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

O. Okobiah[1], S. P. Mohanty[2], E. Kougianos[3], Oleg Garitselov[4], and Geng Zheng[5]

NanoSystem Design Laboratory (NSDL, http://nsdl.cse.unt.edu)
University of North Texas, Denton, TX 76203, USA.[1,2,3,4,5]

Email: oo0032@unt.edu[1], saraju.mohanty@unt.edu[2], eliask@unt.edu[3], omg0006@unt.edu[4], and gengzheng@my.unt.edu[5]

Presented By
Oghenekarho Okobiah
University of North Texas
Email: oo0032@unt.edu

Acknowledgment: This research is supported in part by NSF awards CNS-0854182 and DUE-0942629 and SRC award P10883.

Outline of the talk

- Background and Motivation
- Related Prior Research
- Thermal Sensors
- Proposed Design Optimization Flow Methodology
- Experimental Results
- Conclusions and Future Research Directions

Issues in NanoCMOS Design

- Expensive Computer Simulations
- Pronounced effects of process variations in deep nanometer regions
  - Increase in design parameters
  - Current modeling techniques unable to capture effects of process variation

Background and Motivation

- Complex and High density designs
- Design for low power consumption
- Reliability issues
  - Thermal monitoring

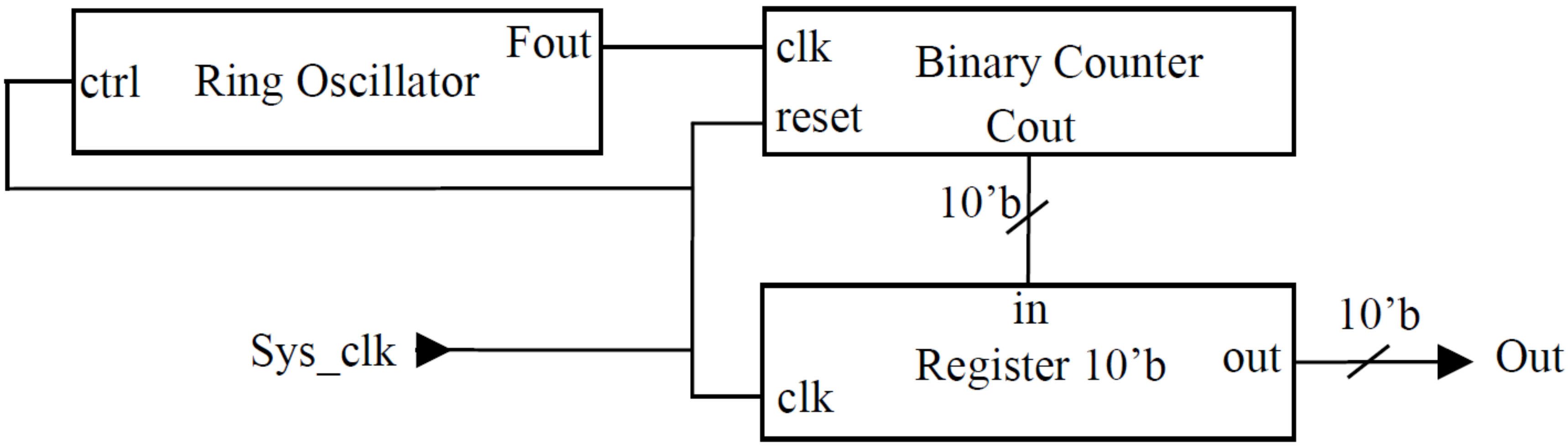
Novel Contributions

- Design flow methodology incorporating SGD for nano-CMOS design optimization
- Modification of SGD algorithm for global optimization.

