Variation-Driven Device Sizing for Minimum Energy Sub-threshold Circuits

Joyce Kwong
jyskwong@mtl.mit.edu

Anantha P. Chandrakasan
anantha@mtl.mit.edu

Massachusetts Institute of Technology
50 Vassar St., Room 38-107
Cambridge MA 02139, USA

ABSTRACT

Sub-threshold operation is a compelling approach for energy-constrained applications, but increased sensitivity to variation must be mitigated. We explore variability metrics and the variation sensitivity of stacked device topologies. We show that upsizing is necessary to achieve robustness at reduced voltages and propose a design methodology to meet yield constraints. The need for upsizing imposes an energy overhead, influencing the optimal supply voltage to minimize energy. Finally, we characterize performance variability by summing delay distributions of each stage in an arbitrary critical path and achieve results accurate to within 10% of Monte Carlo simulation.

Categories and Subject Descriptors: B.8.1 [Reliability, Testing, and Fault-Tolerance]

General Terms: Performance, Design, Reliability

Keywords: Sub-threshold circuits, Minimum energy point, Delay model

1. INTRODUCTION

In sub-threshold circuits, the power supply is set below the transistor threshold voltage V_T to obtain energy savings when speed is not the primary constraint [1]. Authors of [2][3] derived analytical expressions for the optimum V_{DD} to minimize energy in sub-threshold and showed its dependence on major circuit parameters. Sub-threshold circuits rely on leakage currents that are exponentially dependent on V_T and are therefore more sensitive to process variation than traditional above-threshold designs.

It was suggested in [4] that minimum size devices are theoretically optimal for minimizing energy in sub-threshold. However, minimum size devices have increased sensitivity to V_T variation because σ_{V_T} is roughly proportional to $(WL)^{-\frac{1}{2}}$. If a minimum size circuit does not function at the optimum V_{DD} due to degraded logic output swing, it is necessary to upsize devices to improve robustness at the expense of increased energy consumption. Therefore, variability must be considered when analyzing the minimum energy operating point.

Previous work in [5] addresses intra-die variation by providing statistical models for energy and delay of an inverter chain in sub-threshold. An empirical expression for the optimum voltage is shown as a function of logic depth, assuming complete functionality at V_{min}. Work in [6] presents a unified delay variability expression for strong- and weak-inversion and applies it to a NAND gate. Researchers have also proposed various approaches to optimize delay yield by tuning V_{DD}/V_T or choosing gates of different drive strengths, for example in [7]. However, functional yield was not considered until [8][9], which address unsatisfactory V_{OH} and V_{OL} in sub-threshold inverters whose output levels are degraded by leaking devices, such as in a register file. Body biasing is another option for mitigating variation in sub-threshold [10] when a triple-well process is available.

We address inter- and intra-die variation and show that functionality in sub-threshold circuits may be compromised without proper design for variations. We first explore variability metrics for the inverter and logic gates with stacked devices, and propose a metric to size logic gates for a fixed failure rate under process variation. We then examine the energy versus V_{DD} profile given the failure rate constraint and find the optimum sizing and supply voltage. We present an efficient methodology to model delay variability of a chain of logic gates and characterize the effect of yield-based sizing constraints on performance variability.

2. VARIABILITY METRICS AND DEVICE SIZING

A commonly used expression for sub-threshold current is given by [11]

$$I_{sub} = I_o e^{\frac{V_{GS}-V_T+\eta V_{DS}}{nV_{th}}}(1-e^{\frac{-V_{DS}}{V_{th}}}) \quad (1)$$

$$I_o = \mu_o C_{ox}\frac{W}{L}(n-1)V_{th}^2 \quad (2)$$

where n is the sub-threshold swing factor, V_{th} the thermal voltage, and η the DIBL coefficient. The nominal current scales linearly with W/L, while standard deviation of V_T distribution reduces with $(WL)^{-\frac{1}{2}}$, thus lowering sub-threshold current variation. This section explores how sizing affects

ISLPED'06, October 4–6 2006, Tegernsee, Germany

variability in output swing and active current in the inverter and stacked device topologies.

