

UC Santa Cruz

UC Santa Cruz Previously Published Works

Title

Energy Savings and Performance Improvement in Subthreshold Using Adaptive Body Bias

Permalink

<https://escholarship.org/uc/item/1dq8892w>

ISBN

978-1-4503-4972-7

Authors

Sankaranarayanan, Rajsaktish
Guthaus, Matthew R

Publication Date

2017-05-10

DOI

10.1145/3060403.3060421

Peer reviewed

Energy Savings and Performance Improvement in Subthreshold Using Adaptive Body Bias

Rajsaktish Sankaranarayanan
University of California Santa Cruz
rsankara@ucsc.edu

Matthew R. Guthaus
University of California Santa Cruz
mrg@ucsc.edu

ABSTRACT

In subthreshold operation, circuits are more sensitive to the impact of parametric variation due to reduced supply voltages. To meet timing specification and ensure reliable operation, circuits require compensation techniques that mitigate variation. We developed a design methodology to use adaptive forward body bias and reduce worst-case 3σ active energy, delay and standby energy caused by threshold voltage variation. We validated this methodology on the ISCAS85 benchmarks and improved the worst-case metrics in each case, with no loss of performance. Our approach reduces worst-case standby energy and worst-case active energy by up to 21.06% and 18.80%, respectively, on average.

1. INTRODUCTION

Designing circuits for subthreshold operation is challenging as the impact of process variation is more significant at reduced voltages due to the exponential dependence of drain current on gate voltage. Variation in analog circuits like current mirrors and voltage reference circuits is minimized by using large geometry devices. In subthreshold digital design, however, minimum sized devices are optimal for energy efficiency [3]. For a low target frequency, body biasing was more energy efficient than supply voltage scaling in mitigating process and temperature variation [9]. Reverse bias was found to worsen drain current mismatch and forward bias reduced the mismatch [5]. Gate-level clustering with cluster specific body bias was found to improve leakage power [7]. However, these techniques do not adapt to variation, which could limit their effectiveness. Threshold voltage variation could affect transistor delays on critical paths, which in turn affects active energy, circuit performance and thus overall energy efficiency. Critical path replication [4] or approaches like block-based and path-based statistical analyses [10] have been used to measure this. To accurately determine the impact of variation, the exact critical path and its delay need to be estimated.

In this work, we reduce standby energy caused by threshold voltage variation. We developed a design methodology to use forward bias with a body bias regulator [1]. We use a previous regulator, but, develop a digital design methodology to mitigate the impact of variation. This paper presents the methodology of: finding the optimal number of regula-

tors, determining their placement and assigning cells to be biased by these regulators.

The rest of this paper is organized as follows: We present an overview of proposed approach along with a motivational example in Section 2. In Section 3, we describe details of our design methodology and LP formulation. We discuss experimental methods including variation model, static timing analysis and energy measurement in Section 4. We present results in Section 5 and conclusions in Section 6.

2. METHODOLOGY OVERVIEW

We use the regulator circuit shown in Figure 1 which has an NMOS and PMOS transistor in diode-connected mode that are connected at the drains with their wells shorted. This improves device matching in a circuit (an inverter in this case) by producing an output voltage at which the leakage currents match. Even small variations in the doping profile of FETs can lead to exponentially mismatched drain currents. By shorting the wells, a pathway for the minority carriers exists between wells.

2.1 Variation-aware Adaptive Regulation

Consider an inverter with equal rise and fall times, biased by an equal sized regulator as shown in Figure 1. The regulator is designed to output a voltage of $V_{dd}/2$ since the two FETs in cutoff act like a voltage divider. In the presence of threshold voltage variation, the actual regulator output voltage will be higher or lower than $V_{dd}/2$ depending on the specific FET threshold voltages.

We assume the threshold voltages of NFETs (and PFETs) of both circuits to be normally distributed around the nominal value. According to Pelgrom's model the variance of the parameter mismatch between transistors is directly proportional to the distance between them and inversely proportional to the area of the devices [8]. Given the distance between two transistors, variance of threshold voltage mismatch can be computed using Pelgrom's model. Using variance of mismatch, variance of threshold voltage distribution and uncorrelated random samples from this distribution, we determine the correlation between random variables.

