THERMAL EFFECTS IN DEEP SUB-MICRON VLSI INTERCONNECTS AND IMPLICATIONS FOR RELIABILITY AND PERFORMANCE

by

Kaustav Banerjee

Memorandum No. UCB/ERL M99/48

22 September 1999

COVER

THERMAL EFFECTS IN DEEP SUB-MICRON VLSI INTERCONNECTS AND IMPLICATIONS FOR RELIABILITY AND PERFORMANCE

by

Kaustav Banerjee

Memorandum No. UCB/ERL M99/48

22 September 1999

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Thermal Effects in Deep Sub-Micron VLSI Interconnects and Implications for Reliability and Performance

by

Kaustav Banerjee

B.S. (St. Xavier's College, University of Bombay, India) 1990
M.S. (Marquette University, Wisconsin, USA) 1993

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering
and Computer Sciences

in the

GRADUATE DIVISION

of the

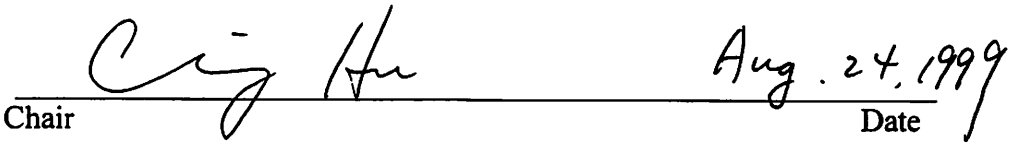
UNIVERSITY OF CALIFORNIA AT BERKELEY

Committee in charge:

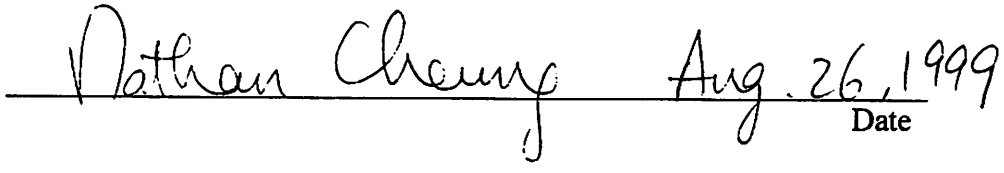
Professor Chenming Hu, Chair
Professor Nathan W. Cheung
Professor Tim Sands

Fall 1999

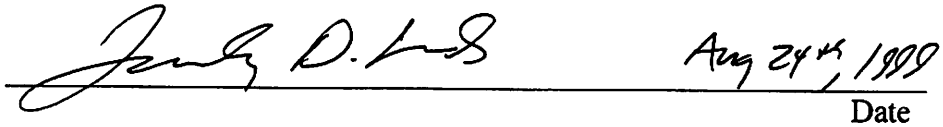
The dissertation of Kaustav Banerjee is approved:



Chair Date



Date



Date

University of California at Berkeley

Fall 1999

Thermal Effects in Deep Sub-Micron VLSI Interconnects and Implications for Reliability and Performance

by

Kaustav Banerjee

Abstract

Thermal Effects in Deep Sub-Micron VLSI Interconnects and
Implications for Reliability and Performance

by
Kaustav Banerjee

Doctor of Philosophy in Engineering-Electrical Engineering and Computer Sciences
University of California at Berkeley
Professor Chenming Hu, Chair

An investigation of thermal effects in deep sub-micron VLSI interconnect structures has been carried out and their implications on reliability and performance have been analyzed. Thermal effects are becoming an increasingly important issue in all high performance circuits due to increasing current density and power consumption. They are emerging as a key factor determining the design, performance, and reliability of future ICs.

This study begins with the development of suitable characterization and thermometry techniques to supplement the theoretical modeling and simulation work with experimental data. A novel thermometry technique, namely Transient Resistive Thermometry (TRT), together with a variation utilizing the Kelvin technique has been developed to facilitate systematic studies of thermal characterization in various interconnect structures in a modern VLSI technology. These techniques provide contactless temperature measurements by monitoring the temperature induced change of the resistivity of the material under study. They provide accurate estimation of spatially averaged temperatures with excellent temporal resolution within nanosecond timescales and also allow precise identification of thermal failures in the interconnect structures, thus helping the optimization of interconnect process design and reliability. A second thermometry technique called Scanning Joule Expansion Microscopy (SJEM) has also been advanced to facilitate thermal mapping of deep sub-micron VLSI structures with sub-100 nm spatial resolution.

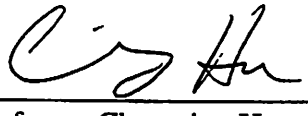
The TRT technique has been used to characterize and model thermal effects in VLSI metallization arising due to ultra short-duration high-current pulses during peak current and electrostatic discharge events. A model that takes into account heat diffusion under non-steady state conditions has been formulated to generate design guidelines for metal interconnects. Several IC manufacturers including TI, IBM, and INTEL have incorporated this model in their design process. Additionally, a new latent interconnect failure mode in AlCu lines has been discovered that significantly degrades their electromigration performance.

The TRT technique has also been shown to be an effective characterization tool to study the dynamic thermal response of low dielectric-constant materials used as inter-layer insulators in advanced interconnect systems. It offers a unique way of comparing the thermal characteristics (including thermal conductivity) of various dielectric films. The high-current behavior of thin Titanium and Cobalt silicide films used in advanced CMOS technologies has also been characterized and modeled using the TRT technique.

Using the TRT technique a new failure mode in deep sub-micron silicided barrier contacts has been identified for the first time.

The impact of thermal effects on the silicide-Si contact-resistance of small-geometry contact structures has been quantified. It has been found that contact-resistance sensitivity to temperature and current is a strong function of the silicide thickness, and is independent of the contact-plug material.

This research ends with an extensive study of circuit and interconnect performance implications arising due to thermal effects. The increasing influence of thermal effects on the design (driver sizing) and optimization of the interconnect length between repeaters for signal and power lines has been demonstrated.



Professor Chenming Hu
Dissertation Committee Chair

To

Ma, Bapi, and Sheetal

Acknowledgments

My years at *Cal* will always have a special place in my memories. What an amazing place! From the early hours of a misty Sunday morning in mid-August 1993, when I first landed in Berkeley after an interesting drive from the San Francisco airport through University Avenue, I must admit that I have been captivated by this place. Apart from its academic excellence, its unique landscapes with the SF Bay on the west and the Berkeley Hills to the east, wonderful array of restaurants from around the world, cafes that are busier during the nights, musicians, fortunetellers, and hawkers, on Telegraph Avenue, the old bookstores, the activists of Sproul Plaza and People's Park, along with numerous radical characteristics of Berkeley make it a haven for so many world renowned academicians, including several Nobel Laureates, Turing Award and Field's Medal winners, and top-notch researchers and students from across the world. It is truly a University beyond compare, and I couldn't have picked a better place for my doctoral education and research.

During my exciting and eventful student life that began in St. Xavier's College, Bombay, I have met several outstanding students and professors, some of whom have left most indelible marks on my memory and have significantly influenced my thoughts and my personality. However, one individual who has inspired me the most is Professor Chenming Hu, my research advisor at Berkeley, an exceptionally gifted person with outstanding technical achievements and with an unusual flair for fine arts. He has truly been my mentor, philosopher, and guru. I would like to express my most sincere appreciation for him for giving me complete freedom to explore new ideas and for taking keen interest in my research projects and personal well being. I am grateful for all his support, encouragement, and invaluable advice.

I would also like to thank Professor Nathan Cheung and Professor Tim Sands, for serving in my dissertation committee. They have provided several constructive criticisms, which have in turn, helped me formulate my dissertation in its present form. I have always been very impressed with Professor Cheung's sense of humor that helped lighten up the atmosphere during several research meetings and discussions. Professor Sands is one of the most approachable faculties in Berkeley, and I have always enjoyed talking to him.