Prior Related Research

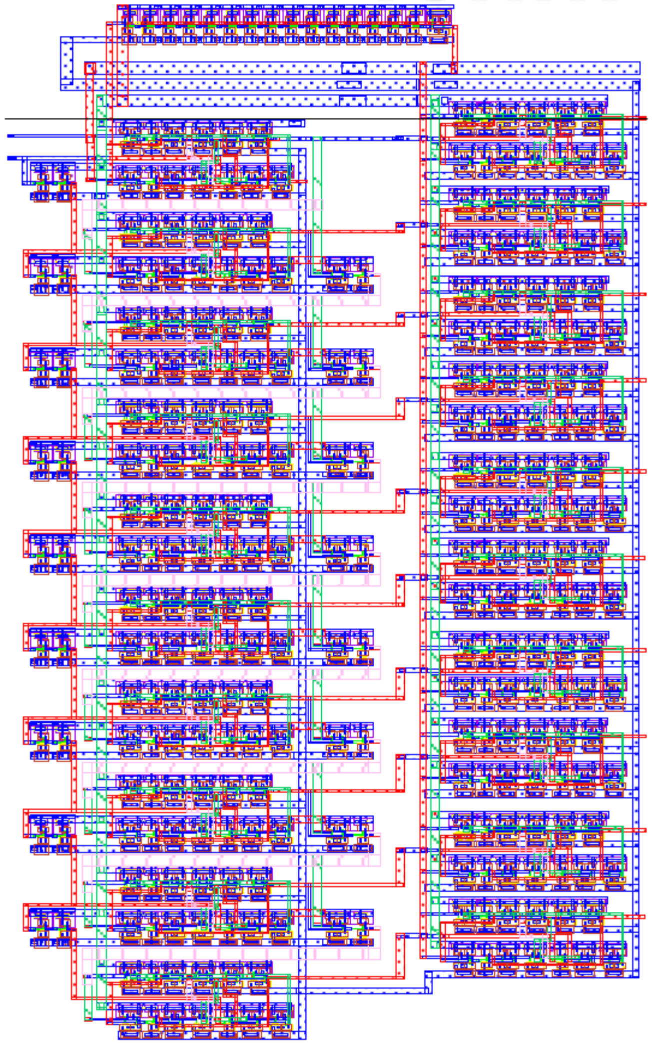
- Exploration of optimization algorithms for NanoCMOS designs
  - Simulated annealing, swarm intelligence
- Efficient designs for thermal sensors

