1-1-2003

The Role of Temperature in Testing Deep Submicron CMOS ASICs

Ethan Schuyler Long
Portland State University

10.15760/etd.34

THESIS APPROVAL

The abstract and thesis of Ethan Schuyler Long for the Master of Science in Electrical and Computer Engineering were presented October 24, 2003, and accepted by the thesis committee and the department.

COMMITTEE APPROVALS: _____
W. Robert Daasch, Chair

_____
James McNames

_____
Aslam Khalil
Representative of the Office of Graduate Studies

DEPARTMENT APPROVAL: _____
James E. Morris, Chair
Department of Electrical and Computer Engineering

ABSTRACT

An abstract of the thesis of Ethan Schuyler Long for the Master of Science in Electrical and Computer Engineering presented on October 24, 2003.

Title: The Role Of Temperature In Testing Deep Submicron CMOS ASICs.

Among the many efforts to improve the IC test process are tests that attempt to differentiate between healthy and defective or low reliability ICs by manipulating the operating conditions of the IC being tested. This thesis attempts to improve the common understanding of multiple and targeted temperature testing by evaluating work published on the subject to date and by presenting previously unpublished empirical observations.

The empirical observations are made from SCAN and LBIST based MinVDD measurements, Static IDD measurements, as well as parametric measurements of transistor characteristics. The test vehicles used are 0.25µm and 0.18µm CMOS ASICs fabricated by LSI Logic.

An IC's performance is bound by a three dimensional space defined by VDD, frequency, and temperature. A model is presented to explain the boundaries of the performance region in terms of the ability of the IC's constituent transistors to provide power and the Zero-Temperature-Coefficient (ZTC). Also, it is determined that multiple temperature testing can add new tests to current test suites to improve the resolution between healthy and defective ICs.

THE ROLE OF TEMPERATURE IN TESTING DEEP SUBMICRON CMOS ASICS

by

ETHAN SCHUYLER LONG

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
ELECTRICAL AND COMPUTER ENGINEERING

Portland State University
2003

Acknowledgements

This work would not be what it is without the generous support and guidance of W. Robert Daasch and Robert Madge. Their work to establish and maintain such a close relationship between Portland State University and LSI Logic was integral not only to making this work a success but to making this work a possibility.

Also, many thanks to Chris Schuermyer for all of his help in troubleshooting test setups and test program development; to Chuck Lundegard and Electroglas for providing and supporting the EG4|200 prober in the IC Design & Test Lab at PSU; to Tricia Justice, Alan Aoki, and Credence for providing and supporting the Quartet tester in the IC Design & Test Lab at PSU.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

For the past few years the testing process has widely been considered the Achilles heel of semiconductor manufacturing. There are ongoing efforts to make the test process less expensive while maximizing the number of healthy integrated circuits (ICs) and minimizing the number of non-functional and low reliability ICs shipped to the customer. Among the efforts to improve the test process are tests that attempt to differentiate between healthy and defective or low reliability ICs by manipulating the operating conditions of the IC being tested. Power supply voltage and operating frequency are two operating conditions that have been studied in detail. A third operating condition, temperature, has only been investigated superficially.

When multiple or targeted temperatures are used as test conditions, the distinctive behavior of healthy ICs as compared to defective ICs may provide a critical improvement in test resolution. This thesis attempts to improve the common understanding of multiple and targeted temperature testing by evaluating work published on the subject to date and by presenting previously unpublished empirical observations of deep submicron application specific integrated circuits (ASICs).

Chapter 2

Background

2.1 Transistor Characteristics

2.1.1 Threshold Voltage (V_t)

The threshold voltage, V_t, of a MOSFET is generally described as the voltage at which the transistor turns on, though there is no definitive point at which this occurs as evidenced by the finite nature of the transconductance of a transistor. As such, there are several definitions for V_t. One common and practical method of finding the V_t is referred to as the transconductance or g_m method, shown in Figure 2.1. The transconductance, $g_m = \frac{dI_D}{dV_G}$, is the slope of the drain current, I_D, versus gate voltage, V_G, curve. The I_D versus V_G curve is hereafter referred to as a transistor's IV curve. The g_m method involves finding the point on the IV curve where g_m reaches a maximum. A line is fit to the IV curve at the point of maximum g_m and extrapolated to the point where I_{DS} is zero where the V_G is taken as the V_t ([1], pp. 243-244).

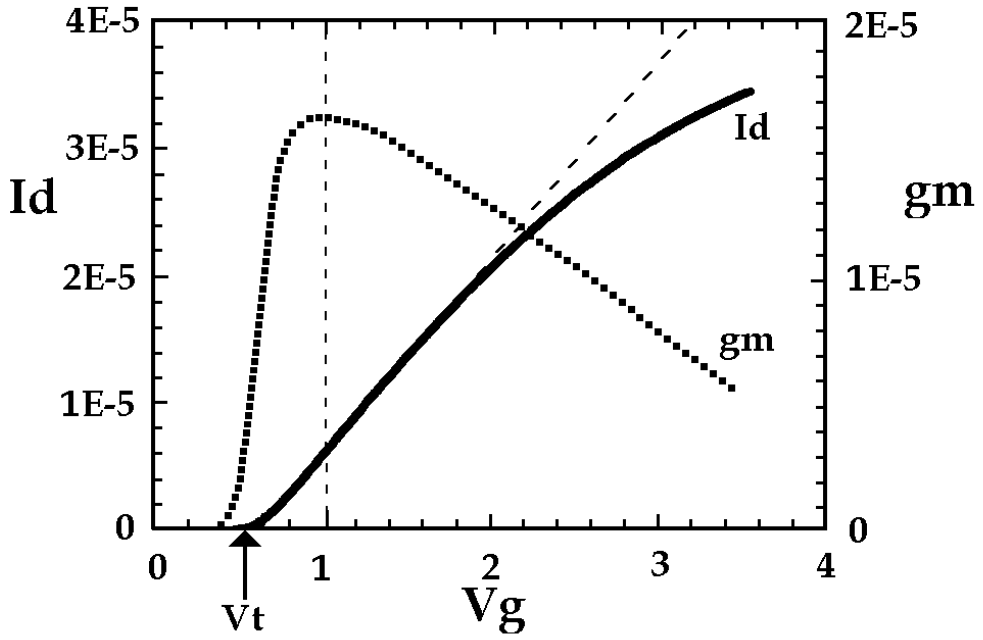


Figure 2.1 (after [1], Figure 4.24) - Transistor IV Curve With Transconductance

2.1.2 dVt/dT

The threshold voltage of a transistor has long been recognized as being temperature sensitive. Both [2] and [3] approach their analysis of temperature induced V_t variations by giving a theoretical description of V_t in the form of:

$$V_t = \Phi_{MS} - \frac{Q_{SS}}{C_O} + \Phi_B \pm \frac{\sqrt{2\varepsilon q N \Phi_B}}{C_O}$$

Equation 2.1

where Φ_{MS} is the metal to semiconductor work function difference, Q_{SS} is the surface-state charge density, C_O is the gate oxide capacitance, N is the impurity dopant

concentration, ε is the permitivity of the semiconductor, Φ_B is commonly equated to $2\Phi_F$, and Φ_F is the Fermi potential in the channel. The last term is positive for NMOS transistors and negative for PMOS transistors. Those terms in Equation 2.1 that are taken as being independent of temperature are Q_{SS}, C_O, N, and ε. In [2] Φ_{MS} is assumed to be independent of temperature and in [3] Φ_{MS} is empirically determined to be:

$$\Phi_{MS} = -0.61 - \Phi_F$$

Equation 2.2

The study presented in [3] uses transistors with aluminum gates whereas modern transistors use poly-silicon gates. Though the two studies take different approaches to solving for dV_t/dT, in either case the temperature dependence of V_t rests on that of Φ_F. In [2] the solution is found through differentiation as:

$$\frac{dV_t}{dT} = \frac{d\Phi_F}{dT}\left[2 - \frac{\sqrt{2\varepsilon q N \Phi_B}}{C_O \Phi_B}\right]$$

Equation 2.3

where

$$\frac{d\Phi_F}{dT} = \pm\frac{1}{T}\left[\frac{E_G}{2q} - \left|\Phi_F\right|\right]$$

Equation 2.4

In [3], "the dependence of the Fermi potential on temperature ($\Phi_F[T]$) is obtained by assuming charge neutrality and using Boltzmann statistics." These results were used along with empirically determined relationships, Equation 2.1, and a computer to find dV_t/dT.

Both of these studies find that $V_t(T)$ is linear or very nearly linear between −73°C and 125°C as determined by both experimental data and their respective theoretical approaches. Other important observations made in [2] and [3] are that transistors with more heavily doped channel regions and those with thicker gate oxides have V_t's that are more strongly dependent on temperature. These findings are supported in [4] where it is claimed that the approach given above, with some minor modifications to Equation 2.1, is appropriate for determining dV_t/dT for long channel (>1μm) transistors with heavily doped poly-silicon gates.

A very similar method for finding dV_t/dT for long channel transistors begins by recognizing that V_t depends primarily on two factors, the Fermi potential and the band gap. For *P-type* silicon, the difference between the valence energy level and the Fermi energy level, E_V-E_F, becomes smaller as temperature is lowered [4]. The temperature dependence of the band-gap, E_G, in silicon is given as:

$$E_G(T) = E_G(0) - \frac{\alpha T^2}{(T+\beta)}$$

Equation 2.5 ([5], pp. 15)

where $E_G(0)$=1.170, α=4.73E-4, and β=636 ([5], pp. 15). As temperature is lowered E_V-E_F becomes smaller and E_G becomes larger. Both of these factors contribute to an increase in V_t with lower temperatures.

The V_t is described as:

$$V_t = -\frac{E_G}{2q} + \Phi_B + \frac{\sqrt{4\varepsilon q N_A \Phi_B}}{C_O}$$

Equation 2.6 ([6], pp. 131)

where N_A is the acceptor dopant concentration for an NMOS transistor and Φ_B is the difference between the Fermi potential and the intrinsic potential, defined as:

$$\Phi_B \equiv \left|\Phi_F - \Phi_I\right| = \frac{kT}{q}\ln\left(\frac{N_A}{n_I}\right)$$

Equation 2.7 ([6], pp. 25)

where Φ_I is the intrinsic carrier potential, n_I is the intrinsic carrier concentration given as:

$$n_I = \sqrt{N_C N_V}\,\exp\left(-\frac{E_G}{2kT}\right)$$

Equation 2.8 ([6], pp. 12)

where N_C and N_V are the conduction band and valence band effective densities of state. While Φ_B is defined as in Equation 2.7 it is commonly equated to $2\Phi_F$. Derivation of Equation 2.6 yields:

$$\frac{dV_t}{dT} = -\frac{1}{2q}\frac{dE_G}{dT} + \left(1 + \frac{\sqrt{\varepsilon q N_A}}{C_O\sqrt{\Phi_B}}\right)\frac{d\Phi_B}{dT}$$

Equation 2.9 ([6], pp. 131)

or:

$$\frac{dV_t}{dT} = -\left(1 + \frac{\sqrt{\varepsilon q N_A}}{C_O\sqrt{\Phi_B}}\right)\frac{k}{q}\left[\ln\left(\frac{\sqrt{N_C N_V}}{N_A}\right) + \frac{3}{2}\right] + \frac{\sqrt{\varepsilon q N_A}}{qC_O\sqrt{\Phi_B}}\frac{dE_G}{dT}$$

Equation 2.10 ([6], pp. 131)

The threshold voltage is simply a description of a critical concentration of free electrons in the transistor channel. So, it is reasonable that the rate of change of threshold voltage with respect to temperature depends directly on those factors that control the availability of electrons in the conduction band of the channel. Lower overall thermal energy in the lattice resulting from lower temperatures leads to fewer

electrons with enough energy to occupy the conduction band states. This is reflected in Equation 2.7 where lower temperatures give smaller values of Φ_B. As the Fermi potential approaches the intrinsic potential N_V grows and N_C shrinks. The larger band gap energy, E_G, at lower temperatures indicates that there is a larger energy required for electrons to surmount before they can occupy conduction band states. The dopant concentration, N_A, directly skews the number of electrons available for conduction by effectively inserting extra electrons into the lattice. The gate oxide capacitance, C_O, and the permitivity of silicon, ε, control the gate induced electric field strength in the channel.

It should be emphasized that Equations 2.9 and 2.10 are long channel transistor models and do not consider short channel effects or other parasitic V_t shifting effects. Any temperature sensitive leakage mechanism that contributes to drain current will contribute to dV_t/dT. If the leakage mechanism is relieved by lower temperatures it will also induce an increase in $|V_t|$ at lower temperatures. That is, lower subthreshold currents (discussed in section 2.4.6) equate to higher $|V_t|$. Hot carrier injection (HCI, discussed in section 2.4.10) will contribute nominally to the concentration of conduction electrons in the channel as well as injecting fixed charge into the gate oxide. Both of these effects will lower the threshold voltage and will be aggravated by lower temperatures in short channel transistors with large drain voltages. Lower temperatures will have the same qualitative effect on drain induced barrier lowering, DIBL (discussed in section 2.4.7), as they do on V_t but DIBL's effect on dV_t/dT is not

described by the long channel model. It can be expected that dV_t/dT will increase with short channel devices.

The Performance of a healthy IC is dictated by the performance of its constituent transistors. Having established that transistor performance is a function of temperature (see also sections 4.1 and 4.2), a discussion of an IC's performance as a function of temperature (see sections 2.1.4 and 4.2) is justified.

2.1.3 Negative Bias Temperature Instability (NBTI)

Another parasitic, temperature sensitive phenomenon is Negative Bias Temperature Instability, NBTI. When a transistor, either PMOS or NMOS, is held at an elevated temperature for an extended period of time with the gate at a large negative bias, mobile positive charges are attracted to the gate oxide and the Si-SiO_2 interface ([1], pp. 365). If the transistor is cooled to room temperature while under this large gate bias and the V_t remeasured, these parasitic positive charges induce a negative shift in the V_t. The PMOS V_t shifts further from zero and the NMOS V_t shifts towards zero [7].

NBTI is caused by the presence of water or hydrogen in the fabrication process [7,8]. Nitrogen has also been reported as a source of NBTI in [9]. When a wafer is removed from a very high temperature oxidation chamber the ambient air will release water onto the oxide layer [7]. Water may also drift from the intermetal dioxide or

PSG (phosopho-silicate glass) layers to effect NBTI [10]. The hydrogen commonly used in post metal anneals may also act as a catalyst for NBTI [11].

Typical conditions under which this NBTI effect can be induced are temperatures of 150°C to 250°C and gate voltages sufficient to cause oxide electric fields of about 10^6 V/cm ([1], pp. 365). The temperature and bias stress times vary widely, from a few minutes to hundreds of hours ([1] pp. 365, [7,8,9,11,12,13,14]). Longer bias times, higher bias temperatures, and more negative bias voltages will all lead to larger shifts in V_t [12]. The effects of NBTI are also exaggerated in transistors with thin oxides [9] and shorter gate lengths [10]. NBTI is worse for PMOS than it is for NMOS [9]. In [11] this discrepancy is shown to result in distinct performance variations. Rise times, where PMOS transistors dominate, show significant changes as a result of NBTI. Fall times, where NMOS transistors dominate, do not show such significant shifts [13]. NBTI is not such a concern for NMOS transistors not only because they are not as profoundly affected by it, but also because NMOS transistors are only subjected to negative gate biases for short durations during transitions.

A complementary phenomenon called Positive Bias Temperature Instability, PBTI, would seem to present a more significant concern for NMOS transistors. PBTI occurs under the same stress conditions with a positive, instead of a negative, gate bias. NMOS transistor V_t's have been shown to become larger, further from zero, after PBTI stressing [13]. However, both PMOS and NMOS transistors have also been shown to be insensitive to PBTI stressing [14]. In any case, it is the effect of NBTI on surface channel PMOS transistors and not PBTI, NBTI for NMOS

transistors, or NBTI for buried channel PMOS transistors that is considered to be a major reliability concern.

This is all a bit of a mute point with regard to this study as neither NBTI or PBTI will have an impact on the results of this study. The reasons for this are three fold. First, the temperatures used are at or below 85°C, well below the activation temperatures for bias temperature instability. Secondly, the devices tested in this study are powered up for very short times, typically less than one minute. The tertiary reason involves an effect reported in [13] and [14]. These studies report that, after NBTI stressing, if a transistor is held at an elevated temperature with no gate bias stress the V_t will recover to near the prestress value. This is a result of the positive charges at the Si-SiO_2 interface diffusing throughout the device. In this study, when an IC or site on a wafer is tested at an elevated temperature the entire wafer is heated. So, every IC or site tested experiences a healing soak time at an elevated temperature and no gate bias prior to testing, with the last IC or site being tested having the longest soak time.

2.1.4 Performance/speed and Zero-Temperature-Coefficient

The performance of an IC as defined by its operating speed is a critical parameter in IC testing. Signs of the importance of IC performance in IC testing are transition delay fault (TDF), maximum frequency (f_{max}), ring oscillator speed tests, at-

speed testing, and built-in-self-test (BIST). Stated concisely in [15], if simplistically, "Lower temperature increases the transistor switching speed and reduces its leakage current." In changing the operating temperature from 100°C to -50°C, performance improvements of between about 20% and 60% have been reported ([6], pp. 288). Performance improvements are due primarily to improvements in carrier mobility at lower temperatures. The increase in the absolute value of PMOS and NMOS transistors' threshold voltages with lower temperatures acts as a compensating factor in IC performance [16]. The said 20% and 60% performance improvements correspond to mobility improvements for CMOS ICs of about 40% and 200% respectively ([6], pp. 286). Considering the contrary effects of threshold voltage shift and mobility changes with temperature, there exists a VDD above which lower temperatures improve performance and below which lower temperatures degrade performance [17]. For a 0.25μm CMOS technology operating with VDD of 2.5V, the temperature induced 0.25V threshold voltage drop caused an increase in drain current of about 10%. When the VDD is set at 0.5V the drain current increased by about 55% due to the temperature increase induced threshold voltage drop. In the former case, the drain current reduction due to mobility degradation is more profound than the drain current increase due to a lower threshold voltage and the IC's performance is degraded. In the latter case, the drain current increase due to lower threshold voltages dominates leading to improved performance [18]. These shifts in threshold voltage and saturation current with temperature help define the influence of temperature on a transistor's IV curve as shown in Figure 2.2.

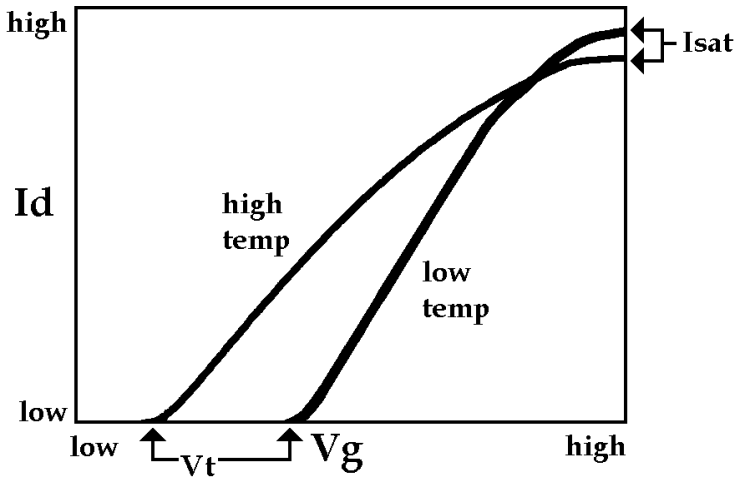


Figure 2.2 (after [4], Figure 4) - Effect Of Temperature On Transistor IV Curve

As temperature is lowered both V_t and I_{sat} are increased. However, the variation in drain current is a function of both temperature and VDD as is demonstrated in Figure 2.3.

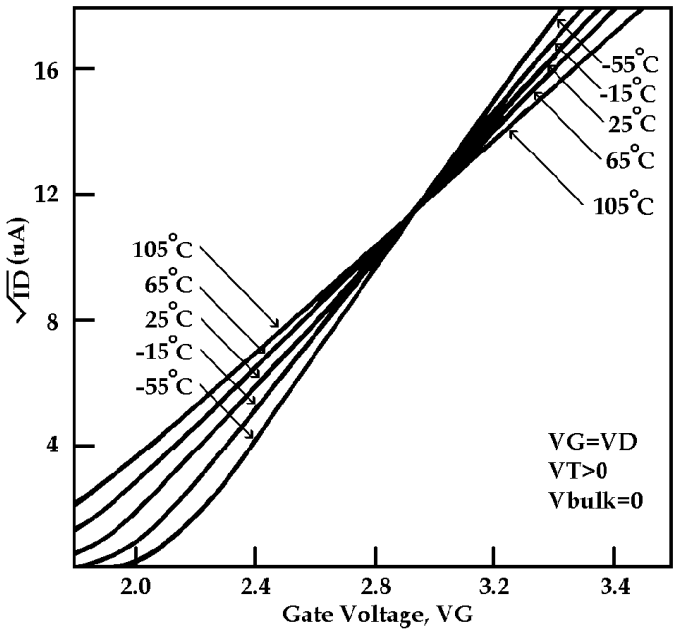


Figure 2.3 (after [19], Figure 4.21) - Detail Of Temperature Induced Change In I_D

The gate voltage at which the temperature induced shifts in threshold voltage and mobility exactly compensate one another is commonly reffered to as the Zero-Temperature-Coefficient (ZTC) bias point. ZTC biasing induces no change in drain current with a shift in temperature and is discussed in detail in [58], [59], and [60].

A rough theoretical analysis of these temperature driven performance characteristics is given in [17], [18], and [60]. Delay is given by:

$$t_d = \frac{C_L V_{DD}}{2 I_{av}}$$

Equation 2.11

where C_L is a temperature independent load capacitance and I_{av} is the average drain to source current. I_{av} is in turn:

$$I_{av} \propto \mu [V_{DD} - V_t]^{1.5}$$

Equation 2.12

where the mobility, μ, and threshold voltage, V_t, are functions of temperature given by:

$$\mu = \mu_0 \left(\frac{T}{T_0} \right)^{-M}$$

Equation 2.13

and

$$V_t = V_{t0} - KT$$

Equation 2.14

where V_{t0} and μ_0 are taken at a nominal temperature, T_0=27°C. K and M are the V_t temperature coefficient and mobility-temperature exponent respectively and, though nearly constant for any given technology, are technology dependent. In [17], the author solves for the rate of change of t_d with respect to temperature and finds the

VDD value at which t_d is independent of temperature. The author concludes that where:

$$\frac{\partial t_d}{\partial T} = 0$$

Equation 2.15

the temperature insensitive VDD is:

$$V_{DD} = V_{t0} + 1.5\frac{K}{M}T$$

Equation 2.16

The conclusion as given by the author of [17] in Equation 2.16 is not entirely accurate though. When the derivation was redone with Equations 2.11 to 2.15 the temperature insensitive VDD is found to be:

$$V_{DD} = V_{t0} + 1.5\frac{K}{M}T - KT$$

Equation 2.17

This is an important distinction as M is given as 1.5 in both [17] and [18] which leads to the temperature insensitive VDD being equal to V_{t0}. Such a situation would make operating in the low VDD region unfeasible. However, the simulation results given in [17] support a temperature insensitive VDD that is greater than V_{t0}.