2.1 Logic Gate Output Swing

In the sub-threshold regime, the ratio of active to idle currents in a logic gate is much lower than in strong inversion. If, for example, process variation strengthens NMOS relative to PMOS, a pull-up network will not be able to drive the logic gate output fully to V_{DD} because of idle leakage in the pull-down network. This degradation in gate output swing is illustrated in Figure 1(a). The solid line shows the voltage transfer characteristic (VTC) of a minimum size inverter in a 65nm technology at skewed global process corner. Dashed lines plot the VTCs when random local V_T mismatch is applied to the inverter. One case shows a severely degraded V_{OL}, which can cause functional error if it is above the input low threshold (V_{IL}) of the succeeding gate. Therefore, V_T variation significantly impacts circuit functionality in deeply scaled technologies.

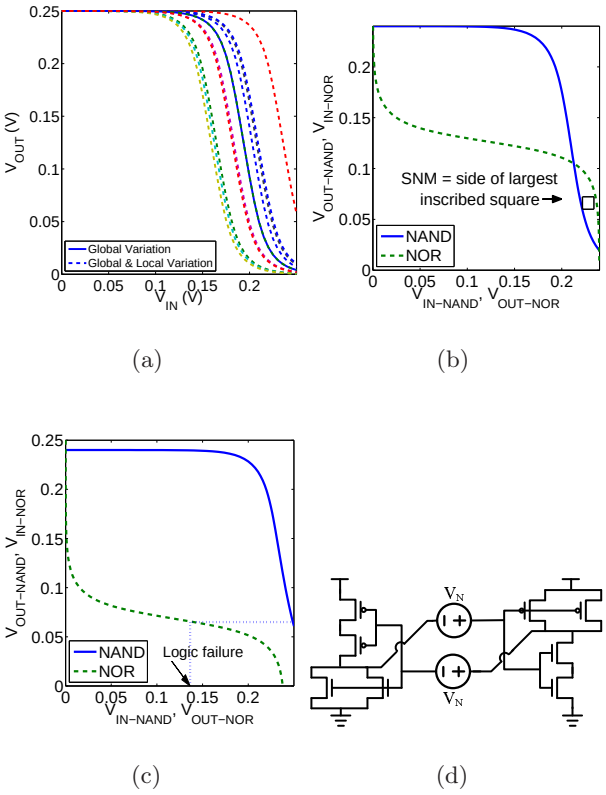


Figure 1: (a) Inverter VTCs at skewed process corner with random V_T mismatch. (b) Butterfly plot of NAND/NOR gates with functional output levels. (c) Butterfly plot of NAND with failing V_{OL}. (d) Example circuit for verifying logic gate output levels.

A consistent metric is necessary to determine whether a logic gate has sufficient V_{OL} and V_{OH} levels. Arbitrary limits, such as 10% and 90% of V_{DD}, do not scale well across global process corners. For example, at the strong-PMOS weak-NMOS corner, strong leakage through PMOS raises V_{OL} of all gates above ground. This also shifts VTCs to the right, and thus logic gates can tolerate higher V_{OL} in the preceding gate. Instead of arbitrary limits, we propose using butterfly plots to verify output voltage levels, specifically in the context of standard cell design.

2.1.1 Use of the Butterfly Plot

To verify V_{OL} of a given gate, we superimpose its VTC with the mirrored VTC of NOR, since the latter has the most stringent V_{IL} requirement from stacked devices in the pull-up network and parallel devices in the pull-down. Similarly, we verify V_{OH} using the NAND VTC, which has the worst case V_{IH}.

In Figure 1(b), a NAND gate has sufficient output swing such that $V_{OL-NAND}$ produces a logic high output in a succeeding NOR gate. In contrast, the NAND gate in Figure 1(c) exhibits $V_{OL-NAND}$=65mV and produces a NOR output of 136mV, close to mid-rail and thus causing logic failure.

A gate with failing output levels is analogous to a 6T SRAM cell displaying negative static noise margin (SNM), in that the butterfly plots for both cases do not contain an inscribed square. Therefore, we can also apply [12] to find the side of the largest inscribed square, illustrated in Figure 1(b). Figure 1(d) shows an equivalent circuit for this measurement on two back-to-back logic gates. Because the VTC is input-dependent, all inputs are varied simultaneously to obtain the worst case V_{IH} and V_{IL}.

