# Reliable Minimum Energy CMOS Circuit Design

Sean Keller*, Siddharth S. Bhargav†, Chris Moore* Alain J. Martin*

*Department of Computer Science
California Institute of Technology
Pasadena, CA 91125, USA
{sean,cc,alain}@async.caltech.edu

†Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089, USA
ssbharga@usc.edu

***Abstract*—Low voltage operation can markedly reduce the energy requirements of digital circuits; however, in the face of variability it also greatly reduces the *reliability* and *yield*. In order to mitigate this effect, device dimensions can be increased and adaptive body biasing can be employed, but at the cost of potentially increasing energy per operation. This paper presents a new framework for determining the minimum energy operating point of digital CMOS circuits with precise guarantees on reliability and parametric yield.**

## I. Introduction

Current trends show that minimizing power consumption is paramount in the design of digital integrated circuits. In order to achieve ultra-low energy operation it is necessary to reduce the supply voltage below the process nominal $V_{DD}$. As the supply voltage is reduced, the *dynamic* component of energy decreases monotonically, but the *leakage* component of energy has a global minimum well above the functional minimum voltage [1], [2]. As such, there exists a non-zero minimum-energy operating voltage. This minimum-energy operating point depends largely upon the ratio of *dynamic* to *leakage* energy, but it generally occurs near the device threshold voltage [2].

Numerous articles discuss this fundamental problem of determining the minimum energy operating point, *e.g.* [2]–[5]; however, these papers do not address the critical issue of reliability in the face of random parameter variation. All digital CMOS circuits eventually fail to switch as $V_{DD}$ is lowered. Moreover, CMOS circuits become less reliable as $V_{DD}$ is lowered. That is, the circuits become more susceptible to noise sources; intuitively this is closely related to the $I_{on}/I_{off}$ ratio. $I_{off}$ is a weak function of $V_{DD}$ due to DIBL (drain induced barrier lowering), and in modern technologies $I_{off}$ has a range of approximately one order of magnitude from Meindl's minimum $V_{DD}$ to the process nominal $V_{DD}$. On the other hand, $I_{on}$ varies by many orders of magnitude over this same range [6].

As $V_{DD}$ is lowered, nominal $I_{on}/I_{off}$ diminishes, and at the "minimum energy voltage" in the sub-threshold operating region $I_{on}/I_{off}$ is typically in the range of $10^3 - 10^4$ [7]. This range is deceptively small, as both $I_{on}$ and $I_{off}$ are exponentially dependent on the device threshold voltage ($V_t$). In fact, it is only necessary to consider $V_t$ variation out to the $3-4\sigma$ tail — in modern technologies (90nm and smaller) — in order to observe the $I_{on}/I_{off}$ ratio of some minimum-size devices approach unity [8]. The circuit failure rate can be reduced by employing a variety of techniques, *e.g.* by raising $V_{DD}$, increasing device dimensions, body biasing, or by using MTCMOS (Multi-Threshold CMOS). However, it is difficult to determine which techniques should be applied and how best to apply them, in order to achieve the goal of minimizing energy while maintaining functionality.

Quantification and modeling of the energy per operation of a digital circuit is a well-understood problem. On the other hand, quantification of the variation-induced device failure rate is a considerably newer problem that requires the use of noise margins [9]. Static Noise Margins (SNMs) are used extensively in the analysis of memory cells, and [10], [11] use SNMs to address the functional yield problem in sub-threshold circuits. Kwong [8] builds upon this and proposes an SNM analysis method for sizing logic cells; Verma [7] expounds on this work, and Bol [12] builds upon it by considering MTCMOS, body biasing, and device length adjustment. These works make considerable contributions but fail to address several important issues. A common thread throughout [7], [8], [10]–[12] is the assumption that the minimum energy operating point — both with and without functional yield constraints — typically occurs sub-threshold. Without a functional yield constraint, Harris [6] shows that for moderate and low activity factors, the minimum energy supply voltage is actually near-threshold. Furthermore, the standard analytical expression for subthreshold current non-trivially underestimates the minimum energy operating point. With a functional yield constraint, it is even more necessary to consider near-threshold operating points, so analyses using sub-threshold models have considerable error. Another limitation of the *constant yield device sizing method* (used in [7], [8], [12]) is the reliance on butterfly curves for calculating the SNM; it makes exact modeling "not possible for a standard cell design where the