2.2 Test

If an IC's performance is a function of temperature, how can the performance be measured? The simple answer is to test the IC. What follows is a discussion of how ICs are commonly tested.

There are two fundamental forms of test, parametric and functional. Functional type testing relies on fault models to test an IC whereas parametric tests look for defects directly. The distinction between a fault and a defect is subtle. A defect is a physical deviation from the design of an IC. A defect may be a metal short, a misaligned implant, an impurity in the gate oxide, etc., but a defect is always something physically where it should not be or part of the IC missing from where it was designed to be. A fault is a symbolic abstraction of a defect representing an electrical or functional characteristic of the defect in the context of the function of the circuit ([21], pp. 57-58). Fault models are designed to reflect the behavior of a defect and to demonstrate how a defect might affect the performance of a circuit.

Functional and parametric testing are generally described as,

> (1) *Parametric Tests*. DC parametric tests include shorts test, opens test, maximum current test, leakage test, output drive current test, and threshold levels test. AC parametric tests include propagation delay test, setup and hold test, functional speed test, access time test, refresh and pause time test, and rise and fall time test. These tests are usually technology-dependent. CMOS voltage output measurements are done with no load …..

(2) *Functional Tests*. These consist of the input vectors and the corresponding responses. They check for proper operation of a verified design by testing the internal chip nodes. Functional tests cover a very high percentage of modeled (e.g., stuck type) faults in logic circuits ….. Often, functional vectors are understood as verification vectors, which are used to verify whether the hardware actually matches its specification. However, in the ATE world, any vectors applied are understood to be functional fault coverage vectors applied during manufacturing test. These two types of functional tests may or may not be the same.

([21], pp. 21)

2.2.1 Fault Model Testing

2.2.1.1 Functional

Functional testing is one of the more fundamental forms of VLSI testing, the generation and use of which is based on the stuck-at fault model. This stuck-at fault model became popular in the 1980's for the relative simplicity with which test patterns, a collection of test vectors, could be generated [22]. A stuck-at fault models a defect at any given node within the circuit design that induces that node to be stuck at ground or stuck at VDD, that is, stuck at 0 (*SA0*) or stuck at 1 (*SA1*).

Any given circuit design will have a series of primary inputs and primary outputs. These are the outside world's access to the circuit or IC. Functional testing involves applying a combination of 1's and 0's, called a test vector, to the primary inputs and testing for the expected combination of 1's and 0's, an output vector, on the

primary outputs. An output vector is determined through an understanding of the logic and function of the circuit. The premise of this testing is that when the expected output vector does not appear on the primary outputs a stuck-at fault exists in the circuit that will cause the circuit to malfunction in the hands of the customer.

It is not possible to test the functionality of an IC by testing all possible test vectors. The number of possible input vectors is 2^n where n is the number of primary inputs. Modern VLSI chips have ever increasing numbers of primary inputs. For instance, the Sematech IC was an IBM ASIC with 249 signal IO's [23]. With 249 inputs this circuit would have 9.05E74 possible test vectors. With an ATE (advanced test equipment) that runs at 100MHz testing all possible vectors would require 2.87E59 years. Since the age of the universe is estimated at 1.2E10 to 2.0E10 years ([24], pp. 617) testing all possible test vectors is clearly not realistic.

Testing with the stuck-at fault model is made practical by selecting a small number of test vectors that will achieve high fault coverage. In one test efficacy study about 1E7 functional vectors were applied to the DUT (device under test) [25]. An additional technique for maximizing test efficiency involves selecting the test vectors with the highest fault coverage to test first. Most production testing is done on a stop on first fail basis which is to say that the testing of the DUT stops as soon as it fails a test. This stop on first fail testing insures that tester time is not wasted on suspected bad parts.

With modern IC's exceeding fifty million transistors or five million logic gates [26] the process of test pattern generation for high fault coverage must be automated.

Automatic test pattern generators (ATPG) are software tools that are used to go through the process of test vector generation. Even so, the semiconductor industry will continue to shrink the length and pitch of transistors ([27], pp. 153) thereby increasing the number of transistors per chip and the number of possible stuck-at faults per chip. The *2002 International Technology Roadmap for Semiconductors* gives the current maximum number of transistors per chip as about 899 million and increasing to 2041 million in 2007 ([27], pp. 160). With such large numbers of transistors and logic gates on a VLSI (very large scale integration) IC it is not possible for even an ATPG tool to simulate and generate all of the possible test vectors in an attempt to generate a vector set with 100% fault coverage. So, designers must work with the ATPG to intelligently simulate some vectors and add to those some number of random vectors in an attempt to get as close to 100% fault coverage as possible ([21], pp. 85). A very high fault coverage for extremely complex ICs is considered to be 99% [26].

It should be clear that functional testing and the stuck-at fault model are not perfect. Though the question applies to all tests that are in current use, when considering functional testing it is critical to ask [26], "Can devices with 10M logic gates and 16MB of embedded RAM be tested to a very high fault coverage within 5 seconds?" When answering this question it is not sufficient to determine whether or not it is possible. The difficulty with which the solution is achieved and the associated costs are integral to the answer to the stated question. In fact, as compared to other test methods, stuck-at fault coverage has been shown to be a poor predictor of the efficacy of testing a DUT [22]. One study showed that, after testing 20,000 ICs with a

vector set based on the stuck-at fault model, 89% of the vectors within the vector set detected no defective ICs [26]. The same data shows that those vectors that did detect defective ICs were spread throughout the vector set [26]. If vector ordering based on fault coverage was performed, this is indicative of the inability of the stuck-at model to model defects or of the ATPG to produce effective vector sets. Such problems may arise from the fact that there has been no standardized method for calculating stuck-at fault coverage. Different ATPG tools may give very different stuck-at fault coverage numbers for the same circuit [22]. One of the primary motivations for moving away from functional testing is the large expense that comes with achieving high stuck-at fault coverage [22]. Furthermore [23], "At speed functional testing, which requires full pin contact and complex timing support, is difficult and expensive to support…" The cost of ATE is directly related to the number of pins the ATE can support and the speed at which the tester can apply a test pattern to a DUT.

An additional problem with functional testing involves the proprietary nature of many IC designs. It is the responsibility of the IC manufacturer to test the parts that they sell to their customers, though, it is not always the responsibility of the manufacturer to design the products. The manufacturer may not be given access to information regarding the functional logic of the product that is required for generating the functional test patterns based on the stuck-at fault model. This leaves the burden of generating these test patterns on the customer which is always expensive and sometimes outside the capabilities of the customer. A variation of this problem arises when ASIC manufacturers purchase designs from other companies to integrate

into their own products. The ASIC manufacturer may have limited information on the design of these third party or IP core circuits making them difficult or impossible to test. Considering this, manufacturers have strong motivations for eliminating functional testing from their test schemes. Elimination of functional test would allow the manufacturer to offer the customer a product at a considerably lower cost. It would also give the manufacturer complete control over the test scheme for which they are ultimately responsible anyway. This control facilitates test coverage improvements, test time reduction efforts, failure analysis, yield enhancement efforts, and defect mode studies. All of the temperature based testing proposed herein will contribute to the goal of eliminating functional testing from ASICs test suites.

2.2.1.2 SCAN

SCAN testing is based on the same fundamental principles as functional testing. Both SCAN and functional testing are based on the stuck-at fault model and involve applying test vectors and looking for the appropriate vectors on the outputs. The distinction between SCAN and functional testing lies in the fact that SCAN utilizes the flip-flops in a circuit's design as well as additional test specific circuit components to simplify the problem of testing a circuit. The flip-flops have hardware added to them to turn them into dual purpose registers that are accessed by either functional inputs or SCAN inputs as dictated by a test control input. Such registers

can act as a functional part of the circuit or as a SCAN register devoted to testing the circuit for defects. These SCAN registers are then connected in a chain from an input to an output. Any given circuit design may have many independent SCAN chains. The Sematech IC had 5280 SCAN registers divided into 8 SCAN chains [23].

SCAN can be run in a shift register mode in which this series of registers is fed a test vector which can then be read out through a SCAN devoted output. Each bit of the test vector is loaded at the SCAN input and is passed from one register to the next until it is read out through the SCAN output. Failure to read out the expected vector indicates a Boolean fault that occurred within the SCAN chain ([21], pp. 467-469). Using a 00110011 pattern exercises the four possible transitions, 0-1, 1-1, 1-0, and 0-0 ([21], pp. 471).

A more significant test mode for SCAN involves using ATPG tools to generate test patterns that test for stuck-at faults in the logic much as is done for functional testing ([21], pp. 471-473). In this mode the SCAN chain can act as a primary input to specific blocks of a design. This allows test vectors to bypass large blocks of combinational logic effectively partitioning the DUT into smaller, simpler testable circuits and thereby simplifying the test process. The ATPG tool may be applied to small blocks of logic instead of the entire design. This mode also alleviates the problem of IP cores that may be included in an ASIC manufacturer's design by simply circumventing the IP core with one or more SCAN chains.

2.2.1.3 Built In Self Test (BIST)

The method termed built-in-self-test (BIST) is a relatively new technique for more efficient and effective testing. BIST involves designing ICs with test specific hardware that fully implements testing on the IC. BIST was originally designed for testing systems with many PCB boards each with many ICs. Such systems are difficult, if even possible, to test through the primary inputs. ASICs and system-on-chip (SOC) designs mimic such systems by including several different circuit functions on a single chip. Some of the motivations for using SCAN are shared with BIST; manufacturers having limited information on the design of some blocks at the gate level and the ability to partition increasingly high transistor and gate count designs. BIST also enables at-speed testing without the high per pin cost of high speed ATE ([21], pp. 489-491).

There are two primary forms of BIST, LogicBIST and MemoryBIST. At the most rudimentary level, both of these require some test specific hardware in the design. The hardware that is required to add BIST to a design includes a test controller, some form of pattern generator, and an output compactor and comparator/analyzer. The test controller will initiate the IC self test at the prompt from a primary input. The test pattern generator will apply a series of test vectors to the circuit under test. In the LogicBIST case, this pattern generator can include pseudo-random pattern generators, predetermined patterns stored in ROM (a seldom used method), exhaustive test pattern generators that produce all possible

combinations of inputs, or combinations and variations of these ([21], pp. 496-499). In the case of MemoryBIST the pattern generating hardware and methods are appropriate to memory testing. The output compactor and comparator/analyzer receive the outputs and convert circuit response to a GO/NOGO signal sent to the primary output.

2.2.1.4 Delay Testing

The most recent method for improving the sensitivity of the test methods in use is to look for delay faults. The stuck-at fault model may not catch subtle defects because it does not concern itself with the punctuality of an output vector. The stuck-at fault model is only concerned with logic levels with lenient timing scenarios. With ICs exceeding 1GHz [26] the ATE in use by most IC manufacturers is incapable of testing the devices at their operational speeds. So, testing schemes are limited to lenient timing scenarios. Considering these issues, [26] expresses the need for timing based testing schemes, "Clearly, timing-related defects of small magnitude (e.g., <500ps) that were previously 'benign' may cause failures in the future." It is common for manufacturers to use ring oscillators built into each IC to gauge the speed capabilities of each IC as a result of normal statistical process variations. However, this does not address particular defect induced path delays that may cause an IC to malfunction.

Of the many forms of delay testing one of the more basic forms is the transition delay fault (TDF) test. A TDF may be tested for in a similar way to the stuck-at fault test. Where in the case of the stuck-at fault a node is stuck or permanently in a 1 or 0 state the TDF simply causes the node to be slow in reaching the applied 0 or 1. Unlike the stuck-at test the TDF test requires vector pairs. The first vector sets up the defective node and the second sensitizes the node and output to the fault ([21], pp. 428).

The TDF test is best suited for testing for large delay inducing defects. Longer signal paths have inherently longer delays making path length a critical factor in vector generation. Additionally, a long signal path will be more susceptible to the accumulation of many individually small delays. Advantages to the TDF test include its compatibility with stuck-at oriented testing. The same ATPG tools used for stuck-at fault test generation can be used to generate test patterns for TDF tests. The stuck-at fault coverage for any given design will be similar to the TDF fault coverage for that design because of the similarity between the two fault models ([21], pp. 428-429). TDF testing has been demonstrated as being effective at identifying ICs that would otherwise be test escapes and fail in the customer's system [28].

2.2.2 Parametric Testing

2.2.2.1 IDDQ

The quiescent IDD (IDDQ) measurement is a popular test that takes advantage of defect induced currents to highlight ICs with defects. The IDDQ test involves loading a test vector into a SCAN chain then holding that vector in the SCAN chain while the current drawn by the part is measured on either the ground rail or the power rail. After the vector is loaded, the part is held in a static or quiescent state for a period of time to allow the part to settle. Then, while no additional switching is occurring the current consumption is measured. A typical test scheme could include IDDQ measurements for 10 or 100 vectors. The IDDQ measurement is designed to take advantage of the high current consumption caused by shorts or opens which contrasts itself against the low current consumption caused by the subthreshold current (described in section 2.4.6) of the circuit's constituent transistors ([21], pp. 439-440).

IDDQ testing is effective at identifying stuck-at faults, stuck on or off transistors, gate oxide shorts, interconnect shorts and opens ([21], pp. 441). This test detects defects that are not caught by SCAN and functional tests, tests that are aimed at the stuck-at fault model ([21], pp. 439). It has been demonstrated that fewer defective ICs are labeled as good or passing ICs by test schemes that include IDDQ than by those test schemes that do not include IDDQ testing. This is true over a wide range of stuck-at fault coverage values [26]. IDDQ has been demonstrated as an

effective predictor of ICs that are reliability risks as determined by BURN-IN testing [26]. The Sematech data shows that ICs with relatively high IDDQ measurements have two times the probability of failing post BURN-IN testing as those ICs with low IDDQ measurements [23,29].

Parametric IDDQ measurements can be used as a powerful tool in identifying defective ICs by their abhorrent behaviors. The fundamental principle in this is that ICs with explicit defects will behave differently because the mechanisms responsible for their static current consumption is different than the leakage mechanisms responsible for static current consumption on defect free ICs. A resistive metal short has distinctly different electrical properties than the mechanisms that induce subthreshold leakage in a transistor. Some of the methods used in [26] to identify defective ICs with IDDQ measurements include looking for outliers in plots of minimum vs. maximum IDDQ, IDDQ vs. speed, pre vs. post high voltage stress IDDQ, pre vs. post BURN-IN IDDQ, or in [30] IDDQ vs. maximum frequency.

However, with shrinking transistors the static current consumption of the CMOS circuit is increasing and becoming comparable to that of the defects that IDDQ is supposed to identify. One solution to this problem may be to reduce the temperature at which the test is run. The IDDQ of a healthy IC is a function of temperature just as the subthreshold current is (see section 2.4.6), whereas the IDDQ of a defective IC is controlled by the defect. As such, the healthy IC's IDDQ will be lowered by lower temperatures while the defective IC's IDDQ is likely to be

insensitive to changes in temperature. IDDQ's dependence on temperature is discussed in more detail in section 2.4.12.

2.2.2.2 Minimum VDD and Maximum Frequency Testing

A more recently developed method for testing for subtle defects is minimum VDD (MinVDD) testing. MinVDD testing involves searching for the minimum supply voltage at which an IC will function correctly as determined by functional, SCAN, LBIST, MBIST, or delay fault testing [31]. This test method is based on the closely related Very-Low-Voltage (VLV) testing in which a single, lower than nominal, supply voltage is used for running fault model based tests. Very-Low-Voltages are defined as being on the order of 2 or 2.5 times the lower of the NMOS and PMOS threshold voltages for any given technology [32]. VLV testing is reported as being effective at identifying ICs that will function according to specifications but which contain performance degrading or lifetime shortening defects. VLV testing is particularly effective at identifying two such defect types, resistive shorts and hot carriers [33].

The primary drawback to MinVDD testing is the test time increase involved in running the same test pattern over and over until the MinVDD is found [33]. A variant of MinVDD testing presented in [34] alleviates this problem by achieving much reduced test times. The method involves running MinVDD on a small or

reduced vector set (RVS). The resulting minimum supply voltage is then used to conduct VLV testing with the full vector set (Feed Forward VLV testing). Shorter test times are achieved by conducting MinVDD testing only on a small vector set while high defect coverages are maintained using the Feed Forward VLV testing.

The function of an IC is strongly dependent on both supply voltage and operating frequency (when discussing IC performance the speed at which a part operates is alternately referred to as frequency, f, or propagation delay, t_p, where $f = 1/t_p$). When the supply voltage is lowered the maximum frequency (f_{max}) at which the IC will function is also lowered. This concept is demonstrated in Figure 2.4 as the delay of a CMOS inverter is compared to the supply voltage.

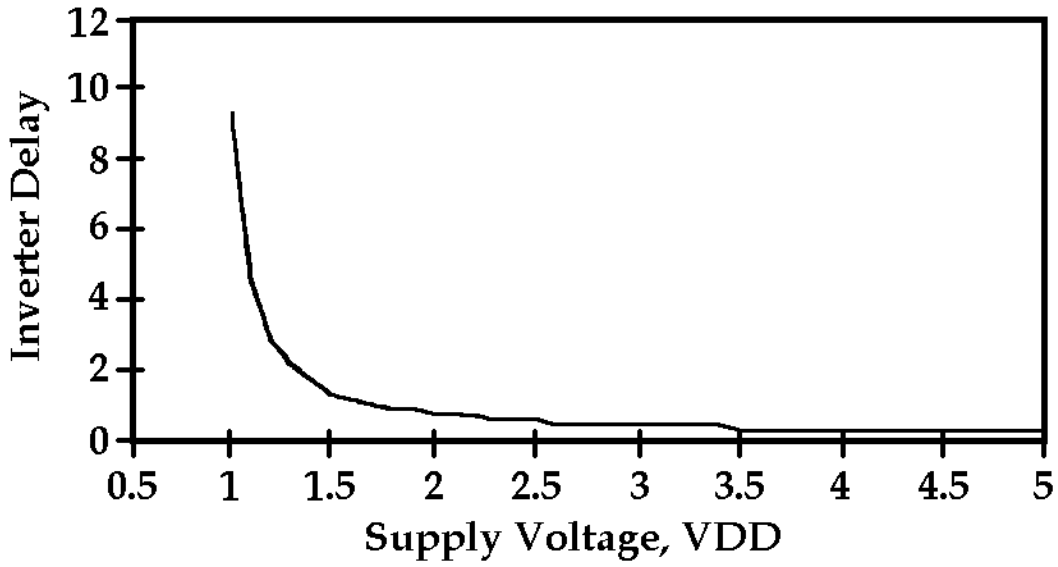


Figure 2.4 (after [32], Figure 9) - Inverter Delay As A Function Of Supply Voltage

While MinVDD testing is done by setting an operating frequency and searching for a minimum VDD, f_{max} testing may be done by setting the VDD and searching for the maximum operating frequency. Both methods, MinVDD and f_{max}, are demonstrated to be adept at detecting delay fault inducing defects in [31] and [35].

A common method for characterizing the performance of an IC that captures both MinVDD and f_{max} information is called Shmoo plotting. A Shmoo plot is generated by testing an IC at many different frequencies and VDDs and plotting the pass/fail results. The Shmoo plot, with one axis being VDD and the other frequency, helps to define the region in the VDD-frequency space in which an IC will function, see Figure 4.14 for an example of a Shmoo plot.

2.2.3 BURN-IN

The traditional method for testing ICs for reliability issues is BURN-IN. This method involves testing or operating a part continuously under elevated temperature and supply voltage for a long period of time. These stress conditions are meant to accelerate the aging of the IC such that ICs that would have abnormally short lifetimes in the customers' hands are screened out. The often mentioned bathtub curve shows the relationship between IC lifetime and the failure frequency for a population of ICs. An example of a bathtub curve is given in Figure 2.5.

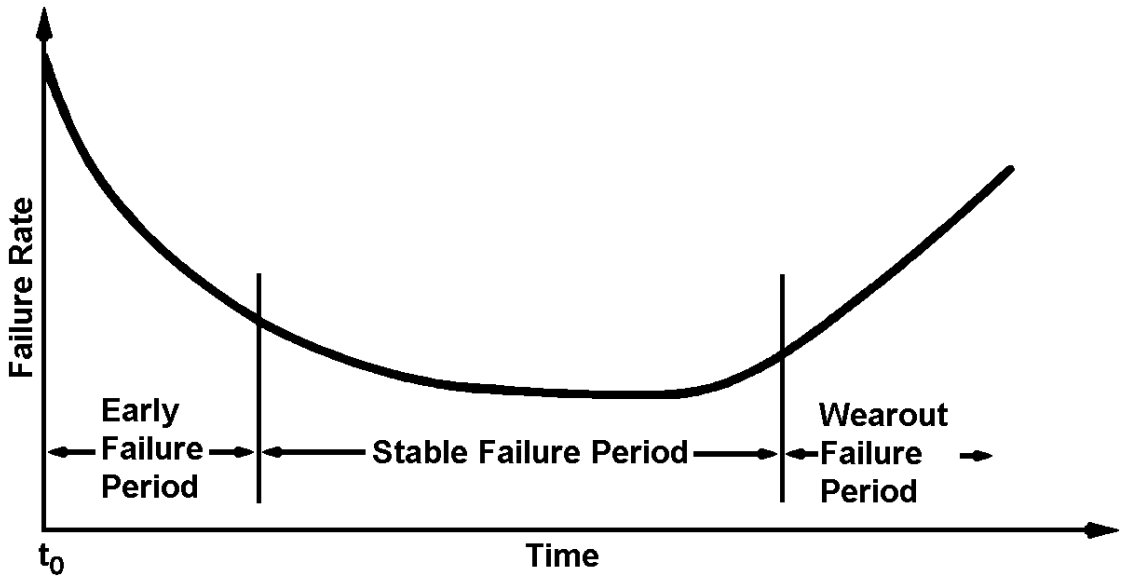


Figure 2.5 (after [57], Figure 2.2) – Bathtub Curve Showing Failure Rates Over Time

There is a high frequency of failures during the early part of the ICs' lifetime followed by a long period of low failure rate and concluded by another period at the normal and healthy end of life for the ICs where the failure rate sharply increases ([21], pp. 20). The efficacy of BURN-IN was shown in the Sematech data where it had a profound impact on such fundamental IC characteristics as IDDQ and speed [29].

Where a manufacturer's test suite at the wafer level might take 30 seconds or a minute for each IC, BURN-IN testing can take 10 or 1000 hours to complete. Such long test times and the resulting low throughputs are obvious obstacles to the use of BURN-IN ([21], pp. 20). Another striking limit to BURN-IN is its use of supply voltage stressing. Part of the process of making transistors smaller involves reducing the supply voltages. As the supply voltage is scaled for shrinking devices the lifetime accelerating aspect of voltage stress is reduced. Also, when these smaller devices are

subjected to voltage stressing, their lifetimes are shortened significantly. This is a critical limit on the efficacy of BURN-IN and demands the development of new methods for predicting reliability problems [26]. Targeted and multiple temperature testing will provide new test methods that will complement currently popular test suites and help eliminate the need for BURN-IN.