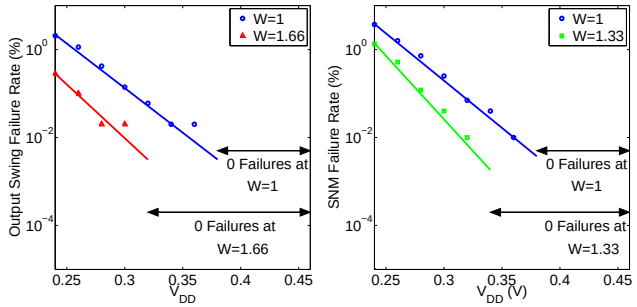
It was shown in [13] that the SNM of two back-to-back gates G1 and G2 is equal to the maximum noise that can be applied to all gates in an infinitely long chain of alternating G1 and G2, before logic failure occurs. Thus when verifying a standard cell G using the butterfly plot, we essentially assume that all logic paths in a synthesized circuit are composed of alternating G and NAND3 gates with the same two skewed VTCs. To accurately model the failure rate of a custom-designed logic path, we would plot VTCs of all gates and trace the signal propagation through the path. Exact modeling is not possible for standard cell design where the target circuit is unknown. Therefore, although the butterfly plot does not reflect the exact mismatch conditions in a circuit, it does provide a guideline for sizing standard cells consistently to account for local variation.

2.1.2 Failure Rate From Insufficient Output Swing

We now define logic failure as having no inscribed square in the butterfly plot and measure how the failure rate varies with V_{DD} and device sizing. To consider logic gates with up to three stacked devices, we verify the INV, NAND2, and NOR2 gates against NAND3 and NOR3, which give the most stringent V_{IH} and V_{IL} requirements respectively. Sizing of NAND3 and NOR3 are fixed to provide a starting point for designing the remaining gates.

The failure rate is estimated from a 5k-point Monte Carlo simulation at worst case temperature. V_T of transistors in the gate under test and global (inter-die) process conditions are randomized such that the Monte Carlo runs are analogous to sampling logic gates across multiple dies. Figure 2(a) shows the failure rate versus V_{DD} of an inverter at various widths normalized to minimum size. Simulated values in markers are fitted to an exponential function ae^{bx}, drawn as a solid line. Note that the failure rate decays more quickly when W=1.66 compared to W=1. Furthermore, zero sam-

ples failed in the 5-k point run at higher voltages, as indicated by arrows on the graph.



(a) (b)

Figure 2: Failure rate of (a) inverter and (b) static register vs. V_{DD}, plotted for various NMOS and PMOS widths (normalized to minimum size).

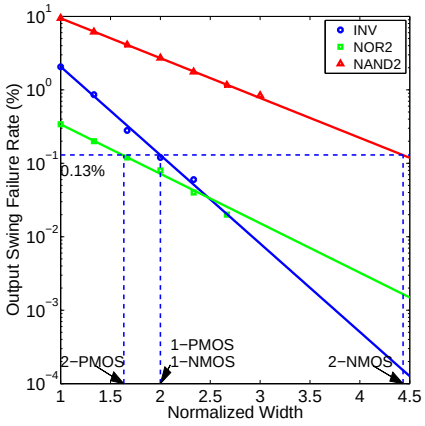


Figure 3: Output swing failure rate of the inverter, NAND2, and NOR2, plotted against device width (normalized to minimum size). V_{DD} is set at 240mV for demonstration.

Figure 3 plots the failure rate versus normalized device width of INV, NAND2, and NOR2. In the inverter, both device sizes are varied simultaneously. In NAND2 and NOR2, the critical two-transistor stack is changed while the two parallel devices are kept constant. The failure rates also decay exponentially with widths. By increasing the device width or V_{DD}, the failure rate can be made to approach 0.

2.2 Noise Margin in Registers

The concept of noise margin is also relevant in sub-threshold register design, where data retention is a particular challenge. Dynamic registers suffer from charge leakage, which worsens in sub-threshold due to slow circuit speeds. Therefore, we consider the static transmission-gate based register. Similar to SRAM cells, the data retention capability of the register is reflected in the hold static noise margin of its cross-coupled inverters. Figure 4 shows the equivalent circuit for measuring the register SNM, accounting for the voltage drop across T2 and the worst case leakage across T1. This circuit is used in a Monte Carlo simulation while varying the V_T of each transistor and inter-die process conditions. Figure 2(b) plots the resulting failure rate in the cross-coupled inverters. Similar to the case of logic gates, the failure rate decreases exponentially to zero when either width or V_{DD} is increased.

