# Chapter 3
# Effect of Variations and Variation Tolerance in Logic Circuits

**Swaroop Ghosh**

**Abstract** Variations in process parameters affect the operation of integrated circuits (ICs) and pose a significant threat to the continued scaling of transistor dimensions. This fluctuation in device geometries might prevent them from meeting timing and power criteria and degrade the parametric yield. Process limitations are not exhibited as physical disparities only; transistors experience temporal device degradation as well. On top of it, power management techniques like voltage scaling, dual $V_{TH}$, further magnify the variation-induced reliability issues. On the other hand, conventional resiliency techniques like transistor upsizing and supply voltage boosting typically increase the power consumption. Low-power dissipation and process variation tolerance therefore impose contradictory design requirements. Such issues are expected to further worsen with technology scaling. To circumvent these non-idealities in process parameters, we describe two approaches: (1) variation-tolerant circuit designs and (2) circuits that can adapt themselves to operate correctly under the presence of such inconsistencies. In this chapter, we first analyze the effect of process variations and time-dependent degradation mechanisms on logic circuits. We consider both die-to-die and within-die variation effects. Next, we provide an overview of variation-tolerant logic design approaches. Interestingly, these resiliency techniques transcend several design abstraction levels – however in this chapter, we focus on circuit level techniques to perform reliable computations in an unreliable environment.

## 3.1 Introduction

Successful design of digital integrated circuits has relied on optimization of various design specifications such as silicon area, speed, and power dissipation. Such a traditional design approach inherently assumes that the electrical and physical

S. Ghosh (✉)
Intel Corporation, Portland, OR, USA
e-mail: swaroopad.ghosh@intel.com

S. Bhunia, S. Mukhopadhyay (eds.), *Low-Power Variation-Tolerant Design in Nanometer Silicon*, DOI 10.1007/978-1-4419-7418-1_3,

properties of transistors are deterministic and hence predictable over the device lifetime. However, with the silicon technology entering the sub-65 nm regime, transistors no longer act deterministically. Fluctuation in device dimensions due to manufacturing processes (sub-wavelength lithography, chemical mechanical polishing, etc.) has emerged as a serious issue in nanometer technologies. In deep sub-micron (approximately up to 0.35 μm) technology nodes, process variation was inconsequential for the IC industry. Circuits were mostly immune to minute fluctuations in geometries because the variations were negligible compared to the nominal device sizes. However, with the growing disparity between feature size and optical wavelengths of lithographic processes at scaled dimensions (below 90 nm), the issue of parameter variation is becoming severe. Variations can have two main components: inter-die and intra-die. Parametric variations between dies that come from different runs, lots, and wafers are categorized into *inter-die variations* whereas variations of transistor strengths within the same die are defined as *intra-die variations*. Fluctuations in length ($L$), width ($W$), oxide thickness ($T_{OX}$), flat-band conditions, etc. give rise to inter-die process variations, whereas line edge roughness (LER) or random dopant fluctuations (RDF) cause intra-die random variations in process parameters [1–7]. Systems that are designed without consideration to process variations fail to meet the desired timing, power, stability, and quality specifications. For example, a chip designed to run at 2 GHz of speed may run only at 1 GHz (due to increased threshold voltage of the devices) making it unworthy to be useful. However, one can certainly follow an overly pessimistic design methodology (using large guard-bands) that can sustain the impact of variations in all process corners. Such designs are usually time and area inefficient to be beneficial. There can be two ways to address the issue of process variation – by controlling the existing process technology (which result in less variation) or by designing circuits and architectures that are immune to variations. In this chapter, the emphasis has been given to the second choice because controlling the process is expensive and in some cases may not be viable without changing the devices itself (it should be noted that some of the emerging devices may suffer from increased variations as well [8–13]). Therefore, process variation awareness has become an inseparable part of modern system design.

### *3.1.1 Effect of Power and Process Variation in Logic Circuits*

Apart from process variations, modern circuits also suffer from escalating power consumption [14]. Not only dynamic power but leakage power has also emerged as a dominant component of overall power consumption in scaled technologies. This is also described in Chapter 2. Power management techniques, e.g., supply voltage scaling, power gating, and multiple-$V_{DD}$, $V_{TH}$ designs further magnify the problems associated with process-induced variations. For example, consider the supply voltage scaling. As the voltage is scaled down, the path delay increases worsening the effect of variations. This is elucidated in Fig. 3.1. It clearly shows that lower $V_{DD}$

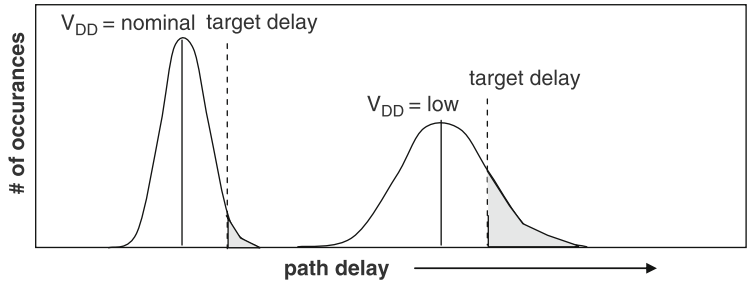



**Fig. 3.1** Impact of supply voltage scaling on path delay distribution. Both mean and sigma of delay distribution as well as the number of paths failing to meet target frequency increases

not only increases the mean but also the standard deviation (STD) of the overall path delay distribution. Therefore, the number of paths failing to meet the target speed also increases thereby degrading the timing yield. On the other hand, upscaling the supply voltage reduces variation at the cost of power consumption. Power and process variation resilience are therefore conflicting design requirements and one comes at the cost of other. Meeting a desired power specification with certain degree of process tolerance is becoming an extremely challenging task.

### *3.1.2 Sources of Variations (Spatial vs. Temporal)*

Figure 3.2 shows the taxonomy of variations in nanoscaled logic circuits. Inter- (die-to-die) and intra-die (within-die) variations fall under *spatial* (or $t = 0$) process variation. These types of variations induce speed marginalities in static logic whereas in dynamic logic they reduce the noise margin. *Temporal* variations can be categorized into two parts – environmental and aging related. Temperature and

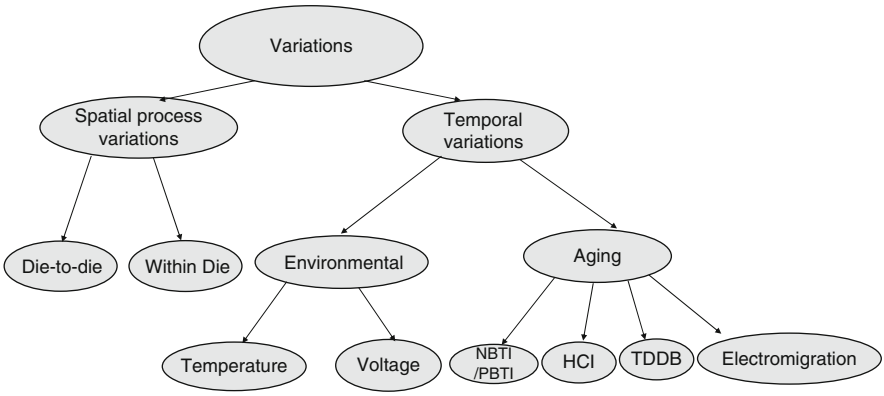



**Fig. 3.2** Taxonomy of process variation

voltage fluctuations fall under the category of *environmental* variations whereas Negative Bias Temperature Instability (NBTI) [15–33], Positive Bias Temperature Instability (PBTI) [34, 35], Hot Carrier Injection (HCI) [36–38], Time-Dependent Dielectric Breakdown (TDDB) [39–41], and electromigration [42] give rise to so-called *aging* variation. Environmental variation like voltage fluctuation is the outcome of circuit switchings and lack of strong power grid. Similarly, temperature fluctuation is the result of excessive power consumption by a circuit that generates heat whereas the package lacks capability to dissipate it. Such variations degrade the robustness of the circuit by increasing speed margins inducing glitches, creating non-uniform circuit delays and thermal runaways. The aging-related variations degrade device strength when a device is used for a long period of time. For example, an IC tested and shipped for 2 GHz of speed may fail to run at the rated speed after 1 or 2 years. Along with device-related variations, the interconnect parameter variations [43] also add to the adversity of the situation.

### *3.1.3 Impact of Variation on Logic*

The variation in $L$, $W$, $T_{OX}$, dopant concentration, etc. modifies the threshold voltage of the devices. One manifestation of statistical variation in $V_{TH}$ is variation of speed between different chips. Hence, if the circuits are designed using nominal $V_{TH}$ of transistors to run at a particular speed, some of them will fail to meet the desired frequency. Such variation in speed would lead to parametric yield loss. In dynamic logic, the parametric variations reduce the noise margin. Another detrimental effect of process variation is leakage. Statistical variation in transistor parameters results in significant spread in different components of leakage. A side effect of increased dynamic and leakage power is localized heating of the die – known as hot-spot. The hot-spots are one of the primary factors behind reliability degradation and thermal runaways.