2.3 IC Failures & Defects

2.3.1 The Origin of Defects

When the high degree of complexity of ICs and the almost unimaginably small size of the constituent components are considered it is amazing that they function at all. This achievement is made so much more impressive when it is recognized that during the IC manufacturing process there are innumerable opportunities for a defect to manifest itself on an IC. Every aspect of the process presents an opportunity for an aberration to occur. An operator could mistakenly process a lot at an implant step twice. A chemical mechanical polishing (CMP) step could be run for too little time on a wafer leaving poor or no contact between a metal layer and the vias. A lithography tool could run low on photoresist leaving wafers with incomplete patterns over much of the wafer. A lithography tool could give poor alignment or poor leveling of a photo mask. The photo mask itself could be damaged resulting in defects being printed on

the wafer in the same way as the intended circuit pattern. The temperature of an anneal chamber could drift too high. The laser scribing of wafer IDs can leave silicon debris near the flat or notch. The edge bead removal process could extend too far into the wafer causing failures about the edge of the wafer. The walls of plasma chambers can develop detritus build-up which can cause scratches or dislodge from the chamber wall, landing on the wafer. A misplaced decimal in a voltage stress test section of a test program can easily cause weak ICs to fail. The design itself may have a defect that goes unnoticed until the IC is manufactured or until customers return the ICs. In one well publicized case, Intel had a design error in it's Pentium IC that caused an obscure malfunction affecting calculations with floating point numbers. The defect was not caught for some time and went unnoticed until one customer proved it's existence and spread word of it through the Internet.

2.3.2 A Defect's Relationship to IC Test

Defects can occur anywhere and their impact can be severe, obscure, or non-existent. Concisely stated in [29], "The precise defective circuit behavior is dependent on the exact circuit and defect characteristics. For example, a metal short in the same physical location could cause either a logical failure, a timing-only failure or an IDDQ-only failure (depending on the exact resistance of this defect). The same

physical defect could cause a logical failure in one circuit and a timing-related failure in another."

The same defects that cause failures in functional testing may cause parametric failures as well. As described in [16], parametric "failures are either due to variation of IC process parameters, or due to defect-related sensitivity to environmental parameters such as power supply, temperature, and clock frequency." The latter being termed an extrinsic failure and the former an intrinsic failure. The idea that an individual IC may have defects causing multiple tests to fail is supported by the Sematech data set [23]. Some of these results are displayed in the Venn diagram in Figure 2.6.

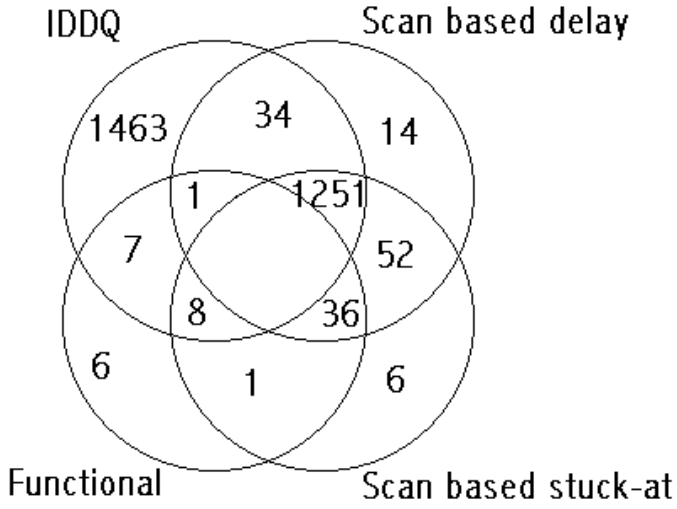


Figure 2.6 [23] - Venn Diagram Of Sematech Test Data

Each circle contains numbers of ICs that failed that test, so for instance, there were 36 ICs that failed functional, SCAN based stuck-at, and the SCAN based delay tests. These 36 ICs fail both functional and parametric type tests but do not cause elevated IDDQ measurements. What is most significant is that functional testing is the least effective of the four test methods.

It is apparent from Figure 2.6 that IDDQ is the most effective at detecting defects. It should be noted, though, that the IDDQ limit was a hard limit and was set somewhat arbitrarily and as such may include many ICs that are healthy. These healthy ICs are functionally sound, making them perfectly saleable, but simply have intrinsic characteristics that cause them to exhibit higher leakage and to run faster than most parts. If so, they are what is referred to as overkill, ICs that are marked as bad that are actually good, functioning ICs. Some of these IDDQ only failures are undoubtedly extrinsic failures. When failure analysis was conducted on these IDDQ only failures, some of the extrinsic defects found were resistive bridges from:

(1) gate to drain,

(2) poly to poly,

(3) poly to nwell,

(4) metal to metal [29].

One defect that did not result in a functional failure but did result in high IDDQ measurements was in an unused gate array [29]. Presumably, had the gate array been used and tested, the IC would also have failed functional testing.

The SCAN based delay testing is the second most effective test. Many of the IDDQ only failures may become delay test failures with more robust delay testing or in more sensitive or smaller feature size technologies that operate at higher speeds [29]. Failure analysis revealed that some of the defects inducing timing related failures are:

(1) drain to source leakage or shorts,

(2) low PMOS V_t causing low Idsat,

(3) poly to poly, ground, or VDD shorts,

(4) other resistive shorts and leakage mechanisms [29].

The similarity between IDDQ and delay related failures is striking and is duly noted in [29], "In general, we did not see a fundamental difference between IDDQ and timing-related defects."

2.3.3 Test Escapes

It should be obvious from Figure 2.6 that there is no test that can catch all defects. It shouldn't be a surprise that some defects are not caught by any tests. Faulty ICs that are not detected by any of the tests in a manufacturer's test suite are referred to as test escapes.

Untested faults or defects are among these test escapes. As described in [26], "Very high fault coverages are typically greater than 99%. This means that 1% of the modeled faults are not tested." An Intel study presented failure analysis of one defect

mode on their Pentium II IC that arose as a test escape because the test suite had less than 100% fault coverage. The defect was a missing via that caused faulty behavior with one customer's specialized use of the design. Other customers may never be affected by this defect. It was estimated that the defect accounted for 60DPM (defects per million) [25].

Most stuck-at based test methods rely on single stuck-at faults. The presence of multiple stuck-at faults in a circuit may create a situation where the faults are aliased and undetectable, a situation that would result in a test escape. Test escapes might also include unmodeled faults, those defects whose effects are not described by the popular stuck-at fault model. This would include reliability failures, ICs whose latent defects are aggravated by use of the IC. BURN-IN was designed to screen such ICs [26].

Many parametrically faulty ICs are not detectable with stuck-at, functional, or lenient delay fault testing and are destined to become test escapes. Such ICs could be screened with aggressive delay fault testing or f_{max} tests [16]. This would include those ICs that were manufactured with an unfortunate combination of processing conditions leaving them in a far corner of the processing/performance window [26]. Those test escapes that may be detectable with delay testing may be caused by defects such as:

(1) highly resistive contacts, vias, or interconnect,

(2) resistive opens or shorts,

(3) gate oxide shorts [26].

2.4 Temperature Sensitive Defects & Mechanisms

2.4.1 Environmental Variables

One example of environmentally sensitive failures are those caused by soft errors. These test escapes are sensitive to operating environments with high concentrations of high energy particles, typically from radioactive material, solar radiation, or cosmic radiation [16,26,36]. More often, though, environmentally sensitive or parametric failures are sensitive to much more mundane conditions. Raising or lowering operating temperatures, voltages, or frequencies is more likely to aggravate defects enough to make them testable [26]. High temperature testing of embedded DRAM can identify some defect mechanisms [37].

Of particular interest to this study are those defects that are sensitive to the operating temperature of the device.

2.4.2 Resistive Vias

Resistive vias are extrinsic defects that are detectable when aggravated by low temperature testing [16]. There are many defect modes that can cause resistive vias, a few of which are:

(1) Etch ash residue,

(2) Residual moisture in the via,

(3) Insufficient filling of the via,

(4) Via misalignment,

(5) Aluminum pushup,

(6) Random particles [38].

Failure analysis by one manufacturer showed their failing ICs had vias with gaps or holes causing bad contact between the vias and adjacent metal lines. As the IC temperature drops the metal in the vias contracts and the gaps and holes expand causing elevated via resistance values. These cold aggravated resistive vias cause the ICs to fail a single test vector. The fault is described as a delay fault that shows up as a *SA0* with cold testing and high frequency testing. Simulations of the defective IC with the defect as a series resistor supports these results. This cold aggravated defect was found to account for about 20DPM [25].

2.4.3 Metal Slivers

Another temperature sensitive defect mechanism and a good example of a reliability risk is the presence of metal slivers as a product of CMP. When exposed to high temperatures the slivers will undergo thermal expansion and could grow enough to short the metal interconnect. As noted in [16], BURN-IN or voltage stress may cause "rupture of the high resistance oxide surface of the metals, bonding the metal

elements." This would leave a metal short at all temperatures. It is curious to note that copper has a lower thermal coefficient of expansion than aluminum ([39], 14-1 to 14-2). This fact may lead to an alleviation of the temperature sensitivity of defects as copper replaces aluminum in IC interconnects.

2.4.4 Silicide Defects

Silicide open defects may be a very significant defect mechanism affecting poly-silicon/silicide interconnects. The poly-silicon used for transistor gates and some interconnect is resistive enough to have become a limiting factor in the speed of ICs. It is common practice to lower the sheet resistance of the poly-silicon interconnect by adding a layer of silicide to the top of the polysilicon. Some common silicide materials are WSi_2, $TaSi_2$, $TiSi_2$, and $CoSi_2$. The metal is deposited on the polysilicon and then a two step anneal process is used to bond the polysilicon with the metal. The second anneal uses temperatures and times that can be high or long enough to lead to partial or total disassociation between the grains, an open defect. This is a particular problem with designs that incorporate both large and small areas of silicide. The larger areas take longer to agglomerate than the smaller areas making the smaller areas more susceptible to silicide open defects [35,40].

The use of temperature to test for these silicide open defects relies on the resistive characteristics of polysilicon versus those of silicide. The sheet resistance of

silicide drops with lower temperatures, referred to as a positive resistance temperature coefficient (RTC). Polysilicon, however, has a negative RTC, the sheet resistance rises with lower temperatures. The mechanism for conduction in polysilicon relies on the thermal energy of free carriers in the material. When temperatures drop, the thermal energy drops and free carriers are trapped at the grain boundaries resulting in a negative RTC for polysilicon. The positive RTC of the silicide typically outweighs the negative RTC of the polysilicon leaving a silicide/polysilicon resistor with a positive RTC. The opposite is also possible though, particularly in cases where silicide open or partially open defects are present. This discrepancy makes cold temperature testing particularly effective in detecting silicide open defects. This is even more significant because of BURN-IN's inability to catch these silicide open defects as they are not particularly sensitive to voltage or heat stressing [35,40].

Effective testing for resistive or open silicide defects was demonstrated in [35,40]. The study showed that when 18,965 ICs were tested functionally at 100°C and 0°C, 53 ICs passed at 100°C and failed at 0°C. MinVDD testing at 0°C was moderately successful at identifying these defects. For example, 16 ICs from the same wafer were subjected to MinVDD testing at 0°C at both 333Mhz and 100Mhz. One of these 16 ICs was a member of the population of 53 cold functional rejects mentioned above. This cold reject was only distinguishable from the good ICs with 333Mhz MinVDD testing. The 100Mhz MinVDD test did nothing to identify the cold reject [35,40].

An earlier study, [25], presents failure analysis results on a single sample with a temperature sensitive silicide open defect. In describing the test results for this sample the authors of [25] say, "A shmoo plot confirmed that the failure would be seen at cold, would be marginal at room, and would pass at hot. At room temperature, the shmoo plot shows that the failure would be worst case at high frequency and low VCC. From this temperature shmoo, we can predict that the defect was probably due to a poly-silicon break." It was found that the state and behavior of the rest of the circuit had a profound impact on the pass/fail status of the IC. In the authors' estimation, "Temperature, voltage, and tight frequency testing were determined to be the best screens for this defect type." Their conclusions about the utility of temperature, frequency, and power supply voltage in testing are more broadly applicable than just to the case of silicide breaks. Data from a 0.25µm 32 bit adder shows that delay is more strongly dependent on temperature at lower VDD values than at higher VDD values [18].

2.4.5 Reverse bias pn diode

The current density in a reverse biased pn diode is described as:

$$J_R = q\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^{\,2}}{N_D} + \frac{qn_iW}{\tau_e}$$

Equation 2.18 ([5], pp. 91)

Where D_p is the hole diffusion coefficient, n_i is the intrinsic carrier density, τ_p is the minority (hole) carrier lifetime, τ_e is an effective lifetime, W is the base width, N_D is the donor concentration, and q is the electron charge. Such equations are of limited practical utility without reliable, empirically determined values for the first principles variables. When presented, Equation 2.18 is followed with data showing a discrepancy of 2 to 3 orders of magnitude between the ideal and real reverse bias *pn* diode leakage ([5], pp. 91). It is also significant to note that D_p, τ_p, τ_e, and n_i are all functions of temperature and the reverse bias current is strongly dependent on temperature. In [4] the reverse bias current of an *n*+-*p* diode was evaluated at several temperatures between 23°C and -197°C (the temperature of liquid nitrogen). The measured data shows significant reductions in reverse bias leakage. Between 23°C and -197°C the leakage dropped by three orders of magnitude. The theory, however, predicted a much larger change, as noted in [4], "The anticipated degree of improvement for an ideal *p-n* junction is proportional to the change in the intrinsic carrier density n_i, a change of about 30 orders of magnitude." A difference explained by the authors of [4], "In a practical situation, however, junction curvature, local defects, and other effects greatly increase the leakage current over its ideal value." For the purposes of this study it is sufficient to understand that reverse bias *pn* junction leakage is strongly tied to temperature and that the leakage current is reduced by a reduction in temperature.

2.4.6 Subthreshold Current

Similar to reverse bias *pn* junction leakage, the effect of temperature on subthreshold current can be confounded when attempting to describe it using theory. In [41], the author describes the subthreshold drain to source current, I_{DS}, as:

$$I_{DS} = \frac{\mu C_{ox} W}{L_E}(\eta - 1)\Phi_t^2 \exp\left(\frac{V_{GS} - V_t}{\eta \Phi_t}\right)$$

Equation 2.19 [41]

where $\Phi_t = kT/q$ [41], "μ is the carrier mobility, C_{ox} is gate oxide capacitance, W is transistor width, L_E is effective channel length and $\eta = 1 + C_{DEP}/C_{ox}$, where C_{DEP} is the channel depletion capacitance." The author of [41] then refers to Equation 2.19 and states that, "it can be seen that the leakage current grows exponentially with rising temperature." In fact, μ is proportional to about $T^{-2.2}$ or $T^{-2.42}$ depending on dopant type and concentration [42], V_t is linearly related to temperature [4,30], L_E's dependence on temperature depends on the drawn channel length as well as the doping profiles and concentrations in the transistor, and in a subthreshold condition V_{GS} is less than V_t. This leaves a relationship between temperature and subthreshold current that is strongly dependent on technology. The data presented in [30] does show a roughly exponential relationship between subthreshold current and temperature.

When considering the effect of temperature on subthreshold current there are two factors that must be considered. First, the subthreshold slope of the transistor's IV curve, and second, the shift in this curve [30]. The shift in the IV curve is synonymous with a shift in V_t.

The inverse of the subthreshold slope, S_t, is derived from Equation 2.19 and given as:

$$S_t = 2.3\frac{kT}{q}\eta$$

Equation 2.20 ([6], pp. 128)

Some qualitative observations can be made here. As temperature decreases the slope increases. This characteristic of the IV curve is favorable as it means that the transistor will turn off over a smaller range of V_G which facilitates the operation of CMOS circuits ([6], pp. 287). It also means that the subthreshold current is lower for a greater range of V_G.

The second factor in temperature's influence on subthreshold current is the threshold voltage, V_t. For the transistors used in [30] the temperature coefficient for V_t was empirically determined to be ~0.8mV/°C. This increase in V_t with cooler temperatures serves to increase the range in V_G over which the transistor is turned off and has a subthreshold I_{DS}.

By changing the temperature from 25°C to -50°C the combination of the changes in subthreshold slope and V_t reduced the subthreshold current by a factor of 356 [30]. Qualitatively similar results were shown in [4] for a 9μm transistor.

2.4.7 Drain Induced Barrier Lowering (DIBL) & Punchthrough

The potential barrier that prevents conduction of electrons or holes between the drain and source of an off transistor is, in the ideal case, constant across the length of the channel. However, there exists a parasitic effect induced by the voltage placed on the drain of the transistor. The drain voltage reduces the potential barrier near the channel to drain junction effecting a variable potential barrier across the channel length. For long channel lengths the approximation of a constant potential barrier is reasonably accurate ([6], pp. 143-144). This approximation is not suitable, however, for short channel transistors. As demonstrated in Figure 2.7, when the channel length is reduced the maximum value of the potential barrier is reduced and the effect of the drain on the potential barrier becomes significant across the entire channel length. Increasing the drain voltage has an analogous effect on the potential barrier, the peak value decreases and the barrier becomes more graded across the channel length.

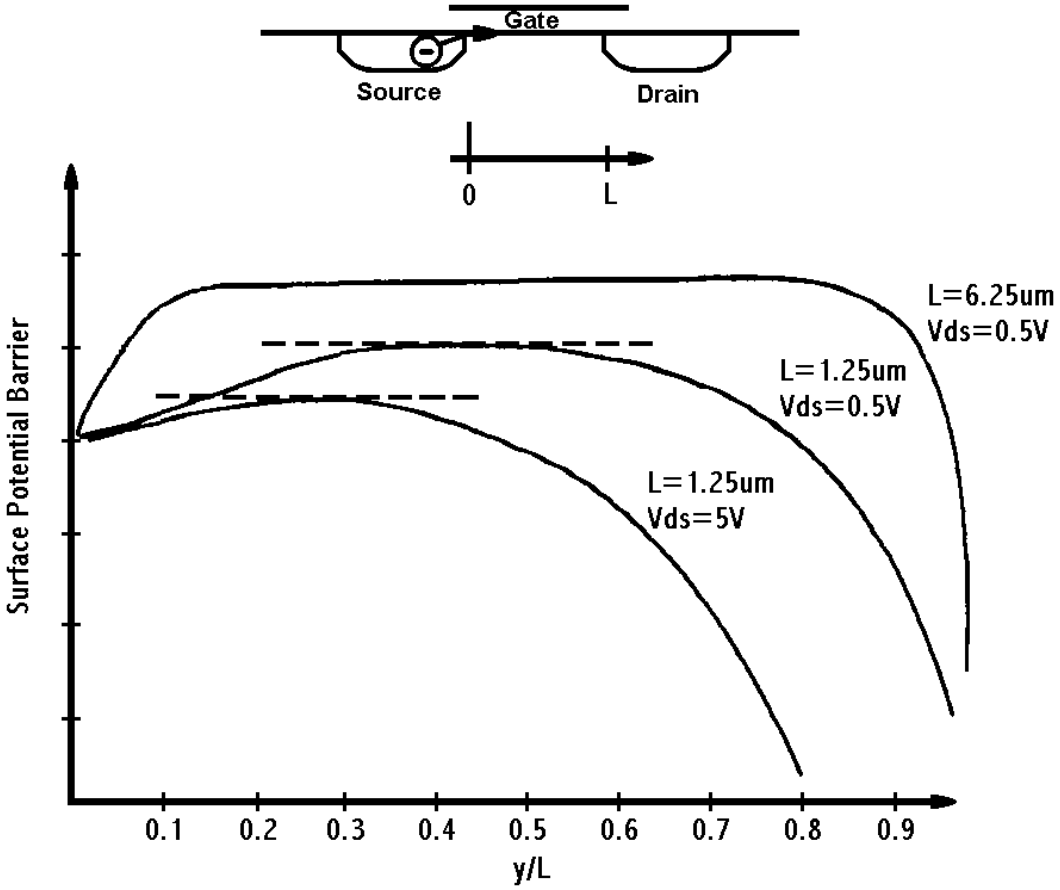


Figure 2.7 (after [43], Figure 1) - DIBL And The Channel Surface Potential Barrier

Punchthrough is simply an extreme case of drain induced barrier lowering (DIBL). When the drain induced depletion region reaches the source induced depletion region a space charge is created providing a current path between source and drain [30]. In a punchthrough condition the gate has no control over drain to source current ([6], pp. 144).

Taking a superficial approach to the influence of temperature on the severity of DIBL would indicate that DIBL will change for the same reasons that the threshold voltage changes. Lower temperatures will induce higher potential barriers in the

channel and shrink the depletion region created by the drain bias. So, as the temperature is lowered, the potential barrier in the channel is raised, the V_t increases, the subthreshold current is lowered, and DIBL will be relieved. Conversely, as temperature is increased, DIBL will be aggravated.

It is important to recognize that the drain to source current due to DIBL is closely tied to the subthreshold current. Although the magnitude of the DIBL induced current will be lower with lower temperatures, the DIBL current as a percentage of the total subthreshold current may or may not change. The study in [44] shows data supporting the notion that the DIBL coefficient is temperature sensitive, particularly for shorter channel lengths. An insensitivity to temperature of DIBL as a percentage of subthreshold current is shown in [54] for deep submicron MOSFETs between -223°C and 27°C and in [52] for 0.25μm PMOS transistors between -50°C and 125°C. Further complicating the picture of DIBL's temperature dependence or independence, a phenomenon wherein the polarity and magnitude of the change in DIBL induced leakage current due to temperature shifts is dependent on channel length is demonstrated in [55]. The author of [55] states, "There is a range of channel lengths between the longer channel lengths and the very short channel lengths, where DIBL is worse at 77K than at 300K." Their findings are supported by both experimental data and simulations. This phenomenon is attributed to the DIBL current at 27°C being subsurface current whereas the DIBL current at -196°C is surface current. This current path variation from surface to subsurface is influenced by the use of boron channel implants in an NMOS transistor and, as the author of [55] describes it, is a

result of, "larger surface potential bending (downward in NMOS device case) caused by the freeze-out effect in the substrate, resulting in more electrons being attracted towards the channel surface." Though the temperature at which freeze-out occurs is a function of doping concentrations this effect should not be a concern for most ASIC applications as freeze-out only occurs below about -173°C for a relatively modest donor impurity concentration of $10^{15}/cm^3$ ([5], pp. 25-26).

Although lowering the operating temperature will delay the onset of punchthrough, once the punchthrough condition is established temperature should have little or no impact on the magnitude of the drain to source current. The Child-Langmuir law shows that the current in a space charge region is determined by the potential between the anode and cathode:

$$I_{DS} = kV_{DS}^{3/2}$$

Equation 2.21

where k is some temperature independent constant.

2.4.8 Gate Induced Drain Leakage (GIDL)

The gate induced drain leakage phenomenon occurs in the region where the gate overlaps the drain. When, for an NMOS transistor, the gate is biased at or below zero and the drain has a positive bias, the gate to drain overlap region may become

inverted as shown in Figure 2.8 ([6], pp. 99-100). When GIDL occurs a current path is created between the drain and substrate.