target circuit is unknown" [8]. Lastly, these works make use of an oversimplified method for failure quantification. A critical operating voltage is determined by choosing an arbitrary failure rate for a pair of representative worst-case gates, *e.g.* NOR3 and NAND3; wherein a failure corresponds to a non-positive SNM. With this method of analysis there is no clear way to relate the representative gate failure rate to the functional yield of an actual circuit design.

This paper presents a new methodology for cell library characterization and circuit analysis that is valid in all operating regions. The method determines the minimum energy operating point of any circuit with a guaranteed minimum SNM bound on every gate and a corresponding parametric yield.[1] The library can be generated and characterized independently of the target design, and the guaranteed minimum SNM and parametric yield are user defined. The algorithm picks the optimal sizing of devices, the optimal supply voltage, and the optimal body biases (if desired).

## II. Models

### A. Variation

In modern CMOS technologies, device parameters such as channel length, oxide thickness, threshold voltage ($V_t$), etc. can have significant deviations from their nominal values due to process-induced and random dopant variation. Parameter variations can be classified as either inter-die or intra-die. Inter-die variation can be readily mitigated by using techniques such as ABB (adaptive body biasing) [14]. Since ABB is so effective at mitigating inter-die variation, this paper only considers intra-die variation and assumes ABB is effectively employed; however, inter-die variation could easily be added to the analysis.

In the sub-threshold and near-threshold operating regions, intra-die variation is dominated by random uncorrelated $V_t$ variation (due to random dopant fluctuation (RDF)) [15]. Random device threshold variation tends to be normally distributed with a standard deviation that varies with device area approximately as $\sigma \propto \frac{1}{\sqrt{WL}}$ [16]. As such, increasing device dimensions reduces the likelihood of variation induced failures, at the cost of increased dynamic energy consumption; the effect on the leakage component of energy depends on which dimension is scaled (W or L) and whether short-channel or reverse-short-channel effects are dominant.

### B. Device Models

Analytical device models have classically provided insight into circuit problems, but it is difficult to use these analytical device models to solve the minimum energy supply voltage and sizing problem. First, depending on the particular technology and the circuit activity factor, the optimal supply voltage ranges from deep sub-threshold to the upper end of near-threshold, necessitating the use of a trans-regional $I_{ds}$ model such as described in [4], [5], [17], [18]. Second, the full analytical $I_{ds}$ equations become inaccurate outside of very narrow ranges, so even finding a closed form solution to a much simpler problem requires the use of $I_{on}$ and $I_{off}$ equations in place of a general $I_{ds}$ model [5]. Third, modern technologies (65nm and smaller) exhibit significant short-channel and narrow-channel effects, thus making the trans-regional analytical modeling grossly inaccurate w.r.t. both length and width scaling around minimum-dimensioned devices. Finally, the currents through stacked devices operating sub-threshold and near-threshold exhibit serious technology and sizing dependencies due both to DIBL and the body effect [6]. For these reasons, analytical device models are not used in this work, and results are garnered by way of circuit simulation via HSPICE using foundry-provided BSIMv4.5 [19] device models along with foundry-provided variation distributions in the form of HSPICE variation blocks.

### C. Noise Margin

In order to characterize gate reliability, the voltage transfer characteristic (VTC) is employed, and the unity gain points are used to determine $V_{OH}$, $V_{IH}$, $V_{OL}$, $V_{IL}$ as in Fig. 1.