During the course of my stay at Berkeley I got the opportunity to interact with several other distinguished faculties, which helped me formulate new ideas and also look at problems from different perspectives. Some of these Professors at Berkeley with whom I have had several fruitful discussions in classrooms or during technical meetings are: Professor Jeff Bokor, Professor Richard White, Professor Michael Liebermann, Professor Jan Rabaey, Professor Andy Glaeser, Professor Jim Morris, Professor Eicke Weber, Professor Ron Gronsky, and Professor Arun Majumdar.

I would also like to thank Professsor Ken Goodson from Stanford University for several interesting discussions on thermal problems and for providing some thermal conductivity data. I must also acknowledge Dae-Yong Kim from Stanford for lending his TEM expertise. I would also thank Dr. Sven Rzepka, visiting scholar from the Technical University of Dresden, Germany, for several interesting discussions on thermal problems.

I would like to thank the U.S. Semiconductor Research Corporation for providing necessary funding at Berkeley during 1995-1998. I must also thank Texas Instruments Inc., Dallas, for supporting much of my research activities through summer internships and for a $20,000 award during 1993-1997. In particular, I would like to thank Dr. Ajith Amerasekera and Dr. William Hunter for providing endless guidance and encouragement during various phases of this work. In fact, Ajith has been very closely involved in my projects, and I have really enjoyed co-authoring several research publications with him. He also provided several constructive comments after reading this entire manuscript.

I would like to acknowledge Dr. Joe W. McPherson, Dr. Charvaka Duvvury, Dr. Robert Havemann, Dr. Ping Yang, and Dr. Tim Rost, all from TI, for their enthusiastic support and collaborative involvement in some of my projects. I would also like to acknowledge several interesting technical discussions with Dr. Ih-Chin Chen, Dr. Girish Dixit, Dr. Amitava Chatterjee, Dr. Mark Rodder, Dr. Jorge Kittl, Dr. Sridhar Ramaswamy, Dr. Jerold Seitchik, Dr. Larry Ting, Dr. Alwin Tsao, and Paul Nicollian, from TI. It has indeed been a pleasure interacting with them.

I have also been very fortunate to be surrounded by several outstanding students at Berkeley, some of whom have become my closest friends and associates: Amit Mehrotra, Shaz Qadeer, Mukul Prasad, Amit Narayan, Premal Buch, Gurmeet Singh Manku, Alok Aggarwal,

Rajeev Murgai, Harmeet Singh, Sriram Rajamani, Amoolya Singh, Igor Kouznetsov, Jon Gudmundsson, Alexei Marakhtanov, Kedar Patel, Philip Chong, Michael Shilman, Subarnarekha Sinha, Heloise Hwawen-Hse, Ravi Gunturi, Luca Carloni, Marco Sgroi, Edoardo Charbon, Sunil Khatri, Sekhar Narayanaswami, George Varghese, and Nikhil Jakatdar to name a few.

I would also like to acknowledge my past and present colleagues and friends from the Device group and the Plasma lab: James Chen, Michael Orshansky, Jone Chen, Dennis Sylvester, Kai Chen, Qiang Lu, Barry Linder, William Wong, Sunder Kumar Iyer (SKI), Erin Jones, Bill En, Jiang Tao, and Xiang Lu. I would also like to thank Guanghua Wu and Masanobu Igeta from the mechanical engineering department for their help and collaboration in thermal microscopy project.

Infrastructure and support are important components of any organization. In this regard, I would like to thank Ruth Gjerde and Mary Byrnes of the EECS graduate office in 231 Cory Hall for all their help with administrative matters. I would also like to specially thank our group administrator Judy Fong for all her help with administrative issues, travel arrangements and reimbursements, and in dealing with funding agencies. She took care of all these important issues most efficiently and always with a smile. Another person at Berkeley who deserves special mention is Mr. Ted Goode, the International Student Advisor. He is one of the most knowledgeable and helpful legal staff in the International Student Center at the I-House and helped me a number of occasions during international travel, summer internships, and even during my change of Visa status after graduation.

I would also like to acknowledge a couple of friends outside the University but residents of the great city of Berkeley: Pradeep Kumar and Rizwan Rahmani. Pradeep always amazed me with his great sense of humor, and his culinary and entrepreneurial skills. I would like to thank Rizwan Bhai (along with Shaz) for all the good times we had in Berkeley and San Francisco, in particular, the late nights at *Café Strada, and Au Coquelet Café* in Berkeley and at *Steps of Rome* and *Shalimar* in San Francisco, appreciating *Meer*, *Ghalib*, *Iqbal*, *Faiz*, *Momin*, and other Urdu poets. These extracurricular activities and the discussions on all aspects of life (and some *EE* problems....) at *Café Nefeli* near Cory Hall made my Berkeley experience worthwhile.

On the family front, I want to thank my parents for all their love and attention and for providing constant encouragement and challenge to achieve "greater things" in life. I would also like to thank my wife's parents, my brother, and my sister for their love and solidarity. Finally, I want to express my deepest gratitude to my wife, Sheetal, for her endless love, and for gracefully enduring these *lost years* of our lives. She has provided constant support, help, and encouragement during the writing of this thesis.

August 24, 1999

Berkeley, California

dam-e-har-mauj mein hai, halqa-e-sad-kam-e-nahang
dekhein kya guzray hai qatray pe guhar hone tak

.....Ghalib

(A noose in every wave, infinite dragons in every noose
Yearning to be a pearl, will I win or will I lose?)

Contents

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

LIST OF FIGURES

List of Tables

Chapter 1

Introduction

1.1 Thermal Effects in VLSI Interconnects: An Overview

As VLSI technology scales, interconnects are becoming the dominant factor determining system performance and power dissipation [1], [2]. Thermal effects are an inseparable aspect of electrical power distribution and signal transmission through the interconnects in VLSI circuits due to self-heating (or Joule heating) caused by the flow of current. Thermal effects significantly impact interconnect design and reliability.

The ever-increasing demand for speed and functionality of Si-based advanced high performance microprocessors, digital signal processing (DSP) chips, application specific integrated circuits (ASICs), and the increasing density of memory (DRAM) devices has caused aggressive scaling of ICs beyond 0.5-μm minimum feature size [3], [4]. These technology nodes, commonly referred to as *deep sub-micron*, (DSM) allow VLSI circuits to meet the required device density and various circuit performance specifications [5] – [7]. This trend has resulted in a dramatic reduction of the interconnect metal pitch and increased the number of metallization

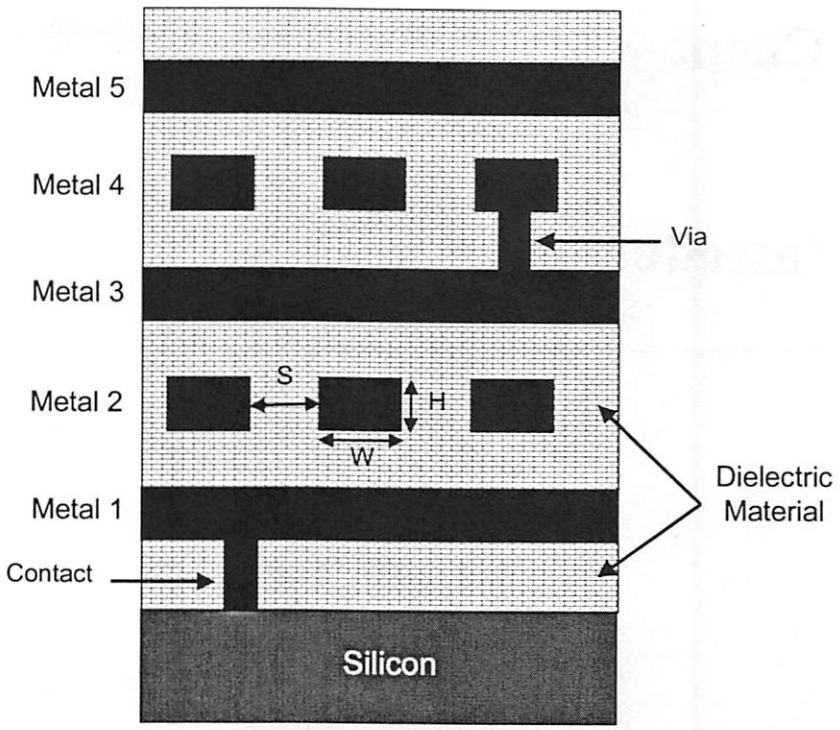


Figure 1.1 A schematic cross section of a multi-level interconnect scheme employed in present VLSI circuits. The metal pitch is defined as ($W + S$) and the aspect ratio is defined as (H/W).

levels to accommodate the increasing number of wired circuits per chip. A schematic cross-section of a multilevel interconnect scheme is shown in Figure 1.1.