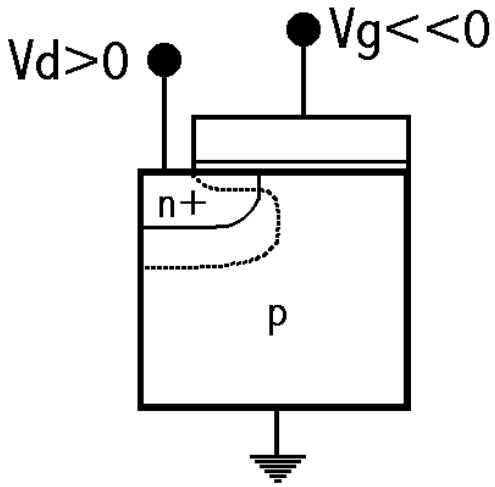


Figure 2.8 (after [6], Figure 2.46c) - Detail Of GIDL

The GIDL effect is aggravated by an increased gate to drain potential difference, thinner gate oxides that accompany deep sub-micron devices, and lightly doped drain (LDD) technology [30]. The GIDL phenomenon depends on the inversion of the gate to drain overlap region and thus depends on the band gap energy and the difference between the intrinsic Fermi potential and the Fermi potential. With this in mind, it should be apparent that GIDL has a temperature dependence similar to that of the threshold voltage. As temperature drops, the gate to drain potential difference required for GIDL to occur is increased. As temperature increases, leakage current due to GIDL is aggravated, and as operating temperature is lowered, GIDL is relieved. The sensitivity of GIDL to temperature changes is also a function of the electric field strength, higher field strengths reduce the sensitivity of GIDL to

temperature. It is shown in [56] that with a gate voltage of 7V and a temperature shift from ~27°C to ~200°C the GIDL is increased by about two orders of magnitude. Whereas with a gate voltage of 1V the GIDL increases by about four orders of magnitude with the same temperature shift.

2.4.9 Oxide leakage

The gate oxide of a MOSFET is used as a dielectric to separate the polysilicon gate from the transistor channel. The SiO_2 used for gate oxides has a high but finite resistivity, $\sim 10^{15}$ Ωcm. Any bias placed on the gate will induce current to flow through the gate oxide resistor. The resistivity of the oxide is not constant though, and is lower when the electric field in the oxide is higher as a result of either higher gate potentials or thinner gate oxides.

Two fundamental mechanisms for gate current are Fowler-Nordheim tunneling and direct tunneling ([1], pp. 389-390). Direct tunneling is the dominant leakage mechanism for smaller gate oxide thickness and for lower oxide electric field strengths. Fowler-Nordheim tunneling does not make a significant contribution to the gate oxide current under normal operating conditions [30].

Fowler-Nordheim current is a function of the effective potential barrier height between the silicon and the gate oxide and will vary with temperature accordingly. As temperature drops the barrier height will be raised and the Fowler-Nordheim current

will decrease. The precise nature of the temperature dependence of Fowler-Nordheim tunneling is explored in depth in [45].

Direct tunneling is closely related to Fowler-Nordheim tunneling though the details of the conduction mechanisms are unique. Direct tunneling involves electrons tunneling directly from the semiconductor, through the entire oxide thickness, to the gate polysilicon, or from gate polysilicon to semiconductor. The temperature dependence of oxide leakage due to direct tunneling is reported in [46] as being exponential. The gate current for both NMOS and PMOS capacitors with oxide thickness of 37Å was reported as decreasing by about three orders of magnitude as temperature was decreased from 175°C to 25°C. Similar results were found with the capacitors in accumulation and inversion at gate voltages of 1.8V, 2.1V, and 2.5V. The authors conclude that the exponential temperature dependence of the gate leakage is related to the thermionic emission process for conduction in insulators [46].

2.4.10 Hot Carrier Injection (HCI)

Hot carrier injection (HCI) occurs when electrons gain large kinetic energies and are injected or propelled into the gate oxide. This occurs as a direct result of the electric field that lays perpendicular to the gate oxide that is induced by a gate potential. These hot carriers can contribute to the overall device current consumption in several ways. They can flow through the gate oxide contributing to gate current.

They can become trapped charge within the oxide or at the Si-SiO_2 interface that will alter the threshold voltage and the subthreshold slope. Impact ionization can create photons that will in turn create more electron-hole pairs that contribute to parasitic device currents ([1], pp. 248-249). They can become substrate current which will also cause "a voltage drop in the substrate, forward biasing the source-substrate junction, leading to further impact ionization" ([1], pp. 249). HCI effects are relieved somewhat by LDD technology. An LDD adds series resistance to the drain to channel to source current path and alleviates areas of high electric field intensity. The primary drawback to LDD technology is the lowering of saturation current capabilities of the transistor which reduces IC performance ([6], pp. 161).

It has long been presumed that lower temperatures will aggravate HCI. This is due to the fact that lattice atoms have less thermal energy which increases the mean free path of electrons. With a longer mean free path, mobile electrons within the Si lattice have larger kinetic energies making electron collisions with the Si lattice or impact ionization events more severe. The longer mean free path coupled with larger drain to source currents at lower temperatures result in aggravated HCI [47]. Experimental results show aggravated HCI at lower temperatures with drain current degradation for both NMOS and PMOS transistors. Though the HCI aggravation at low temperatures was worse with NMOS transistors [48]. Some studies have shown the rate of impact ionization is either independent of temperature or is reduced as the temperature is lowered, contrary to the expected behavior. This contrary behavior is characteristic of lower drain voltages and may or may not be a function of gate length

[49]. In [50] drain current degradation was found to be aggravated at 100°C as compared to 23°C. This improvement in HCI at lower temperatures has been verified with simulation tools and transistor modeling in [51]. Another characteristic used as an indicator and measure of HCI is the substrate current, I_{sub} [47]. When considering I_{sub}, it has been observed that there exists a set of technology dependent operating conditions (gate voltage and drain voltage) that separates regions of positive temperature coefficients from negative temperature coefficients. A low VDD value, 1.5V, shows lower I_{sub} values at 25°C than at 125°C. At a VDD of 2.5V the I_{sub} values are equivalent for both 25°C and 125°C [47]. In [50], 0.15µm NMOS transistors demonstrate higher I_{sub} at lower temperatures for VDD greater than 2.3V. The said transistors show lower I_{sub} at lower temperatures for VDD less than 2.3V. The transition point for PMOS transistors was 3.1V. It is critical to recognize that I_{sub} and drain current degradation do not share the same temperature behavior [47]. Drain current degradation and substrate current measure different HCI mechanisms with unique temperature dependent characteristics.

2.4.11 Speed Temperature Coefficient (STC)

Both resistive vias with holes or gaps and silicide open defects have a negative RTC that would lead to ICs being slower at lower temperatures. This behavior will hereafter be referred to as a positive speed temperature coefficient or positive STC.

There are, however, IC performance characteristics that result in a negative STC, the IC is faster at lower temperatures. The existence of both positive and negative STCs is noted by another study [26], "Note that the two defective devices have different characteristics – one gets slower at low temperatures (below 80 deg. C), one gets faster at low temperatures." Although the data in [18] shows a stronger temperature dependence in the VDD range where the 32 bit adder has a positive STC, there is a VDD range where the device shows a negative STC. It is significant to note that this negative STC occurs with VDD above 1V, well below the 2.5V that a 0.25µm technology is likely to operate.

This negative STC relationship between frequency and temperature was investigated further in [16]. With 20,000 ICs, a relationship between the maximum operating speed of an IC, f_{max}, and temperature was established. Intrinsically fast ICs had an f_{max} to temperature ratio of –13.1KHz/°C while the ratio for intrinsically slow ICs was –12.3KHz/°C. So, lower temperatures resulted in faster ICs, though intrinsically slow ICs were less sensitive to temperature than fast ICs [16].

2.4.12 IDDQ's Temperature Dependence

The use of temperature in testing has also been shown to be effective in detecting defects that cause high IDDQ measurements (IDDQ is introduced in section 2.2.2.1). Sandip Kundu, an advocate of the use of temperature variations in IDDQ

testing, explains in [41], "IDDQ current is dominated by transistor sub-threshold leakage and if it is measured at two different temperatures, then the growth of leakage current from lower to higher temperature must be exponential. An assumption is made that the current due to defects will not have exponential growth with temperature. This is predicated on the fact that the two predominant leakage mechanisms, the resistive bridges between metal lines and permanently ON transistor currents, show little susceptibility to temperature variation."

With a 0.18μm technology an 83°C (from 110°C to 27.2°C) drop in temperature effected a reduction in IDDQ by a factor of about 40. Between 110°C and -55°C there was difference in IDDQ of three orders of magnitude. Defective ICs have much smaller shifts in IDDQ between temperatures [15]. An exponential relationship between IDDQ and temperature is predicted for healthy ICs. Those ICs that do not follow this relationship are shown to be defective as they also fail other tests presumed to be good indicators of defective ICs [41]. The discrepancy between the IDDQ versus temperature behavior of healthy ICs and that of defective ICs is supported by data in [26] as well as simulations using resistive shorts from VDD to ground in [15].

The effect of temperature on the intrinsic component of IDDQ can be explained by the mechanisms that cause IDDQ. Such intrinsic sources of IDDQ common to all designs include:

(1) Reverse bias P-N junction leakage,

(2) Subthreshold current,

(3) Drain Induced Barrier Lowering (DIBL),

(4) Channel Punchthrough,

(5) Gate Induced Drain Leakage (GIDL),

(6) Oxide leakage,

(7) Hot Carrier Injection (HCI) [30].

A few qualitative conclusions may be made about the temperature dependence of the IDDQ of an IC or the off state current of the IC's constituent transistors. As the operating temperature is reduced the current induced by six of the mechanisms described above is reduced. Only HCI induced IDDQ or off state current will be aggravated by lower temperatures and this only occurs in some technologies operating under certain conditions. With this said, it is important to ask how much do each of these mechanisms contribute to the IDDQ at various operating temperatures and what is the cumulative effect of temperature on IDDQ. Data for a 0.25μm PMOS transistor shows the off state current, I_{off}, being reduced by four orders of magnitude when the operating temperature is reduced from 125°C to -50°C [52]. This I_{off} is typically dominated by the subthreshold current [30,53]. These conclusions must be carefully qualified, though. The total I_{off}, the relative contribution of each of these leakage mechanisms to I_{off}, as well as the temperature dependence of each of the leakage mechanisms is strongly dependent on the device technology being used. In [30] it is shown that I_{off} for 0.18μm transistors is about five orders of magnitude higher than for 1.0μm transistors. This study also shows that I_{off} for 0.36μm NMOS and PMOS transistors with either LDD or HDD (highly doped drain) implants vary widely.

Another study, [53], gives a detailed examination of the relative contributions of several leakage mechanisms to I_{off} for NMOS and PMOS transistors made for low voltage, low power, or ultra low power applications in 0.13µm, 0.15µm, and 0.18µm technologies. They find that subthreshold current can contribute as much as 99.75% of the total I_{off} for 0.18µm low voltage PMOS transistors to as little as 40% for both 0.13µm and 0.15µm ultra low power PMOS transistors. In contrast, GIDL contributes no more than 26.1% to I_{off} in the case of 0.13µm low power NMOS and as little as 0.0004% in the case of 0.13µm low voltage PMOS transistors. The study also finds that subthreshold current can be significantly more temperature sensitive than other mechanisms. As a result, the contribution of subthreshold current to I_{off} is strongly exaggerated at 125°C as compared to 25°C. For the 0.13µm low voltage PMOS transistors subthreshold current is dominant for the entire temperature range, 125°C to 25°C, but for the 0.13µm ultra low power PMOS transistors both the magnitude of subthreshold current is comparable to other leakage mechanisms. The authors of [53] also observe that the stand-by current (analogous to IDDQ) for an SRAM cell is a function of the boron implant dosage for that cell. SRAM cells, using 0.13µm low power transistors, with higher boron implant dosages have stand-by currents with higher percentage GIDL current which in turn makes stand-by current less temperature sensitive.

The temperature dependence of IDDQ due to extrinsic defects on an IC are even less predictable than the intrinsic sources of IDDQ. The temperature dependence of the defect induced IDDQ could be as diverse as the defects themselves and will

certainly be strongly dependent on the quantity, location, and severity of the defects. Many resistive shorts may be presumed to have a linear relationship to the IDDQ current that they induce, though this is likely not always true. What is certain is that there are many extrinsic defects that will cause an IC to have an IDDQ temperature dependence that can not be attributed to intrinsic leakage mechanisms. So, by lowering the testing temperature, the contribution of intrinsic leakage mechanisms to an IC's total IDDQ will be reduced. This reduction in the background IDDQ will help highlight defect induced IDDQ signatures.

Chapter 3

Experiment

3.1 Experimental Design

The motivation of this experiment is to evaluate the potential of using multiple temperature or targeted temperature testing for identifying both non-functional and low reliability ICs. This is possible because most sort tests as well as parametric measurements will have healthy, intrinsic shifts in response to changes in temperature. This healthy behavior is distinct from the behavior of ICs with functionality or performance limiting defects. When the sort testing is conducted at multiple temperatures, the change in a measured value, as opposed to a single measured value, may be used to separate defective ICs from healthy ICs. Alternatively, a single temperature may be chosen to run a test at based on the contrast that that temperature produces in the test results between healthy and defective ICs. The IDDQ measurements, for example, of defective ICs and those of healthy ICs may be indistinguishable at high temperatures but clearly separable at low temperatures. So, sort testing at targeted temperatures may be useful in separating defective ICs from healthy ICs.

Two ASIC designs fabricated in two different technologies were used as experimental vehicles and are referred to herein as ASIC1 and ASIC2. Some defining characteristics of the two ASICs are given in Table 3.1.

	Technology node	Gate Count (millions)	Die size X,Y (mm)	Metal Layers
ASIC1	0.25μm	Not Available	7.8,7.6	3
ASIC2	0.18μm	5.3	15.7,15.7	5

Table 3.1 - Design Characteristics Of Test Vehicles

One key to distinguishing the behavior of defective ICs from that of healthy ICs is clearly identifying the expected behavior of healthy ICs. The behavior of healthy ICs will, of course, vary with the process variations inherent to producing semiconductor ICs. Such process variations are reliably reflected in the electrical characteristics of the IC's constituent transistors. Said transistor characteristics are routinely measured in parametric test structures arranged in the scribe lines across the wafer, referred to herein as Etest structures and measurements. Etest measurements, with Etest structures occurring only once in each reticle field, are not measurements of transistor and interconnect characteristics on each IC. Rather, Etest structures are commonly assumed to be a good approximation of the transistor and interconnect characteristics on those ICs spatially close to the Etest structure. ASIC1 is used to more directly associate sort measurements to an IC's constituent transistor characteristics by characterizing an NMOS and a PMOS transistor for each IC tested at wafer sort.

The healthy IC behavior is then compared to and distinguished from the behavior of defective ICs. Both ASIC1 and ASIC2 are used to compare and contrast the utility of targeted and multiple temperature testing with a variety of sort tests in identifying defective and low reliability ICs.

3.2 Test Vehicles

3.2.1 ASIC1

ASIC1 data is collected from two wafers with a total of 912 ICs fabricated in a 0.25μm CMOS technology. Each ASIC1 IC contains a single PMOS and a single NMOS transistor that are not part of the functional IC but do occur on the IC. These PMOS and NMOS transistors are located in opposite corners of each IC and as such will hereafter be referred to as corner transistors to distinguish them from the transistors in the Etest structures and the constituent transistors of the IC. Each of these corner transistors has its own set of contact pads devoted to the gate, drain, source, and substrate contacts.

Two ASIC1 wafers were tested at 30˚C and at 85˚C with two different test sets,

1) Corner transistors,
2) Wafer sort.

All corner transistors on each of the two ASIC1 wafers were tested using a Keithley 4200 semiconductor characterization system and a Cascade Microtech 12000 semi-automated probe station. The tests run on the corner transistors of each IC include two fundamental tests, threshold voltage and drain current measurements. The threshold voltages were determined using a technology dependent, preset drain current. A binary search routine that forces a voltage on the gate and measures the drain current was used to search for this preset drain current. Upon the sixteenth measurement in the binary search the gate voltage is recorded as the threshold voltage. Drain current measurements were taken with each of the conditions given in Table 3.2.

Test #	Gate (V)	Drain (V)	Source (V)	Substrate (V)
1	2.5	2.5	0	0
2	1.25	2.5	0	0
3	0.625	2.5	0	0
4	0.3125	2.5	0	0
5	0	2.5	0	0
6	1.25	1.25	0	0
7	0.625	0.625	0	0
8	0.3125	0.3125	0	0

Table 3.2 - ASIC1 – Corner Transistor Measurement Conditions

The wafer sort testing included MinVDD testing on three separate logic blocks and Static IDD testing on the core circuitry.

3.2.2 ASIC2

ASIC2 is a relatively large area ASIC fabricated using 0.18μm technology. The sample set for ASIC2 was composed of 314 packaged parts which were primarily final test rejects. ASIC2 has 804 IO pins and is normally tested on an expensive, high speed, high pin count tester, the Schlumberger IX9000 tester. A load board and test program were created to conduct limited testing of the part on a low pin count tester, the Credence Quartet tester. As ASIC2 was tested in packaged part form and disassociated from the wafers, no Etest data is available.

The test suite was run on each of the 314 ASIC2 packaged parts at three separate temperatures, 75˚C, 22˚C, and –40˚C. A Thermonics T-2420 Precision Temperature Forcing System was used to control the temperature of the IC during testing at 75˚C and –40˚C. Prior to testing, each IC had hot or cold air forced over it by the Thermonics unit for between sixty and ninety seconds. The 75˚C temperature was chosen because this is the temperature at which the final testing of the IC is conducted before the IC is shipped to the customer. The MinVDD testing of the LBIST at 75˚C on the Credence Quartet may then be compared directly to the final test results from the Schlumberger IX9000 tester. The 22˚C is the ambient or room temperature of the area in which the parts were tested. The -40˚C temperature was chosen to be near the specification limit of the IC's operation temperature range.

The testing of ASIC2 includes MinVDD testing of two LogicBIST patterns at several different operating frequencies. The standard operating frequency for ASIC2

is 110Mhz, though, the maximum test frequency on the Credence Quartet tester is 100Mhz. The testing conducted on ASIC2 is summarized in Table 3.3 with an X indicating that the test was run under the listed conditions.

LogicBIST Pattern	Temperature (˚C)	Operating Frequency (Mhz)			
		20	66.6	83.3	100
L1	-40			X	
	22			X	
	75			X	
L2	-40	X	X	X	X
	22	X	X	X	X
	75	X	X	X	X

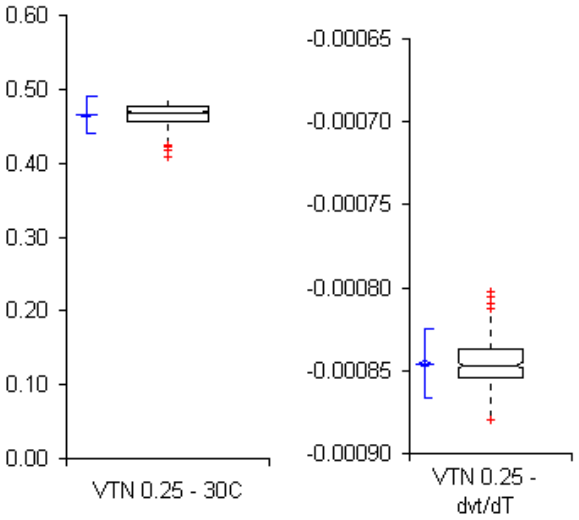
Table 3.3 - ASIC2 – MinVDD Search Conditions

Chapter 4

Results

4.1 dVt/dT

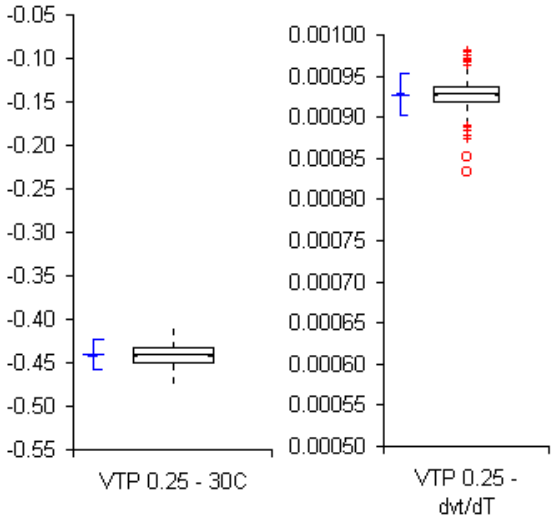
As noted in section 2.1.2, for any given transistor, the V_t changes linearly with temperature; lower temperatures induce V_t values further from zero. The rate of change of V_t with respect to temperature, dV_t/dT, is a constant for any given transistor but will vary for different technologies and for different transistors within the same technology. The distributions for the values of dV_t/dT from ASIC1 wafers are given in Figures 4.1 and 4.2 along with distributions for the V_t. These distributions confirm that lower temperatures drive a transistor's V_t away from zero. For instance, an NMOS transistor with a V_t of 0.468V at 30˚C will increase by 47 millivolts to 0.515V at 85˚C. PMOS transistors are slightly more temperature sensitive with a 51 millivolt shift in V_t when subjected to the same temperature shift.



	n	Mean	SD	SE	95% CI of Mean
VTN 0.25 - 30C	452	0.465	0.015	0.001	0.464 to 0.466
VTN 0.25 - dvt/dT	452	-0.00085	0.00001	6.0E-07	-0.00085 to -0.00084

	Median	IQR	95% CI of Median
VTN 0.25 - 30C	0.468	0.021	0.467 to 0.470
VTN 0.25 - dvt/dT	-0.00085	0.00002	-0.00085 to -0.00084

Figure 4.1 - ASIC1 VTN And dVTN/dT Distributions



	n	Mean	SD	SE	95% CI of Mean
VTP 0.25 - 30C	880	-0.441	0.010	0.000	-0.442 to -0.440
VTP 0.25 - dvt/dT	880	0.00093	0.00001	5.0E-07	0.00093 to 0.00093

	Median	IQR	95% CI of Median
VTP 0.25 - 30C	-0.441	0.016	-0.442 to -0.440
VTP 0.25 - dvt/dT	0.00093	0.00002	0.00093 to 0.00093

Figure 4.2 – ASIC1 VTP And dVTP/dT Distributions

The transistors within the technology are well enough behaved that, given a V_t measurement at 30°C or 85°C, the value of the V_t at 85°C or 30°C can be predicted to a high degree of certainty. This is demonstrated in Figures 4.3 and 4.4 where the R^2 value for the distribution of 456 NMOS transistors' V_ts is 0.9987 and of 880 PMOS transistors' V_ts is 0.9957. Such reliable behavior will make a significant contribution to the reliability of any predictions of temperature induced IC performance based on the analysis presented in section 4.2.