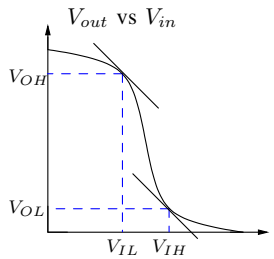


Fig. 1: **Gate VTC.**

Given any two gates $G_x$, $G_y$, where $G_x$ is driving $G_y$, the SNM is defined as

$$NM_H = V_{OH}(G_x) - V_{IH}(G_y), \quad (1)$$

$$NM_L = V_{IL}(G_y) - V_{OL}(G_x), \text{ and} \quad (2)$$

$$SNM = min(NM_H, NM_L). \quad (3)$$

Due to parameter variation, the VTC of a gate shifts away from the nominal VTC and thus moves the corresponding $V_{OH}$, $V_{OL}$, $V_{IH}$, and $V_{IL}$, which are treated as random variables (RVs).

The choice of SNM as a reliability metric and the method used to calculate it was chosen deliberately: to decouple library characterization from reliability analysis of real circuits. First, as the name suggests, the SNM of a gate corresponds to a *static* form of analysis. That is, SNM is calculated by way of a DC sweep, so it is independent of output load. Second, the calculation of $V_{OH}$, $V_{OL}$, $V_{IH}$, and $V_{IL}$ for each gate makes

[1] Die yield loss can be partitioned into a functional yield loss component and a parametric yield loss component where functional yield loss corresponds to dice that fail to function, and parametric yield loss corresponds to dice that function but not to *specification*. [13]. In this paper, the parametric yield *specification* is a constraint on SNMs. Circuits that do not satisfy this constraint may not function, or they may function less reliably. Moreover, when this SNM constraint is reduced to zero, circuits that do not satisfy the constraint will not function. In this paper, the term parametric yield simply refers to the component of die yield governed by the SNM constraint.

it possible to subsequently determine $NM_H$ and $NM_L$ for any gate pair.

### D. Energy

The total energy per operation in a digital circuit can be broken into an dynamic component that accounts for net-switching capacitance, and a leakage component that accounts for all parasitic currents. That is,

$$E_{tot} = \alpha E_{dyn} + E_{leak}, \text{ where} \quad (4)$$

$$E_{dyn} = C_L \cdot V_{DD}^2, \text{ and} \quad (5)$$

$$E_{leak} = I_{leak} \cdot \tau V_{DD}. \quad (6)$$

The activity factor, $\alpha$, accounts for the fact that for any operation, only a subset of the gates typically switch; $C_L$ is the total load capacitance, and $\tau$ is the time to complete an operation, *i.e.* cycle time. In the face of parameter variation, $C_L$, $I_{leak}$, and $\tau$ can be treated as random variables, but RDF affects $C_L$ negligibly, and as the number of gates in a circuit increases $I_{leak}$ approaches the mean. Similarly, $\tau$, is often treated as an RV in other analyses, but in the context of minimum-energy architectures, long path lengths can be assumed. Additionally, if modern average-case timing techniques are employed as in [20] or [21], nominal values for $\tau$ can be used.

## III. Library Characterization

Standard-cell library characterization is an important step in digital system design. At the simplest, propagation delays are determined for each gate at process nominal $V_{DD}$ and the results are recorded in tables. Modern tool-flows now include complex timing and energy models, so additional gate characteristics are needed, *e.g.* switching energy, leakage power, and slew-rate. Moreover, these characterizations are routinely performed for several process corners and several $V_{DD}$ values.

In order to guarantee the reliability of a design, statistical noise margin data must also be collected. If a fixed standard-cell library must be used (such a library is often provided by the foundry), then the noise margin data only needs to be collected for each cell in the library. In this paper, standard-cell library creation is considered, so a much wider range of device lengths and width are considered and a simplified propagation delay model is employed; however the algorithm presented in Section IV-B is valid regardless of whether or not a fixed library is used.