This aggressive interconnect scaling has resulted in increasing current densities [8] and associated thermal effects. Furthermore, low dielectric constant (Low-k) materials are being introduced as an alternative insulator to reduce interconnect capacitance (therefore delay) and cross-talk noise to enhance circuit performance [9]. These materials can further exacerbate thermal effects owing to their poor thermal properties.

Apart from normal circuit conditions, ICs also experience high current stress conditions, the most important of them being the electrostatic discharge (ESD), that causes accelerated thermal failures [10]. Semiconductor industry surveys indicate that ESD is the largest single cause of failures in ICs [11]. Figure 1.2 shows an ESD failure in an NMOS output transistor of an advanced CMOS process after ESD stress at 3.5 kV (Human Body Model) HBM stress. Interfacing between multiple power supply chips, as well as between multiple power supply blocks within a chip, also causes high current conditions at the I/O circuitry [12].

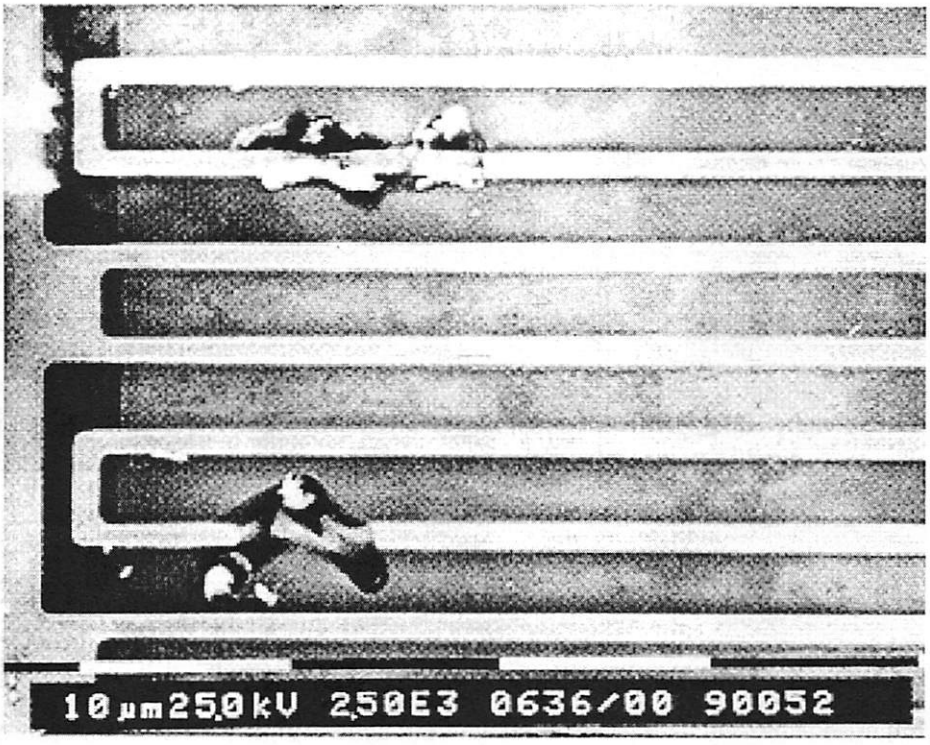


Figure 1.2 An ESD failure in an NMOS output transistor of an advanced CMOS process after a 3.5 kV HBM ESD stress. (Courtesy of Dr. Ajith Amerasekera, Texas Instruments Inc.)

Interconnects are also known to experience similar stress conditions during testing for latchup robustness. The need to understand high current behavior of VLSI circuits has increased in importance due to their continuous scaling. Technology scaling of interconnects have also evolved rapidly in recent years, leading to new interconnect and dielectric isolation materials. There is an increasing need to comprehend high current, ESD and thermal effects in these structures in order to provide robust design guidelines.

1.2 Trends in VLSI Scaling and Implications on Thermal Effects

In this section various interconnect technology scaling trends will be discussed to illustrate their impact on thermal effects. As VLSI circuits continue to be scaled aggressively, a rapid increase in functional density and chip size is observed as shown in Figure 1.3. This has resulted

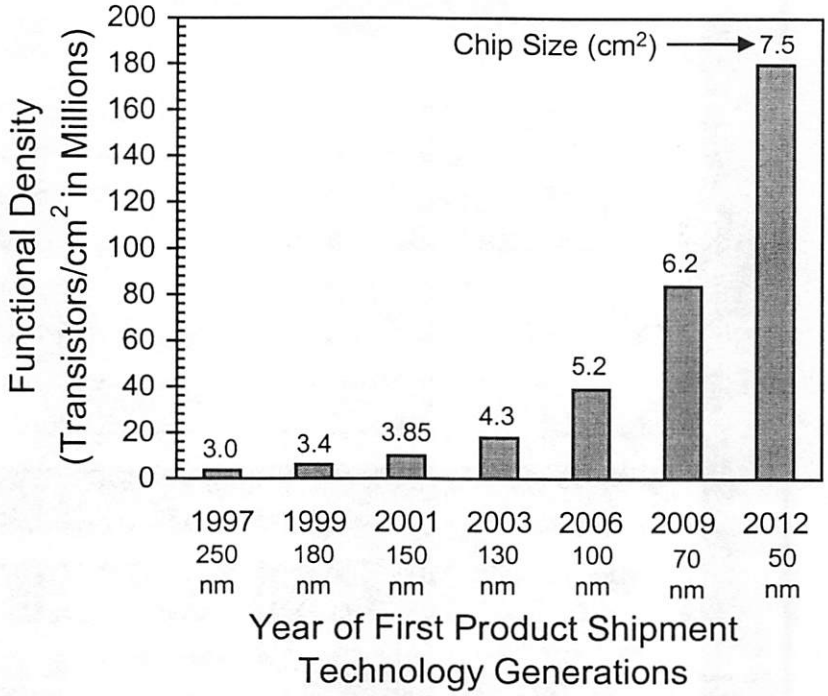


Figure 1.3 Present and projected functional density and chip size of microprocessors for different technology nodes and year of first product shipment as per NTRS [13]. The functional density is expected to increase from 3.7 million in 1997 to 180 million in 2012.

in increasing number of interconnect levels and reduction in interconnect pitch in order to realize all the inter-device and inter-block communications. Figure 1.4 shows this trend with interconnect levels increasing further in the near future, from 6 levels at the 250 nm node to 9 levels at the 50 nm node. This increase in the number of interconnect levels causes the upper most interconnect layers to move further away from the Si substrate making heat dissipation more difficult. Furthermore, the critical dimensions of contacts and vias are also decreasing with scaling as shown in Figure 1.5 resulting in higher current densities in these structures. Compounded with the introduction of low-k dielectrics as alternative insulators, whose thermal conductivities are also much lower than that of silicon dioxide (Figure 1.6), it is envisioned that thermal effects in interconnects can potentially become another serious design constraint.