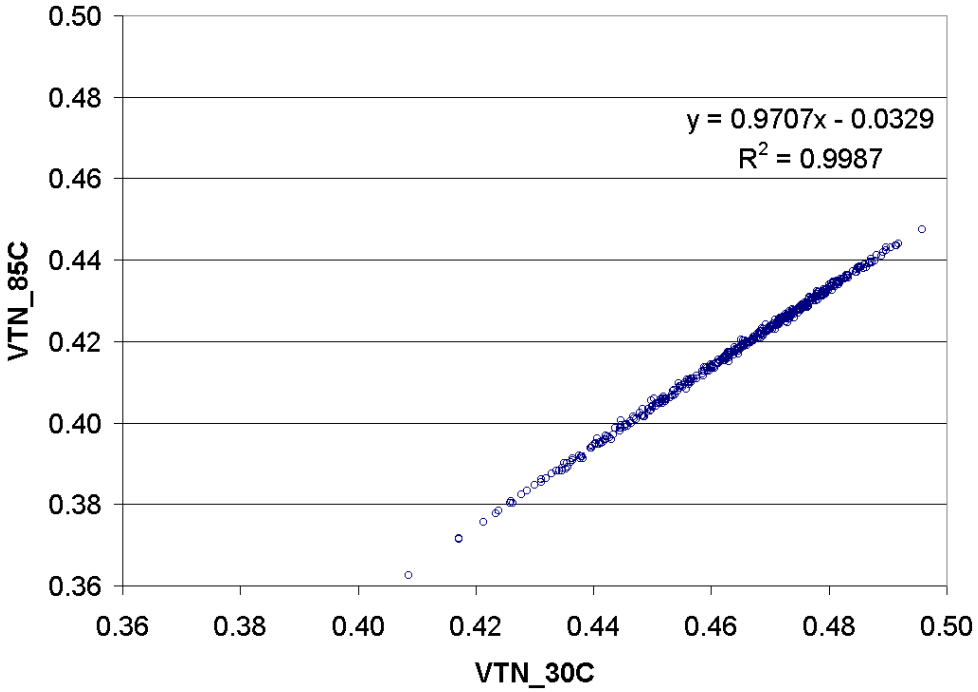


Figure 4.3 – ASIC1 Corner Transistor VTN At 85°C As A Function Of That At 30°C

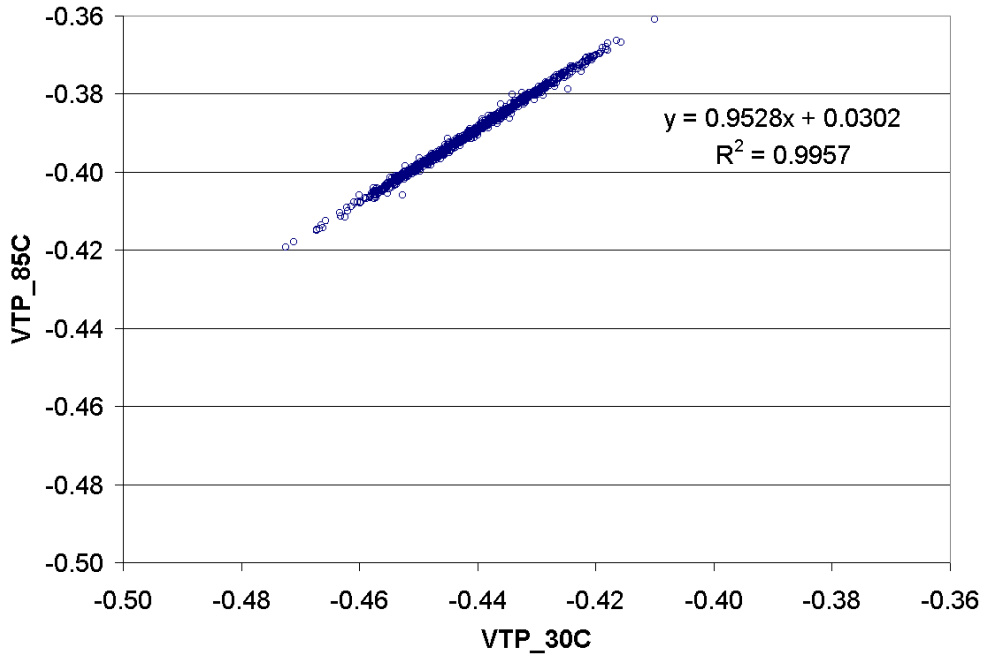


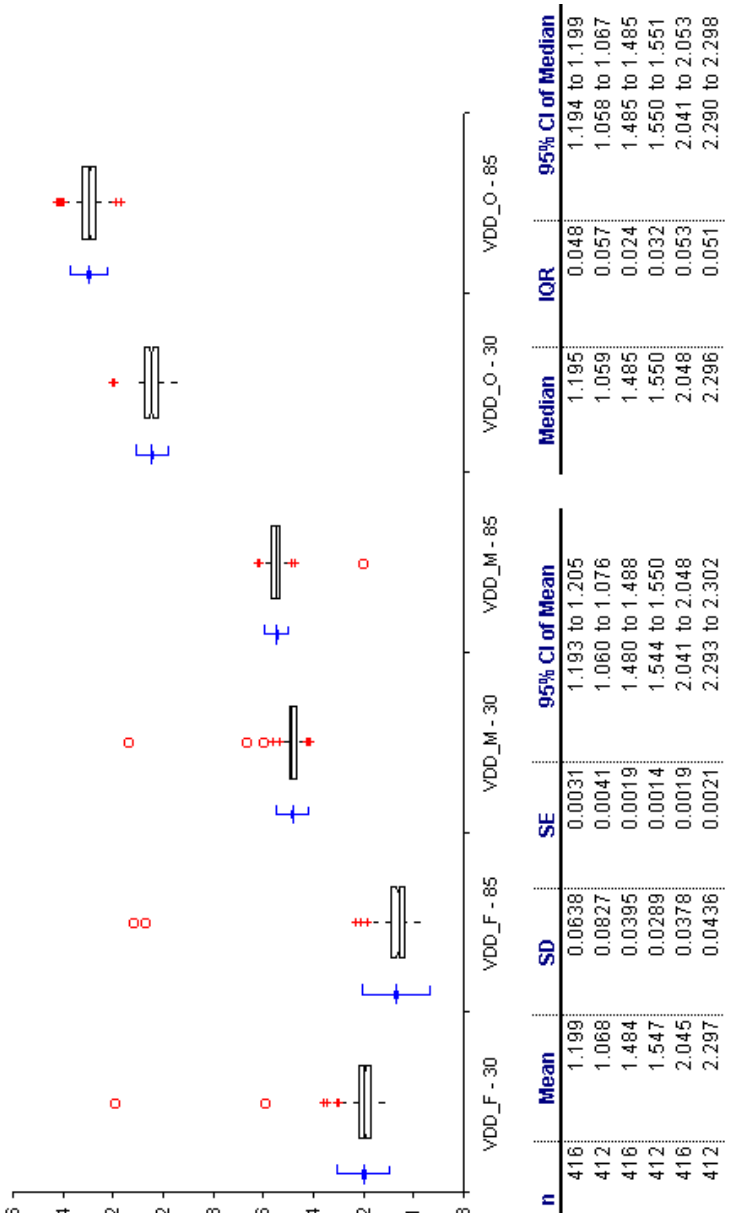
Figure 4.4 - ASIC1 Corner Transistor VTP At 85°C As A Function Of That At 30°C

4.2 Performance Temperature Response

MinVDD testing was conducted on three different blocks of ASIC1. The MinVDD measurements for the three blocks resulted in the distributions shown in Figure 4.5.

The most outstanding characteristic of these MinVDD measurements lies in the fact that when the temperature is increased from 30°C to 85°C the MinVDD distribution for block F is lowered. It is not particularly intuitive that when the temperature is increased the IC should perform better or at lower supply voltages.

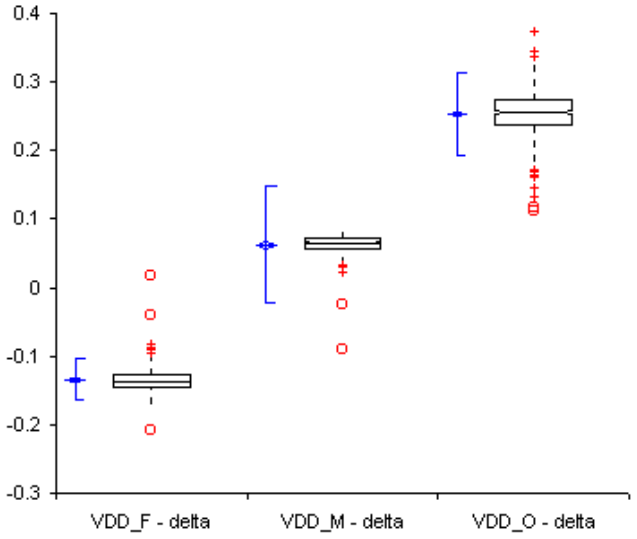
Blocks M and O show more intuitive behavior, that is, with higher temperatures the minimum supply voltage for correct operation is increased.



	n	Mean	SD	SE	95% CI of Mean	Median	IQR	95% CI of Median
VDD_F - 30	416	1.199	0.0638	0.0031	1.193 to 1.205	1.195	0.048	1.194 to 1.199
VDD_F - 85	412	1.068	0.0827	0.0041	1.060 to 1.076	1.059	0.057	1.058 to 1.067
VDD_M - 30	416	1.484	0.0395	0.0019	1.480 to 1.488	1.485	0.024	1.485 to 1.485
VDD_M - 85	412	1.547	0.0289	0.0014	1.544 to 1.550	1.550	0.032	1.550 to 1.551
VDD_O - 30	416	2.045	0.0378	0.0019	2.041 to 2.048	2.048	0.053	2.041 to 2.053
VDD_O - 85	412	2.297	0.0436	0.0021	2.293 to 2.302	2.296	0.051	2.290 to 2.298

Figure 4.5 – ASIC1 MinVDD Distributions

This behavior is summarized by looking at the temperature induced MinVDD delta. The said delta was found by subtracting the MinVDD taken at 30°C from that taken at 85°C, $\Delta MinVDD = MinVDD(85°C) - MinVDD(30°C)$, and is summarized in Figure 4.6. Notice that the MinVDD delta distribution for block F is negative while those for blocks M and O are positive.



	n	Mean	SD	SE	95% CI of Mean
VDD_F - delta	401	-0.135	0.0183	0.0009	-0.137 to -0.133
VDD_M - delta	401	0.063	0.0519	0.0026	0.058 to 0.068
VDD_O - delta	401	0.253	0.0360	0.0018	0.250 to 0.257

	Median	IQR	95% CI of Median
VDD_F - delta	-0.137	0.018	-0.137 to -0.136
VDD_M - delta	0.065	0.015	0.064 to 0.067
VDD_O - delta	0.257	0.038	0.252 to 0.258

Figure 4.6 – ASIC1 MinVDD Temperature Shift Distributions

Although this discrepancy between the MinVDD deltas of block F and blocks M and O may not be intuitive, it is not a fluke. Evidence presented in the literature

supports this change in the sign of temperature induced IC performance shifts as a function of supply voltage (see sections 2.1.4 and 2.4.11).

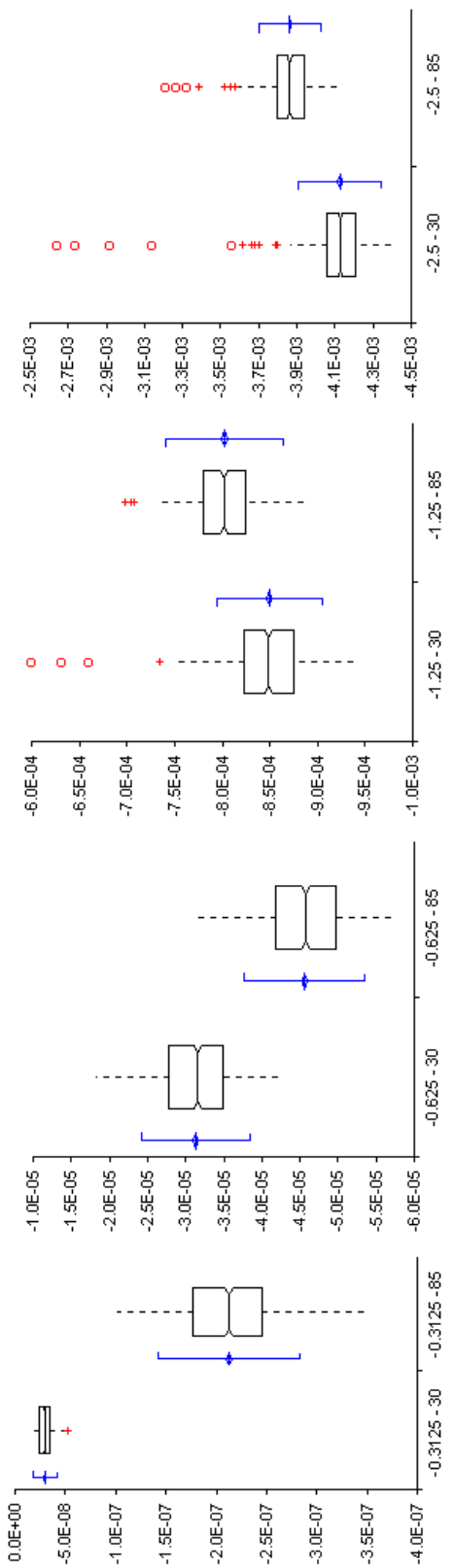
IC performance is governed by the IV curves, as in Figures 2.1 and 2.2, of the IC's constituent transistors. Temperature shifts modify these IV curves by changing the leakage current, the threshold voltage, the transconductance, and the saturation current, that is, nearly every aspect of the curve (see sections 2.1.2 and 2.1.4). Furthermore, the change in the threshold voltage and saturation current due to a temperature shift tend to compensate one another. Building a relationship that accurately reflects MinVDD as a function of temperature, MinVDD(T), by calculating the temperature dependence of each of these IV curve characteristics, V_t(T), I_{sat}(T), and *ILK*(T), from common transistor models, such as in Equation 2.1, and first principles is not a trivial task.

Certain qualitative conclusions can, however, be reached by making some approximations and assumptions. Some such attempts, as noted in section 2.1.4, have been made. Considering the relationship presented in Equation 2.11 a fundamental observation may be made. Higher drive currents in the constituent transistors of an IC enable the IC to perform at lower voltages and higher frequencies.

$$Higher\_I_{sat} \Leftrightarrow LowerMinVDD \Leftrightarrow HigherPerformance$$
$$Lower\_I_{sat} \Leftrightarrow HigherMinVDD \Leftrightarrow LowerPerformance$$

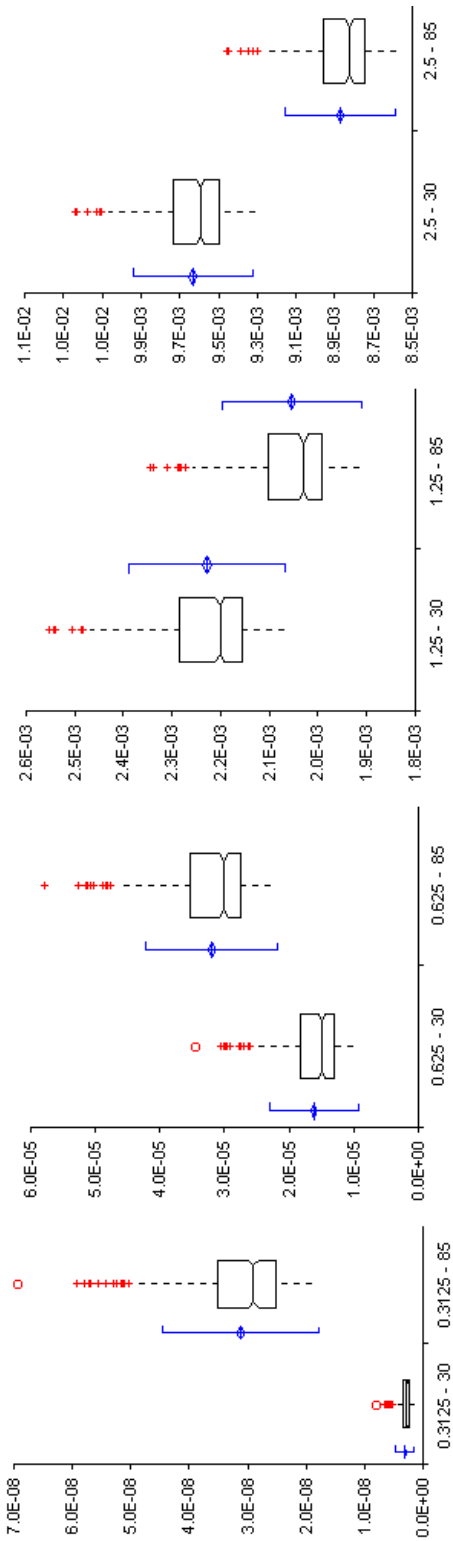
The behavior of the saturation current of the corner transistors does in fact reflect the MinVDD measurements. The MinVDD values for block F are around 1.1V or 1.2V while those of blocks M and O are about 1.5V and 2.1V respectively. Figures 4.7 and 4.8 show the drain current distributions for the ASIC1, diode tied, NMOS and

PMOS corner transistors for four separate supply voltages. For both NMOS and PMOS transistors with supply voltages of ±0.3125V and ±0.625V, lower temperatures induce lower drive currents. The same transistors with supply voltages of ±1.25V and ±2.5V have higher drive currents when the temperature is lowered. These temperature induced shifts in the saturation currents are shown directly in Figures 4.9 and 4.10.



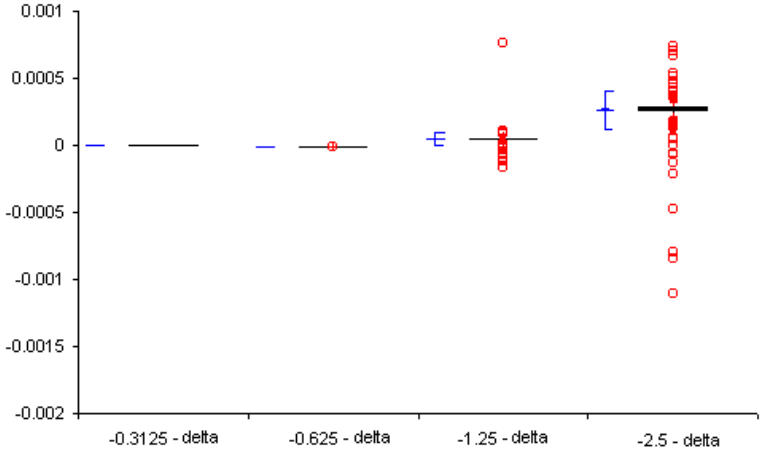
Diode Tied PMOS ID	n	Mean	SD	SE	95% CI of Mean	Median	IQR	95% CI of Median
30C _ -0.3125V	880	-2.9E-08	7.2E-09	2.4E-10	-3.0E-08 to -2.9E-08	-2.9E-08	1.1E-08	-3.0E-08 to -2.9E-08
85C _ -0.3125V	880	-2.1E-07	4.3E-08	1.4E-09	-2.2E-07 to -2.1E-07	-2.1E-07	6.9E-08	-2.2E-07 to -2.1E-07
30C _ -0.625V	880	-3.1E-05	4.3E-06	1.5E-07	-3.2E-05 to -3.1E-05	-3.2E-05	7.1E-06	-3.2E-05 to -3.1E-05
85C _ -0.625V	880	-4.6E-05	4.8E-06	1.6E-07	-4.6E-05 to -4.5E-05	-4.6E-05	7.8E-06	-4.6E-05 to -4.5E-05
30C _ -1.25V	880	-8.5E-04	3.4E-05	1.1E-06	-8.5E-04 to -8.5E-04	-8.5E-04	5.2E-05	-8.5E-04 to -8.4E-04
85C _ -1.25V	880	-8.0E-04	3.8E-05	1.3E-06	-8.0E-04 to -8.0E-04	-8.0E-04	4.5E-05	-8.1E-04 to -8.0E-04
30C _ -2.5V	880	-4.1E-03	1.3E-04	4.5E-06	-4.1E-03 to -4.1E-03	-4.1E-03	1.5E-04	-4.1E-03 to -4.1E-03
85C _ -2.5V	880	-3.9E-03	1.0E-04	3.4E-06	-3.9E-03 to -3.9E-03	-3.9E-03	1.4E-04	-3.9E-03 to -3.9E-03

Figure 4.7 – ASIC1 PMOS Corner Transistor Saturation Current Distributions



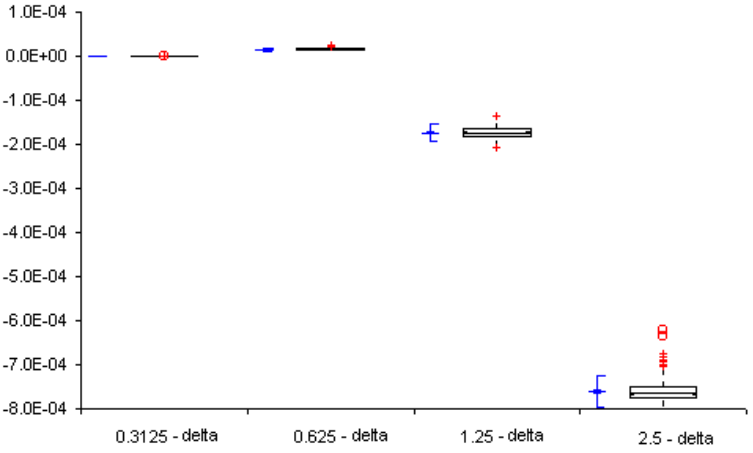
Diode Tied NMOS ID	n	Mean	SD	SE	95% CI of Mean	Median	IQR	95% CI of Median
30C _ 0.3125V	420	3.1E-09	9.6E-10	4.7E-11	3.0E-09 to 3.2E-09	2.8E-09	1.1E-09	2.7E-09 to 2.9E-09
85C _ 0.3125V	420	3.1E-08	8.1E-09	4.0E-10	3.0E-08 to 3.2E-08	2.9E-08	9.9E-09	2.8E-08 to 3.0E-08
30C _ 0.625V	420	1.6E-05	4.2E-06	2.0E-07	1.6E-05 to 1.7E-05	1.5E-05	5.2E-06	1.5E-05 to 1.5E-05
85C _ 0.625V	420	3.2E-05	6.2E-06	3.0E-07	3.1E-05 to 3.3E-05	3.0E-05	7.9E-06	3.0E-05 to 3.1E-05
30C _ 1.25V	420	2.2E-03	9.8E-05	4.8E-06	2.2E-03 to 2.2E-03	2.2E-03	1.3E-04	2.2E-03 to 2.2E-03
85C _ 1.25V	420	2.1E-03	8.7E-05	4.3E-06	2.0E-03 to 2.1E-03	2.0E-03	1.1E-04	2.0E-03 to 2.0E-03
30C _ 2.5V	420	9.6E-03	1.9E-04	9.1E-06	9.6E-03 to 9.7E-03	9.6E-03	2.4E-04	9.6E-03 to 9.6E-03
85C _ 2.5V	420	8.9E-03	1.7E-04	8.5E-06	8.9E-03 to 8.9E-03	8.8E-03	2.1E-04	8.8E-03 to 8.8E-03

Figure 4.8 – ASIC1 NMOS Corner Transistor Saturation Current Distributions



Diode Tied PMOS ID	n	Mean	SD	SE	95% CI of Mean	Median	IQR	95% CI of Median
-0.3125 - delta	880	-1.8E-07	3.6E-08	1.2E-09	-1.9E-07 to -1.8E-07	-1.8E-07	5.7E-08	-1.9E-07 to -1.8E-07
-0.625 - delta	880	-1.4E-05	6.3E-07	2.1E-08	-1.4E-05 to -1.4E-05	-1.4E-05	7.9E-07	-1.4E-05 to -1.4E-05
-1.25 - delta	880	4.7E-05	2.8E-05	9.4E-07	4.5E-05 to 4.9E-05	4.8E-05	8.2E-06	4.7E-05 to 4.8E-05
-2.5 - delta	880	2.6E-04	8.8E-05	3.0E-06	2.6E-04 to 2.7E-04	2.7E-04	2.6E-05	2.7E-04 to 2.7E-04

Figure 4.9 – ASIC1 PMOS Saturation Current Temperature Shift Distributions



Diode Tied NMOS ID	n	Mean	SD	SE	95% CI of Mean	Median	IQR	95% CI of Median
0.3125 - delta	420	2.8E-08	7.2E-09	3.5E-10	2.7E-08 to 2.9E-08	2.6E-08	8.9E-09	2.5E-08 to 2.7E-08
0.625 - delta	420	1.6E-05	2.0E-06	1.0E-07	1.6E-05 to 1.6E-05	1.5E-05	2.6E-06	1.5E-05 to 1.5E-05
1.25 - delta	420	-1.7E-04	1.2E-05	5.8E-07	-1.8E-04 to -1.7E-04	-1.7E-04	1.7E-05	-1.8E-04 to -1.7E-04
2.5 - delta	420	-7.6E-04	2.1E-05	1.0E-06	-7.6E-04 to -7.6E-04	-7.7E-04	2.8E-05	-7.7E-04 to -7.6E-04

Figure 4.10 - ASIC1 NMOS Saturation Current Temperature Shift Distributions

While the supply voltage transition point for the transistor saturation currents is between ±0.625V and ±1.25V, the supply voltage transition point for the MinVDD values is quite close to 1.4V as shown in Figure 4.11.