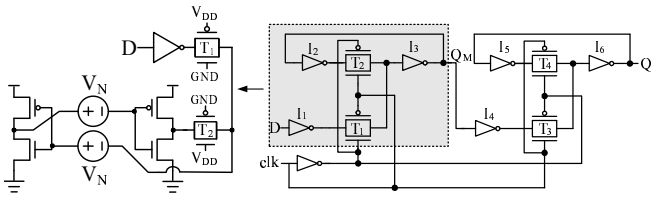


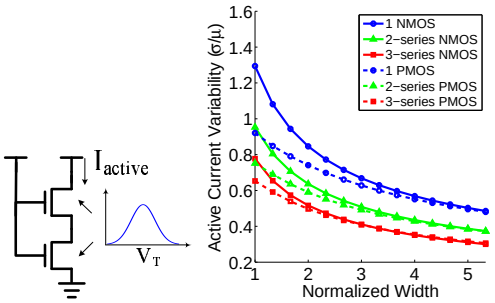
Figure 4: Static register schematic and equivalent circuit for measuring SNM.

2.3 Current Variability

In addition to output swing, active current variability is another metric of interest since it relates directly to variation in propagation delay. With the common assumption that V_T is normally distributed, sub-threshold current can be modeled as a lognormal random variable. From the property of lognormal distributions, the coefficient of variation of active current is given by

$$\sigma_{I_{sub}}/\mu_{I_{sub}} = \sqrt{e^{(\frac{\sigma_{V_T}}{nV_{th}})^2} - 1} \qquad (3)$$

It was observed in [5] that as V_{DD} reduces, the sub-threshold swing factor n decreases. This leads to higher uncertainty in the sub-threshold current through a single device. To examine the impact of topology, Figure 5 plots simulated $\sigma_{I_{sub}}/\mu_{I_{sub}}$ versus device width for static CMOS primitives consisting of one to three devices in series. Variability decreases with larger widths as expected. Stacked device topologies clearly display lower spread in active currents.



(a) (b)

Figure 5: (a) Monte Carlo setup for current variability measurement. (b) Active current variability of different CMOS primitives vs. device width (normalized to minimum size) at V_{DD}=300mV.

2.4 Constant Yield Device Sizing

We now address the issue of device sizing for single and stacked device topologies, given the metrics of output swing and current variability. In above-threshold design, series devices are sized to give equivalent resistance as the inverter. However, in sub-threshold design when the objective is to minimize energy, device sizes should be kept as small as possible while satisfying variability constraints.

Compared to a single device, stacked devices display lower current spread but higher uncertainty in output levels, which may lead to functional errors. Reducing the error rate clearly takes precedence, so output swing rather than current variability should be considered first in sizing decisions.

The output swing failure rate versus width plot of Figure 3 illustrates a sizing methodology for single and stacked devices. Suppose we constrain all topologies to have the same failure rate, or interchangeably, a constant yield. We obtain the required device sizes by drawing a horizontal line at the desired failure rate, then finding where this line intersects the failure curve and the corresponding x-axis value. In Figure 3, a target failure rate of 0.13% requires a single and 2-stack NMOS to be sized at 2 and 4.43 times minimum width respectively. 1-PMOS is sized the same as 1-NMOS as both devices are varied together in simulation. The 2-stack sizing here can be used for any static CMOS gate with two series NMOS, since it was derived from NAND2 where two leaking parallel PMOS give the worst case V_{OL}.

Because the failure rate reduces at higher V_{DD}, the required size for a given yield constraint also decreases. The resulting energy trade-off will be analyzed in Section 3.1. Table 1 lists device widths for a constant failure rate of 0.13% while V_{DD} is varied at 20mV intervals. 0.13% represents the 3σ tail of a normal distribution and is chosen for demonstration. It should be noted that such a target allows sizing logic gates consistently, but does not relate in a straightforward way to the failure rate of a circuit built from these gates. As mentioned previously, this value is a pessimistic estimate because it assumes that every second gate in the circuit is NAND3 or NOR3. Furthermore, failing logic gates tend to cluster on die at process corners.