### *3.1.4 Overview of Variation Insensitive Designs*

In order to address the above-mentioned issues in logic circuits, two approaches can be followed: (1) design time techniques and (2) post-silicon adaptation. The design time techniques include statistical design that has been investigated as an effective method to ensure certain yield criteria. In this approach, the design space is explored to optimize design parameters (e.g., power and reliability) in order to meet timing yield and target frequency. Several gate-level sizing and/or $V_{TH}$ (transistor threshold voltage) assignment techniques [44–55] have been proposed recently addressing the minimization of total power while maintaining the timing yield. Timing optimization to meet certain performance/area/power usually results in circuits with many equally long (critical) paths creating “a wall” in the path delay distribution. Therefore, a number of paths get exposed to process fluctuation induced

delay variation. An uncertainty-aware design technique proposed in [56] describes an optimization process to reduce the wall of equally long paths. A detailed analysis of this approach has been presented in the next chapter.

On the other end of the spectrum, design techniques have been proposed for post-silicon process compensation and process adaptation to deal with process-related timing failures. Adaptive body biasing (ABB) [57, 58] is one such technique, which tries to reduce the frequency and leakage spread, improving the timing yield. Several other research works try to optimize power and yield by mixing gate sizing, multi-$V_{TH}$ with post-Si tuning (e.g., body bias) [59–64]. Due to the superlinear dependence of dynamic power of a circuit on its operating voltage, supply voltage scaling has been extremely effective in reducing the power dissipation. Researchers have investigated logic design approaches that are robust with respect to process variations and, at the same time, suitable for aggressive voltage scaling. One such technique, called RAZOR [65–69], uses dynamic detection and correction of circuit timing errors to tune processor supply voltage. Adaptive (or variable) latency techniques for structures such as caches [70] and combined register files and execution units [71–73] can address device variability by tuning architectural latencies. The fast paths are exercised at rated frequency whereas slower paths due to process variability are operated with extended latency. In [74], Tiwari et al. presented pipelined logic loops for the entire processor using selectable "donor stages" that provide additional timing margins for certain loops. Therefore, extra delay in computation due to process variation in one stage can be mitigated by stealing some extra time from another stage.

It can be noted that pipelined design can experience timing slack either due to frequency or due to supply voltage under process variation. Therefore, both voltage and latency can be selectively picked for each of the pipelined stages for achieving optimal power and performance under variability. In [75–76], the authors describe ReVIVaL – which is a post-fabrication voltage interpolation technique that provides an "effective voltage" to individual blocks within individual processor cores. By coupling variable-latency with voltage interpolation, ReVIVaL provides significant benefits in terms of process resilience and power over implementing variable-latency alone. These architectural techniques for process resilience have been covered in Chapter 7.

Contrary to RAZOR, Ghosh et al. [77] proposed a design time technique called CRISTA to allow voltage over-scaling while meeting the desired frequency and yield. CRISTA isolates the critical (long) paths of the design and provides an extra clock cycle for those paths. The CRISTA design methodology ensures the activation of long paths to be rare, thereby, reducing performance impact. This allows them to drop the supply voltage to expose the timing slack of off-critical (short) paths. Algorithmic noise-tolerance (ANT) [78] is another energy-efficient adaptive solution for low-power broadband communication systems. The key idea behind ANT is to permit errors to occur in a signal processing block and then correct it via a separate error control (EC) block. This allows the main DSP to operate below the specified supply voltage (over-scaled voltage). The important point to note is that proper measures at all levels of hierarchy (from circuits to architecture) is essential

to mitigate the problems associated with parameter variations and to design robust systems.

This chapter presents an overview of the process- and reliability-induced variations. We review several state-of-the-art circuit level adaptive techniques to deal with these issues. The chapter is organized as follows. The impact of process variation and reliability degradations is introduced in Section 3.2. The process variation-tolerant design techniques are presented in Section 3.3 whereas resilience to temporal degradation is discussed in Section 3.4. Finally, the conclusions are drawn in Section 3.5.

## 3.2 Effect of Parameter Variations on Logic – Failures and Parametric Yield

As mentioned before, process variations can be categorized into two broad areas: (a) spatial and (b) temporal. This is further clarified in Fig. 3.3 with an example of chips that are designed for a particular speed. Variations in device strengths between chips belonging to different runs, lots, and wafers result in dies that vary in speed. The figure shows that the speed follows a distribution where each point belongs to chips running at a particular speed. This behavior corresponds to $t$=0 s. Devices also change their characteristic over time due to NBTI (Negative Bias Temperature Stability), PBTI (Positive Bias Temperature Stability), TDDB (Time-Dependent Dielectric Breakdown), HCI (Hot Carrier Injection), etc. For example, the PMOS transistor becomes slower due to NBTI-induced $V_{TH}$ degradation and wire delay may increase due to electromigration of metal. Therefore, the dies become slow and the entire distribution shifts in temporal fashion (Fig. 3.3). The variation in device characteristics over time (i.e., at $t = t'$) is termed as temporal variations.

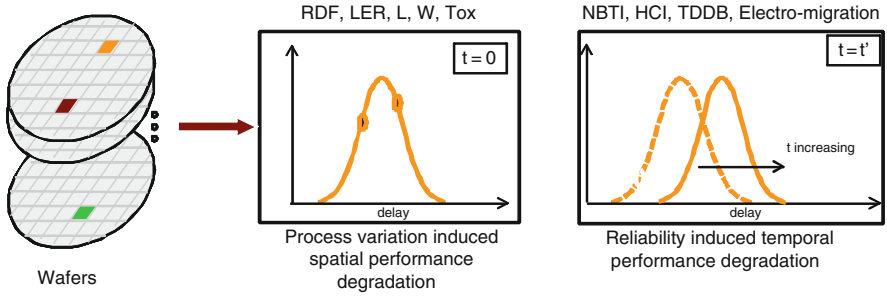



**Fig. 3.3** Spatial and temporal process variation effect on circuit delay. The PDF of chip delay due to temporal degradation has been shown

### *3.2.1 Impact of Spatial Process Variation*

The variation in $L$, $W$, $T_{OX}$, dopant concentration, work function, flat-band condition, etc. modifies the threshold voltage of the devices. These are termed as

die-to-die $V_{TH}$ variation. Another source of variation originates from line edge roughness, line width variation, and random variation in dopant atoms in the channel, and they are known as within-die process variation. In older technology generations, the fraction of within-die variation used to be small compared to die-to-die counterpart. However, with scaling of device dimensions within-die parametric variation has become relatively larger fraction. This is shown in Fig. 3.4. In the following paragraphs, we will discuss the manifestation of various types of spatial process variations.

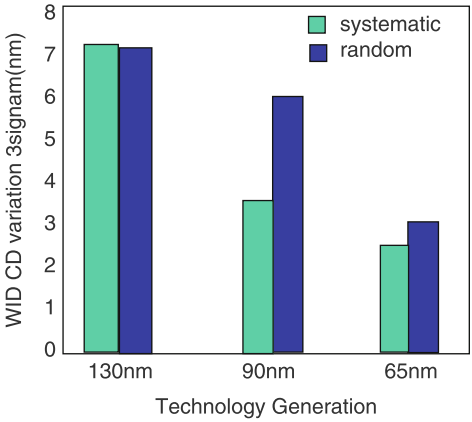



**Fig. 3.4** Growing proportion of within-die process variation with technology scaling [58]

### 3.2.1.1 Increased Delay/Delay Distribution

One manifestation of statistical variation in $V_{TH}$ is variation of speed between different chips. Hence, if the circuits are designed using nominal $V_{TH}$ transistors to run at a particular speed, some of them will fail to meet the desired frequency. Such variation in speed would lead to parametric yield loss. Figure 3.5 shows how the variation in threshold voltage (under the influence of both within-die and die-to-die) translates into speed distribution. The path delay distribution can be represented by normal distribution or by a more accurate model like lognormal distribution. Several statistical static timing analysis techniques have been investigated in the past [79–83] to accurately model the mean and standard deviation (STD) of the circuit delay. The chips slower than the target delay have to be discarded (or can be sold at a lower price). An interesting observation is the impact of within-die and die-to-die process variation. It has been shown in [58] that the effect of within-die process variations tends to average out with the number of logic stages. This is shown in Fig. 3.6. However, the demand for higher frequency necessitates deeper pipeline design (i.e., shorter pipeline depths) making them susceptible to within-die process variation as well.

