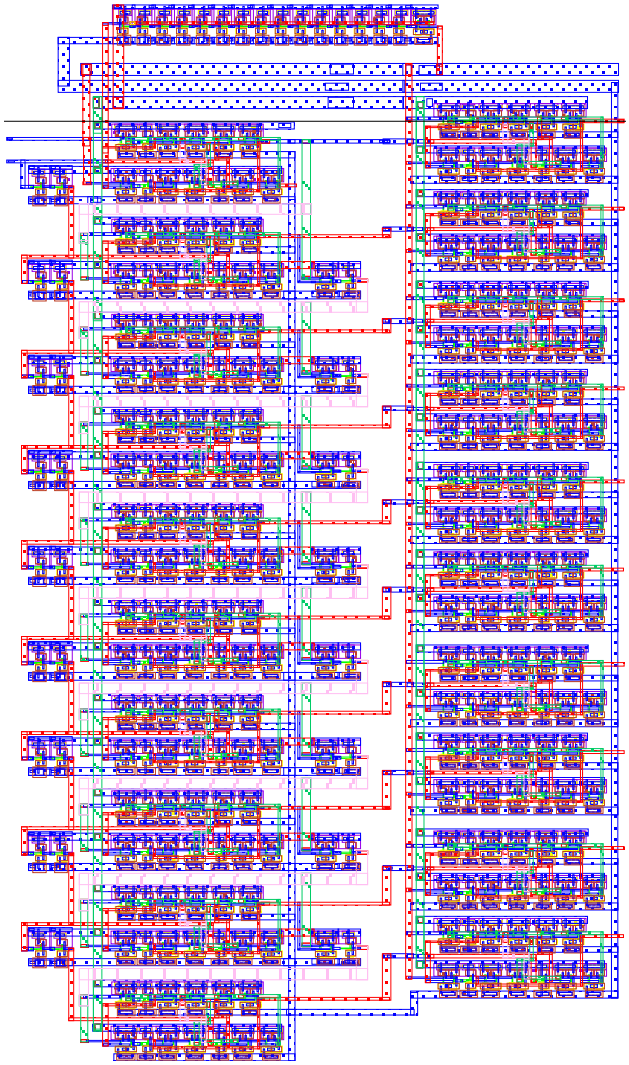

Fig. 5. Physical design of the 45 nm thermal sensor.

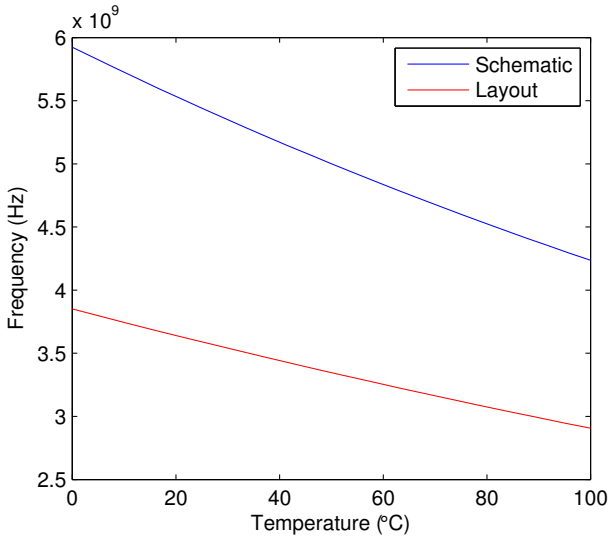

Fig. 6. Ring oscillator frequency response versus temperature for both schematic and physical design of 45nm thermal sensor.

The schematic frequencies range from 0°C= 5.924 GHz to 100°C= 4.236 GHz. Assuming a 6 GHz max clock rate for the ring oscillator, and a 10 bit counter (1024 max count) the effective resolution is calculated by dividing the temperature range by the number count 100°C/1024 bit which gives a 0.097°C/bit resolution. The range of frequency output is also severely degraded as also seen in Fig. 6. The range drops to 3.867 GHz to 2.986 GHz. The resolution can also be specified in terms of GHz/°C to reflect the degrading effect of parasitics from the physical design. There is a 47.8% change is frequency/temperature resolution between the schematic and physical design. The area of the layout is 1221.37 μm^2.