In Figure 2, we compare the worst-case 3σ standby energy of an inverter biased by the regulator, and an inverter biased by a voltage source at $V_{dd}/2$. Well-biases and offsets between wells can have many possible combinations to meet target performance and power. Instead, a simple way to mitigate

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](http://permissions.acm.org).

GLSVLSI '17, May 10-12, 2017, Banff, AB, Canada

© 2017 ACM. ISBN 978-1-4503-4972-7/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3060403.3060421>

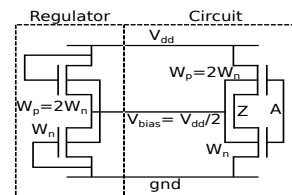


Figure 1: An inverter biased by the bias regulator circuit when the transistors are perfectly matched.

impact of variation is biasing both wells by the same amount. For this work, we use $V_{dd}/2$ because, the substrate current is much smaller than drain current and thus scalable to deep subthreshold voltages. In the case of the inverter biased by the regulator, we see the inverter worst-case standby energy decreases over a range of correlation values which changes over distance between inverter and regulator.

3. ON-CHIP REGULATOR METHODOLOGY

Standard cell design flows use a set of pre-designed library of cells. From Figure 2, we see savings in worst-case 3σ energy when the regulator is nearby and thus correlates better with the cell. The regulator circuit consumes standby energy and incurs area. Our methodology makes a trade-off in the number of regulators and the performance of the circuit. We target filler cell locations for regulator insertion.

3.1 Regulator Design

We use the Nangate Open Cell Library implemented using 45nm FreePDK process technology. The library cells were redesigned for equal rise and fall times in the worst case switching condition. The impact of variation can vary between a cell and its associated regulator, which in turn could affect the leakage current match and hence energy savings. For improved matching, the transistor dimensions of the regulator need be similar to that of the library cells. So, we custom designed a set of regulators based on the functionality and transistor dimensions of the library cells.

3.2 Cell Characterization

The worst-case standby energy of the inverter follows a linear trend over distance between inverter and regulator. So, to use this methodology in a standard cell based flow targeting worst-case energy savings, we characterized all cells in Nangate Open Cell Library for power and performance using cell specific regulators and with the variation model that was used to characterize the inverter circuit described earlier. Using the characterization data, we fit linear models for the worst-case standby energy savings of each cell in the library, as a function of distance.

3.3 Design Implementation

We developed a CAD flow with industry standard tools for synthesis and placement of standard cells and a separate regulator placement and assignment tool. The regulator assignment is formulated as a Linear Program (LP), which is described in the next section. In Figure 3, we show the CAD flow of our proposed design and verification methodology. Starting with Verilog description of the benchmarks, we synthesized them using Synopsys Design Compiler and Nangate Open Cell Library to obtain gate level netlists. These netlists were placed using Synopsys IC Compiler. The placement contains filler cells in the unused areas. Our LP optimizer reads in the placed netlists containing geometric coordinates of the cell and filler instances. Because the regulators consume standby power, we do not want to replace all of the filler cells, but they are candidate locations for

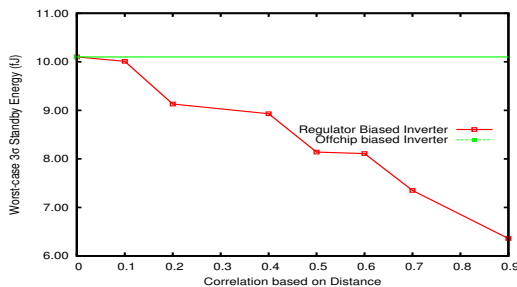


Figure 2: On-chip bias regulators improve worst-case 3σ standby energy better when they are nearby, and more correlated, with the circuit they bias.

