

Compensation for Within-Die Variations in Dynamic Logic by using Body-Bias

Navid Azizi and Farid N. Najm
Department of Electrical & Computer Engineering
University of Toronto, Toronto, Ontario, Canada
{nazizi,najm}@eecg.utoronto.ca

Abstract—We propose a fine-grained scheme to compensate for within-die variations in dynamic logic to reduce the variation in leakage, delay and noise margin through body-biasing. We first show that the amount of body-bias compensation needed depends on the correlation that exists between gates, and then analytically show the possible reduction in the variance of the leakage of both a single and multiple dynamic logic gates. We then design a circuit to implement the system which provides the reduction in the variance of the leakage, delay and noise margin of dynamic logic gates and show that it produces a close match to the analytical results. In our design, the variance of a typical test circuit is reduced by 27% and the variance of the path delay is reduced by 33%.

I. INTRODUCTION

CMOS scaling has been driven by the desire for higher transistor densities and faster devices. Along with the continued CMOS scaling, down into the nanometer regime, has come increased process variations of circuit parameters such as the transistor channel length and transistor threshold voltage [1]. The increased process variations can have a significant effect on circuit performance and power [2].

Historically, in order to cope with intrinsic variability, Integrated Circuit (IC) designers have implemented circuits with the worst-case process variations in mind [3]. However, designing at the worst-case process corner leads to excessive guard-banding, and thus more recent techniques have implemented adaptive circuit techniques by implementing control circuits on-chip that monitor the process variations within circuit devices, and change the characteristic of the devices [4], [5], [6]. These techniques, however, are usually implemented at the chip-level, or block-level.

One specific type of circuit topology that is extremely sensitive to process variations is dynamic logic [7]. Dynamic logic is usually used in high-performance parts of microprocessors and other VLSI circuits [8]. Furthermore in applications such as decoders and register-files, arrays of the same dynamic gate are usually used in a wide-OR configuration.; the wide-OR configuration is further sensitive to variations [8].

In the past, a weak keeper, which did not impact performance significantly, was sufficient to maintain the dynamic node voltage [8]. However with the exponential increase in leakage currents keepers must be made larger to offset for the worst-case leakage through the pull-down network, thus, reducing the performance advantage of dynamic gates over other circuit topologies. Also for dynamic logic, which can

be sensitive to highly local variations, chip or block level techniques cannot provide the required compensation to adjust the leakage and performance.

We propose a fine-grained adaptive circuit technique that trades off noise-margin/leakage for performance post-fabrication to reduce variability. The control scheme is referred to as fine-grained because it is done *locally* in a small neighborhood on the die, and because it is done using a continuous analog signal rather than a discrete digital signal. By reducing the variability, the keeper can be down-sized leading to increased performance.

The rest of this paper is organized as follows: Section II provides an overview of the dynamic compensation scheme. Section III then provides a framework for finding the optimal amount of compensation. The design of circuits to provide the compensation is presented in Section IV. Section V provides the results for the compensation scheme, and finally Section VI concludes.

A. Technology

All simulation results reported in this paper are based on HSPICE, using Berkeley Predictive Technology Models (BPTM) [9] for a $70nm$ technology. The transistor models were extended to include gate tunneling leakage which was modeled using a combination of four voltage-controlled current-sources, as in [10]. All simulations presented were performed on a four-input dynamic NOR. The simulations were performed at $110^{\circ}C$ where leakage, delay, and noise margin are all more critical than at low temperatures.

II. OVERVIEW

To compensate for variations in the leakage, performance, and noise margin of dynamic logic gates, we will use both forward and reverse body biasing to change the characteristics of the pull-down transistors in response to the underlying variations.

Body biasing, via controlled changes to V_{bs} , can compensate for variations by changing the threshold voltage, V_{tn} , of the pull-down transistors [4]. We obtain V_{bs} from a monitoring circuit which *measures* the process variations and produces a body bias that provides the compensation. The monitoring circuit is designed and layed out to look like an actual functioning dynamic gate, thus allowing systematic Within-Die

(WID) variations within the monitor circuit to be correlated to those within the functioning gate.

The monitor circuit produces a change in V_{bs} based on the actual variations in that chip; we refer to the functional dependence of V_{bs} on the variations as the *transfer function*. In order to determine this transfer function, the effect of V_{tn} , V_{tp} , and V_{bs} on the leakage, delay, and noise margin of a dynamic gate was determined through simulation. Leakage has an exponential dependence on V_{tn} and V_{bs} but has very little dependence on V_{tp} (the keeper is ON when the gate is not switching). The delay and noise margin of the dynamic gate both nearly have a linear dependence on V_{tn} , V_{tp} and V_{bs} . V_{tn} has a stronger effect on both the delay and the noise margin compared to V_{tp} .