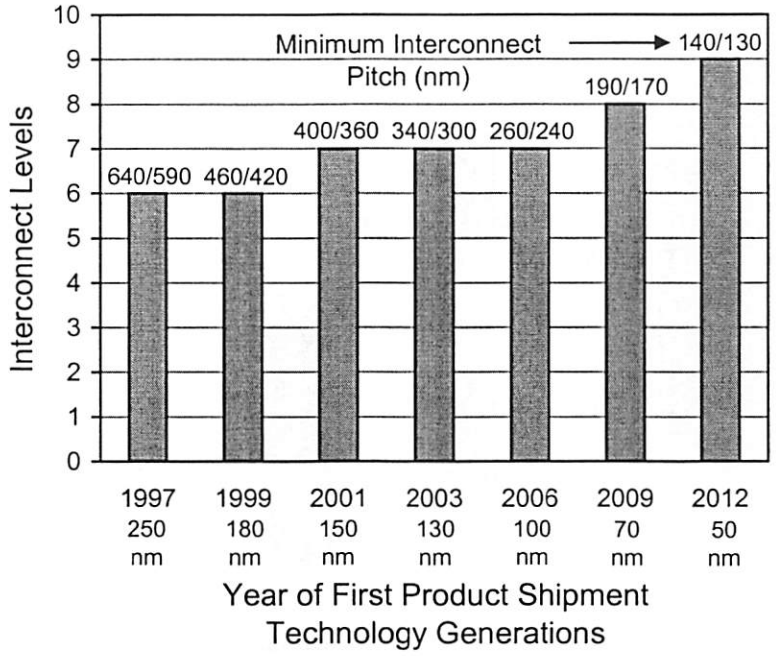


Figure 1.4 Present and projected interconnect levels and minimum contacted/non-contacted interconnect metal pitch of logic circuits for different technology nodes and year of first product shipment as per NTRS [13].

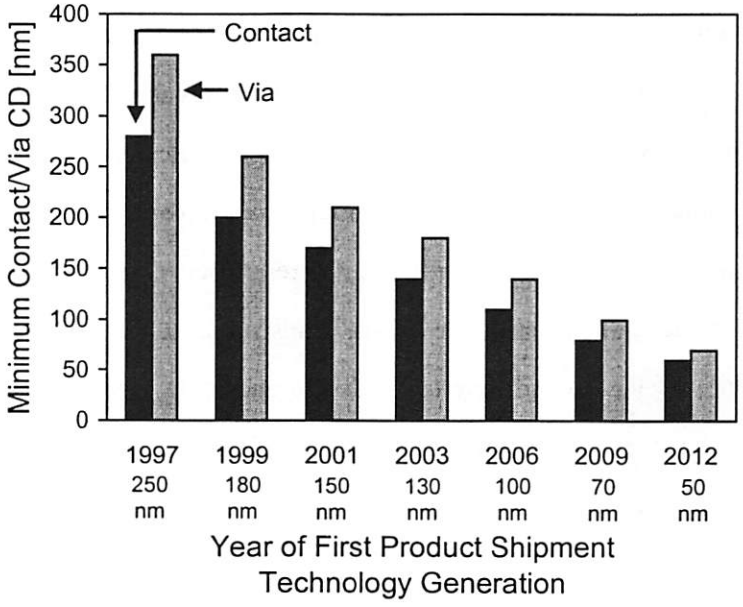


Figure 1.5 Present and projected minimum contact/via critical dimensions (CD) of logic circuits for different technology nodes and year of first product shipment as per NTRS [13].

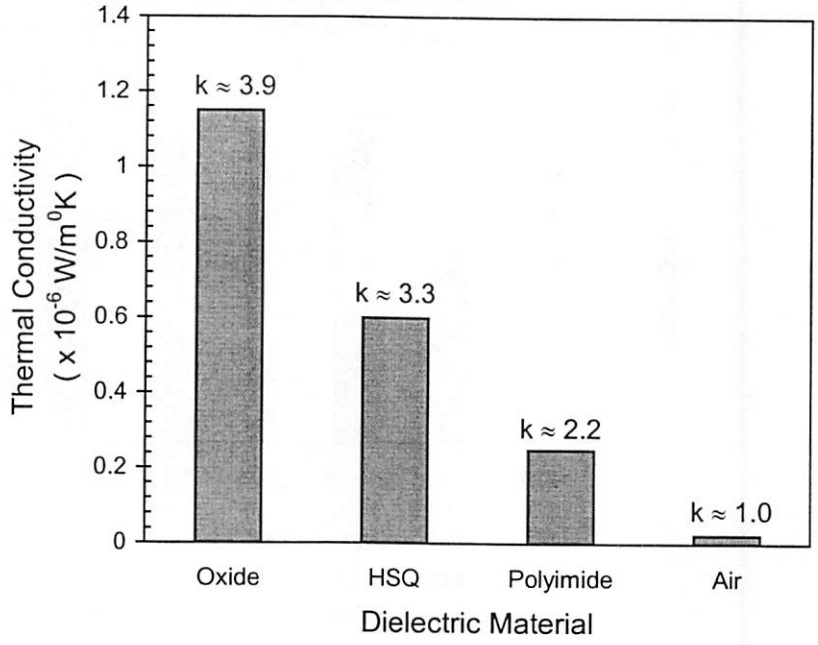


Figure 1.6 Thermal conductivity and approximate dielectric constants (k) of some insulating materials used (or being introduced) in high performance circuits.

A simple analysis to demonstrate the implications of technology scaling on technology performance is now presented. This analysis is instructive since these in turn have important implications on thermal effects and on reliability requirements for interconnects. Consider an interconnect segment of length l between two inverters as shown in Figure 1.7 (a). Figure 1.7 (b) shows the equivalent distributed RC representation. Here the inverter on the left that is driving the interconnect is represented as a voltage source controlled by the voltage V_{st} at the input capacitance. R_{tr} is the equivalent transistor resistance, C_p is the parasitic capacitance composed mainly of the drain capacitance of the transistors. C_L is the load capacitance or the input capacitance of the second inverter. Also, c and r are the capacitance and resistance per unit length of the interconnect line. Using this simple model the performance of a technology can be most simply summarized by the delay time t_d for a logic signal, which is given by equation (1.1) [14].

$$t_d = R_{tr}\left(C_p + C_L\right) + \left(R_{tr}c + rC_L\right) l + \frac{1}{2} rcl^2 \qquad (1.1)$$

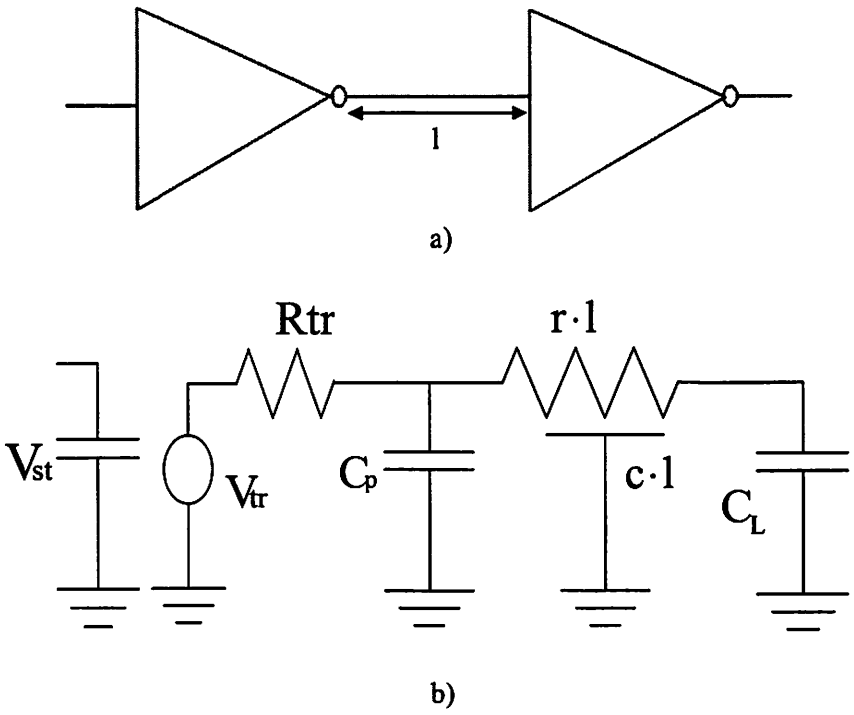


Figure 1.7 a) Schematic representation of an interconnect segment of length *l* between two inverters, and b) an equivalent distributed RC circuit.