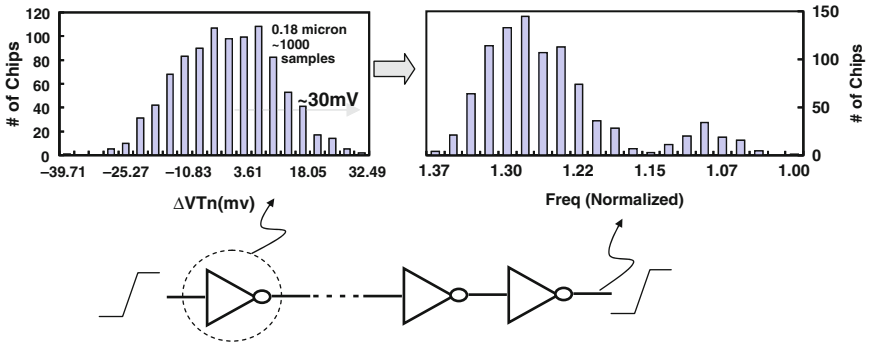
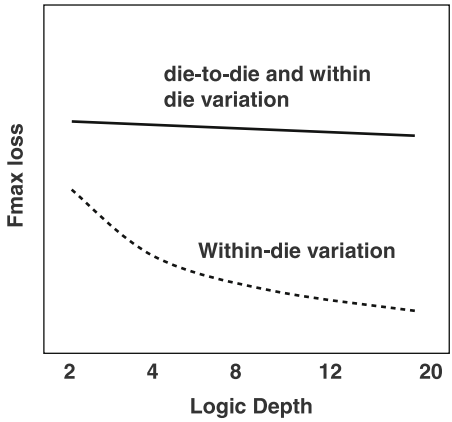



**Fig. 3.5** Variation in threshold voltage and corresponding variation in frequency distribution [58]

**Fig. 3.6** Impact of within-die process variation becomes prominent with increasing pipeline depth [58]




### 3.2.1.2 Lower Noise Margins

The variations have different impacts on dynamic circuits. These circuits operate on the principle of *pre-charge* and *evaluate*. The output is pre-charged to logic "1" in the negative phase of the clock (clk = 0). The positive phase of the clock (clk = 1) allows the inputs decide if the output will be kept pre-charged or will be discharged to ground. Since the information is saved as charge at the output node capacitor, dynamic logic is highly susceptible to noise and timings of input signals. Due to inherent nature of the circuit, a slight variation in transistor threshold voltage can kill the logic functionality. For example, consider a domino logic shown in Fig. 3.7. If the $V_{TH}$ of the NMOS transistors in the second stage is low due to process variation, then a small change in $Out_1$, $IN_4$, or $IN_5$ can turn the pull down path ON and may result in wrong evaluation of $Out_2$.

In register files, increased leakage due to lower $V_{TH}$ dies has forced the circuit designers to upsize the keeper to obtain an acceptable robustness under worst-case

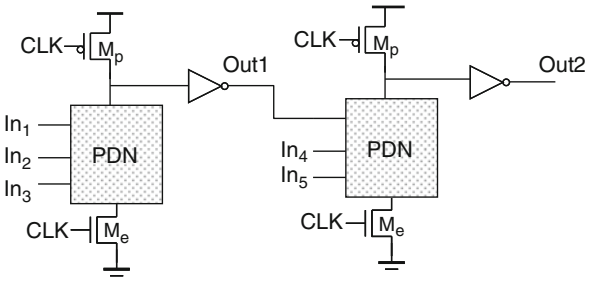



**Fig. 3.7** Example of a dynamic logic circuit. The $V_{TH}$ of NMOS transistor determines the noise margin

$V_{TH}$ conditions. Large variation in die-to-die $V_{TH}$ indicates that (i) a large number of low leakage dies suffer from the performance loss due to an unnecessarily strong keeper, while (ii) the excess leakage dies still cannot meet the robustness requirements with a keeper sized for the fast corner leakage.

#### 3.2.1.3 Degraded Yield in Pipelined Design

Increasing inter-die and intra-die variations in the process parameters, such as channel length, width, threshold voltage, result in large variation in the delay of logic circuits. Consequently, estimating circuit performance and designing high-performance circuits with high yield (probability that the design will meet certain delay target) under parameter variations have emerged as serious design challenges in sub-100 nm regime. In the high-performance design, the throughput is primarily improved by pipelining the data and control paths. In a synchronous pipelined circuit, the throughput is limited by the slowest pipe segment (i.e., segment with maximum delay). Under parameter variations, as the delays of all the stages vary considerably, the slowest stage is not readily identifiable. The variation in the stage delays thus result in variation in the overall pipeline delay (which determines the clock frequency and throughput). Traditionally, the pipeline clock frequency has been enhanced by (a) increasing the number of pipeline stages, which, in essence, reduces the logic depth and hence, the delay of each stage; and (b) balancing the delay of the pipe stages, so that the maximum of stage delays are optimized. However, it has been observed that if intra-die parameter variation is considered; reducing the logic depth increases the variability (defined as the ratio of standard deviation and mean) [58]. Since the pipeline yield is governed by the yield of individual pipe stages, balancing the pipelines may not always be the best way to maximize the yield. This makes the close inspection of the effect of pipeline balancing on the overall yield under parameter variation an extremely important task.

#### 3.2.1.4 Increased Power

Another detrimental effect of process variation is variation in leakage. Statistical variation in transistor parameters results in significant spread in different components of leakage. It has been shown in [5] that there can be ~100X variation in

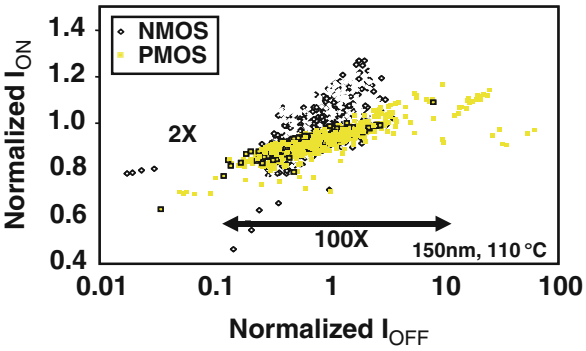



**Fig. 3.8** Leakage spread in 150 nm technology (source: Intel Inc)

leakage current in 150-nm technology (Fig. 3.8). Designing for the worst case leakage causes excessive guard-banding, resulting in lower performance. On the other hand, underestimating leakage variation results in low yield, as good dies are discarded for violating the product leakage requirement. Leakage variation models have been proposed in [84–87] to estimate the mean and variance of the leakage distribution. In [87], the authors provide a complete analytical model for total chip leakage considering random and spatially correlated components of parameters, sensitivity of leakage currents with respect to transistor parameters, input vectors, and circuit topology (spatial location of gates, sizing, and temperature).

#### 3.2.1.5 Increased Temperature

A side effect of increased dynamic and leakage power is localized heating of the die – called hot-spot. The hot-spot is outcome of excessive power consumption by the circuit. The power dissipates as heat and if the package is unable to sink the heat generated by the circuit, then it is manifested as elevated temperature. The hot-spots are one of the primary factors behind reliability degradation and thermal runaways. There have been several published efforts in compact and full-chip thermal modeling and compact thermal modeling. In [88], the authors present a detailed die-level transient thermal model based on full-chip layout, solving temperatures for a large number of nodes with an efficient numerical method. The die-level thermal models in [89] and [90] also provide the detailed temperature distribution across the silicon die. A detailed full-chip thermal model proposed in [91] uses an accurate three-dimensional (3-D) model for the silicon and one-dimensional (1-D) model for the package. A recent work [92] describes HotSpot, a generic compact thermal modeling methodology for VLSI systems. HotSpot consists of models that consider high-level interconnects, self-heating power, and thermal models to estimate the temperatures of interconnect layers at the early design stages.

### 3.2.2 Impact of Temporal Variations

Temporal variations like voltage and temperature fluctuations add to the circuit marginalities. The higher temperature decreases the $V_{TH}$ (good for speed) but reduces the ON current – reducing the overall speed of the design. Similarly, if a circuit that is designed to operate at 1 V works at 950 mV due to voltage fluctuations, then the circuit speed would go below the specified rate. Since the voltage and temperature depends on the operating conditions, the circuit speed becomes unpredictable.

The aging-related temporal variations on the other hand affect the circuit speed systematically over a period of time. In random logic, the impact of NBTI degradation is being manifested as the increase of delays of critical timing paths [19, 22, 23, 26, 27, 29–33] and reduction of subthreshold leakage current [31]. In [25, 32], the authors developed a compact statistical NBTI model considering the random nature of the Si–H bond breaking in scaled transistors. They observed that from the circuit level perspective, NBTI variation closely resembles the nature of RDF-induced $V_{TH}$ variation in a sense that it has a complete randomness even among the transistors that are closely placed in a same chip. Hence, they considered both the RDF- and NBTI-induced $V_{TH}$ variation ($\sigma_{RDF}$ and $\sigma_{NBTI}$, respectively) as follows,

$$\sigma_{V_t} = \sqrt{\sigma_{RDF}^2 + \sigma_{NBTI}^2(t)} \tag{3.1}$$

where $\sigma_{VTH}$ represents the total $V_{TH}$ variation after time $t$.