III. FRAMEWORK FOR OPTIMAL COMPENSATION

In this section we will provide the mathematical framework for determining the optimal compensation for reducing the variation in leakage. To simplify the analysis, the effect of V_{tp} will initially be ignored.

Variations normally have three components: a die-to-die component, a within-die systematic component, and a within-die random component. The term “systematic” refers to the parts of the variations which have some correlation across the die, while the “random” component refers to the parts of the variations that are totally independent. Compensation for die-to-die variations has been discussed in the literature. In this work, we focus on the within-die systematic variations. Compensation for random variations remains a topic of future work.

Threshold voltage variations have a within-die random component arising from random dopant fluctuations, and a within-die systematic component arising from systematic variations in length [11] (and, of course, from any systematic deliberately applied variations in body voltage that are introduced by our monitoring circuits). Thus, focusing on the systematic component, leakage has an exponential dependence on both V_{tn} and V_{bs} (the dependency to V_{tn} is much stronger than the dependency to V_{bs}) and it can be written as $I = I_{nom}e^{b\Delta V_{tn0}+a\Delta V_{bs}}$, where ΔV_{tn0} is the variation of the threshold voltage of the gate of interest and b and a are constants obtained through simulation; they are sensitivity coefficients. This last equation can also be written as $\Delta \ln I = \ln(I/I_{nom}) = b\Delta V_{tn0} + a\Delta V_{bs}$. Let ΔV_{tn1} be the threshold voltage variations in the MOSFETS of the monitoring circuit itself. If ΔV_{tn0} and ΔV_{tn1} are totally correlated (i.e. $\Delta V_{tn0} = \Delta V_{tn1}$), then it is clear that the transfer function that completely eliminates the variation in leakage is $\Delta V_{bs} = -\frac{b}{a}\Delta V_{tn1}$. We will call this transfer function the *basic* transfer function. As we will see below, when the correlation is not total, other transfer functions will be required, effectively providing less compensation than this basic transfer function.

If we assume that the distributions of ΔV_{tn0} and ΔV_{tn1} are Gaussian with means 0 and variances σ_{n0}^2 and σ_{n1}^2 , respectively, then the mean and variance of $\Delta \ln I$ can be determined with and without compensation, as follows. When there is

no compensation, $\Delta \ln I$ is a linear function of ΔV_{tn0} and thus the mean and variance of $\Delta \ln I$ can easily be computed as $E[\Delta \ln I] = 0$ and $\text{Var}[\Delta \ln I] = b^2\sigma_{n0}^2$. When there is compensation, we will define $V_{bs} = -(b/a^*)\Delta V_{tn1}$ instead of $V_{bs} = -(b/a)\Delta V_{tn1}$ to keep the calculations general and allow for a discussion on how different transfer functions effect the distribution of the leakage after compensation. Thus $\Delta \ln I = b\Delta V_{tn0} + \frac{a}{a^*}b\Delta V_{tn1}$ and its mean can easily be computed to be 0 and its variance to be

$$\text{Var}[\Delta \ln I] = b^2\sigma_{n0}^2 + \left(\frac{ab}{a^*}\right)^2 \sigma_{n1}^2 - 2b^2\frac{a}{a^*}\sigma_{n0}\sigma_{n1}\rho_{n0,n1} \quad (1)$$

where $\rho_{n0,n1}$ is the correlation between the dynamic gate and the monitor. Taking the above equation and differentiating with respect to a^* , it is found that there is a minimum at:

$$\frac{a}{a^*} = \frac{\sigma_{n0}}{\sigma_{n1}}\rho_{n0,n1} \quad (2)$$

Thus, depending on the correlation between the variations in threshold voltage in the monitor and the functioning gate, there is an optimal amount of under-compensation from the basic transfer function to minimize the variance of the log of the relative leakage of a gate. Since \ln is a monotonically increasing function, the value a/a^* that minimizes the variance of $\Delta \ln I$ also minimizes the variance of I . When using the optimal amount of under-compensation, the variance of the log of the relative leakage becomes $\text{Var}[\Delta \ln I] = b^2\sigma_{n0}^2(1 - \rho_{n0,n1}^2)$ which is always lower than than the variance of the uncompensated gate.