R_{tr}, C_p, and C_L are functions of the transistor design and the circuit design *W/L* ratios used. The interconnect resistance per unit length, *r*, depends on the resistivity of the interconnect metal while *c* depends on the dielectric constant of the surrounding insulating material, interconnect metal pitch, interconnect geometry, and underlying insulator thickness [15]. The three terms on the right hand side of equation (1.1) represent the intrinsic stage delay, the load delay, and the interconnect delay contributions, respectively. The factor of half in the interconnect delay arises due to the distributed nature of the interconnect [16].

Now if *s* is the scaling parameter, which can be defined as the ratio of the feature size at a newer technology node to the feature size at an older reference technology node, then *s* is always less than one. There are two simplified scaling scenarios for interconnects:

- Scale metal pitch ($W + S$) at constant metal thickness (H) (Figure 1.1)
- Scale metal pitch and the metal thickness

Under the first scenario, the load delay, line delay and current density will scale as $1/s$, $1/s^2$, and $1/s$ respectively. Under the second scenario, the load delay, line delay and current density will scale as $1/s^2$, $1/s^2$, and $1/s^2$ respectively. In either case, it can be seen that interconnect delays begin to dominate technology performance, and that current density increases with scaling.

The interconnect scaling requirements drive several technology enhancements. Use of low-k materials lowers the interconnect capacitance per unit length c (particularly intra-level) in equation (1.1) and therefore lowers the total interconnect delay ($0.5rcl^2$). Lower interconnect capacitance also helps in minimizing cross-talk noise. Furthermore, lower interconnect capacitance also helps in reducing the dynamic power dissipation (P_{dyn}) during the switching of gates in digital circuits, which can be estimated by [16].

$$P_{dyn} = \frac{1}{2}\alpha C V^2 f \tag{1.2}$$

where α is the activity factor or switching probability, C is the total capacitance, V is the power supply voltage and f is the clock frequency of the circuit.

Hence, it is the minimization of interconnect capacitance that is driving the introduction of low-k dielectric materials. Similarly, lower interconnect resistance and higher current density requirements drive the use of new metallization (namely Cu). Since these low-k materials also have lower thermal conductivity than silicon dioxide (Figure 1.6), heat dissipation becomes even more difficult. Thus VLSI technology scaling has important implications on thermal effects as discussed in this section. The various trends in technology scaling that causes increased thermal effects in interconnects can be summarized as:

- Increasing current density
- Increasing number of interconnect levels
- Introduction of Low-k dielectric materials

1.3 Impact on Interconnect Reliability and Design

Thermal effects impact interconnect design and reliability in the following ways. Firstly, they limit the maximum allowable RMS current density, (since the RMS value of the current

density is responsible for heat generation) in the interconnects, in order to limit the temperature increase. Secondly, interconnect lifetime (reliability) which is limited by electromigration (EM) (transport of mass in metals under an applied current density), has an exponential dependence on the inverse metal temperature [17]. Hence, temperature rise of metal interconnects due to self-heating phenomenon can also limit the maximum allowed average current density, since EM capability is dependent on the average current density [18]. Thirdly, interconnect failure under high current stress conditions including electrostatic discharge (ESD) is also a reliability concern [19], [20]. Interconnect failure due to ESD is becoming an important issue as VLSI scaling continues. As the number of wired gates (G) per chip increases (see Figure 1.3) the number of I/O pins (N_P) also increases as per Rent's Rule [16], which is given by

$$N_P = K \cdot G^{\beta} \tag{1.3}$$

where K is the average number of I/Os per gate, and β is the Rent exponent that can vary from 0.1 to 0.7. As a consequence of this increasing I/O pins, the package floor planning is changing from peripheral package connections to array grids in order to accommodate the increased I/O pin count [20]. In the array architecture, the interconnect widths between external pads and the ESD structures must decrease to preserve chip wirability and to prevent timing delays in critical paths and in the receiver and driver networks. This trend can increase the susceptibility of interconnects to ESD failure. Furthermore, technology scaling and the transition to new interconnect and dielectric materials (discussed in section 1.2), necessitates a growing need to comprehend the high current behavior of these structures and analyze their failure mechanisms in order to provide robust design guidelines.

1.4 Dissertation Objective

The broad goal of this research is to develop suitable thermometry and diagnostic techniques, characterize high-current and thermal effects in various interconnect structures, formulate analytical models, and to provide insight into their failure mechanisms and their implications on reliability and performance. The outcome of this work will have significant

implications on the design of deep sub-micron VLSI interconnect structures and on developing electromigration, ESD protection circuit, and I/O buffer interconnect design rules.

Since a modern VLSI interconnect scheme consists of metal lines, contacts, vias and silicide films, this dissertation will analyze thermal effects in all these interconnect structures. It will also introduce novel thermal characterization techniques and present new information on physical mechanisms leading to their reliability degradation or failure under high-current stress conditions. Characterization work will be supplemented with analytical modeling to optimize interconnect and interconnect process design, which will in turn help in optimizing interconnect reliability and performance.

1.5 Organization

After this introductory chapter describing the motivations behind studying thermal effects in VLSI interconnect structures, a general treatment involving physical and mathematical foundations of self-heating and thermal effects in interconnects is provided in *Chapter 2*. This chapter outlines the fundamentals of self-heating in metal lines and introduces the equations describing heat conduction. It also explains the concepts behind the justification of analyzing thermal problems analogous to electrical circuits. The next few chapters are focused on characterization and modeling of high-current and thermal effects in specific components in an interconnect scheme. Due of the discrete nature of the problems addressed in this study, relevant prior work is discussed separately in each chapter wherever applicable, and also a brief summary is provided separately in each of these chapters.

Chapter 3 analyzes self-heating in interconnect lines under DC and short-pulse stress conditions. It introduces the *Transient Resistive Thermometry* technique that is used in several subsequent chapters (Chapter 4 and 6) to study high-current effects and failure mechanisms of interconnect structures. The TRT technique has also been shown to be useful for thermal characterization of thin dielectric materials. This technique can be used to evaluate the effective thermal conductivity of various low-k inter-layer dielectric films used in DSM interconnect processes. *Chapter 4* describes the impact of interconnect scaling and employment of low-k dielectrics on the thermal characteristics of the interconnect lines. The transient technique

developed in chapter 3 is shown to be an excellent methodology to compare thermal characteristics of different insulating dielectric materials. *Chapter 5* describes the *Transient Kelvin Thermometry* (TKT), a variation of the TRT that utilizes the Kelvin technique. The TKT technique has been developed to study the self-heating of small-geometry contacts and vias and in particular to identify their failure mechanisms under high-current stress conditions. *Chapter 6* is an application of the transient resistive thermometry technique to study self-heating effects in thin silicide films and their failure mechanisms.

Chapters 7-10 each addresses a distinct problem related to thermal effects. *Chapter 7* analyzes effects of temperature on the $TiSi_2$-Si contact resistance in small-geometry silicided contacts and also characterizes their high-current behavior and identifies their failure mode using the transient Kelvin thermometry technique.

In *Chapter 8* another novel thermometry technique based on an atomic force microscope, *Scanning Joule Expansion Microscopy*, developed by the microscale thermophysics group at Berkeley, is used to study the thermal characteristics of sub-micron vias and to obtain a spatial temperature distribution over the via.

Chapter 9 analyzes the impact of self-heating and high current density during highly accelerated EM tests on the sensitivity of reliability ranking of different interconnect processes.

Chapter 10 is devoted to analyzing the implications of thermal effects and technology scaling on interconnect design and performance. In this chapter reliability (EM and thermal) driven interconnect design rules are compared to performance driven design rules to quantify the increasing influence of thermal effects on the design and optimization of the signal line length.

Finally, *Chapter 11* draws the major conclusions of this study and also outlines future research directions related to thermal problems in high performance circuits.