Figure 3.9a represents the histogram of a simple inverter delay with/without the impact of NBTI variation assuming 3-year stress (for 32-nm PTM [93] using Monte Carlo and Equation (3.1)). As can be observed from the two curves on the right, the variability of gate delay can increase significantly with an added impact of NBTI. Comparison between 32 and 22 nm node results emphasizes the fact that NBTI variations will grow larger in scaled technology. It is important to note that in reality,

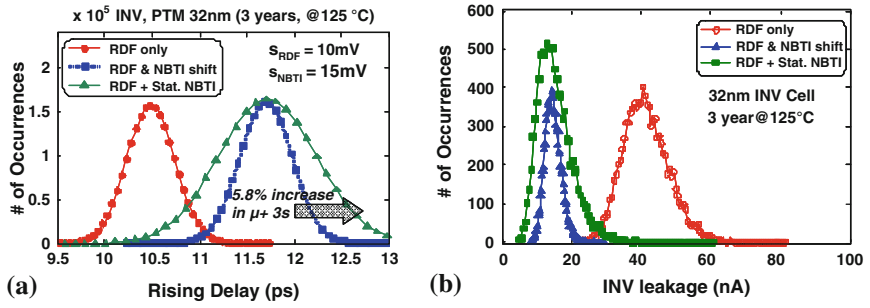



**Fig. 3.9** Variation under NBTI degradation and RDF (**a**) both the mean and spread of inverter gate delay distribution increases due to combined effect of RDF and NBTI. (**b**) inverter leakage is reduced with NBTI however, the spread of leakage can increase due to statistical NBTI effect

the variation of circuit delays are mostly dominated by lower granularity sources such as inter-die variations. And as a result, random $V_{TH}$ variation-induced by NBTI degradation will usually take only a small portion of the overall delay variations [94]. Similar effects can be observed in the subthreshold leakage variation under NBTI degradation. Figure 3.9b represents the histogram plot of inverter leakage. As can be observed, leakage current reduces due to the NBTI. However, the two curves on the left side show that the increased variability of $V_{TH}$ due to NBTI can lead to more variation in leakage.

## 3.3 Variation-Tolerant Design

In the previous section, we discussed failures and yield loss due to variations. This section will provide some advanced circuit level techniques for resilience to parameter variations. These techniques can be broadly categorized as design-time (pre-Si) and run-time (post-Si) techniques. Figure 3.10 shows the taxonomy of variation-resilient schemes. The design time techniques are further categorized into (a) conservative design, (b) statistical design (Gate sizing [52–55], dual $V_{TH}$ assignment [44–51], pipeline unbalancing [95]), and (c) resilient design (CRISTA [77]). Adaptive body biasing [57], adaptive voltage scaling, programmable sizing [96], sensor-based design [97–99], RAZOR [65–69], etc. fall under the category of post-Silicon techniques. In the following subsection, we present these methodologies in detail.

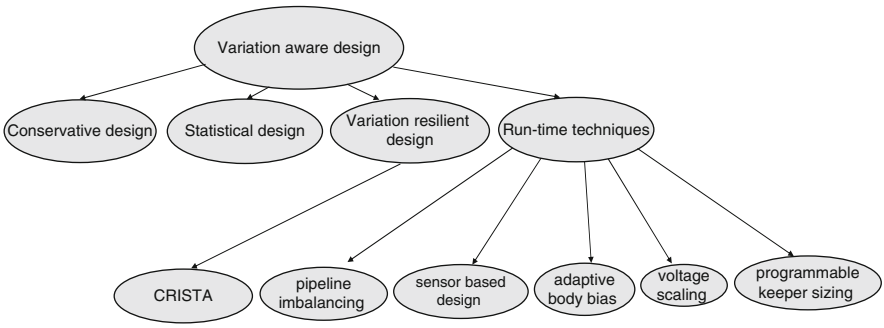



**Fig. 3.10** Taxonomy of variation resilient circuit design

### *3.3.1 Conservative Design*

Conservative approach to design circuits is based on providing enough timing margins for uncertainties, e.g., process variation, voltage/temperature fluctuations, and temporal degradation ($T_m$ as shown in Fig. 3.11a). One possible option is to

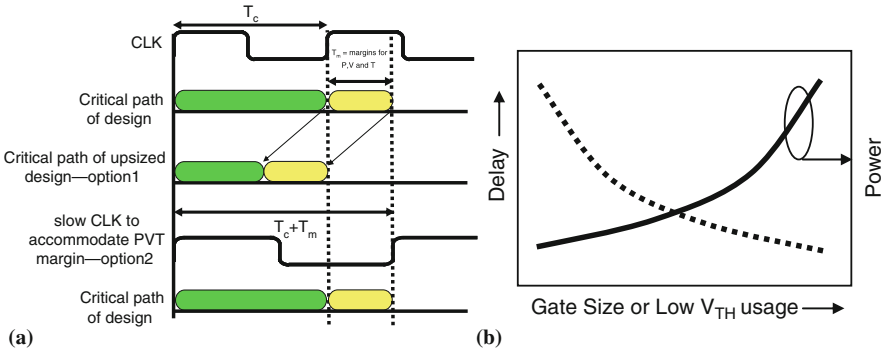



**Fig. 3.11** (**a**) Two possible options to tolerate process, voltage and temperature fluctuation induced timing uncertainties namely, upsizing of devices and slow down of clock frequency, (**b**) tradeoff between power and delay in conservative design approach

conservatively size the circuit to meet the target frequency after adding all margins or boosting up the supply voltage. This is shown in Fig. 3.11a as option-1. However, from Fig. 3.11b, it is obvious that conservative approach is area and power intensive. Both supply voltage and circuit size upscaling increases the power consumption. One can definitely trade power/area by slowing down the clock frequency (option-2). However, this upfront penalty cannot be accepted in today's competitive market where power and operating frequency are dominating factors. From the above discussions, it is apparent that conservative design approach is not very desirable in scaled technologies.

## *3.3.2 Statistical Design*

### 3.3.2.1 Logic Sizing

It is well known that the delay of a gate can be manipulated by modifying its size [52–55]. Either the length or width of the gate can be tweaked to modulate the (*W*/*L*) ratio and control its drive ability. Since process variations may increase the delay of the circuit, many design time transistor sizing algorithms have been proposed in [52–55] to reduce the mean and STD of delay variations. The next chapter presents an in-depth analysis of various sizing techniques.

### 3.3.2.2 Sizing and Dual $V_{TH}$

Threshold voltage of the transistor is another parameter that can be adjusted to achieve a tradeoff between speed and leakage [44–51]. A new statistically aware dual-$V_{TH}$ and sizing optimization has been suggested in [51] that considers both the variability in performance and leakage of a design. Further details on simultaneous gate sizing and dual $V_{TH}$ can be found in the next chapter.

#### 3.3.2.3 Pipeline Unbalancing

Shrinking pipeline depths in scaled technologies makes the path delay prone to within-die variations [81, 95]. By overdesigning the pipeline stages, better yield can be achieved at the cost of extra area and power. A framework to determine the yield of the pipeline under the impact of process variation has been proposed in [81]. This framework has been successfully employed to design high yield pipeline while compromising the area/power overhead [95] by utilizing the concept of pipeline unbalancing. This technique advocates slowing down certain pipeline stages (by downsizing the gates) and accelerating the other stages (by upsizing) to achieve overall better pipeline yield under iso-area. For example, consider a three-stage pipeline design. Let us also assume that the combinational logic of each stage is optimized for minimum area for a specific target pipeline delay and yield. If the stage yield is $y$, then pipeline yield $= (y)^3$. Now if the pipeline is carefully optimized to achieve the pipeline stages yields to be $y_0$, $y_1$, and $y_2$ (under iso-area), one can improve yield if $(y_0 y_1 y_2) > (y)^3$. This can be elucidated by assuming that the sizing of pipeline stages are done such that $y = 0.9$ and $y_0 = 0.98$, $y_1 = 0.95$, and $y_2 = 0.85$. It is obvious that better pipeline yield can be achieved by pipeline unbalancing (yield = 79.1%) rather than balanced pipeline (yield=72.9%).

### *3.3.3 Resilient Design*

#### 3.3.3.1 CRISTA

It is a novel paradigm for low-power, variation, and temperature tolerant circuits and system design, which allows aggressive voltage over-scaling at rated frequency. The CRISTA design principle [77] (a) isolates and predicts the paths that may become critical under process variations, (b) ensures (by design) that they are activated rarely, and (c) avoids possible delay failures in such long paths by adaptively stretching the clock period to two-cycles. This allows the circuit to operate at reduced supply voltage while achieving the required yield with small throughput penalty (due to rare two-cycle operations). The concept of CRISTA is shown in Fig. 3.12a with example of three pipelined instructions where the second instruction activates the critical path. CRISTA can be performed by gating the third clock pulse during the execution of second instruction for correct functionality of the pipeline at scaled supply. The regular clock and CRISTA-related clock-gating is shown in Fig. 3.12a for the sake of clarity. The CRISTA design methodology is applied to random logic by hierarchical Shannon expansion and gate sizing (Fig. 3.12b). Multiple expansions reduce the activation probability of the paths. In the example shown in Fig. 3.12b, critical paths are restricted within $CF_{53}$ (by careful partitioning and gate sizing). The critical paths are activated only when $x_1!x_2x_3 = 1$ with activation probability of 12.5% assuming that each signal can be logic "1" 50% of the time. CRISTA allows aggressive supply voltage scaling to improve power consumption by ~40% with only 9% area and small throughput penalty for a two-stage pipelined ALU

[100] compared to standard approach to designing circuits. Note that CRISTA can be applied to circuits as well as micro-architecture level [101].