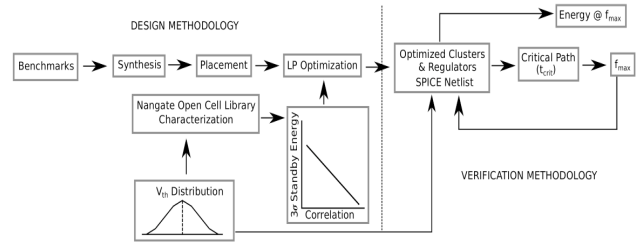


Figure 3: Block diagram showing on-chip bias regulator design and verification methodology.

regulator instances. Our optimizer solves for the optimal clustering of cells biased by regulators, using LP_Solve [2].

3.4 Regulator Placement and Assignment

Insertion of each regulator incurs a standby energy cost. To achieve worst-case energy savings at the circuit level we determine the optimal number of regulators, the placement of these regulators and the cell(s) assigned to a regulator.

We model the problem of assigning regulators to one or more groups of standard cells as an LP. We demonstrate two algorithms to solve this LP problem. The first is an exhaustive approach yielding the optimal solution, while the second is a faster, heuristic solution. The two approaches differ in the number of LP constraints formulated and how they generate the constraints.

In the optimal solution, we consider all filler cells as candidate locations for regulators, each of which can be paired with any cell on chip.

In the heuristic approach, we derive candidate subsets of filler locations from the set of all filler locations, for regulator insertion. We then proceed to determine the regulators to make clusters with cell instances. For these subsets, we consider only rows immediately adjacent to a cell, for each cell in the design. Using this row adjacency constraint, we build LP constraints and attempt to solve the LP. If the LP does not converge, we increase the set of rows from which we derive the subset of filler cells, for each cell in the design. Again we attempt to solve the LP. The goal of the heuristic is to determine the smallest number of rows from which the subsets can be derived leading to a solvable LP. The goal of the LP formulation is to identify a subset of regulators from the available regulators and determine the grouping of standard cells to be assigned to those selected regulators.

The LP formulations of both algorithms have identical cost functions and differ only in the number of constraints. So, for brevity we present the equations of the optimal formulation and indicate the areas where they differ.

Both algorithms have similar inputs, namely, placed file containing fillers and cells and the cell-specific characterization data. The output from both algorithms is the gate clustering and regulator assignment. For brevity, we present the heuristic and indicate where it differs from the optimal solution. Lines 2,5 and 6 constitute the iterative part of the heuristic solution only and are not applied in the optimal solution. All other aspects are common to both algorithms.

To physically implement our methodology we leverage features available in place and route tools. We first create relative placement groups for each cell cluster connected to a regulator and associate a power domain with that group. We then designate the output net of each regulator as a supply net and connect those nets to cell bias pins.

- Let m be the number of cell instances in the design.
- Let i be an index variable such that $i = 0, 1, \dots, (m-1)$.
- Let C represent the set of all cells indicated by C_i where $i = 0, 1, \dots, (m-1)$.
- Let n represent the number of regulators in the design.
- Let j be an index variable such that $j = 0, 1, \dots, (n-1)$.

Algorithm 1: Heuristic algorithm finds clustering of cells and assigns bias regulators

Define : $RSS =$ Row Search Space,
 $RSS_{min} = 1$, $RSS_{curr} = RSS_{min}$,
 $RSS_{max} =$ Number of rows in placed chip

Input : Placed File, Cell characterizing information

Output: Gate clustering, Regulator assignment

```

1 while  $RSS_{curr} \leq RSS_{max}$  do
2   write ILP constraints using  $RSS_{curr}$ ;
3   solve_ILP();
4    $RSS_{curr}++$ 
5 end
```

- Let R represent the set of regulators indicated by R_j where $j = 0, 1, \dots, (n-1)$.
- Let a be a constant representing the standby energy cost of a single regulator.
- Let X_{ij} represent a decision variable taking values $\{0, 1\}$ indicating whether a specific cell C_i is assigned to a specific regulator R_j .
- Let Y_i represent an auxiliary variable denoting the cost of a specific cell being assigned a specific regulator from the available regulators.
- Let e_{ij} represent the energy savings coefficient for a given cell C_i regulator R_i assignment. This is obtained from the linear energy savings model discussed earlier, with distance of a fixed regulator to a given cell as input to this model.