Chapter 2

Basic Concepts for Analyzing Self-Heating Effects in Metal Interconnects

In this chapter some fundamental concepts in Joule heating (self-heating) will be reviewed along with an introduction to the heat conduction equations for interconnects. The basis for analyzing thermal problems as electrical circuits will also be discussed in detail. Various electrothermal analogs will be defined and their relationships to interconnect geometry and thermal properties of the materials involved in ICs will be established.

2.1 DC Electrical Conductivity of a Metal

It is useful to understand the various parameters related to electrical conductivity of metals at a micro level since they are needed in analyzing joule heat generation. Consider a metal wire of length L, cross sectional area A, and resistivity ρ, at uniform temperature. The resistivity is defined to be the proportionality constant between the electric field E at a point in the metal and the current density j that it induces.

$$E = \rho j \tag{2.1}$$

In general E and j are vector quantities and they need not be parallel, in which case ρ is a tensor. The current density j is parallel to the flow of charge, whose magnitude is the amount of charge per unit time crossing a unit area perpendicular to the flow. Thus for a uniform current I flowing through the wire, $j = I/A$. Since the potential drop across the wire will be $V = EL$, equation (2.1) gives $V = \rho IL/A$, and hence the resistance of the wire (from Ohm's law, $V = IR$) is given by

$$R = \rho \frac{L}{A} \tag{2.2}$$

At any point in a metal, electrons are always moving in a variety of directions with a variety of thermal energies. If n electrons per unit volume move with an average velocity v, then in the absence of an electric field they are as likely to be moving in one direction as in any other, and hence $v = 0$. However, in the presence of an electric field there will be an average electronic velocity v, directed opposite to the field, which can be computed as follows.

Consider a typical electron at time *zero*. Let t be the time elapsed since its last collision. Its velocity at time *zero* will be its velocity v_0 immediately after that collision plus the additional velocity $-eEt/m$ it has subsequently acquired. Since it is assumed that an electron emerges from a collision in a random direction, there will be no contribution from v_0 to the average electronic velocity, which must therefore be given entirely by the average of $-eEt/m$. Since the average of t is the relaxation time τ, the average electronic velocity is given by

$$v = -\frac{eE\tau}{m} \tag{2.3}$$

Now in a time dt the electrons will advance by a distance vdt in the direction of v. Thus $n(vdt)A$ electrons will cross an area A perpendicular to the direction of flow. Since each electron carries a charge $-e$, the charge crossing A in time dt will be $(-nevAdt)$, and hence the current density is $(j = -nev)$. Hence from equation (2.3) the current density is given by [21],

$$j = \left(\frac{ne^2\tau}{m}\right)E \tag{2.4}$$

This relation is usually stated in terms of the electrical conductivity σ (inverse of the resistivity, $\sigma = 1/\rho$) as [21],

$$j = \sigma E; \qquad \sigma = \frac{ne^2\tau}{m} \tag{2.5}$$

2.2 Physics of Joule Heat Generation

The basic physics behind Joule-heating or self-heating is presented in this section. Consider that the metal wire of length L and cross sectional area A, described in the previous section is in a static electric field, E. An electron in the metal wire experiences several collisions. The average energy lost to the ions per electron per collision can be calculated as follows. The force acting on an electron would be $F = -eE$, and if the average distance traveled between collisions is x, the average energy loss, will be $U_{avg} = (-eE)x$, where x = (average electron velocity, v) . (average time between collisions, τ). Therefore, from equation (2.3) it follows that

$$x = \left[-\frac{eE\tau}{m}\right]\tau \tag{2.6}$$

and hence [21],

$$U_{avg} = \frac{(eE\tau)^2}{m} \tag{2.7}$$

Therefore the average rate of energy lost to the ions per unit volume is given by

$$u_{avg} = \frac{ne^2\tau}{m}E^2 \tag{2.8}$$

Now the power loss, P, is simply the average energy loss per second, and using equation (2.1) (2.4) and (2.5) it can be expressed as

$$P = \left(\frac{ne^2\tau}{m}\right)E^2(LA) = \sigma(\rho j)^2(LA) \tag{2.9}$$

And since $\sigma = 1/\rho$ the power loss is given by [21],

$$P = I^2R \tag{2.10}$$

This is the electrical power dissipation in a metal wire of resistance R when a current I flows through it. This electrical energy gets converted to thermal energy (Joule heat) which raises the temperature of the metal interconnects. Hence, for uniform current flow in a metal, Joule heating can be expressed as

$$P = IV = I^2R = \frac{dQ}{dt} \tag{2.11}$$

Here Q is the heat energy generated.

2.3 Fourier's Law and Thermal Conductivity

The fundamental law of heat conduction (also known as Fourier's law) is consistent with the second law of thermodynamics. It states that the quantity of heat (Q) conducted through a homogeneous solid is directly proportional to the cross-sectional area (S) normal to the path through the solid and the temperature gradient along the path [22], (see Figure 2.1).

Hence the rate of heat flow (q) through the slab can be expressed as,

$$q = \frac{dQ}{dt} \sim S\frac{(T_1 - T_2)}{L} \sim -S\frac{dT}{dx} \tag{2.12}$$

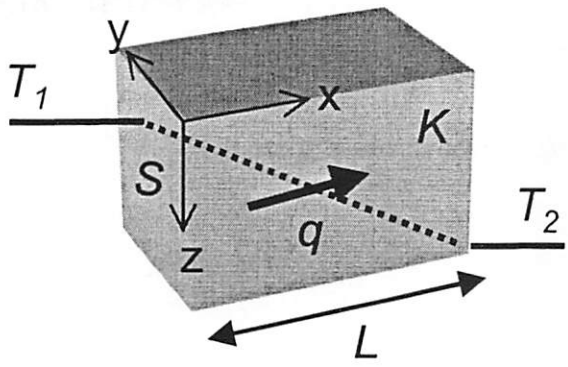


Figure 2.1 Heat transfer by conduction through a plane slab of cross-sectional area S and length L. The temperature on the left hand surface is T_1 and on the right hand surface is T_2, with $T_1 > T_2$, q is the rate of heat transfer.

where the minus sign indicates that a flow of heat can occur only in the presence of a falling temperature gradient, consistent with the second law of thermodynamics. Insertion of a proportionality constant yields the Fourier's law, which also serves to define the thermal

$$q = \frac{dQ}{dt} = -KS\frac{dT}{dx} \tag{2.13}$$

conductivity K of the material. Therefore,

$$K = \frac{dQ/dt}{S\ dT/dx} \tag{2.14}$$

and has the units of $Wm^{-1}C^{-1}$. This equation is used to formulate the heat conduction equation in a Joule heated metal line in the next section. Strictly speaking K is not constant for the same material but depends upon the temperature. However in the ordinary mathematical analysis it is assumed that K is invariant with temperature. The thermal diffusivity κ is related to K by [22],

$$\kappa = \frac{K}{\delta c} \tag{2.15}$$

where δ is the average density of the material and c is the specific heat. κ has units of m^2s^{-1}.

2.4 Heat Conduction Equation in a Metal Line Heated by an Electric Current

Consider a metal line of length L, width w, and thickness d as shown in Figure 2.2. If this line is Joule heated by a current of magnitude I, the simple 1-D heat flow equation can be formulated as follows. In time Δt, the increase in volumetric heat stored must equal the volumetric Joule heat generated and the net heat dissipation. Therefore, if the temperature rise of the metal is ΔT, the heat balance can be expressed as

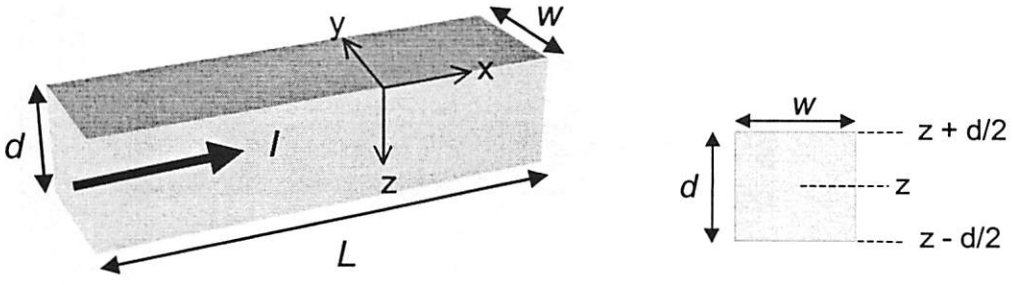
$$mc\Delta T = P\Delta t + \frac{dQ}{dt}\Delta t \tag{2.16}$$

Here m is the mass and c_m is the specific heat of the metal respectively. Now, if δ_m is the metal density, it follows that

$$mc\Delta T = \delta_m (Lwd) c_m \left[T(z,t+\Delta t) - T(z,t)\right] \tag{2.17}$$

and,

$$P\Delta t = j(z,t)^2 \rho Lwd\Delta t \tag{2.18}$$



a) b)

Figure 2.2 a) Schematic view of a metal line carrying current I. b) The cross sectional view.