Table 1: Required widths (normalized to minimum size) vs. V_{DD} for constant failure rate=0.13%

$V_{DD}(V)$	0.24	0.26	0.28	0.30	0.32	0.34
1-NMOS	2	1.67	1.33	1	1	1
2-NMOS	4.43	2.93	2.3	2.27	1.3	1
1-PMOS	2	1.67	1.33	1	1	1
2-PMOS	1.63	1	1	1	1	1

3. MINIMUM ENERGY OPERATION

The total energy per operation consumed by an arbitrary circuit is modeled in [2] as

$$E_T = E_{DYN} + E_L = C_{eff}V_{DD}^2 + W_{eff}I_{leak}V_{DD}t_dL_{DP} \quad (4)$$

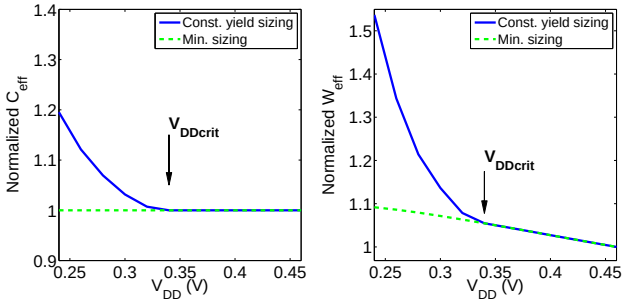
E_{DYN} and E_L model the dynamic switching and leakage energy per cycle respectively. C_{eff} and W_{eff} denote the average total switched capacitance and normalized width contributing to leakage current. t_d and I_{leak} represent the delay and leakage current of a characteristic inverter, while L_{DP} is the logic depth in terms of the inverter delay. As V_{DD} decreases, E_{DYN} is lowered quadratically. The leakage current reduces because of DIBL, but t_d goes up exponentially at sub-threshold voltages and causes a similar increase in leakage energy. The two opposing trends give rise to an optimal supply voltage V_{DDopt} at which total energy is minimized, assuming the circuit is functional.

Section 2 has shown that functionality is no longer guaranteed at low supply voltages when V_T variation is significant. Reducing the probability of logic failure requires either upsizing devices or increasing V_{DD}, which must be considered when finding V_{DDopt}. This can be accounted for within the framework of [2] by treating C_{eff} and W_{eff} as a function of V_{DD}. The resulting energy versus V_{DD} characteristic of an inverter chain and 32-bit Kogge-Stone adder are simulated in a 65nm process and presented as examples.

3.1 Minimum Energy Point with Yield Constraint

Figure 6 plots C_{eff} and W_{eff} versus V_{DD} for the Kogge-Stone adder under two sizing schemes. The solid line plots energy of designs satisfying an upper bound on the output swing failure rate, derived from constant yield sizing of Table 1. The dashed line indicates an adder with only minimum size devices. Note that W_{eff} is obtained by normalizing the adder leakage current to that of a characteristic inverter [2]. DIBL affects leakage through the two circuits differently as V_{DD} decreases, causing a slight increase in W_{eff} in this case. V_{DDcrit} denotes the critical operating voltage at which minimum size devices can be used to satisfy the yield constraint. When $V_{DD} \geq V_{DDcrit}$, the circuit under both schemes are identical.

It should be noted that once the yield constraint is set, V_{DDcrit} can be found immediately from Table 1 and the topology of a given circuit. For example, a circuit without stacked devices does not require upsizing when $V_{DD} \geq V_{DDcrit} = 300$mV. In contrast, a circuit with stacks of two NMOS has $V_{DDcrit} = 340$mV.



(a) (b)

Figure 6: (a) C_{eff} and (b) W_{eff} for adder with constant yield (CY) and minimum sizing (MS).