Every gate type (*e.g.* INV, NAND2, NOR2, NAND3, NOR3 etc. ) is characterized using SPICE with a range of lengths and widths for NFETs and PFETs, a wide range of $V_{DD}$ operating points, and several body biases. Each characterization is written to a database and includes:

- $t_{pdr}$
- $t_{pdf}$
- $I_{leak}$
- $E_{dyn}$
- $\mu(V_{OH})$
- $\sigma(V_{OH})$
- $\mu(V_{OL})$
- $\sigma(V_{OL})$
- $\mu(V_{IH})$
- $\sigma(V_{IH})$
- $\mu(V_{IL})$
- $\sigma(V_{IL})$

where the RVs are each computed with 2.5K Monte Carlo trials, $E_{dyn}$ is calculated for *logically worst-case* input rising and falling edges, and similarly $I_{leak}$ is computed for *logically worst-case* static high and low input. For $E_{dyn}$, the *logically worst-case* input corresponds to the input pattern that typically maximizes propagation delay, *e.g.* for the NOR2 $t_{pdf}$ one input is tied low and the other input transitions from low-to-high so only one of the parallel NFETs actively pulls the output low. Similarly, for $I_{leak}$, the *logically worst-case* static input corresponds the input pattern that maximizes leakage. The SNM data (*e.g.* $\mu(V_{OH})$, $\sigma(V_{OH})$, $\mu(V_{OL})$, etc.) are also calculated with the *logically worst-case* input pattern. For example, the *logically worst-case* input pattern for the NOR2 SNM is identical to that used to calculate $t_{pdf}$.

## IV. Reliable-Circuit Minimum Energy Algorithm

Given a digital circuit netlist, the goal of the *reliable-circuit minimum energy algorithm* is to determine the gate sizing, body bias voltages ($V_{DDB}$ and $GND_B$), and $V_{DD}$ values that minimize energy while guaranteeing that every connected gate-pair meets or exceeds a particular static noise margin target, $SNM_T$, with a parametric yield that meets or exceeds $YIELD_T$. Since intrinsic noise sources are proportional to $V_{DD}$, $SNM_T$ is specified as a percentage of $V_{DD}$, but an absolute $SNM_T$ could be used if desired. In the context of this experiment, the parametric yield corresponds to the percentage of circuits that satisfy the $SNM_T$ constraint. For example, if the designer of a microprocessor chooses $SNM_T = 10\%V_{DD}$ and $YIELD_T = 95\%$, the *reliable-circuit minimum energy algorithm* guarantees that 95% of the microprocessors will contain no connected gate-pairs with static noise margins less than $10\%V_{DD}$ in spite of parameter variation.

### A. Statistics

For any two connected gates, $(G_x, G_y)$ with $G_x$ driving $G_y$, let $P(PASS(G_x, G_y))$ represent the probability that $SNM(G_x, G_y)$ equals or exceeds $SNM_T$. In order to make global statistical guarantees, $P(PASS(G_x, G_y))$ must be calculated for each connected gate-pair. This calculation is straightforward, because $V_{IH}$ and $V_{IL}$ are normally distributed and $V_{OH}$ and $V_{OL}$ are constant (see Section V for details), so the Gauss error function, $erf$ [2], can be used directly to calculate $P(PASS(G_x, G_y))$.

Using Equation 3, for gate $G_x$ driving gate $G_y$, the $NM_H$ and $NM_L$ components of $SNM(G_x, G_y)$ are separated and

$$P(PASS(G_x, G_y)) = P_H \cdot P_L, \quad (7)$$

where

$$P_H = P(NM_H(G_x, G_y) \geq SNM_T), \text{ and} \quad (8)$$

$$P_L = P(NM_L(G_x, G_y) \geq SNM_T). \quad (9)$$

With the one-sided error function,

$$P(NM_{H/L}(G_x, G_y) \geq SNM_T) = 1 - \frac{1}{2}\left(1 - erf(\frac{n_{H/L}}{\sqrt{2}})\right), \quad (10)$$