The LP solver solves for decision variables X_{ij} which determines the optimal grouping of cells biased by regulators. Our formulation is described below:

$$\text{Minimize } \sum_{i=1}^m Y_i + m \cdot a \quad \text{such that} \quad (1)$$

$$\sum_{j=1}^n [e_{(i-n)+j} \cdot X_{(i-n)+j}] \leq Y_{(i+1)} \quad \forall i = 0, 1, 2, \dots, (m-1) \quad (2)$$

$$\sum_{j=1}^n X_{(i-n)+j} = 1 \quad \forall i = 0, 1, 2, \dots, (m-1) \quad (3)$$

$$X_p = \{0, 1\} \quad \forall p = 1, 2, \dots, (m \cdot n) \quad (4)$$

The cost function in Eq.1 describes the goal of this formulation, which is to minimize the number of clusters into which all standard cells can be grouped, such that each cluster is connected to a regulator. Here Y_i denotes the linearized worst-case 3σ standby energy cost of a cell i from amongst the set of all cell instances C_i , when biased by a regulator. Summed over the set of all cell instances, our goal is to minimize the cost of clustering all the cell instances in the design. The second term indicates the standby energy cost of the regulators needed to cluster all the cell instances.

The set of constraints denoted by Eq.2 determine which of the cells get clustered together to be biased by a common regulator and get assigned a regulator. The coefficient term is the energy cost of driving a particular cell with a particular regulator. This term is obtained by precharacterizing the cells of the standard cell library and the distance between the regulator and the cell.

The LHS of constraint Eq.2 represents the energy cost of biasing a given standard cell by each of the available regulators. The RHS of constraint Eq.2 ensures this cell is biased by one of the available regulators only.

The constraint Eq.3 ensures each cell has to be driven by a regulator and enables grouping of cells to a common regulator. The constraint Eq.4 indicates X_{ij} is a binary decision variable taking values $\{0, 1\}$.

The above described constraints apply to the Algorithm-1. For Algorithm-2, the indices of the variables and the limits of the summation are not constant n and instead take variable values based on the number of regulator candidates available in the rows adjacent to each cell.

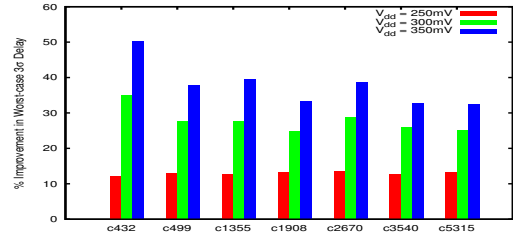


Figure 4: On-chip regulator methodology improves worst-case 3σ delay compared to an unbiased circuit.

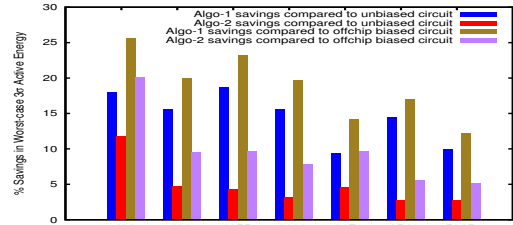


Figure 5: On-chip regulator methodology improves worst-case 3σ active energy using $V_{dd}=350mV$

4. EXPERIMENTAL METHODS

We use Nangate Open Cell Library implemented using the 45nm FreePDK process as our standard cell library. We designed the bias regulators using this process. The nominal threshold voltages of the transistor models in this process were $V_{th0N} = 0.4106$ V and $|V_{th0P}| = 0.3842$ V.

4.1 Variation Model

We consider parametric variation of threshold voltage caused by Random Dopant Fluctuation (RDF). Our input variation model is based on Pelgrom's model given by,

$$\sigma^2(V_{th}) = (S_{V_{th}}^2 \cdot Distance^2) + \left(\frac{A_{V_{th}}^2}{W \cdot L}\right). \quad (5)$$

$S_{V_{th}}^2$ and $A_{V_{th}}^2$ are technology specific constants in the range of 0.01 mV/ μm and 0.001 mV respectively. We assume the threshold voltage is normally distributed around the nominal value, with 1σ variance of 20 mV [11]. All instances of standard cells and regulators are subject to threshold voltage variation from the distribution. Given spatial separation between a cell instance and regulator instance and the variance of mismatch between them, we proceed to find the respective threshold voltages of the FETs. For a given mismatch variance, which is a function of the spatial separation, and the threshold voltage distribution, the correlation between the random variables is computed. Two randomly picked threshold voltage values are transformed into correlated values using the computed correlation coefficient, by applying Cholesky decomposition [6].