The switching, leakage, and total energy of the inverter chain and adder are then calculated according to Equation 4. Figure 7(a) plots the energy versus V_{DD} characteristic of the inverter chain at nominal process and temperature. Total energy in both constant yield and minimum sized chains are dominated by the dynamic component. Therefore, the optimum supply voltage of the minimum size chain (dashed

line) is the lowest V_{DD} at which yield constraints are met. By definition, this is equal to V_{DDcrit}. In the constant yield sizing scheme (solid line), reducing the supply below V_{DDcrit} necessitates an increase in device widths. The resulting rise in C_{eff} dominates total energy. In this situation, there is no benefit from upsizing in order to operate at lower V_{DD}. The optimum operating point is with minimum sizing at the lowest V_{DD} permitted by the failure rate constraint.

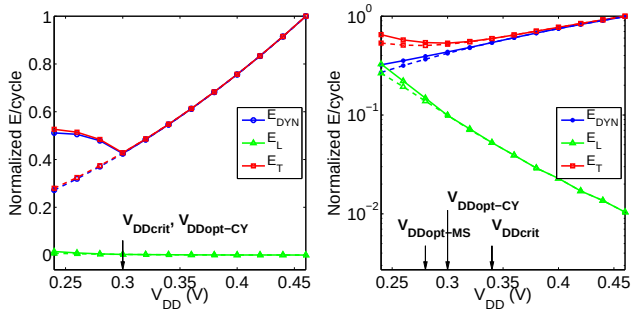
When the minimum size circuit does have a local minimum in its energy characteristic, three scenarios exist depending on the relationship between V_{DDcrit} and the optimum V_{DD} of the constant yield ($V_{DDopt-CY}$) and minimum sizing ($V_{DDopt-MS}$) schemes.

Case 1) $V_{DDopt-MS} > V_{DDcrit}$: No upsizing is required to operate at the minimum energy point, therefore a minimum sized circuit at $V_{DDopt-MS}$ yields optimum energy.

Case 2) $V_{DDopt-MS} < V_{DDopt-CY} < V_{DDcrit}$: A minimum size circuit cannot operate at $V_{DDopt-MS}$ without violating failure rate constraints. A circuit suitably upsized to operate at $V_{DDopt-CY}$ yields optimum energy while satisfying yield requirements.

Case 3) $V_{DDopt-MS} < V_{DDopt-CY} = V_{DDcrit}$: At V_{DDcrit}, the circuit under both sizing schemes are identical. Therefore a minimum size circuit operating at V_{DDcrit} provides minimum energy.

An example of case 2 is seen in Figure 7(b) for a synthesized 32-bit Kogge-Stone adder with interconnect parasitics extracted from layout. Ignoring failure rate constraints, the minimum size adder (dashed line) has an optimum supply voltage of $V_{DDopt-MS} = 280$mV. When we account for failure rate constraints, the effect of constant yield sizing (solid line) is to add energy overhead when $V_{DD} < V_{DDcrit}$. This shifts the local minimum to the right, hence $V_{DDopt-CY} > V_{DDopt-MS}$. Here $V_{DDopt-CY}$ is also $< V_{DDcrit}$, therefore the adder with constant yield sizing at $V_{DDopt-CY} = 300$mV consumes 10.1% less energy than a minimum size adder at $V_{DDcrit} = 340$mV. In this example, constant yield sizing results in a small reduction in energy due to the shallow minimum of the energy versus V_{DD} curve.



(a) (b)

Figure 7: Energy vs. V_{DD} of (a) 11-stage inverter chain and (b) 32-bit adder. Solid and dashed lines indicate CY and MS sizing respectively.

4. PERFORMANCE VARIABILITY

4.1 Delay Variability Modeling

Circuits in sub-threshold display significantly higher delay variability than in above-threshold, therefore proper modeling is essential for timing verification. This section presents a methodology to efficiently model the delay distribution of a chain of logic gates. Using this model, we characterize the delay variability of the Kogge-Stone adders of Section 3.1.

From [2], the delay of a sub-threshold logic gate can be modeled as

$$t_d = \frac{KC_gV_{DD}}{I_o e^{\frac{V_{GS}-V_T}{nV_{th}}}} \quad (5)$$

where K is a delay fitting parameter, C_g is the output capacitance, and the denominator models the gate active current. Both the active current and t_d are lognormally distributed with the same σ parameter. Therefore, delay variability is also given by Equation 3. It depends on σ_{V_T}, which decreases as $(WL)^{-\frac{1}{2}}$, and the sub-threshold swing n, which decreases with V_{DS}. To the first order, σ/μ does not depend on input slew or load capacitance.