When introducing V_{tp} variations into the analysis it is important that the monitor should also be fairly insensitive to V_{tp} variations since the leakage of the dynamic gate is quite insensitive to V_{tp} .

A. Optimum Under-Compensation for Groups of Gates

When a monitor controls a group of gates, the under-compensation that provides the minimum variance for the total leakage must be determined. Let the total leakage of a group of gates, I_T , be defined as $I_T = I_0 + I_1 + \dots + I_{N-2} + I_{N-1}$ where I_i is the leakage of a single gate defined above. Since $\Delta \ln I$ is a normal distribution with mean $E[\Delta \ln I] = 0$ and variance $\sigma_{\Delta \ln I}^2$ as shown in (1), I_i is a lognormal distribution with $\mu_{I_i} = I_{nom}e^{\frac{1}{2}\sigma_{\Delta \ln I}^2}$ and $\sigma_{I_i}^2 = 2I_{nom}^2(e^{\sigma_{\Delta \ln I}^2} - 1)e^{\sigma_{\Delta \ln I}^2}$. Furthermore the variance of I_T can be determined to be:

$$\text{Var}[I_T] = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sigma_{I_i} \sigma_{I_j} \rho_{i,j} \quad (3)$$

To simplify the analysis of (3) we will make the following reasonable simplifying assumptions¹:

- 1) $\sigma_{ni} = \sigma_n$ for all i (i.e. the variance of the underlying V_{tn} variations in all transistors is the same).

¹These assumptions, however, do not reduce the generality of the approach. Using different assumptions would just provide a different numerical solution.

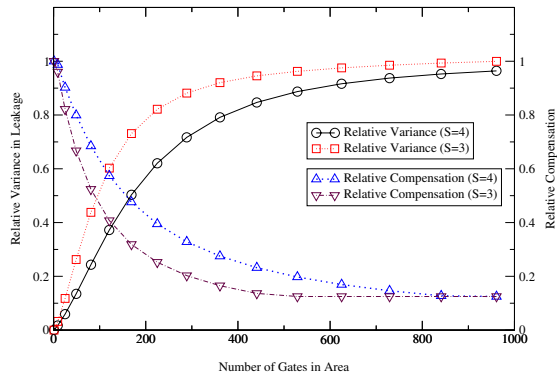


Fig. 1. Optimum Undercompensation and Reduction in Variance of Leakage

- 2) The correlation between the threshold voltage variations approaches 1 as the distance between two transistors is lowered, and approaches 0 as the distance gets larger.
- 3) The monitor is placed at the centre of a square of gates.
- 4) All other gates i are placed around the monitor.

Now we replace $\rho_{i,j}$ in (3) to be $f_\rho(d(i,j))$ where $f_\rho(\cdot)$ is the correlation function described in item 2 above and d is the distance between two gates.

1) *Solving for the Optimal Undercompensation:* We have chosen to model the correlation function as $f_\rho(x) = e^{-\frac{x^2}{2S^2\delta^2}}$ where x is the distance between the two logic gates in question, δ is a measure of the separation (or pitch) between two adjacent gates, and S is a measure of how quickly the correlation between gates decreases as the distance between them increases. Notice that $f_\rho(\cdot)$ looks very much like the Gaussian distribution, but it is not being used as a distribution. For practical purposes, one can think of $3S$ as the largest distance for which correlation between two transistors is not negligible. Again, the usefulness of analysis is not limited by using this specific function, but it allows us to obtain a numerical solution.

Fig. 1 shows the optimum undercompensation needed and the corresponding reduction in total leakage for an area of gates for different number of gates controlled by the monitors for both $S = 3$ and $S = 4$ in $f_\rho(\cdot)$. If the number of gates is low, the compensation that minimizes the variance is near the basic compensation, and the variance of the leakage is almost eliminated compared to an uncompensated system. As the monitor controls more gates a reduced compensation is needed and the variance of the leakage rises, though always being lower than that of an uncompensated system. At 169 gates, which comprises a $13\delta \times 13\delta$ area the optimal amount of undercompensation when $S = 3$ ($S = 4$) is 0.32 (0.48) where the variance of the total leakage is reduced by 27% (50.7%).