4.2 Optimized Circuit

Our optimizer reads in placed netlists containing coordinates of cell and filler instances. Considering filler cells as candidate locations for regulator insertion, our optimizer solves for the clustering of cells biased by regulators and the locations of regulators, using LP_Solve [2]. Using the results of optimization and placed netlists, we obtain SPICE netlists containing clusters of cells biased by regulators. We apply our variation model on these netlists.

4.3 Timing and Energy Measurement

Circuit timing paths are affected when threshold voltage variation causes delay variation in transistors. This affects circuit performance and energy consumption. So, we find the exact critical path delay by performing transistor level static timing analysis using Synopsys Nanotime and find the maximum operating frequency f_{max} of the circuit. Using

Table 1: Savings in worst-case 3σ standby energy using on-chip regulator assignment at $V_{dd}=350mV$

Circuit	Unbiased vs On-chip biased					Offchip biased vs On-chip biased				
	Unbiased (fJ)	Optimal (fJ)	Savings (%)	Heuristic (fJ)	Savings (%)	Offchip bias (fJ)	Optimal (fJ)	Savings (%)	Heuristic (fJ)	Savings (%)
c432	0.92	0.60	41.95	0.87	5.70	1.04	0.60	42.30	0.87	17.12
c499	1.15	0.98	1.64	1.24	-8.13	1.32	0.98	25.75	1.24	6.41
c1355	1.16	0.90	3.30	1.26	-8.61	1.29	0.90	30.23	1.26	2.69
c1908	1.67	1.35	12.52	1.65	1.41	1.74	1.35	22.41	1.65	5.50
c2670	2.00	1.96	4.07	1.98	1.15	2.04	1.96	3.92	1.98	3.06
c3540	1.94	1.80	5.21	1.95	-0.82	2.04	1.80	11.76	1.95	4.03
c5315	2.34	2.16	3.15	2.36	-0.57	2.43	2.16	11.11	2.36	2.97
Avg.			10.26		-1.41	-		21.06		5.96

critical path delay, we then simulate the circuit at f_{max} using random input stimuli and measure active and standby energy per cycle using Synopsys Hspice. We perform this process of applying variation, determining critical path, finding f_{max} and using it to compute energy using 1000 iterations of Monte Carlo simulation. We then compare the benchmark performance with an unbiased circuit and a circuit biased using a voltage source at $V_{dd}/2$, described earlier in Section 2, comparable to an off-chip bias source. Since the regulators are connected to on-chip supply voltage, our results include the energy overhead of the regulators.

5. RESULTS

In this section, we present the results of our optimization process and the impact of optimization driven body bias on energy and performance of ISCAS85 benchmarks.

The run time of our optimization process was measured for several benchmarks using the optimal and heuristic approaches. Across the benchmarks, the improvement in runtime of the heuristic solution, was 68% on average. This is due to the reduced subset of candidate regulators considered for each cell by our optimizer.

We found the active energy, standby energy and delay following log-normal distributions. So, we evaluate the distributions using typical parameters namely mean (μ) and worst-case 3σ . The regulator bias method provides improvement by delivering a forward bias which is adapted to the cell variation in that spatial vicinity. This locally relevant forward bias offers a better matching of the FET off-currents and lowers the cell threshold voltage, thus improving its performance. This increased performance enables scaling the circuit to lowered supply voltages, thus saving worst-case active energy and standby energy.

Since each regulator could output a bias voltage different from another regulator, we consider inter-well spacing for wells at different potential to measure the impact on circuit area. This area overhead spans a range of 17.5% at the minimum to 24.5% at the maximum across all benchmarks with an average of 19.4%.