Also, the rate of (1-D) heat flow through the metal of thermal conductivity K_m is expressed as

$$\frac{dQ}{dt} = K_m(wL)\left[\left.\frac{\partial T}{\partial z}\right|_{z+d/2,t} \quad -\left.\frac{\partial T}{\partial z}\right|_{z-d/2,t}\right] \tag{2.19}$$

Thus the 1-D heat balance equation is given by

$$\delta_m(Lwd)c_m[T(z,t+\Delta t)-T(z,t)] = j(z,t)^2\,\rho Lwd\Delta t + K_m(wL)\Delta t\left[\left.\frac{\partial T}{\partial z}\right|_{z+d/2,t} \quad -\left.\frac{\partial T}{\partial z}\right|_{z-d/2,t}\right] \tag{2.20}$$

Dividing equation (2.16) by ($Lwd\Delta t$) it follows that

$$\delta_m c_m \frac{\partial T_m}{\partial t} = K_m \frac{\partial^2 T_m}{\partial z^2} + j(z,t)^2\,\rho_m \tag{2.21}$$

For an interconnect line embedded in a dielectric (oxide) as shown in Figure 2.3, this 1-D heat equation for the metal is valid only when the width of the metal line is much greater than the oxide thickness above and below, $w >> t_{ox1}, t_{ox2}$.

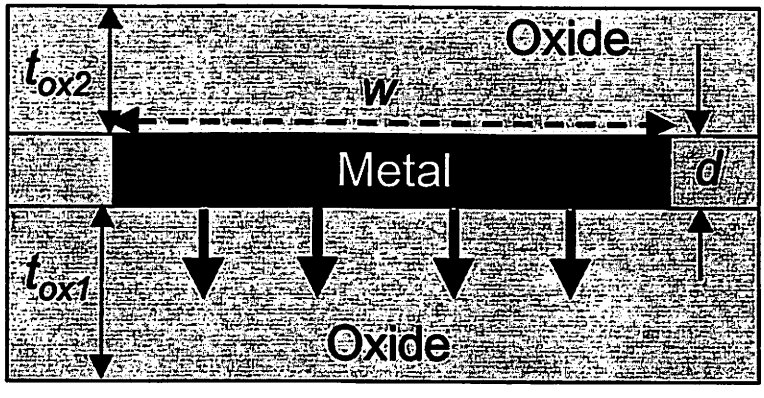


Figure 2.3 A 1-D example of heat flow, Joule heated interconnect line embedded in oxide, $w >> t_{ox1}, t_{ox2}$.

2.5 Electrothermal Analog

In order to obtain the solution to a problem in conduction heat transfer it is often converted to an equivalent electrical circuit, which is then analyzed using traditional circuit concepts. In this section, the basis for employing this equivalence will be examined in detail. Understanding these equivalent electrical parameters is important since they have been widely used in the analysis of thermal problems in various interconnect structures in this study.

2.5.1 The Transmission Line Equation

The transmission line equation [23] serves as the basis for the electrothermal analog. It consists of two parallel conductors, which comprise a network with distributed parameters. Analysis of the transmission line involves four distributed parameters, three of which have analogous thermal parameters. These are resistance, capacitance, and, conductance. Inductance is an electrical parameter that does not have an analogous thermal quantity. In the present analysis two of these parameters will be considered present, namely the resistance and the capacitance. They are present as series and shunt parameters respectively. The entire transmission line may be thought of as a cascaded network of lumped parameter "Tee" sections, one of which is shown in Figure 2.4. Here the series impedance Z has been divided in half for network symmetry and equals $R/2$. The shunt admittance, Y, is the capacitance, C. These parameters are considered invariant with time and also independent of current and voltage. Here R and C are in terms of Ohms per unit length and Farads per unit length respectively. If the

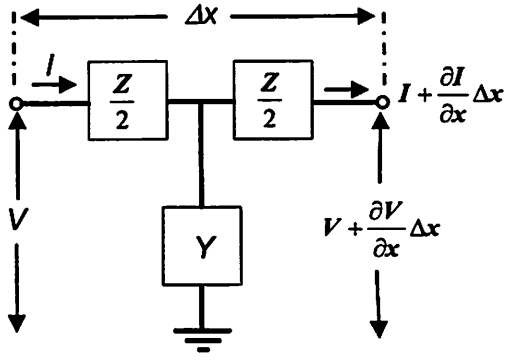


Figure 2.4 A single lumped parameter "Tee" section of a transmission line.

length of the transmission line between x and $x + \Delta x$ is considered (as in Figure 2.4), the drop in voltage along the line in element x will be

$$V(x+\Delta x,t) - V(x,t) = -(R\Delta x)I \tag{2.22}$$

Similarly, the change in current along the element will be

$$I(x+\Delta x,t) - I(x,t) = -(C\Delta x)\frac{\partial V}{\partial t} \tag{2.23}$$

Dividing equation (2.22) and (2.23) by Δx, one obtains, in the limiting case as Δx approaches zero, the equations

$$\lim_{\Delta x \to 0} \frac{V(x+\Delta x,t) - V(x,t)}{\Delta x} = \frac{\partial V}{\partial x} = -RI \tag{2.24}$$

and

$$\lim_{\Delta x \to 0} \frac{I(x+\Delta x,t) - I(x,t)}{\Delta x} = \frac{\partial I}{\partial x} = -C\frac{\partial V}{\partial t} \tag{2.25}$$

Differentiating equation (2.24) with respect to x it follows that

$$\frac{\partial^2 V}{\partial x^2} = -R\frac{\partial I}{\partial x} \tag{2.26}$$

and substituting this result in equation (2.26) yields the transmission line equation [23]:

$$\frac{\partial^2 V}{\partial x^2} = RC\frac{\partial V}{\partial t} \tag{2.27}$$

Now the heat conduction equation for a metal line carrying zero current, i.e., $j = 0$ in equation (2.21), is given by

$$\delta_m c_m \frac{\partial T_m}{\partial t} = K_m \frac{\partial^2 T_m}{\partial z^2} \tag{2.28}$$

Comparison of equation (2.27) and (2.28) reveals that these equations are strikingly similar. The analogous quantities in the electrothermal analog are therefore electric potential or voltage to temperature and the reciprocal of the resistance-capacitance product to thermal diffusivity, i.e.,

$$V \leftrightarrow T; \quad and \quad \frac{1}{(RC)l^2} \leftrightarrow \kappa_m = \frac{K_m}{\partial_m c_m} \tag{2.29}$$

where l is the length of the transmission line.

2.5.2 Limitations of the Analog

The resemblance of the transmission line equation to the Fourier equation shows that transient heat flow problems can be handled by using transmission line theory. However, the transmission line representation will yield an exact solution only when an infinite number of Tee sections are considered. This, in reality, requires that the length of each Tee section be infinitesimal. Fortunately, for most cases an infinitesimal Tee, or at least a Tee that yields a solution of sufficient accuracy, is approached quite readily.

2.5.3 Concept of Thermal Resistance (or Impedance)

Intuitively, just as a potential difference between two points in a conductor sets up a current flow, a temperature difference between two points in a material sets up heat flow. The thermal resistance of a configuration (like Figure 2.1) can be determined by comparing the 1-D steady state heat conduction equation (2.12) to Ohm's law.