[2](Note that for normally distributed samples, $erf(\frac{n}{\sqrt{2}})$ gives the proportion of values that fall within $n$ standard deviations of the mean.)

where

$$n_H = \frac{[\mu(V_{OH}(G_x)) - \mu(V_{IH}(G_y))] - SNM_T}{\sigma(V_{IH}(G_y))} \text{ and} \quad (11)$$

$$n_L = \frac{[\mu(V_{IL}(G_y)) - \mu(V_{OL}(G_x))] - SNM_T}{\sigma(V_{IL}(G_y))}. \quad (12)$$

Finally, each of these gate-pair probabilities can be combined to give the net parametric yield, expressed as

$$YIELD = \prod_{G_x, G_y \in K} P(PASS(G_x, G_y)), \quad (13)$$

where $K$ is the set of all connected gates $(G_x, G_y)$ with $G_x$ driving $G_y$. The minimization algorithm guarantees correctness by ensuring that $YIELD \geq YIELD_T$.

Equation 13 does not entirely account for correlation effects between the inputs of multiple-input gates. Consider $G_3$ in the center of Fig. 2. The library characterization phase only uses a single *logically worst-case* input pattern for the SNM characterization of each gate, so $NM_{H/L}(G_1, G_3)$ and $NM_{H/L}(G_2, G_3)$ are not independent. However, Equation 13 treats them as independent and thus underestimates $YIELD$.[3] For a fan-out, the situation is much simpler. Using the approximation that $V_{OH}$ and $V_{OL}$ are constant (from Section V): $NM_{H/L}(G_3, G_4)$, $NM_{H/L}(G_3, G_5)$, and $NM_{H/L}(G_3, G_6)$ are independent and Equation 13 remains accurate.

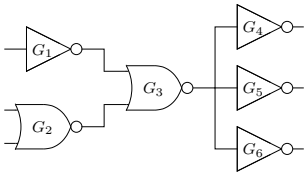


Fig. 2: **Fan-in and fan-out correlation example circuit.**

### B. Algorithm

The search algorithm shown in Fig. 3 iterates through every possible operating point ($V_{DD}$ and body bias), and for each operating point, `minE()` and `reliable()` is executed. The

```
for vdd = minVdd to maxVdd
   for vddb = minVddb to maxVddb
     for GNDb = minGNDb to maxGNDb
        minE()
        reliable()
        record totalEnergy and sizing
return lowest totalEnergy sizing
```

Fig. 3: **The reliable-circuit minimum energy algorithm.**

`minE()` function (see Fig. 4) finds the minimum energy sizing without reliability guarantees. With typical activity factors, the dynamic component of energy exceeds the leakage component, so using entirely minimum-sized gates will achieve minimum total energy. However, for sufficiently low activity factors $E_{leak}$ can exceed $E_{dyn}$; as such, $E_{tot}$ can be reduced by lengthening some devices to reduce leakage. `minE()` begins with minimally sized gates and then iterates through the gate list computing the change in $E_{tot}$ for lengthening each transistor in turn. The change in length that maximizes the energy savings is chosen at each iteration, and the function returns when further up-sizing increases $E_{tot}$. The `reliable()` function (see Fig. 5) starts with the minimum energy sizing and judiciously up-sizes gates in order to meet the reliability requirement. The function continuously iterates through the gate list, and at each iteration, the gate-size increase which improves $YIELD$ the most for the smallest increase in $E_{tot}$ is chosen. The function returns when $YIELD \geq YIELD_T$. The validity of `reliable()` follows from Assumption 1 about gate sizing in the face of variability.

```
set each gate N/P_W and N/P_L to minimum size
lastE = E_tot
while E_tot <= lastE
   for each gate
      for each FET
         calculate d(E_tot)
   choose largest d(E_tot) and resize gate
   lastE = E_tot
   recalculate E_tot
undo last gate upsize
return all gate sizing changes
```