IV. Proposed Design Optimization Flow

In many of the reviewed circuits proposed, optimization of the circuits for minimal power consumption has been the goal. In striving to achieve optimal power consumption, the accuracy or sensitivity of the circuit is often compromised. A new design flow methodology that uses a stochastic gradient descent based algorithm is proposed, as shown in Fig. 7.

A. Design Optimization Flow

The first step in the design flow process is to create the baseline schematic design of the circuit that meets the given design specifications. For the case study circuit we implement, the thermal sensor, common design objectives include power consumption, temperature resolution, and temperature range. After the schematic design has been created, a set of performance objectives are identified (Figures-of-Merit, FoMs) and a functional simulation is performed to ensure that the circuit meets initial specifications. If the design specifications are not met, the schematic is reiteratively designed until the specifications are met. The next step is to create the physical layout design of the circuit. The physical layout is validated with Design Rule Checks (DRC), and Layout

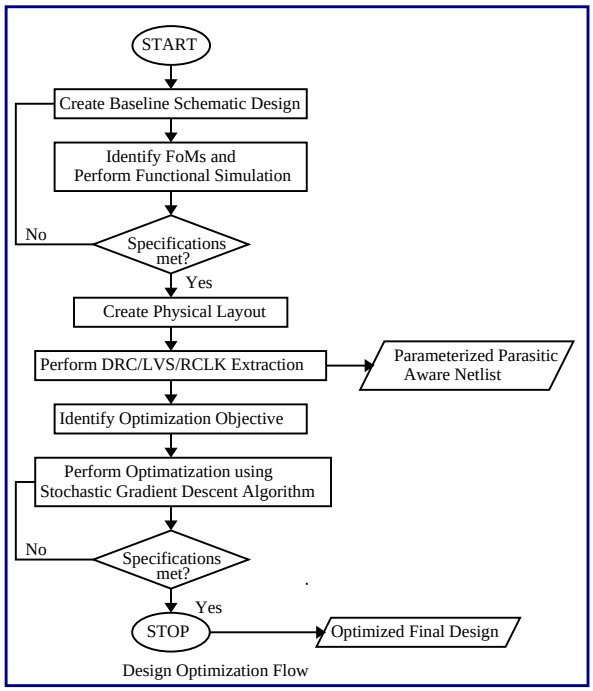

Fig. 7. The Proposed design optimization flow.

vs. Schematic (LVS) tests. From the physical layout, a fully parasitic netlist - resistance, capacitance and self and mutual inductance (RLCK) is extracted to ensure the simulation model is as silicon accurate as possible. The parasitic netlist is then parameterized with design and process parameters, including the length and width of the transistors (L, W), threshold voltage (V_t), oxide thickness (T_{ox}), etc. It is only after the optimization is complete that the physical design is redrawn using the parameters obtained from the optimization process. This ensures that the manual design of the physical layout is done at most twice, once before the parasitic extraction of the netlist and after the optimization process is complete.

With a fully parameterized parasitic aware netlist and a chosen performance objective, a stochastic gradient descent based algorithm is used to optimize the circuit to obtain the final optimized design. The stochastic gradient takes in as input the parameterized netlist, the design objective and the range of parameter values for the design. The output of the optimization algorithm is the design variable points that give the optimal performance objective. The optimization process is reiterated until the target specifications are met as seen in Fig. 7. Upon completion of the optimization process, the final parameter values are used to manually redesign the physical layout. In using the parasitic extracted netlist, the process ensures that the design flow is parasitic aware, and the final physical design is implemented to reflect more silicon accurate results. A detailed discussion of the SGD based algorithm are presented in Section IV-B.