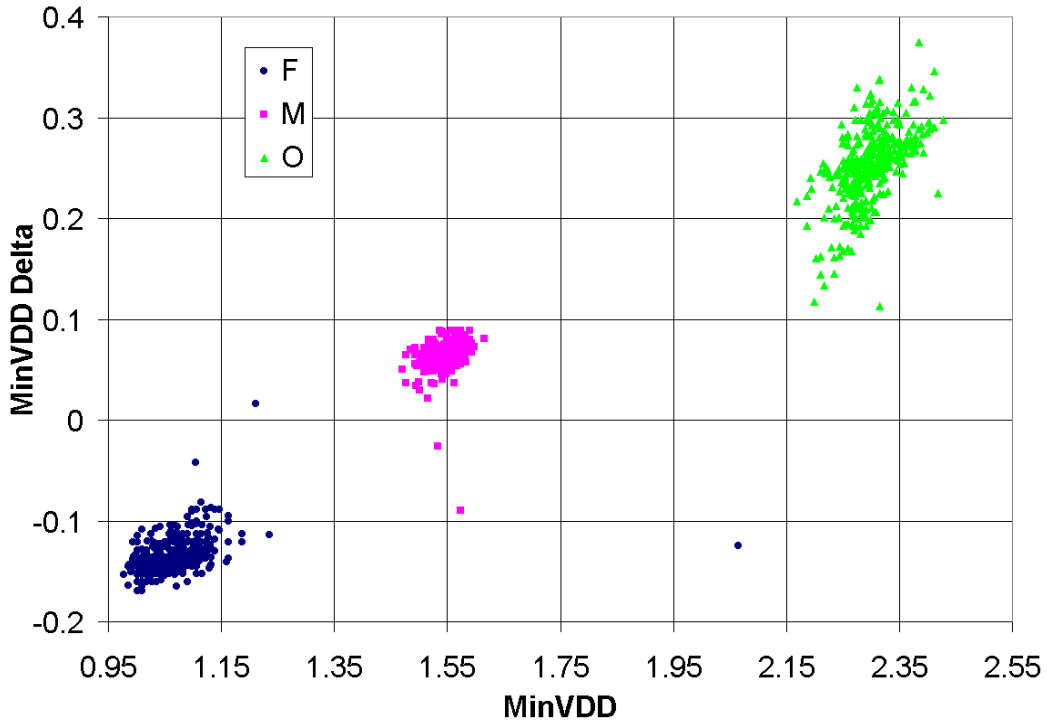


Figure 4.11 – ASIC1 MinVDD Temperature Shift As A Function Of MinVDD

This small discrepancy between the transition point of the saturation currents and the MinVDD measurements is not surprising. The behavior of temperature induced shifts in MinVDD are tied to the saturation current of the IC's constituent transistors, but, as demonstrated in Equations 2.11 and 2.12, are also tied to the threshold voltage. Equations 2.11 and 2.12 detail that the IC performance depends on three fundamental variables, VDD, V_t, and μ. In particular, Equation 2.12 shows that performance is a function of the gate overdrive, (VDD-V_t). This term appears in the denominator

indicating that when the gate overdrive decreases, the delay is pushed up and the IC performance is pushed down. The two ways to decrease the gate overdrive are to decrease the temperature and hence increase the V_t, or to decrease the VDD.

It is clear from the discussion and data above as well as that presented in section 2.1.4 that the performance of an IC depends on the IV characteristics of the IC's constituent transistors. However, the existing theoretical models for incorporating the temperature dependence of V_t, I_{sat}, and *ILK* into the temperature dependence of the IC performance have not been well developed.

A clear and fairly simple picture of MinVDD's dependence on temperature is possible if it is recognized that an IC's performance is driven by its constituent transistors' ability to provide power.

The power consumption of a CMOS circuit is described as:

$$P = C_{load} * VDD^2 * f$$

Equation 4.1 ([20], pp. 244)

For any given, fixed frequency there exists a minimum power, P_{min}, that a circuit requires to function correctly, which, in turn, sets the MinVDD. This relationship is reflected in Equation 4.2.

$$P_{\min} = C_{load} * MinVDD^2 * f$$

Equation 4.2

The maximum frequency, f_{max}, test gives a complimentary approach to the power to performance relationship. For a given value of VDD there exists a maximum power, P_{max}, that a circuit is able to support. In turn, this P_{max} sets the f_{max} at which the IC is able to perform. This relationship is reflected in Equation 4.3.

$$P_{\max} = C_{load} * VDD^2 * f_{\max}$$

Equation 4.3

A MinVDD test runs a functional test at a fixed frequency and incrementally lowers VDD. Lower values for VDD act to lower the value of P_{max} and, in turn, f_{max}. The VDD at which the circuit's f_{max} equals the operating frequency is determined as the MinVDD of the circuit.

The value of P_{max} is set by the power provided by the IC's constituent transistors which is determined by the transistors' IV curves as shown in Figures 2.1 and 2.2. These IV curves are described by the long channel model for drain current:

$$I_D = \frac{W}{L}\mu C_{ox}\frac{(V_G - V_t)^2}{2}$$

Equation 4.4 ([19], pp. 130)

When the channel length of the transistor enters the deep submicron region the drain current begins to be controlled by the velocity saturation effect. As channel length is reduced, the exponent of the gate overdrive term, $(V_G\text{-}V_t)^2$, is reduced from 2 to 1 such that the drain current becomes linearly dependent on the gate overdrive ([19], pp. 180). For very small gate lengths the drain current becomes:

$$I_D = W\mu C_{ox}\varepsilon_C(V_G - V_t)$$

Equation 4.5 ([19], pp. 180)

where ε_C is a critical value for the electric field parallel to the channel length ([19], pp. 176).

The saturation currents for the ASIC1 corner transistors were measured with V_D=V_G at 0.3125V, 0.625V, 1.25V, and 2.5V. The results were consolidated into mean values which are presented in Figure 4.12. The NMOS and PMOS transistors

show very nearly linear IV curves. The bias conditions and measured drain currents for the PMOS transistors are negative but are reported here as absolute values for easy comparison to the NMOS transistors.

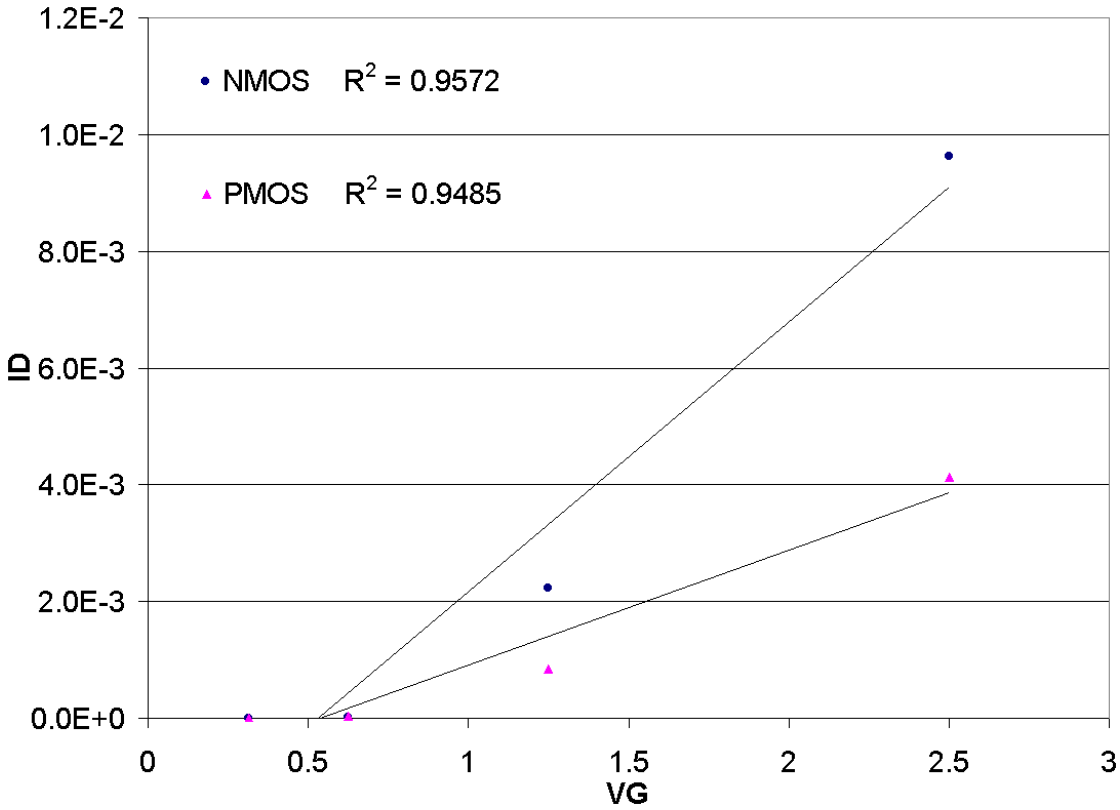


Figure 4.12 – Linearity Of ASIC1 Deep Submicron Transistor IV Curves

If a chain of inverters is used to simulate the transitions in an IC the relationships between drain current, drain voltage, and instantaneous power may be observed. The performance limiting node in an IC will have a comparatively large capacitance for the widths of the transistors that compose the inverter. In such a circumstance the input voltage transition will approach a step input. As gate lengths and gate overdrives shrink the transistors in the inverter will spend increasing

proportions of their transition time in saturation. For an output node of an inverter with such a step input the drain current of the charging or discharging transistor will remain constant for the majority of the transition. The constant drain current results in a drain voltage as a linear function of time with a slope, *s*. The instantaneous power is then given as:

$$P(t) = I_{sat} V_{drain}(t) = I_{sat} st$$

Equation 4.6

The average power for a transition occuring between times t_1 and t_2 is then:

$$P_{avg} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} P(t)dt = I_{sat} \frac{s}{t_2 - t_1} \int_{t_1}^{t_2} tdt$$

Equation 4.7

and:

$$P_{avg} = I_{sat} \frac{s}{2} \frac{t_2^2 - t_1^2}{t_2 - t_1} = I_{sat} \frac{s}{2} (t_2 + t_1)$$

Equation 4.8

when the average power is evaluated for the time period in question:

$$st_1 = V_t$$
$$st_2 = V_{DD}$$

Equation 4.9

Remembering that the saturation current for deep submicron devices is described as:

$$I_{sat} = g_m (V_{DD} - V_t)$$

Equation 4.10

substitution yields the average power:

$$P_{avg} = I_{sat} \frac{(V_{DD} + V_t)}{2} = g_m \frac{V_{DD}^2 - V_t^2}{2}$$

Equation 4.11

Remembering that both V_t and g_m increase as temperature is reduced, the average power of a weak node may be modeled as a function of VDD for two temperatures as in Figure 4.13.

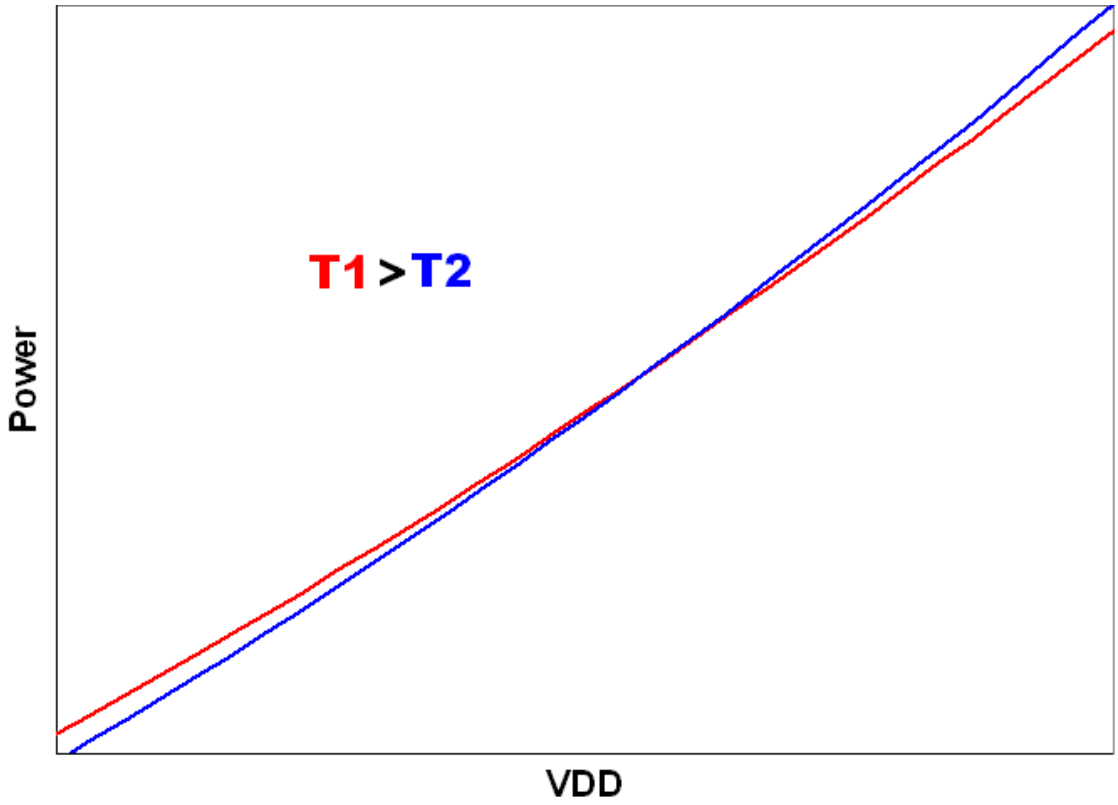


Figure 4.13 – Average Power As A Function of VDD At Two Temperatures

The values of P_{avg} for the same node at two different temperatures may be found by evaluating Equation 4.11 using appropriate values for both V_t and g_m. The dimensions of the transistors' IV curves that define P_{avg} change as the temperature is reduced from T1 to T2. There are three significant conditions for this temperature induced change in IV characteristics, $P_{tr}1 < P_{tr}2$, $P_{tr}1 = P_{tr}2$, and $P_{tr}1 > P_{tr}2$. From the observation

that $P_{avg} \propto P_{\max}$ and the discussion of Equation 4.3, these three conditions correspond to lower temperatures inducing a higher f_{max}, lower temperatures inducing no change in f_{max}, and lower temperatures inducing a lower f_{max}, respectively. As the temperature induced change in P_{avg} is controlled by VDD, the f_{max} may also be controlled by shifting VDD up or down.

The critical point in the power characterization is the VDD that gives no temperature induced change in power, $P_{tr}1 = P_{tr}2$, and will hereafter be referred to as VDD_{TI} or the temperature independent VDD. This VDD_{TI} sets the critical value for the power, P_{TI}, at which the performance of the IC is insensitive to the change in temperature. Setting the values for P_{avg} at two temperatures equal:

$$g_{m1}\frac{(VDD^2 - V_{t1}^2)}{2} = g_{m2}\frac{(VDD^2 - V_{t2}^2)}{2}$$

Equation 4.12

It follows that:

$$\frac{(VDD^2 - V_{t1}^2)}{(VDD^2 - V_{t2}^2)} = \frac{g_{m2}}{g_{m1}}$$

Equation 4.13

Solving for VDD_{TI}:

$$VDD_{TI} = \sqrt{\frac{V_{t1} - V_{t2}(g_{m2}/g_{m1})}{(1-(g_{m2}/g_{m1}))}}$$

Equation 4.14

To find the values of V_t and g_m for Equation 4.14 consider that for any given transition in a CMOS circuit there are roughly equal numbers of NMOS transistors pulling nodes to zero as there are PMOS transistors pulling nodes to VDD. The circuit

performance is limited by the weaker of the two transistor types, in this case the PMOS transistors. So, the VDD_{TI} of ASIC1 is calculated with the means of the absolute values of the PMOS corner transistor characteristics as:

$$VDD_{TI} = \sqrt{\frac{0.390 - 0.441(1.99E-3/1.85E-3)}{(1-(1.99E-3/1.85E-3))}} = 1.06V$$

Equation 4.15

A linear regression applied to the data in Figure 4.11 gives the line:

$$MinVDD\_Delta = 0.429MinVDD - 0.6156$$

Equation 4.16

Setting MinVDD_Delta to zero in Equation 4.16 gives a VDD_{TI} of 1.43. The VDD_{TI} for ASIC1 predicted by Equation 4.15 shows good agreement with the empirically determined VDD_{TI}.

It is clear from Equation 4.1 that both VDD and frequency contribute to the power consumption of a circuit. As such, there exists a region of space within the frequency-VDD plane in which an IC will function. This fact is widely recognized and demonstrated by creating Shmoo plots. A Shmoo plot is shown in Figure 4.14 where the means of the MinVDD data taken for the L2 LogicBIST pattern for ASIC2 is plotted. The green area (higher VDDs and lower frequencies) represents the region in frequency-VDD space where the IC should function correctly while the red area (lower VDDs and higher frequencies) represents the failing region.

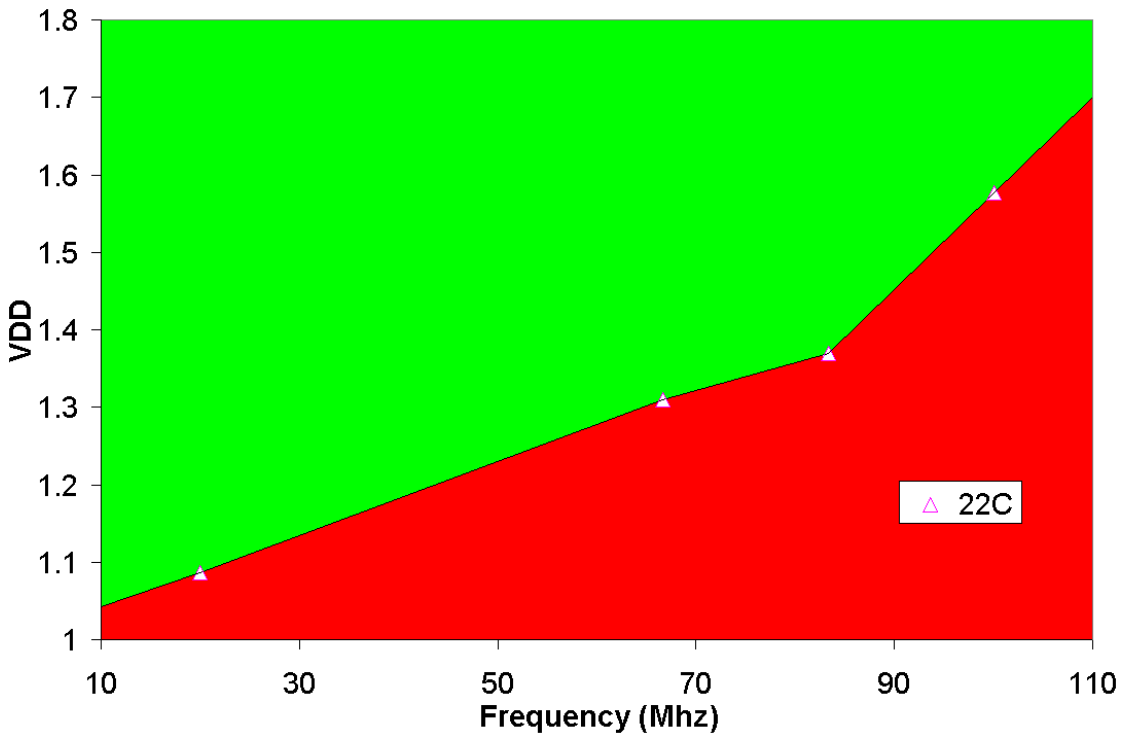


Figure 4.14 – An Improvised Shmoo Plot From ASIC2 LBIST MinVDD Data

It has been established by theory and data from ASIC1 that a technology has a critical value for the power, P_{TI}, at which the performance is insensitive to changes in temperature. Below P_{TI} lower temperatures induce lower performance while above P_{TI} lower temperatures induce higher performance. These same effects are demonstrated in Figure 4.15 which displays the same data as in Figure 4.14 for three separate temperatures. Temperature, in addition to VDD and frequency, is a third condition that controls the function of an IC. Temperature, VDD, and frequency are the three variables in an IC's operating environment and a Shmoo plot may be created to completely define the performance space of an IC.