Fig. 4: **The `minE()` function pseudocode.**

```
while YIELD < YIELD_T
   for each pair of connected gates
      for each FET
         calculate d(YIELD)/d(E_tot)
   choose largest d(YIELD)/d(E_tot) and resize gate
   recalculate YIELD
```

Fig. 5: **The `reliable()` function pseudocode.**

**Assumption 1.** *For any gate, G, with NFET and PFET lengths and widths given as, $N_L$, $N_W$, $P_L$, and $P_W$, increasing the values of some subset of $\{N_L, N_W, P_L, P_W\}$ increases the "reliability" of G. That is, up-sizing will either increase one or more of $\mu(V_{OH}), \mu(V_{OL}), \mu(V_{IH}), \mu(V_{IL})$, or it will reduce one or more of $\sigma(V_{OH}), \sigma(V_{OL}), \sigma(V_{IH}), \sigma(V_{IL})$.*[4]

We do not prove this assumption, but it follows naturally from two facts. First, if a gate has a poor nominal $I_{on}/I_{off}$ due to poorly ratioed devices then proper sizing can improve the nominal $I_{on}/I_{off}$. Second, $\sigma(V_t) \propto \frac{1}{\sqrt{WL}}$, so up-sizing will reduce the spread of $V_t$.

This assumption is necessary, because with a large standard cell library the full search space for the optimal operating point can be enormous, and *Assumption 1* allows for a substantial reduction of the search space. In fact, the computational complexity of the *reliable-circuit minimum energy algorithm* is only $O(N \cdot V \cdot B)$, where $N$ is the number of gates (it is assumed that the number of connected gate pairs is some small

[3] This effect should be accounted for and is left as future work.

[4] Validity is assumed up to some technology dependent bound on device dimensions.

constant times N), $V$ is the number of $V_{DD}$ steps, and $B$ is the number of body bias points. Furthermore, the implementation is performant, because the algorithm only needs to iterate over every unique type of gate-pair, not every gate pair, *e.g.* every NAND2 driving an INV can be lumped into a single calculation.

## V. Preliminary Results

Characterization of a handful of gates (INV, NAND2, NOR2) in a low-power 65nm process and a low-power 40nm process has been completed, and in total more than 100K experiments each of 2.5K Monte Carlo trials have been completed. Two important results have emerged. First, $V_{IH}$ and $V_{IL}$ are always normally distributed; for each trial an Anderson-Darling normality test was performed, and no significant departure from normality was found in any experiment. Second $V_{OL}$ and $V_{IL}$ are nearly constant, and can be treated as such for statistical analysis. For example, the 65nm inverter characterization shows that the average ratio of $\frac{\sigma(V_{IH})}{\sigma(V_{OH})} = 65$ and $\frac{\sigma(V_{IL})}{\sigma(V_{OL})} = 68$. The characterization time is dominated by circuit simulation time, and on a modern high-end server (4 x AMD Opteron 6168 with 64GB of RAM) the row generation rate for the characterization database (see Section III) is 25 rows-per-minute using the 65nm PDK and 50 rows-per-minute with the 40nm PDK.

For the 65nm process, $V_t = 400mV$ nominally, and a Monte Carlo analysis of a minimum size NFET (60nm x 120nm) yielded a normally distributed $V_t$ distribution with $\frac{3\sigma}{\mu} = 30\%$ at the global TT corner. The gates have not yet been characterized at other global corners or with an inter-die variation distribution; additionally, the gates have not yet been characterized with a body bias. Despite these limitations, results have been generated with the *reliable-circuit minimum energy algorithm*.

Fig. 6 depicts a simple test circuit: 1M inverters organized as linear chains of gates with path lengths of 50 gates. Fig. 7 shows the *reliable-circuit minimum energy algorithm* results for this test circuit with all inverters sized identically using $\alpha = 5\%$, $YIELD_T = 95\%$, and $SNM_T = 10\%V_{DD}$.[5] For this simple test circuit, the *reliable-circuit minimum energy algorithm* completed in under 10 seconds.