IV. MONITOR DESIGN

Now that the basic amount of compensation and the optimum undercompensation is known, the monitor can be designed to produce the required transfer function. The transistor level design of the monitor must meet some requirements including (1) a similar topology to that of a dynamic gate to maximize correlations; (2) a transfer function equal or close

TABLE I
CHARACTERISTIC OF MONITOR GENERATING V_{bs}

	V_{tn}	V_{tp}
Ideal Compensation	5.35	0
Undercompensation Needed	0.318	N/A
Resulting Slope Needed	1.7	0
Slope Obtained	1.71	0.0416
Ideal Bias	0	0
Bias Obtained	-0.105	-0.105

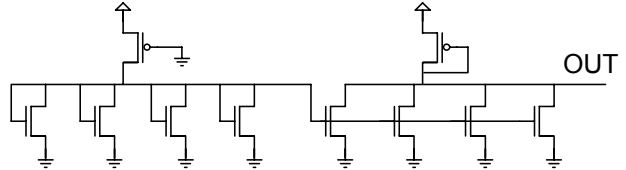


Fig. 2. Monitor producing V_{bs}

to the required one; (3) an output average level in the correct operating range (near 0V for V_{bs}); (3) a minimal amount of power consumption.

To find the monitor that provided the required transfer function for an area of 169 gates (13×13), a number of circuits were tested and one was found that met the requirements for V_{bs} very well; the required compensation with respect to V_{tn} is matched almost exactly, and there is very little variation in the monitor's output with changes in V_{tp} . The average bias output by the circuit is a little lower than optimal, but the negative body bias produced reduces the average leakage of the gates, with very little performance impact as will be seen below. The requirements for the monitor and the resulting characteristics for the monitors can be seen in Table I. The transistor level schematic of the monitor is shown in Fig. 2.

V. RESULTS

To validate that the designed monitor does provide the reduction in variance that was predicted by the analysis in Section III, Monte-Carlo (MC) analysis was performed on the circuitry. The testbench consisted of one functioning gate and one monitor. The MC analysis was performed with different correlation coefficients between the functioning gates and monitor. Fig. 3 shows the reduction in the variance of leakage when using compensation and compares it to the theoretical reduction in variance. The match is very close under the different MC simulation scenarios. When including the power drawn by the monitor, which is comparable to the leakage of 21 dynamic gates, the mean total leakage power will be reduced if more than 56 gates are controlled by the monitor since the mean leakage of a dynamic gate is reduced with the average negative body bias provided by the monitor. When comparing the worst-case leakage, it is reduced when the monitor controls more than 33 gates.

Fig. 4 shows the effect on the variance of delay of the dynamic gates when the compensation scheme is used. It can be seen that when the correlation is high there is around a 40% decrease in the variance and it nears no decrease when the correlation is 0. Using this curve and applying it to a path

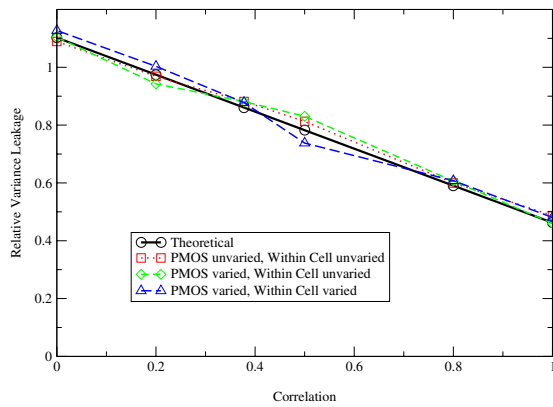


Fig. 3. Reduction in the Variance of the Leakage

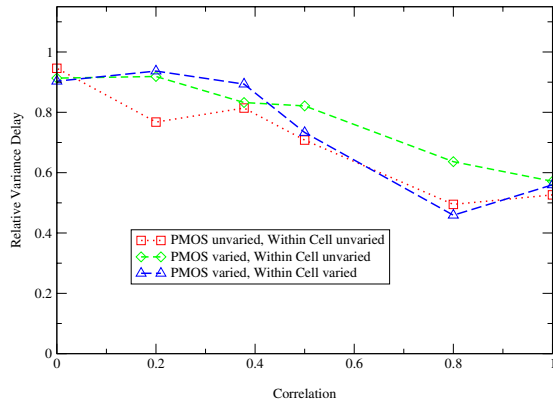


Fig. 4. Reduction in the Variance of the Delay

of 13 gates, the reduction in the variance of the delay of the path is reduced by 33% (37%) when assuming $S=3$ ($S=4$). The mean delay of the compensated system remains virtually identical to the mean of the uncompensated system.

Fig. 5 shows the reduction in variance of the noise margin compared to that of an uncompensated gate. At high correlations there is a near 50% reduction in the variance and in the worst-case there is a 15% reduction in the variance. This reduction in the variance of the noise margin along with the average negative body bias, which reduced the mean leakage and increased the mean noise margin, was used to weaken the keeper, keeping the mean delay of the compensated and uncompensated gate nearly the same.