Figure 3.12c shows application of CRISTA for dynamic thermal management in a pipelined processor. Since execution (EX) unit is statistically found to be one of the hottest components, CRISTA is employed to allow voltage over-scaling in EX stage. The critical paths of EX stage are predicted by decoding the inputs using a set of pre-decoders. At nominal temperature, the pre-decoders are disabled; however, at elevated temperatures the supply voltage of the execution unit is scaled down and pre-decoders are enabled. This ensures cooling down of the system and occasional two-cycle operations in EX stage (at rated frequency) whenever the critical path is activated. If the temperature still keeps rising and crosses the emergency level then conventional dynamic voltage frequency scaling (DVFS) is employed for throttling the temperature down. DVFS is often associated with stalling while the PLL frequency is locked at the desired level. It is interesting to note that CRISTA can avoid triggering of DVFS for all SPEC2000 benchmark programs, saving throughput loss. This is evident from Fig. 3.12d that shows activation count of DVFS with and without CRISTA.

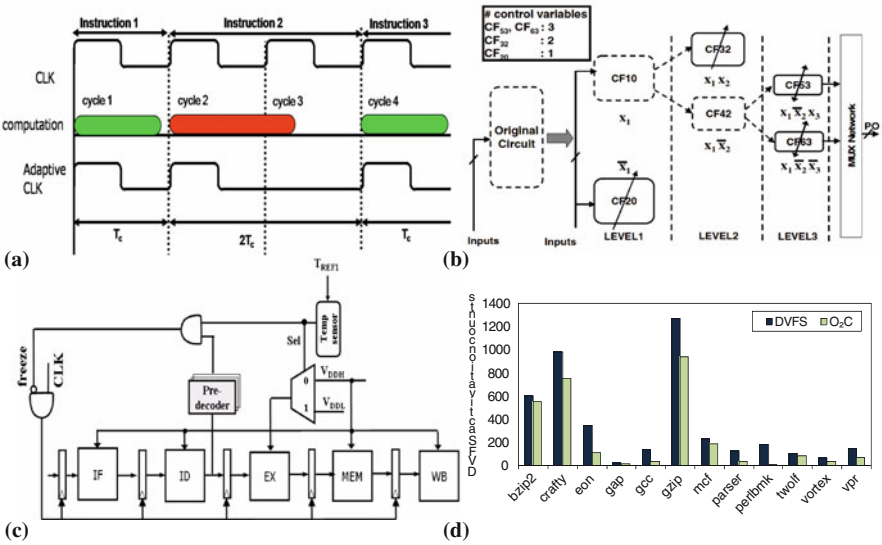



**Fig. 3.12** (**a**) Timing diagram of CRISTA paradigm. Long paths are activated rarely and they are evaluated in two cycles. (**b**) Shannon expansion based CRISTA design methodology for random logic. Multiple expansions reduce the activation probability of the paths. In this example, critical paths are restricted within $CF_{53}$ (by careful partitioning and gate sizing) which is activated when $x_1!x_2x_3 = 1$ with activation probability of 12.5%. (**c**) CRISTA at micro-architectural level of abstraction for adaptive thermal management [77]. (**d**) Activation count of dynamic voltage frequency scaling (DVFS) and CRISTA. CRISTA avoids application of DVFS saving valuable overhead in terms of throughput

## *3.3.4 Run Time Techniques*

### 3.3.4.1 Adaptive Body Bias

As discussed before, threshold voltage of the transistor can be tuned carefully to speed up the design [58]. However, lowering the $V_{TH}$ also changes the on-current. It has been noted that transistor $V_{TH}$ is a function of body to source voltage ($V_{BS}$). Hence body potential of the transistor can be modulated to achieve speed or lower the leakage power. The forward body bias to PMOS transistors improves the performance due to lower $V_{TH}$ while the reverse body bias reduces leakage due to higher $V_{TH}$. In [58], the authors propose adaptive body biasing technique on selective dies to improve the power and performance. The faster and leakier dies are reverse body biased while slower dies are forward biased. This results in tighter delay distribution as shown in Fig. 3.13a.

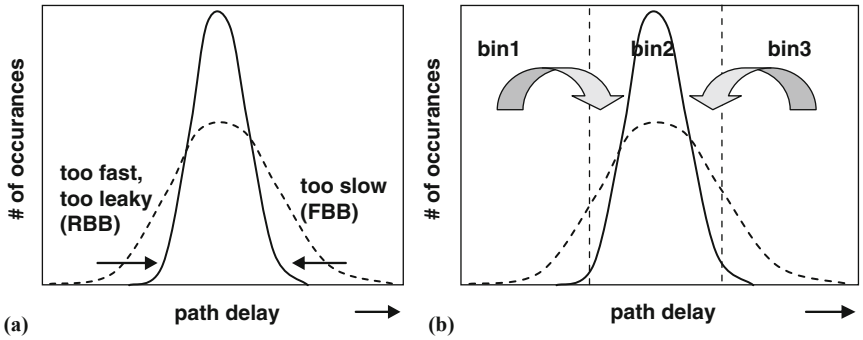



**Fig. 3.13** (**a**) Effect of adaptive body bias – slower paths are forward body biased whereas faster paths are reversed body biased to squeeze the distribution. (**b**) Effect of adaptive voltage scaling – voltage boosted up (down) for slower (faster) dies. The chips in slower/discarded bins move to faster bins

### 3.3.4.2 Adaptive Voltage Scaling

Circuit speed is a strong function of supply voltage [58]. Although increasing the supply voltage speeds up the circuit, it also worsens the power consumption. However, careful tuning of supply voltage can indeed improve the frequency while staying within the power budget. For example, the slower dies already consume low power due to high $V_{TH}$ transistors. Therefore, supply voltage of these dies can be boosted up to improve the frequency. Similarly, the supply voltage of faster and power hungry dies can be scaled down to save power dissipation while affecting the speed minimally. In [58], the authors propose adaptive supply voltage technique to improve frequency and meet power specification. As shown in Fig. 3.13b, The slower dies from bin3 can be moved to nominal speed bin2 and faster dies from bin1 to bin2.

### 3.3.4.3 Programmable Designs

Several post-Silicon techniques have been proposed to tune the design and improve the robustness [96]. One such example is [96], where the authors describe a Process-Compensating Dynamic (PCD) circuit technique for register files. Unlike prior fixed-strength keeper techniques [102], the keeper strength is optimally programmed based on the respective die leakage. Figure 3.14 shows the PCD scheme with a digitally programmable three-bit keeper applied on an eight-way register file local bitline (LBL). Each of the three binary-weighted keepers with respective widths $W$, $2W$, and $4W$ can be activated or deactivated by asserting appropriate globally routed signals b[2:0]. A desired effective keeper width can be chosen among $[0, W, 2\,W, \ldots, 7\,W]$. The programmable keeper improves robustness and delay variation spread by restoring robustness of worst case leakage dies and improving performance of low-leakage dies.

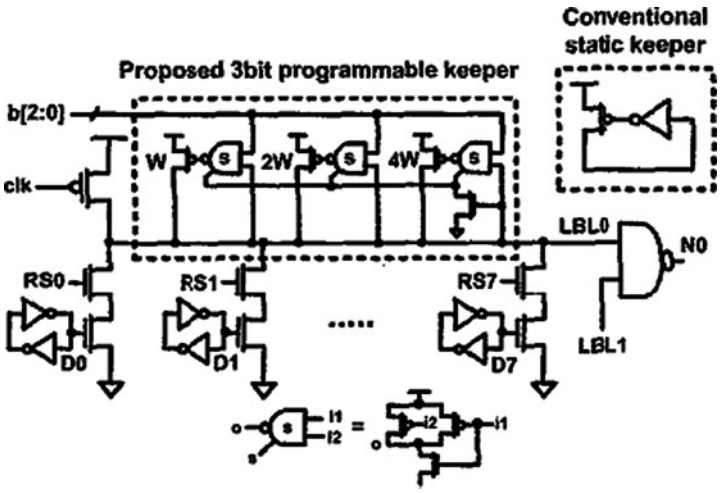



**Fig. 3.14** Register file 8-way LBL with proposed PCD technique [96]

### 3.3.4.4 Sensors Based Design

Design time techniques are prone to errors due to operating PVT conditions [97–99]. Therefore, on-chip sensors can be used to estimate the extent of variations and apply right amount of corrective actions to avoid any possible failures. A variation-resilient circuit design technique is presented in [97] for maintaining parametric yield under inherent variation in process parameters. It utilizes on-chip phase locked loop (PLL) as a sensor to detect process, $V_{DD}$, and temperature (PVT) variations, or even temporal degradation stemming from negative bias temperature instability (NBTI) and uses adaptive body bias for correction. Several similar sensor-based adaptive design techniques can be found in literature. For example, [98] presents an on-chip wearout detection circuit whereas an adaptive

multichip processor based on online NBTI and oxide breakdown detection sensors has been presented in [99]. Details of this approach can be found in following chapters.