## VI. Conclusion

This paper presents a new method for digital CMOS library characterization and a new algorithm for finding the minimum energy operating point and gate sizing with a guaranteed minimum static noise margin for any circuit. Several gates have been characterized and experiments run, but further characterization is required. A number of items remain slated for future work, such as a direct comparison against other published results, incorporation of MTCMOS into the algorithm, and the independent sizing of every FET. Finally, correlation effects should be included in the statistical analysis (although preliminary analysis indicates the effects are small).

[5]Notice that the minimum energy operating point with parametric yield and reliability guarantee occurs above the sub-threshold region, near-threshold.

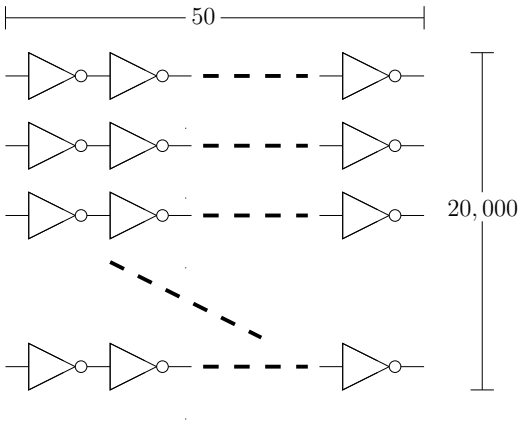


Fig. 6: **1M inverter test circuit.**

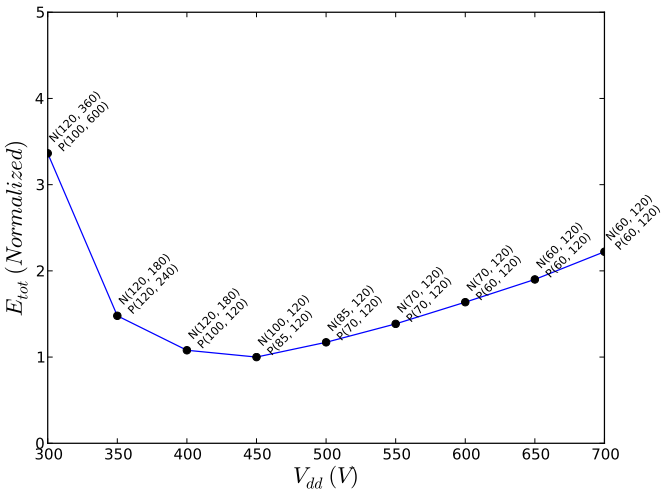


Fig. 7: **Results of the reliable-circuit minimum energy algorithm applied to 1M inverter test circuit with P(L,W) and N(L,W) given for each point in nanometers.**


## Acknowledgment

We would like to thank Synopsys, Inc. for providing all of the HSPICE licenses needed to complete this work. The research described in this paper is in part supported by a grant from the National Science Foundation.


## References


[1] J. Meindl and J. Davis, "The fundamental limit on binary switching energy for terascale integration (tsi)," *Solid-State Circuits, IEEE Journal of*, vol. 35, no. 10, pp. 1515 –1516, Oct. 2000.

[2] B. Zhai, D. Blaauw, D. Sylvester, and K. Flautner, "Theoretical and practical limits of dynamic voltage scaling," in *Design Automation Conference, 2004. Proceedings. 41st*, 2004, pp. 868 – 873.

[3] B. Calhoun, A. Wang, and A. Chandrakasan, "Modeling and sizing for minimum energy operation in subthreshold circuits," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 9, pp. 1778 – 1786, Sep. 2005.

[4] D. Markovic, C. Wang, L. Alarcon, T.-T. Liu, and J. Rabaey, "Ultralow-power design in near-threshold region," *Proceedings of the IEEE*, vol. 98, no. 2, pp. 237 –252, Feb. 2010.