B. Stochastic Gradient Descent

The stochastic gradient descent (SGD) algorithm is a variation of the descent based algorithms that utilize the gradient of functions to search for optimal values. The stochastic gradient descent is a cost function optimization algorithm that has been implemented for many different applications. SGD algorithms could be applied to optimization problems for a function $f(\mathbf{x})$, where $\mathbf{x}$ is the vector of parameters. An example optimization problem is presented as follows:

$$\text{Minimize} P_{TS}(\mathbf{w}), \text{where}(\mathbf{w}) = W_n, W_p, L_n, L_p, V_{th}... \quad (7)$$

The basic form of the SGD algorithm is given as [22]:

$$w_{n+1} = w_n - \gamma_n \nabla P_{TS}(w_n). \quad (8)$$

Where w_n, the parameter which minimizes the objective function, is to be estimated. $\nabla P_{TS}(w_n)$ is the gradient or derivative w.r.t. w_n, of the objective function $P_{TS}(w)$ to be optimized. γ is a user defined factor that controls the step size of the descent. It is also usually referred to as the learning rate. The choice of γ is arbitrary and is commonly set as $\frac{1}{n}$ or some other decaying function with respect to n, where n is the number of iteration steps. A very small γ will result in smaller steps and will increase the convergence time, while a bigger γ may lead to an unstable process.

The SGD is very similar to the gradient descent, the difference being that the gradient of the objective function $P_{TS}(\mathbf{w})$ is computed by an estimation, using a subset of the parameter vector which is randomly chosen in each iteration step. The estimation of the gradient in each iteration step greatly reduces the computation costs and simultaneously speeds up the optimization process. This characteristic makes the SGD very suitable for computational expensive simulations and functions which are not easily differentiable.

The SGD is susceptible at being stuck in a local minimum and is thus effective for local optimization. We propose a technique that reiteratively restarts the algorithm N times, while memorizing the local minima found and the range of parameters traversed. When the algorithm is restarted with a new random point, it checks to make sure it is a new point which has not been searched, thereby eliminating redundant searches. After the algorithm has been run N times, the optimized point is selected from the set of local minima.

The proposed algorithm is shown in Algorithm 1.

Algorithm 1 Stochastic Gradient Descent Based Algorithm

1: $N \leftarrow Max\_Iter$
2: Choose random variable w_0, w_0'
3: Calculate FoM $P_{TS}(w_0)$
4: **while** $||P_{TS}(w_{n+1}) - P_{TS}(w_n)|| > \epsilon$ **do**
5: Choose a decreasing γ_n (generally $\frac{1}{n}$)
6: Estimate $\nabla P_{TS}(w_n)$ using $P_{TS}(w_n')$
7: Compute $x_{n+1} = x_n - \gamma_n \nabla P_{TS}(x_n)$
8: **end while**
9: $W \leftarrow \{w_n, P_{TS}(w_n)\}$
10: Reset w_0, w_0'
11: **if** (w_0) within range of W **then**
12: Reset w_0, w_0'
13: **else**
14: $N \leftarrow N - 1$
15: restart search algorithm
16: **end if**
17: **repeat**
18: algorithm search
19: **until** N = equals 0
20: **return** the lowest couple $w_n, P_{TS}(w_n)$ found.

The algorithm shows the modifications to the traditional SGD in optimizing an objective output $P_{TS}(w)$ as a function of design parameters w. First, the maximum iteration number is set as N, then a random starting point is chosen to start the optimization process. For each iteration step in lines 4-8, a set of solutions is stored in vector W, also marking traversed paths. The algorithm is restarted, .i.e. reiterated through lines 4-16 until the max iteration is reached or some other stop criteria are met. When a new random point is to be picked, it checks to make sure that point has not been searched. At the end of the algorithm, the optimized design objective is chosen from the minimum of values in the vector w. In this algorithm, we improve the efficiency by monitoring the set of random points to limit the range of parameters picked to only those whose paths have not been traversed. This cuts down on the optimization algorithm time by eliminating redundant searches, i.e. searches that will produce already stored optima or discarded results.