#### 3.3.4.5 RAZOR

RAZOR [65–69] uses dynamic detection and correction of circuit timing errors by using a shadow latch to tune processor supply voltage. This technique is quite effective in eliminating the timing margins due to process variation and environmental fluctuations. More details on RAZOR can be found in Chapter 7.

Note that the power management techniques, e.g., supply voltage scaling, power gating, multiple-$V_{DD}$, and $V_{TH}$ designs further magnify the problems associated with process-induced variations. Power and process variation resilience are therefore conflicting design requirements and one comes at the cost of other. Meeting a desired power specification with certain degree of process tolerance is a stiff challenge and topic of further research.

## 3.4 Temporal Variability Management Techniques

In previous section, we discussed various pre-Si and post-Si adaptive techniques to overcome the impact of process-induced spatial variation. This section will summarize the recent techniques to tolerate temporal degradation (voltage, temperature, and NBTI). These techniques will be described in the following paragraphs.

### *3.4.1 Voltage and Temperature Fluctuation Tolerance*

Voltage fluctuation results from a combination of higher switching activity of the underlying circuit and weak power grid. This type of variation can be tolerated by inserting large decoupling capacitors in the power grid. The uneven temperature distribution within the chip is the outcome of excessive power consumption by the circuit. The power dissipates as heat and if the package is unable to sink the heat generated by the circuit, then it is manifested as elevated temperature. Since different parts of the chip experience different switching activity and power consumption, the temperature profile of the chip also varies accordingly over time. The temperature can be managed by reducing the power consumption of the die. For example, the operating frequency can be slowed down or the supply voltage can be scaled down. One can also throttle the heat by simultaneous control of voltage as well as frequency for cubic reduction in power consumption. Several other techniques have been proposed past like, logic shutdown, clock gating, functional block duplication, etc.

### 3.4.2 Gate/TR Sizing

One of the first approaches for reliability-aware design was based on an optimal gate sizing [27, 29]. For example, the method proposed in [27] uses modified Lagrangian Relaxation (LR) algorithm to compute optimal size of each gate under NBTI degradation. The basic idea of this approach is conceptually described in Fig. 3.15. As can be seen, the setup time margin of the design reduces with time due to NBTI, and after certain stress period ($T_{NBTI}$), the design may fail to meet the given timing constraint ($D_{CONST}$). To avoid such failures, the authors over-design (transistor upsizing considering signal probabilities) to ensure right functionality even after the required product lifetime $T_{REQ}$. The results from [27] reported an average area overhead of 8.7% for ISCAS benchmark circuits to ensure 3 year lifetime functionality at 70-nm node. An alternative method of transistor-level sizing was also proposed in [29]. In this approach, rather than sizing both the PMOS and NMOS at the same time, the authors applied different sizing factor for each of them. This method could effectively reduce unnecessary slacks in NMOS network (which is not affected by NBTI) and lower the overall area overhead. The results from [27] reported that the average overhead reduced by nearly 40% compared to [27].

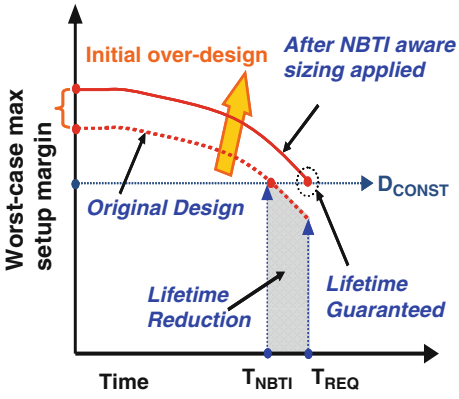



**Fig. 3.15** Transistor sizing for NBTI tolerance

### 3.4.3 Technology Mapping and Logic Synthesis

An alternative method is to consider NBTI at the technology mapping stage of the logic synthesis [19]. This approach is based on the fact that different standard cells have different NBTI sensitivity with respect to their input signal probabilities (probability that signal is logic "1"). With this knowledge, standard cell library is re-characterized with an additional signal probability dependency [20]. During logic synthesis, this new library is applied to properly consider the impact of

NBTI degradation. An average of 10% reduction in area overhead compared to the worst-case logic synthesis was reported using this method.

### *3.4.4 Guard-Banding*

This is a very basic technique where the delay constraints are tightened with a fixed amount of delay guard-banding during sizing. Guard-band is selected as average delay degradations in these circuits for its lifetime years. Though guard-banding ignores the sensitivity of individual gates with respect to NBTI, delay degradation is weakly dependent on how the circuits were originally designed. Interestingly, it has been demonstrated in [33] that for a well-selected delay constraint; guard-banding indeed produces results comparable to the previous two approaches.

### *3.4.5 Self-Correction Using On-Chip Sensor Circuits*

In practice, the guard-banding, sizing, and the synthesis methods introduced above require an accurate estimation of NBTI-induced performance degradation [24, 97, 103, 104]. This estimation may produce errors due to unpredictable temperature, activity (signal probability), or process parameter variations. One way to handle these problems is to employ an active on-chip reliability sensor [24, 103, 104]. Another approach proposed in [97] utilizes the on-chip phase locked loop (PLL) to perform reliability sensing. In these approaches, the actual amount of NBTI degradation can be used for further corrective actions. For example, in [97] the detected signal is efficiently transformed into an optimal body-bias signal to avoid possible timing failures in the target circuit. However, it is also essential to consider the additional design overhead-induced (e.g., increased area, power, and interconnections) by body biasing and the sensor circuits. It should be further noted that self-correcting design techniques lets one design circuits using nominal or optimistic conditions (leading to lower power dissipation). Corrective actions, if required, are taken based on sensing NBTI/parameter degradations.

## 3.5 Conclusions

Parameter variation is becoming an important issue with scaling of device geometries. In this chapter, we described various sources of variations and their impact on logic circuits. We also demonstrated that variation-aware or error-aware design is essential to stay within the power/performance envelop while meeting the yield requirement. We presented various variation insensitive circuit design techniques to address these issues.

## References


1. Asenov A, Brown AR, Davies JH, Kaya S, Slavcheva G (Sept 2003) Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs. IEEE Trans Electron Devices 50(9):1837–1852
2. Hane M, Kawakami Y, Nakamura H, Yamada T, Kumagai K., Watanabe Y (2003) A new comprehensive SRAM soft error simulation based on 3D device simulation incorporating neutron nuclear reactions. In: Proceeding of simulation of semiconductor processes and devices, Boston, MA, pp 239–242
3. Nassif SR (2001) Modeling and analysis of manufacturing variations. In: Proceeding of custom integrated circuit conf., San Diego, CA, pp 223–228
4. Visweswariah C (2003) Death, taxes and failing chips. In: Proceeding of design automation conference, Anaheim, CA, pp 343–347
5. Borkar S, Karnik T, Narendra S, Tschanz J, Keshavarzi A, De V (2003) Parameter variation and impact on circuits and microarchitecture. In: Proceeding of design automation conference, Anaheim, CA, pp 338–342
6. Bhavnagarwala A, Tang X, Meindl JD (2001) The impact of intrinsic device fluctuations on CMOS SRAM cell stability. IEEE J Solid State Circuits 36:658–665
7. Tang X, De V, Meindl JD (1997) Intrinsic MOSFET parameter fluctuations due to random dopant placement. Trans VLSI syst 5:369–376
8. Raychowdhury A, Keshavarzi A (2008) Theory of multi-tube carbon nanotube transistors for high speed variation-tolerant circuits. In: Proceeding of device research conference, Santa Barbara, CA, pp 23–24
9. Nieuwoudt A, Massoud Y (2007) Assessing the implications of process variations on future carbon nanotube bundle interconnect solutions. In: Proceeding of international symposium on quality electronic design, San Francisco, California
10. Patil N, Deng J, Wong HSP, Mitra S (2007) Automated design of misaligned-carbon-nanotube-immune circuits. In: Proceeding of design automation conference, San Diego, California, pp 958–961
11. Zhang J, Patil N, Hazeghi A, Mitra S (2009) Carbon nanotube circuits in the presence of carbon nanotube density variations. In: Proceeding of design automation conference, San Francisco, California, pp 71–76
12. Bobba S et al (2009) Design of compact imperfection-immune CNFET layouts for standard-cell-based logic synthesis. In: Proceeding of Design, Automation & Test in Europe, Nice
13. Borkar S et al (2005) Statistical circuit design with carbon nanotubes. U.S. Patent Application 20070155065
14. Gunther SH, Binns F, Carmean DM, Hall JC (2001) Managing the impact of increasing microprocessor power consumption. Intel Tech J 5, (1):1–9
15. Deal BE, Sklar M, Grove AS, Snow EH (1967) Characteristics of the surface-state charge (Qss) of thermally oxidized silicon. J Electrochem Soc 114:266
16. Nicollian EH, Brews JR (1982) MOS physics and technology. Wiley, New York, NY
17. Blat CE, Nicollian EH, Poindexter EH (1991) Mechanism of negative bias temperature instability. J Appl Phys 69:1712
18. Li MF et al (2004) Dynamic bias-temperature instability in ultrathin $SiO_2$ and $HfO_2$ metal-oxide semiconductor field effect transistors and its impact on device lifetime. Jpn J Appl Phys 43:7807–7814, November
19. Kumar SV, Kim CH, Sapatnekar SS (2007) NBTI-aware synthesis of digital circuits. In: Proceedings of the ACM/IEEE design automation conference, San Diego, CA, pp 370–375
20. Kumar SV, Kim CH, Sapatnekar SS (2006) An analytical model for negative bias temperature instability. In: Proceedings of the IEEE/ACM international conference on computer-aided design, San Jose, CA, pp 493–496