45nm Thermal Sensor

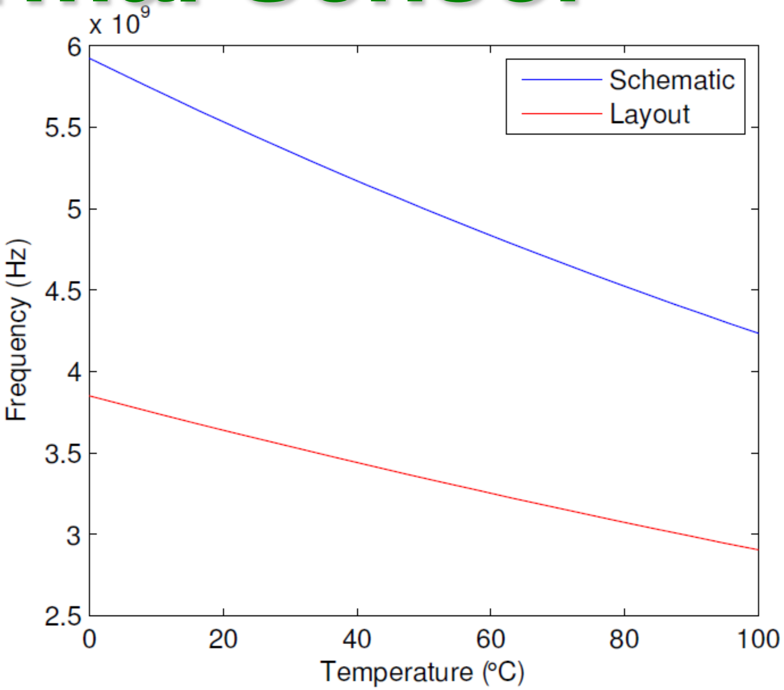


- Ring Oscillators for sensing
- Binary Counter
- Register

45nm Thermal Sensor



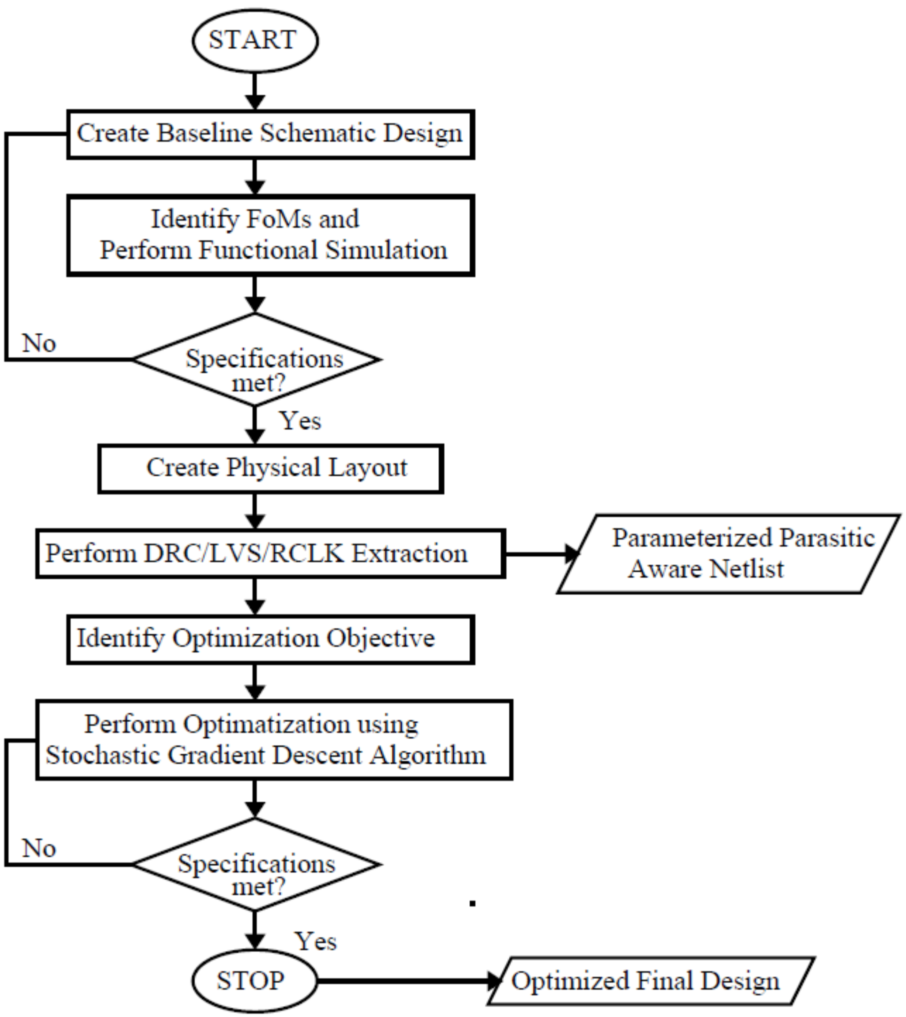
45nm thermal sensor physical design



Frequency response vs temperature for 45nm thermal sensor

Design	Avg Pwr	Sensitivity	Area
Schematic	293.1µW	16.88MHz/°C	-
Layout	379.4µW	9.42MHz/°C	1221.37µm^2
% Change	+29%	-44%	

Proposed Design Flow Diagram



Design Optimization Flow

Stochastic Gradient Descent ...

- Utilizes gradient functions to search for optimal values
- The general form is given by

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

where w_n is vector set of design parameter, P_{TS} is the objective function and γ is an arbitrary factor for controlling the stepping size

Stochastic Gradient Descent ...

- Following the gradient descent approach
  - SGD reiteratively steps through the gradient descent until it converges
  - At each step, a subset of parameter vector is used to estimate the gradient.
  - For each step, the subset is randomly chosen – also referred as training set
- Modified SGD restarts at random points to mitigate local optimum issues

Stochastic Gradient Descent

Algorithm 1 Stochastic Gradient Descent Based Algorithm

$N \leftarrow Max\_Iter$
Choose random variable w_0, w'_0
Calculate FoM $P_{TS}(w_0)$
while $||P_{TS}(w_{n+1}) - P_{TS}(w_n)|| > \epsilon$ **do**
    Choose a decreasing γ_n (generally $\frac{1}{n}$)
    Estimate ∇ $P_{TS}(w_n)$ using $P_{TS}(w'_n)$
    Compute $x_{n+1} = x_n - \gamma_n \nabla P_{TS}(x_n)$
end while
$W \leftarrow \{w_n, P_{TS}(w_n)\}$
Reset w_0, w'_0
if (w_0) within range of W **then**
    Reset w_0, w'_0
else
    $N \leftarrow N - 1$
    restart search algorithm
end if
repeat
    algorithm search
until N = equals 0
return the lowest couple $w_n, P_{TS}(w_n)$ found.

Optimal Design Results

Design	Average Power	Sensitivity	Area
Schematic	293.1µW	16.88MHz/°C	-
Layout	379.4µW	9.42MHz/°C	1221.37µm^2
Final	181.8µW	9.42MHz/°C	1389.31µm^2
% Change	37.97%	-	13.75%

Conclusions

- A novel design flow methodology incorporating Stochastic Gradient design based optimization algorithm was presented.
- Average power consumption was reduced by 38%.
- Current techniques will be extended to multi objective optimization.

Thank you !!!