From Figure 4, we see the regulator method offers improvement in worst-case delay of the circuit for all benchmarks compared to an unbiased circuit. The improvements decrease as the supply voltage is scaled down. This is because, at lower operating voltages, the impact of variation is much higher than the applied bias compensation. The improvement in delay spans a range of 10.89% at the minimum to 50.16% at the maximum. Considering all benchmarks, the average improvement in worst-case delay is 36.93%, 27.85% and 12.74% at 350mV, 300mV and 250mV respectively using Algorithm-2, compared with an unbiased circuit.

From Figure 5, we see the regulator method offers savings in worst-case active energy of the circuit for all benchmarks when compared against an unbiased circuit and an offchip biased circuit. We have verified this using both our algorithms. The active energy savings on average compared to an unbiased circuit across all benchmarks include 14.52% and 4.50% for Algorithm-1 and Algorithm-2 respectively. Compared to an offchip biased circuit, the savings are 18.84% and 9.20% using Algorithm-1 and Algorithm-2 respectively. We can see Algorithm-1 slightly outperforms

Algorithm-2 in both cases of comparison, namely against an unbiased circuit and offchip biased circuit. This is because, the optimal solution algorithm offers a wider choice of regulators to choose from, resulting in improved savings.

In Table 1 we list the savings in worst-case 3σ standby energy using our methodology. We compare this against unbiased circuits and with offchip biased circuits using both algorithms. Algorithm-1 offers savings across all benchmarks, while Algorithm-2 offers savings in some cases. In other cases there is a slight increase in the worst-case standby energy. This corresponds to the cases where the active energy savings are also at the lowest, due to the reduced

6. CONCLUSION

In this work we developed a design methodology to use forward bias and reduced the impact of threshold voltage variation. We formulated optimal assignment of forward bias as a LP optimization in this methodology. We reduce worst-case standby energy and active energy by up to 21.06% and 18.80% on average respectively and reduce worst-case delay by up to 37.8% on average. The improvements are available over a range of deep subthreshold voltages.

7. REFERENCES

- [1] Andres Bryant et al. “Low Power CMOS at $V_{dd} = 4kT/q$ ”. In *Device Research Conference Proceedings*, 2001.
- [2] M. Berkelaar, K. Eikland, and P. Notebaert. *lp_solve 5.5. Open source (Mixed-Integer) Linear Programming System, GNU LGPL*, 2004.
- [3] B. H. Calhoun, A. Wang, and A. P. Chandrakasan. “Device Sizing for Minimum Energy Operation in Subthreshold Circuits”. In *CICC Proceedings*, 2004.
- [4] I. J. Chang, S. P. Park, and K. Roy. “Exploring Asynchronous Design Techniques for Process-Tolerant and Energy-Efficient Subthreshold Operation”. *IEEE JSSC*, 45(2), February 2010.
- [5] M. J. Chen, J. S. Ho, and T. H. Huang. “Dependence of Current Match on Back-Gate Bias in Weakly Inverted MOS Transistors and Its Modeling”. *IEEE JSSC*, 31(2), February 1996.
- [6] P. E. Gill, W. Murray, and M. H. Wright. *Numerical Linear Algebra and Optimization*, volume 1. Addison-Wesley Publishing Company, 1991.
- [7] S. Kulkarni, D. Sylvester, and D. Blaauw. A Statistical Framework for Post-Silicon Tuning through Body Bias Clustering. In *ICCAD Proceedings*, 2006.
- [8] M. J. M. Pelgrom, A. C. Duinmaijer, and A. P. Welbers. “Matching Properties of MOS Transistors”. *IEEE JSSC*, 24(5), October 1989.
- [9] Scott Hanson et al. “Exploring Variability and Performance in a Sub-200mV Processor”. *IEEE JSSC*, 43(4), April 2008.
- [10] A. Srivastava, D. Sylvester, and D. Blaauw. *Statistical Analysis and Optimization for VLSI: Timing and Power*. Springer, 2005.
- [11] Yun Ye et al. “Statistical Modeling and Simulation of Threshold Variation Under Random Dopant Fluctuations and Line-Edge Roughness”. *IEEE TVLSI*, 19(6), June 2011.