21. Kumar SV, Kim CH, Sapatnekar SS (2006) Impact of NBTI on SRAM read stability and design for reliability. In: Proceedings of the international symposium on quality electronic design, San Jose, CA, pp 210–218
22. Kumar SV, Kim CH, Sapatnekar SS (2009) Adaptive techniques for overcoming performance degradation due to aging in digital circuits. In: Proceedings of the Asia-South Pacific design automation conference, Yokohama, pp 284–289
23. Kumar S, Kim CH, Sapatnekar S (2007) NBTI-aware synthesis of digital circuits. In: Proc. design automation conf., Dan Diego, CA, pp 370–375
24. Karl E, Singh P, Blaauw D, Sylvester D (Feb 2008) Compact in-situ sensors for monitoring negative-bias-temperature-instability effect and oxide degradation. IEEE international solid-state circuits conference, 2008 (ISSCC 2008). Digest of technical papers, San Francisco, CA, pp 410–623, 3–7
25. Wang W, Reddy V, Krishnan AT, Krishnan S, Cao Y (2007) An integrated modeling paradigm of circuit reliability for 65 nm CMOS technology. In: Proceeding of custom integrated circuits conference, San Jose, CA
26. Wang W, Wei Z, Yang S, Cao Y (2007) An efficient method to identify critical gates under circuit aging. In: Proceedings of the international conference on computer aided design (ICCAD) San Jose, CA
27. Kang K, Kufluoglu H, Alam MA, Roy K (2006) Efficient transistor-level sizing technique under temporal performance degradation due to NBTI. In: Proceeding of international conference on computer design, San Jose, CA, pp 216–221
28. Kunfluoglu H (2007) MOSFET degradation due to negative bias temperature instability (NBTI) and hot carrier injection (HCI) and its implications for reliability-aware VLSI design. PhD dissertation, Purdue University
29. Paul BC, Kang K, Kuflouglu H, Alam MA, Roy K (2006) Temporal performance degradation under NBTI: estimation and design for improved reliability of nanoscale circuits. In: Proc. design automation and test in Europe, Munich, pp 780–785
30. Paul BC, Kang K, Kufluoglu H, Alam MA, Roy K (2005) Impact of NBTI on the temporal performance degradation of digital circuits. IEEE Electron Device Lett 26(8):560–562
31. Kang K, Alam MA, Roy K (2007) Characterization of NBTI induced temporal performance degradation in nano-scale SRAM array using IDDQ. In: Proc. intl. test conference, Santa Clara, CA, pp 1–10
32. Kang K, Park SP, Roy K, Alam MA (2007) Estimation of statistical variation in temporal NBTI degradation and its impact in lifetime circuit performance. In: Proc. international conference on computer aided design, San Jose, CA, pp 730–734
33. Kang K, Gangwal S, Park SP, Roy K (2008) NBTI induced performance degradation in logic and memory circuits: how effectively can we approach a reliability solution?. In: Proc. Asia and South Pacific design automation conference, Seoul, pp 726–731
34. Jafar S, Kim YH, Narayanan V, Cabral C, Paruchuri V, Doris B, Stathis J, Callegari A, Chudzik M (2006) A comparative study of NBTI and PBTI (charge trapping) in SiO2/HfO2 stacks with FUSI, TiN, Re gates. In: Proceeding of VLSI circuits, Honolulu, HI, pp 23–25
35. Crupi F et al (Jun 2005) Positive bias temperature instability in nMOSFETs with ultra-thin Hf-silicate gate dielectrics. J Microelectron Eng 80:130–133
36. Ning TH, Cook PW, Dennard RH, Osburn CM, Schuster SE, Yu H (1979) 1 μm MOSFET VLSI technology: part IV-Hot electron design constraints. Trans Electron Devices 26: 346–353
37. Abramo A, Fiegna C, Venturi F (1995) Hot carrier effects in short MOSFETs at low applied voltages. In: Proc. intl. electron device meeting, Washington, DC, pp 301–304
38. Taur Y, Ning TH (1998) Fundamentals of modern VLSI devices. Cambridge University Press, New York, NY
39. JEP122-A (2002) Failure mechanisms and models for semiconductor devices. JEDEC Publication, JEDEC solid state technology association

40. Quddus MT, DeMassa TA, Sanchez JJ (2000) Unified model for Q(BD) prediction for thin gate oxide MOS devices with constant voltage and current stress. Microelectron Eng 51–52:357–372
41. MA Alam, Weir B, Silverman A (2002) A future of function or failure. IEEE Circuits Devices Mag 18:42–48
42. D Young, Christou A (1994) Failure mechanism models for electromigration. IEEE Trans Reliab 43:186–192
43. Boning D, Nassif S (2001) Models of process variations in device and interconnect. Design of high performance microprocessor circuits. Wiley, New York, NY.
44. Sirichotiyakul S et al (Apr 2002) Duet: an accurate leakage estimation and optimization tool for dual-Vt circuits. IEEE Trans VLSI Syst 10:79–90
45. Pant P, Roy R, Chatterjee A (2001) Dual-threshold voltage assignment with transistor sizing for low power CMOS circuits. IEEE Trans VLSI Syst 9:390–394
46. Wei L et al (1998) Design and optimization of low voltage high performance dual threshold CMOS circuits. In: Proceeding of design automation conference, San Francisco, CA, pp 489–494
47. Karnik T et al (2002) Total power optimization by simultaneous dual-Vt allocation and device sizing in high performance microprocessors. In: Proceeding of design automation conference, New Orleans, LA, pp 486–491
48. Nguyen D et al (2003) Minimization of dynamic and static power through joint assignment of threshold voltages and sizing optimization. In: Proceeding of international symposium on low-power electronics design, Seoul, pp 158–163
49. Srivastava A et al (2003) Simultaneous Vt selection and assignment for leakage optimization. In: Proceeding of international symposium on low-power electronics design, Seoul, pp 146–151
50. Sundarajan V, Parhi K (1999) Low power synthesis of dual threshold voltage CMOS VLSI circuits. In: Proceeding of international symposium on low-power electronics design, San Diego, CA, pp 139–144
51. Srivastava A, Sylvester D, Blauuw D, Agarwal A (2004) Statistical optimization of leakage power considering process variations using dual-$V_{TH}$ and sizing. In: Proceeding of design automation conference, San Diego, CA, pp 773–778
52. Ketkar M et al (2000) Convex delay models for transistor sizing. In: Proceeding of design automation conference, Los Angeles, CA, pp 655–660
53. Singh J, Nookala V, Luo Z-Q, Sapatnekar S (2005) Robust gate sizing by geometric programming. In: Proceeding of DAC, Anaheim, CA, pp 315–320
54. Choi SH, Paul BC, Roy K (2004) Novel sizing algorithm for yield improvement under process variation in nanometer. In: Proceeding of design automation conf., San Diego, CA, pp 454–459
55. Chen C-P, Chu CCN, Wong DF (1999) Fast and exact simultaneous gate and wire sizing by Lagrangian relaxation. IEEE Trans Comput Aided Des, 18:1014–1025
56. Bai X, Visweswariah C, Strenski PN, Hathaway DJ (2002) Uncertainty-aware circuit optimization. In: Proceeding of design automation conf., New Orleans, LA, pp 58–63
57. Borkar S, Karnik T, De V (2004) Design and reliability challenges in nanometer technologies. In: Proceeding of design automation conference, San Diego, CA, pp 75–75
58. Borkar S, Karnik T, Narendra S, Tschanz J, Keshavarzi A, De V (2003) Parameter variations and impact on circuits and microarchitecture. DAC, Anaheim, CA, pp 338–342
59. Kumar SV, Kim CH, Sapatnekar SS (2006) Mathematically-assisted adaptive body bias (ABB) for temperature compensation in gigascale LSI systems. In: Proceedings of the Asia-South Pacific design automation conference, Yokohama, pp 559–564
60. Kumar SV, Kim CH, Sapatnekar SS (Mar 2008) Body bias voltage computations for process and temperature compensation. IEEE Trans VLSI Syst 16(3):249–262
61. Zhuo C, Blaauw D, Sylvester D (2008) Variation-aware gate sizing and clustering for post-silicon optimized circuits. ISLPED, Bangalore, pp 105–110