The critical path delay in sub-threshold is a sum of lognormal random variables (RVs), typically approximated as another lognormal RV. Authors of [5] derived an expression for the propagation delay of a chain of identical inverters using the Wilkinson approximation. Here we employ the Schwartz-Yeh method [14] to model the sum of non-identically distributed lognormal RVs. The delay of an arbitrary critical path can then be obtained by summing the pre-characterized distributions of each logic gate in the path.

The Schwartz-Yeh method is an iterative algorithm for calculating the sum of lognormal RVs, but requiring much less computation time than Monte Carlo simulation. The modeling methodology using this algorithm is described as follows:

1) Characterize mean delay and standard deviation (μ_{gate}, σ_{gate}) of each logic gate in a cell library, under one input slew and output load condition.

2) Simulate the (N-stage) critical path of interest at nominal process corner and without V_T variation. The delay of the j^{th} stage in the critical path gives μ_{j-path}, for j=1 to N.

3) For each gate j in the critical path, let $\sigma_{j-path} = \sigma_{j-gate} \times \mu_{j-path}/\mu_{j-gate}$, where σ_{j-gate} and μ_{j-gate} are characterized in 1). Since the delay variability σ_j/μ_j is approximately constant across input slew and load conditions, this scales the pre-characterized standard deviation of each gate to the input slew and load conditions in the actual critical path.

4) μ_{j-path} and σ_{j-path} characterize the distribution of each stage, and are input to the Schwartz-Yeh algorithm to generate the delay distribution of the entire critical path.

The above methodology is applied to a three-stage chain consisting of INV-NAND-NOR and to the critical path of a 32-bit Kogge Stone adder at 300mV. Table 2 compares statistical model results with a 1-k point Monte Carlo simulation randomizing V_T of all transistors. The model estimates the mean and standard deviation of the path delay to within a few percent of the Monte Carlo results. This shows that keeping σ/μ constant provides a good approximation.

Table 2: Delay distribution parameters from statistical model and Monte Carlo simulation at 300mV. Values are normalized to FO4 delay.

	Model	Monte Carlo	% Difference
INV-NAND-NOR Chain			
μ	4.957	4.692	5.65%
σ	1.561	1.493	4.51%
Kogge-Stone Critical Path			
μ	36.52	37.13	1.65%
σ	7.038	7.262	3.09%

This method is used to characterize the delay distribution of 1) 32-bit adder with constant yield sizing at $V_{DDopt-CY}$ = 300mV, and 2) adder with minimum size devices at V_{DDcrit} = 340mV. Table 3 shows that the first adder exhibits larger mean and 3σ delay, since $V_{DDopt-CY} < V_{DDcrit}$. However, the delay variability of both adders are comparable, indicating that upsized devices in the first adder offset increased variability from operating at a lower supply voltage.

Table 3: Delay distribution comparison of two adders from Section 3.1. Values are normalized to FO4 inverter delay at $V_{DDopt-CY}$.

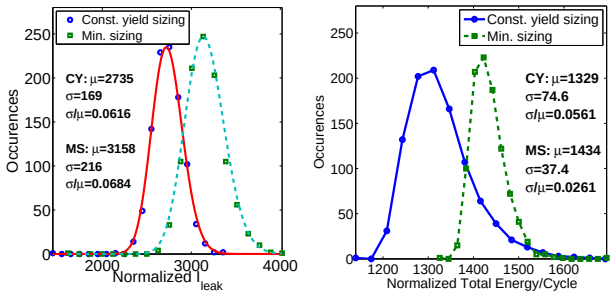
	Const. Yield Sizing	Min. Sizing
μ	90.88	44.92
σ	17.46	8.857
$\mu + 3\sigma$	143.3	71.49
σ/μ	0.1921	0.1972

4.2 Energy Variability

From a 1k-point Monte Carlo simulation, we characterize the energy distribution of the adder with constant yield sizing at $V_{DDopt-CY}$ and the other with minimum size devices at V_{DDcrit}. As suggested in [5], the switched capacitance is verified to vary negligibly with V_T mismatch and is treated as deterministic. Figure 8(a) shows that even though the former adder employs larger devices, it displays lower mean leakage current due to DIBL, and lower variability as an additional benefit. The first adder exhibits lower mean total energy but higher variability in Figure 8(b). The latter effect results from the delay term in leakage energy having larger mean and standard deviation at 300mV compared to 340mV. Note that the leakage component is a product of two dependent lognormal RVs, so E_T is not strictly lognormally distributed.