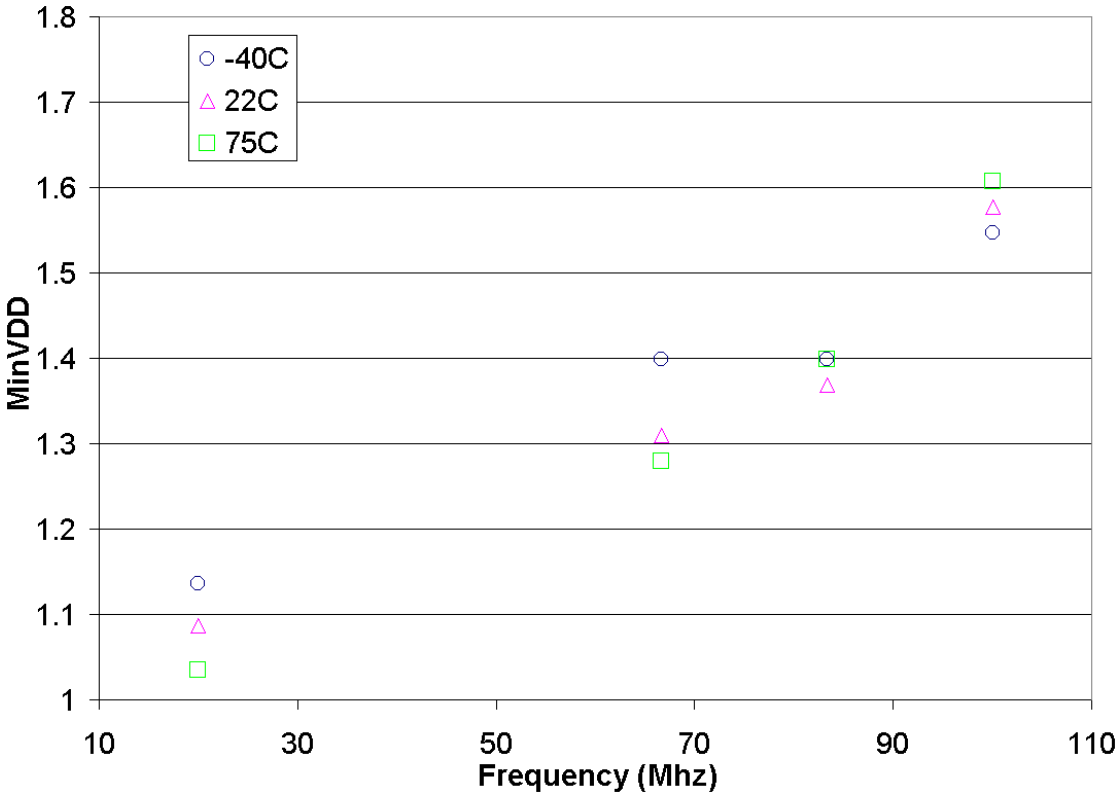


Figure 4.15 – ASIC2 LBIST MinVDD(f_{max}) As Temperature Varies

The relationships in Figure 4.15 may be predicted by stating:

$$\eta P_{avg} = P_{\max} = C_{load} VDD^2 f_{\max}$$

Equation 4.17

where η is a constant determined in part by the gate count of the IC. $P_{avg}(VDD)$ is found as in Equations 4.6 through 4.11.

Substituting Equation 4.11 into Equation 4.17 and solving for f_{max} gives:

$$f_{\max} = \frac{\eta g_m}{C_{load}} \left(1 - \left(\frac{V_t}{VDD}\right)^2\right)$$

Equation 4.18

Though no data is available for ASIC2 to accurately determine the values of g_m, V_t, η, and C_{load}, realistic values would give curves like those in Figure 4.16.

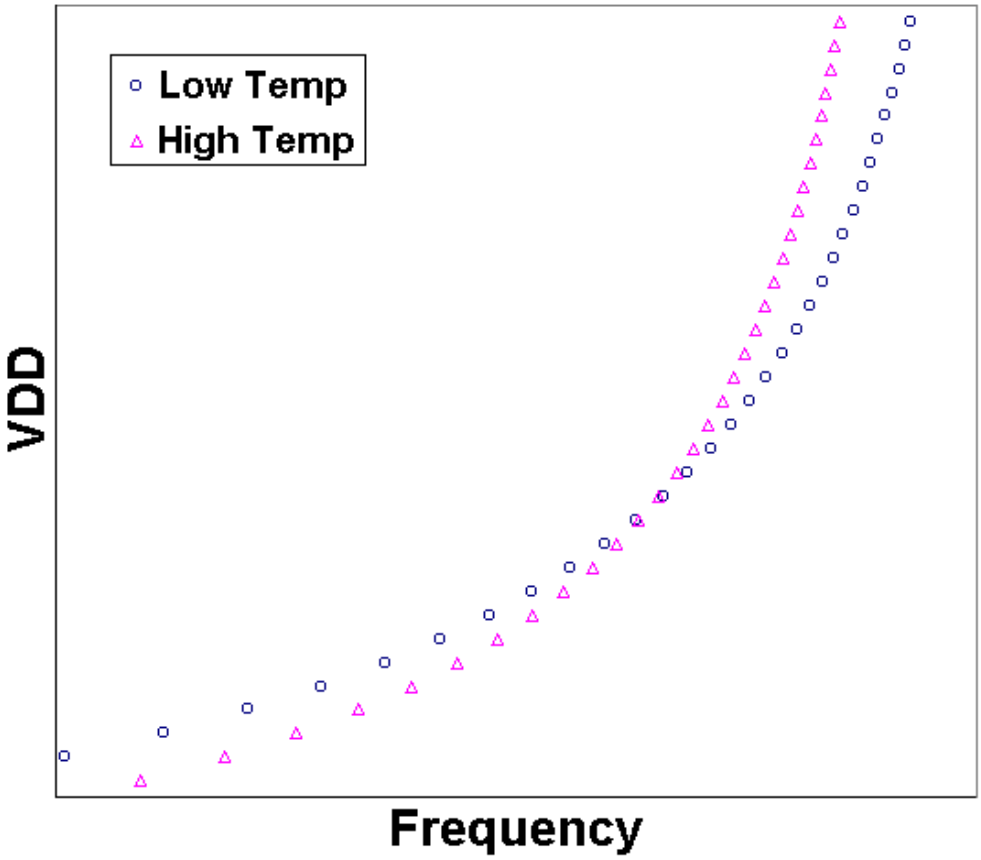


Figure 4.16 – Theoretical MinVDD(f_{max}) As Temperature Varies

Figures 4.16 and 4.15 share the same basic characteristics. Higher values for VDD induce larger available power and higher projected f_{max} values. Also, at higher VDD's lower temperatures induce performance improvements while at lower VDD's lower temperatures induce lower performance.

Since three operating conditions, VDD, frequency, and temperature, define the space in which an IC will function correctly, an IC's operating capabilities may be

viewed from more than one perspective. Figure 4.17 shows MinVDD as a function of temperature and frequency as found by rearranging Equation 4.18. Each red line in Figure 4.17 represents the MinVDD as a function of temperature for a fixed frequency. It should be noticed that if the two temperatures chosen for testing are 20°C and 80°C the VDD_{TI} will occur at just under 0.9V (as indicated by the arrow.)

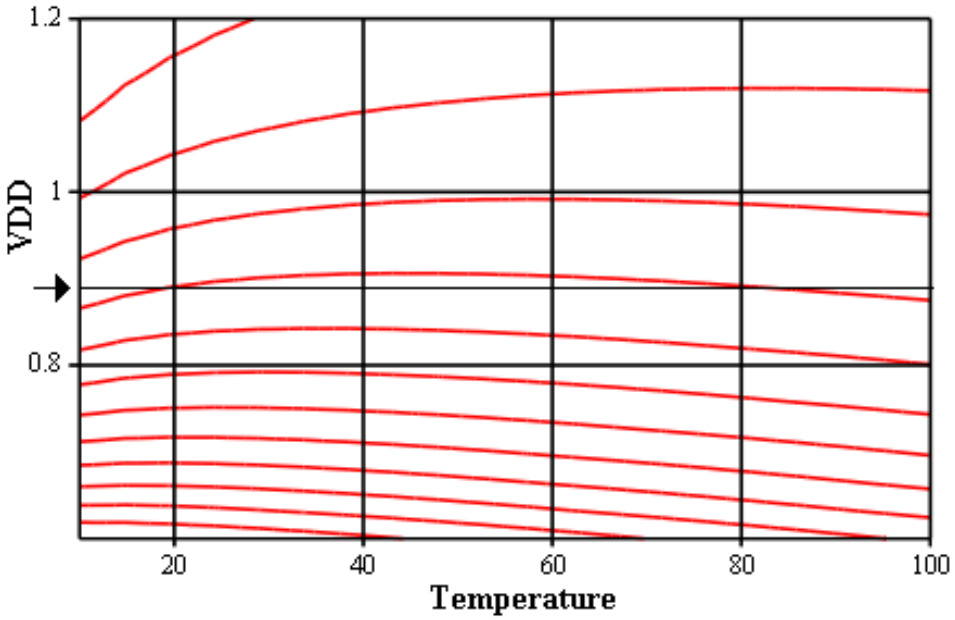


Figure 4.17 – Family Of MinVDD(T) Curves For 12 Fixed Frequencies

The ASIC2 data in Figure 4.15 is replotted in Figure 4.18 to create a family of Shmoo plots in VDD and temperature space.

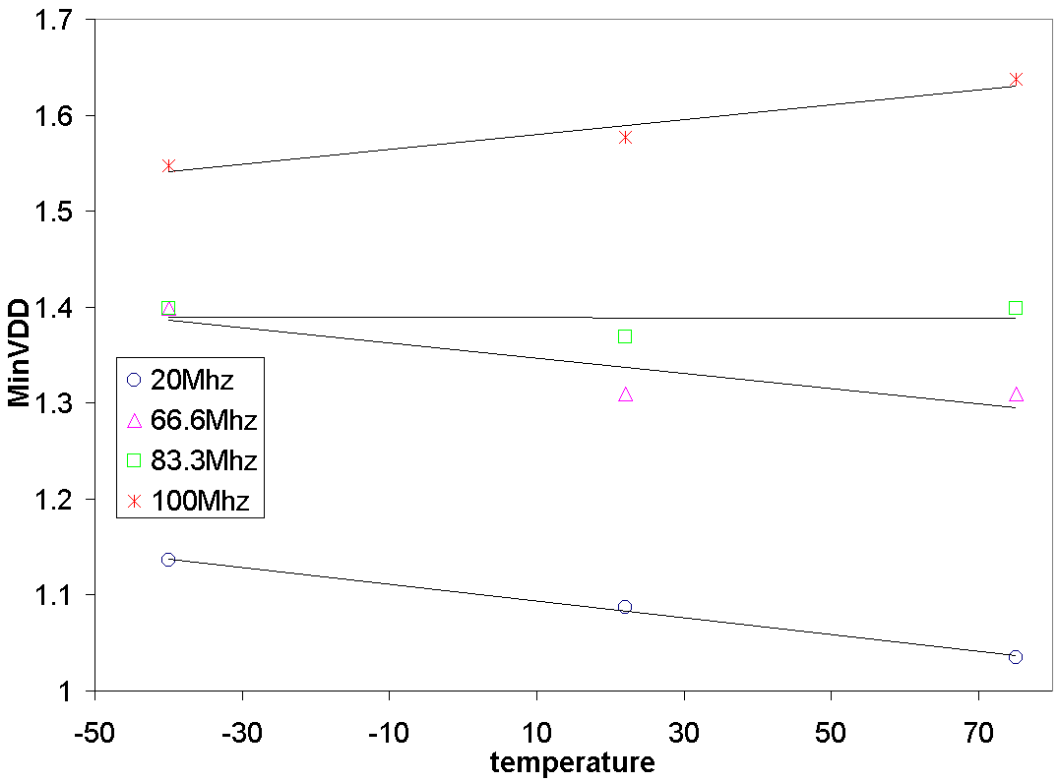


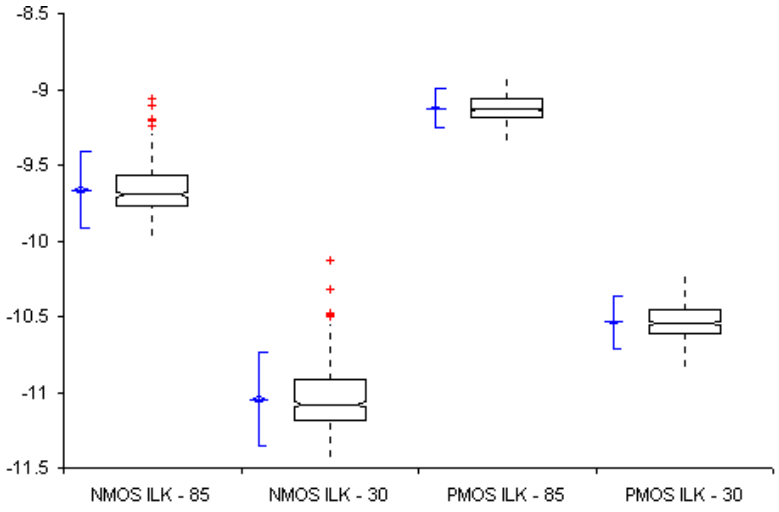
Figure 4.18 – ASIC2 LBIST MinVDD(Temperature) As Frequency Varies

The analysis of the performance to power relationship given above rests on the assumption that the coefficients of the transistor's linear IV curves are independent of VDD. This assumption may be acceptable for small changes in VDD and for some technologies, but it is certainly not always the case. When VDD is changed both the gate bias and the drain bias of the transistor change. Changing the drain bias will change the V_t through the effects of DIBL. Though, this drain bias induced change in V_t will be smaller for lower temperatures due to DIBL's temperature dependence as outlined in section 2.4.7. If the transistor is not operating in the velocity saturation region, a change in VDD will also modify I_{sat} because of its dependence on the electric field parallel to the channel.

4.3 Defect Detection With Temperature Testing

4.3.1 IDDQ and Static IDD

Static IDD (SIDD) measurements taken on ASIC1 are effectively IDDQ measurements with a particular type of test pattern, the "Y" pattern. The quiescent current consumption of a healthy IC is expected to drop as the temperature is lowered in the same way that transistor leakage drops. The ASIC1 corner transistors show a drop in leakage current of just over an order of magnitude as shown in Figure 4.19.

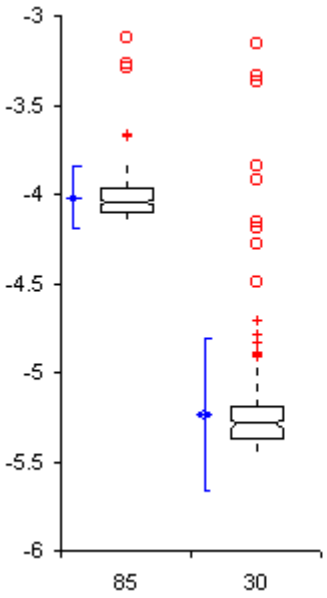


ILK (log(A))	n	Mean	SD	SE	95% CI of Mean
NMOS - 85	420	-9.664	0.1530	0.0075	-9.679 to -9.650
NMOS - 30	420	-11.045	0.1900	0.0093	-11.064 to -11.027
PMOS - 85	880	-9.126	0.0779	0.0026	-9.131 to -9.121
PMOS - 30	880	-10.536	0.1049	0.0035	-10.543 to -10.529

ILK (log(A))	n	Median	IQR	95% CI of Median
NMOS - 85	420	-9.693	0.209	-9.709 to -9.676
NMOS - 30	420	-11.084	0.269	-11.100 to -11.060
PMOS - 85	880	-9.131	0.125	-9.141 to -9.124
PMOS - 30	880	-10.541	0.163	-10.555 to -10.532

Figure 4.19 – ASIC1 Transistor Leakage Currents

The expectation that SIDD should show similar behavior to that of the corner transistors is affirmed by the SIDD measurements on the core of ASIC1 as shown in Figure 4.20.



SIDD_CORE (log(A))

	n	Mean	SD	SE	95% CI of Mean
85	380	-4.019	0.1072	0.0055	-4.030 to -4.008
30	380	-5.237	0.2583	0.0133	-5.263 to -5.211

	Median	IQR	95% CI of Median
85	-4.042	0.128	-4.060 to -4.029
30	-5.286	0.182	-5.310 to -5.269

Figure 4.20 – ASIC1 Static IDD Distributions

It is clear from the data in Figure 4.20 that low temperature SIDD testing will identify more outliers than high temperature testing; outliers are those ICs with SIDD >1.5 IQRs from the upper or lower quartile. The magnitude of the IDDQ or SIDD will be a function of both the transistor leakage currents and the leakage induced by defects. The fact that the leakage due to defects and that due to transistor characteristics respond differently to temperature may be taken advantage of by testing

parts at multiple temperatures. Figure 4.21 shows leakage measurements for a group of ICs that pass all tests at 30°C.

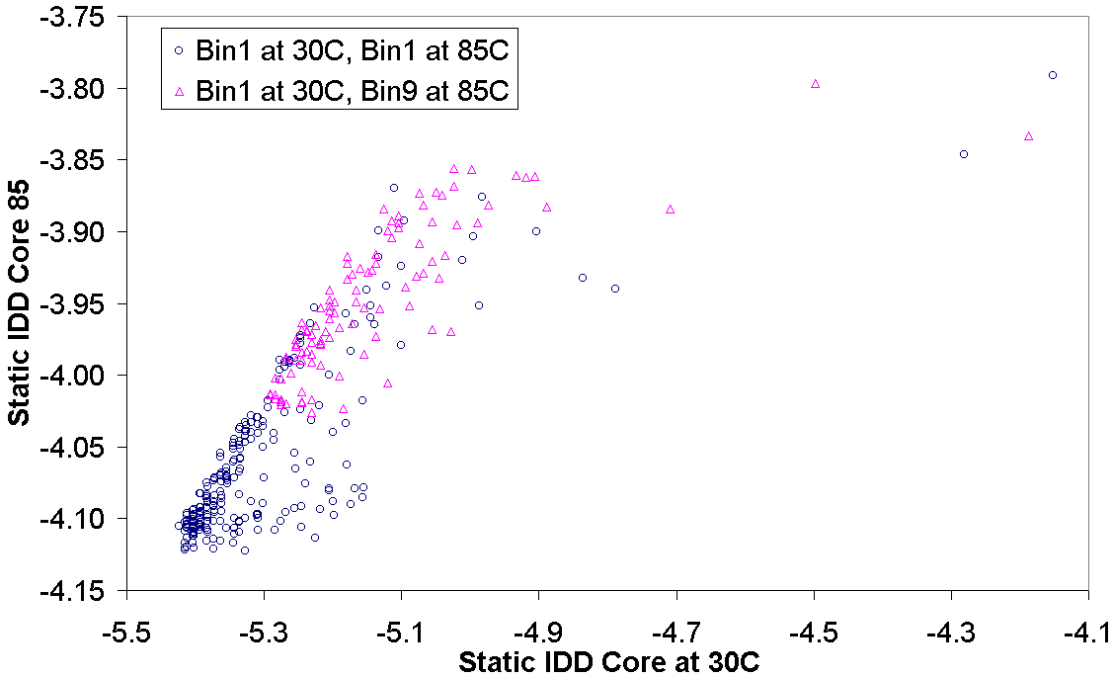


Figure 4.21 – ASIC1 Static IDD At 85°C As A Function Of That At 30°C

There are many ICs that fail *vih/vil* testing at 85°C that are clearly identifiable when the temperature response of their SIDD measurements are viewed as in Figure 21. There are also many ICs with SIDD values that are not lowered by the temperature reduction as much as would be expected based on the behavior of the bulk of the population. This could be an indication that there are temperature insensitive defects on the IC that are contributing to the leakage current. Many such ICs appear as passing all tests at both 30°C and 85°C and may be test escapes. More expansive testing and physical failure analysis could help validate the relationship between the

SIDD behavior and the health of the ICs in Figure 4.21. However, it should be clear that SIDD measurements at multiple temperatures give profound insight into the health of an IC as compared to that gained from a single temperature measurement.

4.3.2 MinVDD and f_{max}

There are several different possibilities for the use of MinVDD or f_{max} testing with temperature variations to detect defects. In any case, the VDD_{TI} for the technology should be determined. For any product, the functional, SCAN, TDF, or BIST tests should be evaluated to characterize the frequency at which the test gives MinVDD values at or very near VDD_{TI}. The ICs may then be tested in three separate regions, where lower temperatures degrade performance, improve performance, and have little affect on performance. Detection of different types of temperature sensitive and temperature insensitive defects is possible depending on the VDD region in which the testing is conducted.

In the VDD region where lower temperatures lower performance, defects that are relieved by lower temperatures, such as metal slivers, will stand out against the defect free distribution. As demonstrated in Figure 4.22, block F of ASIC1 operates in this region. Single temperature testing may be useful but the additional information provided by testing at multiple temperatures will increase the resolution of the testing process. There are two outliers that are not detectable at either 30°C or 85°C alone but can be identified by testing at both temperatures. In one case the defect dominates the MinVDD behavior of the part.

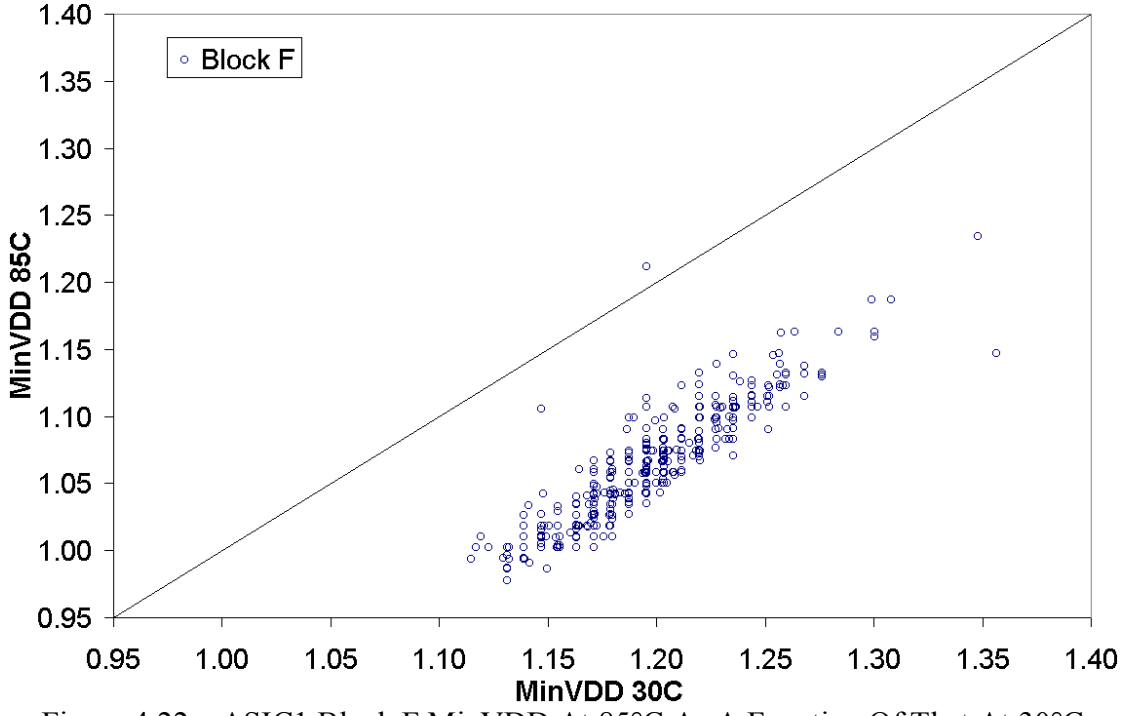


Figure 4.22 – ASIC1 Block F MinVDD At 85°C As A Function Of That At 30°C

In the VDD region where lower temperatures improve performance, defects that are aggravated by lower temperatures, such as via voids and silicide breaks, will stand out against the defect free distribution. Block M of ASIC1 operates in this region where lower temperatures induce higher performance. Figure 4.23 shows that there are two outliers that are not detectable at 85°C and questionably detectable at 30°C but can easily be identified by testing at both temperatures. In both cases the defect dominates the MinVDD behavior of the part.