V. Experimental Results

To demonstrate the efficiency of the proposed design optimization flow, it is applied to the optimization problem of the 45 nm thermal sensor design which was discussed in section III. Initial design parameters for the thermal sensor are as follows: V_{dd} = 1.5 V, and nominal values L of 45 nm and W_n, W_p of 120 nm and 240 nm, respectively, are used. The design temperature range was 0 – 100°C. After the schematic and physical baseline designs were completed to specifications, the netlist was extracted with parasitics and parameterized. Cadence Ocean scripts were written and the simulations were driven by MATLAB using the parasitic netlist extracted from the physical design.

The optimization goal for this experiment was to reduce and optimize the power consumption using the temperature resolution as an optimization constraint. The width of the transistors was used as the design parameter set to be explored. The SGD algorithm in Algorithm 1, was implemented in MATLAB and was used to reiteratively simulate through the design with update inputs of transistor widths. To reduce the possibility of the optimization process being stuck in a local minimum, the algorithm was run with N = 20, each time starting with a random value of transistor width. The results of the optimized designs compared to the baseline design are shown in Table III. The power consumption has been reduced by 38% with an optimal parameter point of W_n =

153 nm. The power consumption for this design is relatively higher because it includes the power consumption from the counter and the register. The designs in [14] have also been implemented in the subthreshold region which significantly reduces the power consumption. A 13.75% increase in the area of the final physical design is incurred. This is because an increase of 27.5% increase in W_n.

TABLE III
EXPERIMENTS OVER THE 45NM CMOS THERMAL SENSOR CIRCUIT.

Design	Average Power (P_{TS})	Sensitivity (T_{TS})	Area (μm^2)
Schematic	293.1 μW	16.88 MHz/°C	-
Layout	379.4 μW	9.42 MHz/°C	1221.37
Final	181.8 μW	44.2%	1389.31
% Change	37.97%	-	13.75%

VI. CONCLUSION

In this paper, a new design optimization flow has been presented. A stochastic gradient descent based optimization algorithm was incorporated into the design flow. The modified SGD algorithm is relatively fast and efficient and eliminates local optima convergence problems. The proposed algorithm was tested on a 45 nm thermal sensor design to optimize its power consumption. The power consumption was reduced by 38% while maintaining the resolution of the thermal sensor at 9.42 MHz/°C. This compares very well to selected optimizations of thermal sensor designs. In future research, the proposed methodology will be extended to multi-objective optimization schemes.

REFERENCES

[1] S. M. Srinivas Devadas, "A Survey of Optimization Techniques Targeting Low Power VLSI Circuits," in *Design Automation, 1995. DAC '95. 32nd Conference on*, 1995, pp. 242–247.

[2] J. Kao and A. Chandrakasan, "Dual-Threshold Voltage Techniques for Low-Power Digital Circuits," *Solid-State Circuits, IEEE Journal of*, vol. 35, no. 7, pp. 1009–1018, Jul 2000.

[3] M. Pedram and S. Nazarian, "Thermal Modeling, Analysis, and Management in VLSI Circuits: Principles and Methods," *Proceedings of the IEEE*, vol. 94, no. 8, pp. 1487–1501, Aug. 2006.

[4] T. Binder, C. Heitzinger, and S. Selberherr, "A Study on Global and Local Optimization Techniques for TCAD Analysis Tasks," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 23, no. 6, pp. 814–822, June 2004.

[5] V. Aggarwal, "Analog Circuit Optimization using Evolutionary Algorithms and Convex Optimization," Master's thesis, Massachusetts Institute of Technology, May 2007.

[6] Y. Li and J. Li, "Swarm Intelligence Optimization Algorithm Based on Orthogonal Optimization," in *Computer Modeling and Simulation, 2010. ICCMS '10. Second International Conference on*, vol. 4, Jan. 2010, pp. 12–16.