All the analysis and simulations performed thus far have been at a high temperature of $110^{\circ}C$ since the leakage is higher, the delay longer, and the noise margin lower at this temperature compared to lower temperatures. As the temperature is decreased the functioning of the compensation system can qualitatively be thought of as an increase in V_{tn} which decreases the leakage; thus the monitor transfer function tries to increase the leakage, and reduce the delay. It, however, does not exactly work so, since the transfer function in the monitor is not purely a function of the subthreshold leakage. After performing the MC simulations at low temperatures the greatest change is the change in the mean of the leakage, delay and noise margin compared to the low temperature mean for the uncompensated gate. At $27^{\circ}C$ the leakage of the compensated gate is 21% larger than that of the uncompensated gate, but since it is only a tenth of the leakage of an uncompensated

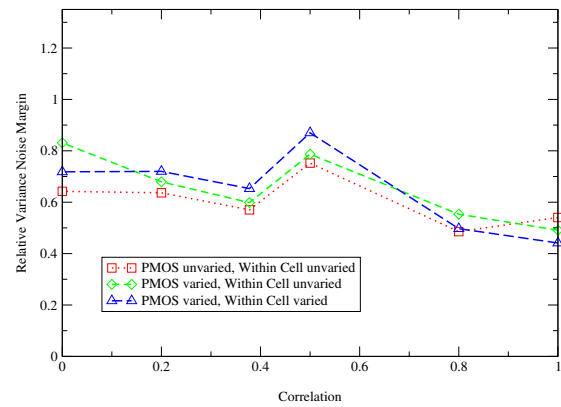


Fig. 5. Reduction in the Variance of the Noise Margin

gate at high temperatures it is not much of an issue. The delay of the compensated gate is faster by 2% than that of an uncompensated gate. The only concern is the noise margin which is decreased from high temperatures to low temperatures but it still is 5% larger than that of the uncompensated gate.

VI. CONCLUSION

We have analytically shown the possible reduction in the variance of leakage of dynamic logic gates that is possible with compensation, and then designed circuits to implement the system. The designed circuits provide a reduction in the variance of the leakage, delay, and noise margin of dynamic logic gates and provide a close match to the analytical results. In our test circuit, which was a generic 169 gate circuit, the variance of leakage is reduced by 27% and the variance of the path delay is reduced by 33%.

REFERENCES

- [1] S. Nassif. Delay variability: Sources, impacts, trends. *International Symposium on Solid-State Circuits*, 2000.
- [2] S. Borkar, T. Karnik, S. Narendra, T. Tschanz, A. Keshavarzi, and Vivek De. Parameter variations and impact on circuits and microarchitecture. *Design Automation Conference*, 2003.
- [3] A. Agarwal, V. Zolotov, and D. Blaauw. Statistical timing analysis using bounds and selective enumeration. *IEEE Transactions on CAD*, September 2003.
- [4] T. Chen and S. Naffziger. Comparison of adaptive body bias (ABB) and adaptive supply voltage (ASV) for improving delay and leakage under the presence of process variation. *IEEE Transactions on VLSI*, October 2003.
- [5] J.W. Tschanz, S. Narendra, and V. De. Effectiveness of adaptive supply voltage and body bias for reducing impact of parameter variations in low power and high performance microprocessors. *IEEE Journal of Solid-State Circuits*, May 2003.
- [6] M. Miyazaki, G. Ono, and K. Ishibashi. A 1.2-GIPS/W microprocessor using speed-adaptive threshold-voltage CMOS with forward bias. *JSSC*, February 2002.
- [7] M. Anis, M. Allam, and M. Elmsary. Impact of technology scaling on CMOS logic styles. *IEEE Transactions on Circuits and Systems II*, August 2002.
- [8] A. Alvandpour, R. Krishnamurthy, K. Soumyanath, and S. Borkar. A sub-130-nm conditional keeper technique. *JSSC*, 37(5):633–638, May 2002.
- [9] <http://www-device.eecs.berkeley.edu/~ptm/>.
- [10] D. Lee, W. Kwong, D. Blaauw, and D. Sylvester. Simultaneous sub-threshold and gate-oxide tunneling leakage current analysis in nanometer CMOS design. *ISQED*, pages 287–292, 2003.
- [11] Siva Narendra, Dimitri Antoniadis, and Vivek De. Impact of using adaptive body bias to compensate die-to-die V_t variation on within-die V_t variation. *ISLPED*, 1999.