[5] D. Harris, B. Keller, J. Karl, and S. Keller, "A transregional model for near-threshold circuits with application to minimum-energy operation," in *Microelectronics (ICM), 2010 International Conference on*, Dec. 2010, pp. 64 –67.

[6] N. H. E. Weste and D. M. Harris, *CMOS VLSI Design, 4th Ed.* Addison-Wesley, 2010.

[7] N. Verma, J. Kwong, and A. Chandrakasan, "Nanometer mosfet variation in minimum energy subthreshold circuits," *Electron Devices, IEEE Transactions on*, vol. 55, no. 1, pp. 163 –174, Jan. 2008.

[8] J. Kwong and A. Chandrakasan, "Variation-driven device sizing for minimum energy sub-threshold circuits," in *Low Power Electronics and Design, 2006. ISLPED'06. Proceedings of the 2006 International Symposium on*, Oct. 2006, pp. 8 –13.

[9] C. F. Hill, "Noise margin and noise immunity of logic circuits," *Microelectronics*, vol. 1, pp. 16– 21, Apr. 1968.

[10] J. Chen, L. Clark, and Y. Cao, "Robust design of high fan-in/out subthreshold circuits," in *Computer Design: VLSI in Computers and Processors, 2005. ICCD 2005. Proceedings. 2005 IEEE International Conference on*, Oct. 2005, pp. 405 – 410.

[11] ——, "Maximum - ultra-low voltage circuit design in the presence of variations," *Circuits and Devices Magazine, IEEE*, vol. 21, no. 6, pp. 12 – 20, Jan.-Feb. 2005.

[12] D. Bol, R. Ambroise, D. Flandre, and J.-D. Legat, "Analysis and minimization of practical energy in 45nm subthreshold logic circuits," in *Computer Design, 2008. ICCD 2008. IEEE International Conference on*, Oct. 2008, pp. 294 –300.

[13] S. Cunningham, C. Spanos, and K. Voros, "Semiconductor yield improvement: results and best practices," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 8, no. 2, pp. 103 –109, may 1995.

[14] J. Tschanz, J. Kao, S. Narendra, R. Nair, D. Antoniadis, A. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage," *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 11, pp. 1396 – 1402, Nov. 2002.

[15] N. Drego, A. Chandrakasan, and D. Boning, "Lack of spatial correlation in mosfet threshold voltage variation and implications for voltage scaling," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 22, no. 2, pp. 245 –255, May 2009.

[16] D. Sylvester, K. Agarwal, and S. Shah, "Variability in nanometer CMOS: impact, analysis, and minimization," *Integration, the VLSI Journal*, vol. 41, no. 3, pp. 319–339, May 2008.

[17] C. C. Enz and E. A. Vittoz, *Charge-Based MOS Transistor Modeling: The EKV Model for Low-Power and RF IC Design*. Wiley, Sep. 2006.

[18] K. Nose and T. Sakurai, "Optimization of vdd and vth for low-power and high speed applications," in *ASP-DAC '00: Proceedings of the 2000 Asia and South Pacific Design Automation Conference*. New York, NY, USA: ACM, 2000, pp. 469–474.

[19] B. Sheu, D. Scharfetter, P.-K. Ko, and M.-C. Jeng, "Bsim: Berkeley short-channel igfet model for mos transistors," *Solid-State Circuits, IEEE Journal of*, vol. 22, no. 4, pp. 558 – 566, Aug. 1987.

[20] A. Martin and M. Nystrom, "Asynchronous techniques for system-on-chip design," *Proceedings of the IEEE*, vol. 94, no. 6, pp. 1089–1120, 2006.

[21] D. Blaauw, S. Kalaiselvan, K. Lai, W.-H. Ma, S. Pant, C. Tokunaga, S. Das, and D. Bull, "Razor ii: In situ error detection and correction for pvt and ser tolerance," in *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*, Feb. 2008, pp. 400 –622.