62. Kulkarni S, Sylvester D, Blaauw D (2006) A statistical framework for post-silicon tuning through body bias clustering. In: Proceeding of ICCAD, San Jose, CA, pp 39–46
63. Mani M, Singh A, Orshansky M (2006), Joint design-time and postsilicon minimization of parametric yield loss using adjustable robust optimization. In: Proceeding of ICCAD, San Jose, CA, pp 19–26
64. Khandelwal V, Srivastava A (2006) Variability-driven formulation for simultaneous gate sizing and post-silicon tunability allocation. In: Proceeding of ISPD, Austin, TA, pp 17–25
65. Ernst D, Kim NS, Das S, Pant S, Pham T, Rao R, Ziesler C, Blaauw D, Austin T, Mudge T (2003) Razor: a low-power pipeline based on circuit-level timing speculation. In: Proceeding of international symposium on microarchitecture, pp 7–18
66. Bowman KA, Tschanz JW, Nam Sung Kim Lee JC, Wilkerson CB, Lu S-LL, Karnik T, De VK (2008) Energy-efficient and metastability-immune timing-error detection and instruction-replay-based recovery circuits for dynamic-variation tolerance. In: Solid-state circuits conference (ISSCC 2008), San Francisco, CA, pp 402–623
67. Blaauw D, Kalaiselvan S, Lai K, Ma W-H, Pant S, Tokunaga C, Das S, Bull D (Feb 2008) RazorII: in-situ error detection and correction for PVT and SER tolerance. In: IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA
68. Austin T, Blaauw D, Mudge T, Flautner K (Mar 2004) Making typical silicon matter with Razor. IEEE Comput 37(3):57–65
69. Ernst D, Das S, Lee S, Blaauw D, Austin T, Mudge T, Nam Sung Kim, Flautner K (Nov-Dec 2004) Razor: circuit-level correction of timing errors for low-power operation. IEEE 24(6):10–20
70. Liang X, Wei G, Brooks D (Dec 2007) Process variation tolerant 3T1D based cache architectures. In: IEEE international symposium on microarchitecture, Chicago, IL
71. Liang X, Brooks D (Dec 2006) Mitigating the impact of process variations on processor register files and execution units. In: IEEE international symposium on microarchitecture, Orlando, FL
72. Ghosh V, Mahapatra D, Karakonstantis G, Roy K (Sep 2010) Low-voltage high-speed robust hybrid arithmetic units using adaptive clocking. IEEE Trans VLSI (accepted) 18:1301–1309
73. Mohapatra D, Karakonstantis G, Roy K (2007) Low-power process-variation tolerant arithmetic units using input-based elastic clocking. In: ISLPED, Portland, OR, pp 74–79
74. Tiwari A, Sarangi SR, Torrellas J (2007) Recycle: pipeline adaptation to tolerate process variation. In: Proceedings of the international symposium on computer architecture, San Diego, CA
75. Liang X, Wei G-Y, Brooks D (Jun 2008) ReVIVaL: a variation tolerant architecture using voltage interpolation and variable latency. In: Proceedings of the international symposium on computer architecture (ISCA-35), Beijing
76. Liang X, Brooks D, Wei G-Y (Feb 2008) A process-variation-tolerant floating-point unit with voltage interpolation and variable latency. In: IEEE international solid-state circuits conference, San Francisco, CA, pp 404–623
77. Ghosh S, Bhunia S, Roy K (2007) CRISTA: a new paradigm for low-power and robust circuit synthesis under parameter variations using critical path isolation. IEEE Trans Comput Aided Des
78. Shanbhag NR (2002) Reliable and energy-efficient digital signal processing. In: Proceeding of design automation conference, New Orleans, LA, pp 830–835
79. Kumar SV, Kashyap C, Sapatnekar SS (2008) A framework for block-based timing sensitivity analysis. In: Proceedings of the ACM/IEEE design automation conference, Anaheim, CA, pp 688–693
80. Blaauw D, Chopra K, Srivastava A, Scheffer L (2008) Statistical timing analysis: from basic principles to state of the art. IEEE Trans Comput Aided Des 27:589–607
81. Datta A, Bhunia S, Mukhopadhyay S, Banerjee N, Roy K (2005) Statistical modeling of pipeline delay and design of pipeline under process variation to enhance yield

in sub-100 nm technologies. In: Proceeding of design automation and test in Europe, pp 926–931
82. Orshansky M, Keutzer K (2002) A general probabilistic framework for worst case timing analysis. In: Design automation conference, New Orleans, LA, pp 556–561
83. Visweswariah C, Ravindran K, Kalafala K, Walker SG, Narayan S, Beece DK, Piaget J, Venkateswaran N, Hemmett JG (2006) First-order incremental block-based statistical timing analysis. IEEE Trans Comput Aided Des Integr Circ Syst 25: 2170–2180
84. Rao R, Srivastava A, Blaauw D, Sylvester D (2004) Statistical analysis of subthreshold leakage current for VLSI circuits. Trans VLSI syst 12:131–139
85. Zhang S, Wason V, Banerjee K (2004) A probabilistic framework to estimate full-chip subthreshold leakage power distribution considering within-die and die-to-die P-T-V variations. In: Proceeding of international symposium on low power electronics and design, Newport Beach, CA, pp 156–161
86. Rao R, Devgan A, Blaauw D, Sylvester D (2004) Parametric yield estimation considering leakage variability. In: Proceeding of design automation conference, San Diego, CA, pp 442–447
87. Agrawal A, Kang K, Roy K (2005) Accurate estimation and modeling of total chip leakage considering inter- & intra-die process variations. In: Proceeding of international conference on computer aided design, San Jose, CA, pp 736–741
88. Wang T-Y, Chen CC-P (Dec 2002) 3-D thermal-ADI: A linear-time chip level transient thermal simulator. IEEE Trans Comput Aided Des Integr Circ Syst, vol. 21, no. 12, pp 1434–1445
89. Su H, Liu F, Devgan A, Acar E, Nassif S (Aug 2003) Full chip estimation considering power, supply and temperature variations. In: Proceeding of international symposium low power electron. design, pp 78–83
90. Li P, Pileggi L, Asheghi M, Chandra R (2004) Efficient full-chip thermal modeling and analysis. In: Proceedings of international conference on computer aided design, Seoul, pp 319–326
91. Cheng Y, Raha P, Teng C, Rosenbaum E, Kang S (Aug 1998) ILLIADS-T: an electrothermal timing simulator for temperature-sensitive reliability diagnosis of CMOS VLSI chips. IEEE Trans Comput Aided Des Integr Circ Syst 17(8):668–681
92. Huang W, Stan MR, Skadron K, Sankaranarayanan K, Ghosh S (May 2006) HotSpot: a compact thermal modeling method for CMOS VLSI systems. IEEE Trans VLSI Syst 14(5):501–513
93. BPTM: Berkeley predictive technology model. http://www-device.eecs.berkeley.edu/~ptm
94. Kang K, Paul BC, Roy K (2006) Statistical timing analysis using levelized covariance propagation considering systematic and random variations of process parameters. ACM Trans Des Autom Electron Syst 11:848–879
95. Datta A, Bhunia S, Mukhopadhyay S, Roy K (2005) A statistical approach to area-constrained yield enhancement for pipelined circuits under parameter variations. In: Proceeding of Asian test symposium, Kolkata, pp 170–175
96. Kim CH, Roy K, Hsu S, Alvandpour A, Krishnamurthy R, Borkhar S (2003) A process variation compensating technique for Sub-90 nm dynamic circuits. In: symposium on VLSI circuits, Kyoto
97. Kang K, Kim K, Roy K (2007) Variation resilient low-power circuit design methodology using on-chip phase locked loop. In: Proceeding of design automation conference, San Diego, CA, pp 934–939
98. Blome J, Feng S, Gupta S, Mahlke S (Dec. 2007) Self-calibrating online wearout detection. In: Proc. 40th intl. symposium on microarchitecture (MICRO), Chicago, IL, pp 109–120
99. Feng S, Gupta S, Mahlke S (2008) Olay: combat the signs of aging with introspective reliability management. The Workshop on quality-aware design (W-QUAD)

100. Ghosh S, Batra P, Kim K, Roy K (2007) Process-tolerant low-power adaptive pipeline under scaled-Vdd. In: Proceeding of custom integrated circuits conference, San Jose, CA, pp 733–736
101. Ghosh S, Choi JH, Ndai P, Roy K (2008) $O^2C$: occasional two-cycle operations for dynamic thermal management in high performance in-order microprocessors. In: Proceeding of international symposium on low power electronics and design, Bengaluru, pp 189–192
102. Krishnamurthy R, Alvandpour A, Balamurugan G, Shanbag N, Soumyanath K, Borkar S (2002) A 130 nm 6-GHz 256×32 bit Leakage-Tolerant Register File, IEEE J Solid State Circuits 37:624–632
103. Kim T, Persaud R, Kim CH (2007) Silicon odometer: an on-chip reliability monitor for measuring frequency degradation of digital circuits. In: Proceeding of VLSI circuit symposium, Kyoto, pp 122–123
104. Agarwal M, Paul BC, Zhang M, Mitra S (2007) Circuit failure prediction and its application to transistor aging. In: Proceeding of VLSI test symposium, Kyoto, pp 277–286