5. CONCLUSION

In this paper, we have examined the effect of variation and sizing on single and stacked device topologies in subthreshold circuits. Compared to a single device, stacked devices exhibit lower current variability but a higher probability of logic failure from insufficient output swing. We introduced the use of butterfly plots to verify logic gates as well as registers against process variation, and showed that upsizing is necessary to mitigate degraded output levels. The need for upsizing to meet a given yield constraint imposes an energy overhead and impacts the optimum sizing and supply voltage at which energy is minimized. We presented a methodology to model delay variation in an arbitrary critical path using the delay distribution of each stage. Finally, we compared the delay and energy variability of the proposed sizing scheme with a minimum size circuit, and showed that energy reduction is possible without compromising yield or performance variability.



(a) (b)

Figure 8: (a) Leakage current and (b) total energy for two adders of Section 3.1, normalized to those of characteristic inverter at $V_{DDopt-CY}$.

6. ACKNOWLEDGEMENTS

This research is supported by DARPA and Texas Instruments. The authors are grateful to N. Verma, B. Ginsburg, and D. Finchelstein for helpful discussions.

7. REFERENCES

[1] A. Wang and A. Chandrakasan, "A 180mV FFT Processor Using Sub-threshold Circuit Techniques," in *ISSCC*, 2004, pp. 292–293.

[2] B. H. Calhoun and A. Chandrakasan, "Characterizing and Modeling Minimum Energy Operation for Subthreshold Circuits," in *ISLPED*, 2004, pp. 90–95.

[3] B. Zhai, *et al.*, "Theoretical and Practical Limits of Dynamic Voltage Scaling," in *DAC*, 2004, pp. 868–873.

[4] B. H. Calhoun, *et al.*, "Device Sizing for Minimum Energy Operation in Subthreshold Circuits," in *CICC*, Oct. 2004, pp. 95–98.

[5] B. Zhai, *et al.*, "Analysis and Mitigation of Variability in Subthreshold Design," in *ISLPED*, 2005, pp. 20–25.

[6] Y. Cao and L. T. Clark, "Mapping Statistical Process Variations Toward Circuit Performance Variability: An Analytical Modeling Approach," in *DAC*, June 2005, pp. 658–663.

[7] S. H. Choi, *et al.*, "Novel Sizing Algorithm for Yield Improvement under Process Variation in Nanometer Technology," in *DAC*, 2004, pp. 454–459.

[8] J. Chen, *et al.*, "Robust Design of High Fan-In/Out Subthreshold Circuits," in *IEEE Int. Conf. on Computer Design (ICCD)*, Oct. 2005, pp. 405–410.

[9] ——, "Maximum-Ultra-Low Voltage Circuit Design in the Presence of Variations," in *IEEE Circuits and Devices Magazine*, Jan.-Feb. 2006, pp. 12–20.

[10] N. Jayakumar and S. P. Khatri, "A Variation-tolerant Sub-threshold Design Approach," in *DAC*, 2005, pp. 716–719.

[11] V. De, *et al.*, "Techniques for Leakage Power Reduction," in *Design of High-Performance Microprocessor Circuits*, A. Chandrakasan, *et al.*, Eds. IEEE Press, 2001, ch. 3, pp. 46–62.

[12] E. Seevinck, *et al.*, "Static Noise Margin Analysis of MOS SRAM Cells," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 5, pp. 748–754, Oct. 1987.

[13] J. Lohstroh, *et al.*, "Worst-case Static Noise Margin Criteria for Logic Circuits and Their Mathematical Equivalence," *IEEE J. Solid-State Circuits*, vol. SC-18, no. 6, pp. 803–807, June 1983.

[14] S. Schwartz and Y. Yeh, "On the Distribution Function and Moments of Power Sums with Log-Normal Components," *Bell Sys. Tech. Journal*, vol. 61, no. 7, pp. 1441–1462, Sept. 1982.