[7] O. Garitselov, S. Mohanty, E. Kougianos, and P. Patra, "Nano-CMOS Mixed-Signal Circuit Metamodeling Techniques: A Comparative Study," in *Proceedings of the International Symposium on Electronic System Design (ISED)*, 2010, pp. 191–196.

[8] M. Aguirre, J. Chavez, A. Torralba, and L. Franquelo, "Analog Design Optimization by Means of a Tabu Search Approach," in *Circuits and Systems, 1994. ISCAS '94., 1994 IEEE International Symposium on*, vol. 1, May–Jun 1994, pp. 375–378.

[9] S. Klein, M. Staring, and J. Pluim, "Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines," *Image Processing, IEEE Transactions on*, vol. 16, no. 12, pp. 2879–2890, Dec. 2007.

[10] P. A. N. Bosman, "On Gradients and Hybrid Evolutionary Algorithms for Real-Valued Multi-objective Optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 16, no. 1, pp. 51–69, Feb. 2012.

[11] M. Cohen, R. Edwards, G. Cauwenberghs, M. Vorontsov, and G. Carhart, "AdOpt: Analog VLSI Stochastic Optimization for Adaptive Optics," in *Neural Networks, 1999. IJCNN '99. International Joint Conference on*, vol. 4, 1999, pp. 2343–2346.

[12] M. Vorontsov, G. Carhart, M. Cohen, and G. Cauwenbergs, "Adaptive Optics Based on Analog Parallel Stochastic Optimization Technique," in *Lasers and Electro-Optics, 2000. (CLEO 2000). Conference on*, 2000, p. 605.

[13] S. Eberhardt, R. Tawel, T. Brown, T. Daud, and A. Thakoor, "Analog VLSI Neural Networks: Implementation Issues and Examples in Optimization and Supervised Learning," *Industrial Electronics, IEEE Transactions on*, vol. 39, no. 6, pp. 552–564, Dec 1992.

[14] S. Park, C. Min, and S.-H. Cho, "A 95nW Ring Oscillator-based Temperature Sensor for RFID Tags in 0.13 μm CMOS," in *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*, May 2009, pp. 1153–1156.

[15] B. Datta and W. Burleson, "Low-Power and Robust On-Chip Thermal Sensing Using Differential Ring Oscillators," in *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on*, Aug. 2007, pp. 29–32.

[16] Y. Zhang and A. Srivastava, "Accurate Temperature Estimation Using Noisy Thermal Sensors," in *Design Automation Conference, 2009. DAC '09. 46th ACM/IEEE*, July 2009, pp. 472–477.

[17] Z. Shenghua and W. Nanjian, "A Novel Ultra Low Power Temperature Sensor for UHF RFID Tag Chip," in *Solid-State Circuits Conference, 2007. ASSCC '07. IEEE Asian*, Nov. 2007, pp. 464 –467.

[18] C. Christoffersen, G. Toombs, and A. Manzak, "An Ultra-Low Power CMOS PTAT Current Source," in *Argentine School of Micro-Nanoelectronics Technology and Applications (EAMTA), 2010*, Oct. 2010, pp. 35–40.

[19] K. Ueno, T. Hirose, T. Asai, and Y. Amemiya, "Ultralow-Power Smart Temperature Sensor with Subthreshold CMOS Circuits," in *Intelligent Signal Processing and Communications, 2006. ISPACS '06. International Symposium on*, Dec. 2006, pp. 546–549.

[20] T. Meng and C. Xu, "A Cross-Coupled-Structure-Based Temperature Sensor with Reduced Process Variation Sensitivity." *Semiconductors, Journal of*, vol. 30, no. 4, pp. 1642– 1648, Apr. 2009.

[21] P. Chen, C.-C. Chen, C.-C. Tsai, and W.-F. Lu, "A Time-to-Digital-Converter-Based CMOS Smart Temperature Sensor," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 8, pp. 1642– 1648, Aug. 2005.

[22] C. Besse, "Why Natural Gradient for General Optimization?" Departement Informatique, Universite Laval Sainte-Foy (Quebec), Canada, Tutorial, Sept. 2009.