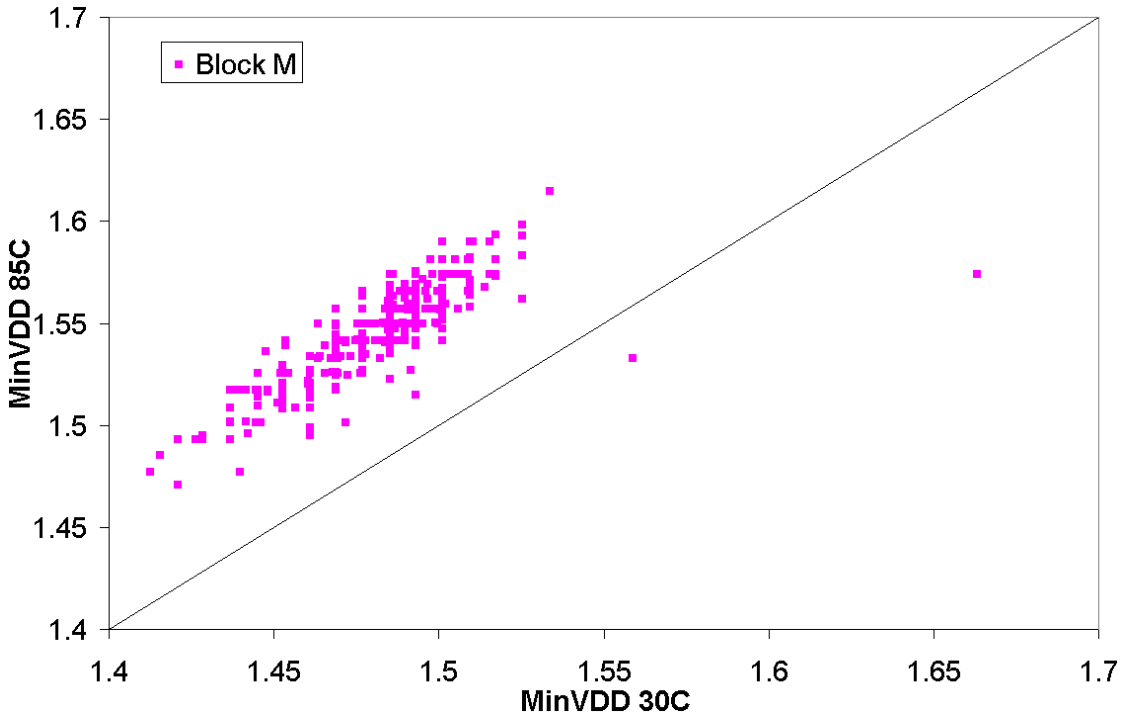


Figure 4.23 - ASIC1 Block M MinVDD At 85°C As A Function Of That At 30°C

Testing in the region where there is little or no change in performance with changes in temperature may be useful in detecting all sorts of temperature sensitive

defects. When testing in this region any large changes in performance can be assumed to be caused by a temperature sensitive defect.

However, any wafer, lot, or product will include individual ICs with varying transistor characteristics which will give variation in the performance characteristics of the tested population. In such a circumstance there will be ICs with both improved and degraded performance making it more difficult to distinguish between defect induced behavior and healthy behavior. Two things may be done to relieve the effects of process variations on the resolution of the test. The use of low temperatures is preferable when testing at either a single temperature or at multiple temperatures. Figure 4.18 shows how lower temperatures have the advantage of creating a situation where parts with different performance characteristics will be squeezed into a smaller MinVDD distribution reducing the obscuring effect of various power ratings within the same wafer, lot, etc.. Secondly, multiple temperature testing is optimized when the difference in temperatures is larger. Larger temperature shifts will exaggerate the differences between defective and healthy behavior.

Temperature plays a critical role in determining the behavior of an IC and should be carefully considered when designing any test scheme. Testing at lower temperatures can increase the ability of currently popular test methods to resolve between defective and healthy ICs. The first way in which this occurs is by lowering leakage due to healthy or normal leakage mechanisms. The second advantage to low temperature testing is the increase in resolution between healthy and defective ICs in MinVDD and f_{max} measurements caused by the shrinking of the distributions for

healthy ICs. Temperature may also be used to create new tests for detecting defects. The comparison of parametric measurements at multiple temperatures can make defective ICs clearly distinguishable from the healthy population where the same defective ICs would be indistinguishable when measured at a single temperature. Furthermore, specific defects may be targeted by careful selection of VDD values for f_{max} or MinVDD measurements. Testing below the VDD_{TI} for a technology will be effective at detecting defects that are relieved by lower temperatures, such as metal slivers, as well as temperature insensitive defects. Testing above the VDD_{TI} for a technology will be effective at detecting defects that are aggravated by lower temperatures, such as via voids and silicide breaks, as well as temperature insensitive defects. Testing at or close too the VDD_{TI} will detect all types of temperature sensitive defects.

Chapter 5

Conclusions and Future Work

The notion that temperature has a profound impact on the operation and performance characteristics of an IC has been explored and verified.

Since the intrinsic performance of an IC relies on its constituent transistors the performance of those transistors is critical in evaluating the health of an IC. To this end three basic characteristics of the performance of transistors fabricated in a 0.25μm technology have been evaluated. The rate of change of threshold voltage with respect to temperature is constant, and for a 55°C change in temperature the change in threshold voltage will be about 10% and 12% of the room temperature NMOS and PMOS values respectively. Transistor leakage is reduced by more than an order of magnitude when the temperature is reduced by 55°C. The same temperature reduction will also change the saturation current of the transistor, but the magnitude and polarity of the change in saturation current is a function of the gate and drain voltages.

The transistor leakage reduction corresponds to a reduction in the intrinsic leakage of the ASIC1 ICs as measured by Static IDD. This temperature induced reduction in leakage also results in higher contrast between intrinsically healthy ICs and defective ICs. Even higher contrast may be obtained by considering the Static IDD or IDDQ at more than one temperature.

The performance of an IC as measured by its maximum operating frequency and its minimum operating voltage depend on the IC's constituent transistors' threshold voltages and saturation currents. It has been demonstrated that an IC's operating temperature has a profound impact on transistor characteristics, which, in turn, affect the IC's performance. There are three critical states for the interaction of temperature and performance. These states, for any given technology, are defined by a critical VDD (VDD_{TI}) at which IC performance is insensitive to changes in temperature. Above this VDD_{TI} performance improves as the temperature is lowered. Below the VDD_{TI} performance is degraded as the temperature is lowered. A model explaining the temperature induced performance shifts in terms of the IC's constituent transistors' capacity to provide power has been presented and empirically validated.

Test efficacy is significantly enhanced when an IC's entire performance space is considered and the technology's VDD_{TI} is used to design MinVDD and f_{max} tests. This performance space is defined by VDD, frequency, and temperature and can be used as a powerful tool in differentiating between healthy and defective or low reliability ICs. F_{max} tests may be run at or below the technology's VDD_{TI} to screen for defects that are alleviated by lower temperatures such as metal stringers. F_{max} tests run at or above the VDD_{TI} may be useful in screening defects that are aggravated by lower temperatures such as resistive vias and silicide breaks. One particularly efficient method would be to tie sort test data (taken at some low temperature, e.g. 0°C) to final test data (taken at some high temperature, e.g. 75°C) using die tracing technology. In

this way, multiple temperature test data is made available for new screening methods without any additional test time.

Further investigation is critical to help evaluate and quantify the efficacy of MinVDD, f_{max}, and IDDQ testing at targeted or multiple temperatures. Some crucial questions that will provide fodder for further research are:

(1) What sorts of reductions in DPM numbers can these methods effect?

(2) Will these methods lower yields considerably?

(3) Can these methods eliminate the need for functional testing?

(4) Can failure and data analysis provide more evidence about the nature of the defects that these methods screen?

(5) Are these methods best implemented with stuck-at based testing or delay fault testing?

(6) Are these methods robust to the broad variations in processing conditions within wafers or lots that induce broad variation in IC performance?

(7) Can low temperature testing accommodate such broad variations in performance better than high temperature testing?

(8) Is there any change to test times due to testing at cold temperatures?

(9) Can any test time increases due to multiple temperature testing be eliminated by using die tracing technology to tie wafer sort data to final test data?

Bibliography

[1] Dieter K. Schroder, *Semiconductor Material and Device Characterization*, second edition. New York: John Wiley & Sons, 1998.

[2] L. Vadasz and A.S. Grove, "Temperature Dependence of MOS Transistor Characteristics Below Saturation," *IEEE Transactions on Electron Devices*, vol. ED13, no. 12, pp. 863-866, Dec. 1966.

[3] Raymond Wang, James Dunkley, Thomas A. DeMassa, Lawrence F. Jelsma, "Threshold Voltage Variations with Temperature in MOS Transistors," *IEEE Transactions on Electron Devices*, pp. 386-388, Jun. 1971.

[4] Fritz H. Gaensslen, V. Leo Rideout, E.J. Walker, John J. Walker, "Very Small MOSFET's for Low-Temperature Operation," *IEEE Transactions on Electron Devices*, vol. ED24, no. 3, pp. 218-229, Mar. 1977.

[5] S.M. Sze, *Physics of Semiconductor Devices*, second edition. New York: John Wiley & Sons, 1981.

[6] Yaun Taur and Tak H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge: Cambridge University Press, 1999.

[7] C.E. Blat, E.H. Nicollian, E.H. Poindexter, "Mechanism of Negative-Bias-Temperature Instability," *Journal of Applied Physics*, vol. 69, no. 3, pp. 1712-1720, Feb. 1991.

[8] A. Suzuki, K. Tabuchi, H. Kimura, T. Hasegawa, S. Kadomura, K. Kakamu, H. Kudo, M. Kawano, A. Tsukune, and M. Yamada, "A Strategy Using a Copper/Low-k BEOL Process to Prevent Negative-Bias Temperature Instability (NBTI) in p-MOSFETs with Ultra-Thin Gate Oxide," *2002 Symposium on VLSI Technology Digest of Technical Papers*, pp. 216-217, 2002.

[9] N. Kimizuka, K. Yamaguchi, K Imai, T. Iizuka, C.T. Liu, R.C. Keller, and T. Horiuchi, "NBTI Enhancement by Nitrogen Incorporation into Ultrathin Gate Oxide for 0.10-μm Gate CMOS Generation," *2000 Symposium on VLSI Technology Digest of Technical Papers*, pp. 92-93, 2000.

[10] K. Sasada, M. Arimoto, H. Nagasawa, A. Nishida, H. Aoe, T. Dan, S. Fujiwara, Y. Matsushita, and K. Yosdoshi, "The Influence of SiN Films on Negative Bias Temperature Instability and Characteristics in MOSFET's," *Proceedings of the*

IEEE 1998 International Conference on Microelectronic Test Structures, vol. 11, pp. 207-210, Mar. 1998.

[11] F. Pagaduan, J.K. Jerry Lee, V. Vedagarbha, K Lui, M. Hart, D. Gitlin, T. Takaso, S. Kamiyama, and K. Nakayama, "The Effects of Plasma-Induced Damage on Transistor Degradation and the Relationship to Field Programmable Gate Array Performance," *Proceedings of the 39th Annual IEEE Reliability Physics Symposium*, pp. 315-318, 2001.

[12] C.H. Liu, M.T. Lee, C.Y. Lin, J. Chen, K. Schruefer, J. Brighten, N. Rovedo, T.B. Hook, M.V. Khare, S.F. Huang, C. Wann, T.C. Chen, and T.H. Ning, "Mechanism and Process Dependence of Negative Bias Temperature Instability (NBTI) for pMOSFETs with Ultrathin Gate Dielectrics," *IEEE International Electron Devices Meeting Technical Digest*, pp. 39.2.1-39.2.4, 2001.

[13] Noboru Shiono and Chisato Hashimoto, "Threshold-Voltage Instability of n-Channel MOSFET's under Bias-Temperature Aging," *IEEE Transactions on Electron Devices*, vol. ED-29, no. 3, pp. 361-368, Mar. 1982.

[14] M. Noyori, J. Yasui, T. Ishihara, H. Higuchi, "Characteristics & Analysis of Instability Induced by Secondary Slow Trapping in Scaled CMOS Devices," *IEEE Transactions on Reliability*, vol. R-32, no. 3, pp. 323-329, Aug. 1983.

[15] A. Keshavarzi, K. Roy, M. Sachdev, C. Hawkins, K. Soumyanath, V. De, "Multiple-Parameter CMOS IC Testing with Increased Sensitivity for I_{DDQ}," *IEEE International Test Conference*, pp. 1051-1059, 2000.

[16] J. Segura, A. Keshavarzi, J. Soden, C. Hawkins, "Parametric Failures in CMOS ICs – A Defect-Based Analysis," *IEEE International Test Conference*, pp. 90-99, 2002.

[17] A. Bellaouar, A. Fridi, M.I. Elmasry, and K. Itoh, "Supply Voltage Scaling for Temperature Insensitive CMOS Circuit Operation," *IEEE Transactions on Circuits and Systems – II: Analog and Digital Signal Processing*, vol. 45, no. 3, pp. 415-417, Mar. 1998.

[18] K. Kanda, K. Nose, H. Kawaguchi, and T. Sakurai, "Design Impact of Positive Temperature Dependence on Drain Current in Sub-1V CMOS VLSI's," *IEEE 1999 Custom Integrated Circuits Conference*, pp. 563-566, 1999.

[19] Yannis P. Tsividis, *Operation and Modeling of the MOS Transistor*. New York: McGraw-Hill Book Company, 1987.

[20] Sung-Mo Kang and Yusuf Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*, second edition. Boston: McGraw-Hill, 1999.

[21] Michael L. Bushnell and Vishwani D. Agrawal, *Essentials of Electronic Testing for Digital Memory & Mixed-Signal VLSI Circuits*. Boston: Kluwer Academic Publishers, 2000.

[22] C. Hawkins, J. Soden, A. Righter, and J. Ferguson, "Defect Classes – An Overdue Paradigm for CMOS IC Testing," *IEEE International Test Conference*, pp. 413-425, Oct. 1994.

[23] P. Nigh, W. Needham, K. Butler, P. Maxwell, R. Aitken, W. Maly. "So What is an Optimal Test Mix? A Discussion of the SEMATECH Methods Experiment," *IEEE International Test Conference*, pp. 1037-1038, 1997.

[24] Jay M. Pasachoff, *Astronomy: From the Earth to the Universe*, fourth edition. New York: Saunders College Publishing, 1995.

[25] W. NeedHam, C. Prunty, and E.H. Yeoh, "High Volume Microprocessor Test Escapes, An Analysis of Defects Our Tests Are Missing," *IEEE International Test Conference*, pp. 25-34, 1998.

[26] P. Nigh, A. Gattiker, "Test Method Evaluation Experiments & Data," *IEEE International Test Conference*, pp. 454-463, 2000.

[27] International Technology Roadmap for Semiconductors 2002 Update, Semiconductor Industry Association, http://public.itrs.net

[28] B. R. Benware, R. Madge, C. Lu, and R. Daasch, "Effectiveness Comparisons of Outlier Screening Methods for Frequency Dependent Defects on Complex ASICs," *Proceedings of the 21st IEEE VLSI Test Symposium*, 2003.

[29] P. Nigh, D. Vallett, A. Patel, J. Wright, F. Motika, D. Forlenza, R. Kurtulik, and W. Chong, "Failure Analysis of Timing and IDDq-only Failures from the SEMATECH Test Methods Experiment," *IEEE International Test Conference*, pp. 43-52, 1998.

[30] A. Keshavarzi, K. Roy, C. Hawkins, "Intrinsic Leakage in Low Power Deep Submicron CMOS ICs," *IEEE International Test Conference*, pp. 146-155, 1997.

[31] Chao-Wen Tseng, Ray Chen, Phil Nigh, and Edward J. McCluskey, "MINVDD Testing for Weak CMOS ICs," *Proceedings of the 19th IEEE VLSI Test Symposium*, pp. 339-344, 2001.

[32] J. Chang and E.J. McCluskey, "Quantitative Analysis of Very-Low-Voltage Testing," *Proceedings of the 14th IEEE VLSI Test Symposium*, pp. 332-337, 1996.

[33] Hong Hao and E.J. McCluskey, "Very-Low-Voltage Testing for Weak CMOS Logic ICs," *IEEE International Test Conference*, pp. 275-284, 1993.

[34] R. Madge, B.H. Goh, V. Rajagopalan, C. Macchietto, R. Daasch, C. Schuermyer, C. Taylor, and D. Turner, "Screening MinVDD Outliers Using Feed-Forward Voltage Testing," *IEEE International Test Conference*, pp. 673-682, 2002.

[35] J. Chang and E.J. McCluskey, "Detecting Delay Flaws by Very-Low-Voltage Testing," *IEEE International Test Conference*, pp. 367-376, 1996.

[36] J.F. Ziegler, et al. "IBM Experiments in Soft Fails in Computer Electronics (1978-1994)," *IBM Journal of Research and Development*, vol. 40, no. 1, 1998.

[37] Zaid Al-Ars, Ad J. van de Goor, Jens Braun, and Detlev Richter, "Simulation Based Analysis of Temperature Effect on the Faulty Behavior of Embedded DRAMs," *IEEE International Test Conference*, pp. 783-792, 2001.

[38] K. Baker, G. Gronthoud, M. Lousberg, I. Schanstra, and C. Hawkins, "Defect-Based Delay Testing of Resistive Vias-Contacts: A Critical Evaluation," *IEEE International Test Conference*, pp.467-476, 1999.

[39] E.A. Brandes and G.B. Brook (editors), *Smithells Metals Reference Book*, seventh edition. Oxford: Butterworth-Heinemann Ltd, 1992.

[40] Chao-Wen Tseng, Edward J. McCluskey, Xiaoping Shao, and David M. Wu, "Cold Delay Defect Screening," *Proceedings of the 18th IEEE VLSI Test Symposium*, pp. 183-188, 2000.

[41] Sandip Kundu, "IDDQ Defect Detection in Deep Submicron CMOS ICs," *Seventh Asian Test Symposium*, pp. 150-152, 1998.

[42] C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quaranta, "A Review of Some Charge Transport Properties of Silicon," *Solid-State Electronics*, vol. 20, pp.77-89, 1977.

[43] Ronald R. Troutman, "VLSI Limitations from Drain-Induced Barrier Lowering," *IEEE Journal of Solid-State Circuits*, vol. SC-14, no. 2, pp. 383-391, Apr. 1979.

[44] H. El Ghitani, "DIBL Coefficient in Short-Channel NMOS Transistors," *Proceedings of the 16th National Radio Science Conference*, 23-25 February, 1999.

[45] G. Pananakakis, G. Ghibaudo, R. Kies, and C. Papadas, "Temperature Dependence of the Fowler–Nordheim Current in Metal-Oxide-Degenerate Semiconductor Structures," *Journal of Applied Physics*, vol.78, no.4, pp.2635-2641, 1995.

[46] A. Yassine and R. Hijab, "Temperature Dependence of Gate Current in Ultra Thin SiO_2 in Direct-Tunneling Regime," *IEEE International Integrated Reliability Workshop Final Report*, pp. 56-61, Oct. 1997.

[47] P. Aminzadeh, M. Alavi, and D. Scharfetter, "Temperature Dependence of Substrate Current and Hot Carrier-Induced Degradation at Low Drain Bias," *Symposium on VLSI Technology Digest of Technical Papers*, pp. 178-179, 1998.

[48] M. Song, K.P. MacWilliams, and J.C.S. Woo, "Comparison of NMOS and PMOS Hot Carrier Effects from 300 to 77K," *IEEE Transactions on Electron Devices*, vol. 44, no. 2, pp. 268-276, Feb. 1997.

[49] F. Balestra, T. Matsumoto, M. Tsuno, H. Nakabayashi, and M. Koyanagi, "New Experimental Findings on Hot Carrier Effects in Sub-0.1μm MOSFET's," *IEEE Electron Device Letters*, vol. 16, no. 10, pp. 433-435, Oct. 1995.

[50] E. Li, E. Rosenbaum, L.F. Register, J. Tao, and P. Fang, "Hot Carrier Induced Degradation in Deep Submicron MOSFETs at 100°C," *IEEE 38th International Reliability Physics Symposium*, pp. 103-107, 2000.

[51] N. Sano, M. Tomizawa, and A. Yoshii, "Temperature Dependence of Hot Carrier Effects in Short-Channel Si-MOSFETs," *IEEE Transactions on Electron Devices*, pp. 2211-2216, vol. 42, no. 12, Dec. 1995.

[52] Swe Toe-Naing and Yeo Kiat-Seng, "Study of Quarter-Micrometre Retrograde Well: Device Characteristics Against Temperature," *Electronics Letters*, vol. 35, no. 4, pp. 345-346, 18 Feb. 1999.

[53] Y.S. Lin, C.C. Wu, C.S. Chang, R.P. Yang, W.M. Chen, J.J. Liaw, and C.H. Diaz, "Leakage Scaling in Deep Submicron CMOS for SoC," *IEEE Transactions on Electron Devices*, vol. 49, no.6, pp. 1034-1041, Jun. 2002.

[54] W. Fikry, G Ghibaudo, and M. Dutoit, "Temperature Dependence of Drain-Induced Barrier Lowering in Deep Submicrometre MOSFETs," *Electronics Letters*, vol. 30, no. 11, pp. 911-912, 26 May 1994.

[55] M. Jamal Deen and Z.X. Yan, "DIBL in Short-Channel NMOS Devices at 77K," *IEEE Transactions on Electron Devices*, vol. 39, no. 4, pp. 908-915, Apr. 1992.

[56] Markus Rosar, Bernard Leroy, and Giorgio Schweeger, "A New Model for the Description of Gate Voltage and Temperature Dependence of Gate Induced Drain Leakage (GIDL) in the Low Electric Field Region," *IEEE Transactions on Electron Devices*, vol. 47, no. 1, pp. 154-159, Jan. 2000.

[57] Paul A. Tobias and David C. Trindade, *Applied Reliability*, second edition. Boca Raton: Chapman & Hall/CRC, 1995.

[58] F.S. Shoucair, "Analytical and Experimental Methods for Zero-Temperature-Coefficient Biasing of MOS Transistors," *Electronics Letters*, vol.25, no. 17, pp. 1196-1198, Aug. 1989.

[59] I.M. Filanovsky and Ahmed Allam, "Mutual Compensation of Mobility and Threshold Voltage Temperature Effects with Applications in CMOS Circuits," *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, vol. 48, no. 7, pp. 876-884, Jul. 2001.

[60] I.M. Filanovsky and L. Najafizadeh, "Zeroing in On a Zero-Temperature Coefficient Point," *IEEE 45th Midwest Symposium on Circuits and Systems*, vol. 1, pp. 271-274, Aug. 2002.