That is, $q = KS\frac{(T_1 - T_2)}{L}$ and $I = \frac{V}{R}$

It can be observed that rate of heat flow (q) and current (I), which is the rate of charge flow, are analogous quantities. The temperature difference $\Delta T = (T_1 - T_2)$ and voltage (V) are also

analogous as established earlier in equation (2.29). Therefore, the thermal resistance (R_θ) analogous to the electrical resistance (R) has the units ^{0}C/W and is defined as,

$$R_\theta = \frac{L}{KS} \tag{2.30}$$

2.5.4 Concept of Thermal Capacitance

The flow of current through a capacitor is proportional to the change of voltage with time, with capacitance (C) being the proportionality constant, defined by

$$I = C\frac{dV}{dt} \tag{2.31}$$

Comparing equation (2.31) with the rate of heat energy gained by a material of mass m, density δ, volume υ, and specific heat c, given by

$$q = mc\frac{dT}{dt} = c\delta\upsilon\frac{dT}{dt} \tag{2.32}$$

one defines the thermal capacitance (C_θ) as,

$$C_\theta = c\delta\upsilon \tag{2.33}$$

in units of Joules/^{0}C. Just as the concept of capacitance in electrical circuits is valid for time varying voltage, similarly, it is meaningful to talk about a thermal capacitance only under transient heat flow conditions (non-steady state).

2.5.5 Thermal Time Constant

In electrical circuits there is an associated electrical time constant given by the product of the resistance (R) and the capacitance (C), that is a measure of the time required to reach steady-state conditions. Similarly, in a thermal circuit, an equivalent thermal time constant (τ_θ) is defined as

the product of the thermal resistance (R_θ) and the thermal capacitance (C_θ). Hence, from equation (2.30), (2.33) and (2.15) and since $\upsilon = LS$, τ_θ is given by

$$RC \equiv \tau_\theta = R_\theta \cdot C_\theta = \left(\frac{L}{KS}\right)(c\delta\upsilon) = \frac{L^2}{\kappa} \tag{2.34}$$

in units of time.

2.5.6 Electrothermal Analogs for VLSI Interconnect

Consider a metal line as described in Figure 2.5. Let P be the power dissipation by the line due to Joule heating under a DC current stress. The steady-state 1-D heat flow equation for this configuration is therefore given by

$$P = q = K_{ox}(wL)\frac{(T_1 - T_0)}{t_{ox}} \tag{2.35}$$

Here K_{ox} is the thermal conductivity of the oxide and (wL) is the surface area of the metal line in contact with the underlying oxide. Hence the relation between power dissipation P, thermal resistance (R_θ) and temperature rise (ΔT) is expressed as

$$\Delta T = (T_1 - T_0) = P \cdot R_\theta \tag{2.36}$$

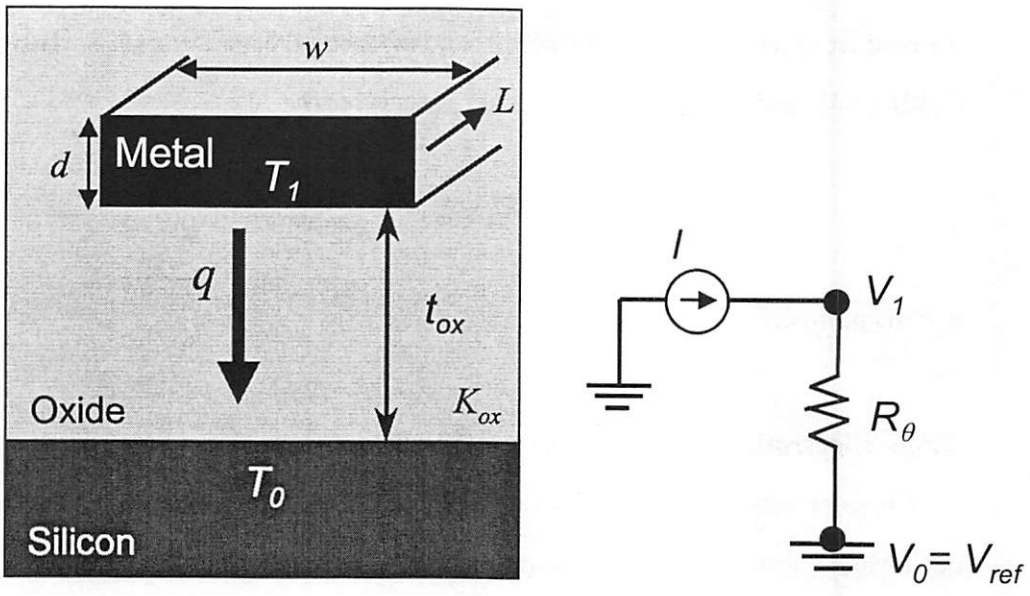
where R_θ is the thermal resistance for this configuration given by

$$R_\theta = \frac{t_{ox}}{K_{ox}(wL)} \tag{2.37}$$

The distributed thermal circuit for the configuration is shown in Figure 2.6. This circuit can be simulated using SPICE [24] and solved for V(t), which is equivalent to solving for temperature.

If the metal line in Figure 2.5 is stressed by an AC or any time varying waveform, capacitive elements must be added to the circuit of Figure 2.6 as shown in Figure 2.7. These capacitances defined in equation (2.33) can be written as

$$C_i = m_i c_i \tag{2.38}$$



a) b)

Figure 2.5 a) Schematic view of an interconnect metal line of length L, width w, and thickness d embedded in oxide. The underlying oxide thickness is t_{ox}. Heat flow (q) is assumed to be 1-D and towards the underlying silicon which acts as the heat sink. The metal line heated by a current is at an average temperature T_1, while the silicon is assumed to be at room temperature (or reference temperature) T_0. b) Equivalent DC lumped thermal circuit. R_θ is the thermal resistance of the interconnect line to the silicon substrate.

where the subscript i stands for the material i.e., metal, oxide, etc. The thermal capacity of the metal, C_m, remains invariant irrespective of the heat flow pattern (1-D, 2-D, or 3-D). Whereas the thermal capacity of the surrounding oxide depends on the heat flow model used. For the simple case of 1-D heat flow (valid for $w >> t_{ox}$) from the metal line to the underlying Si substrate (Figure 2.5), the thermal capacity of the underlying oxide is simply given by

$$C_{ox} = m_{ox}c_{ox} = \delta_{ox}(wLt_{ox})c_{ox} \tag{2.39}$$

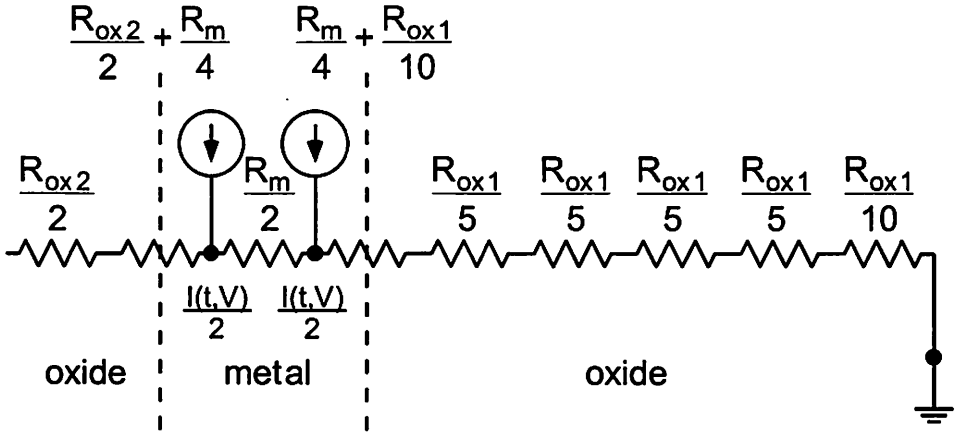


Figure 2.6 Distributed thermal circuit for the metal line of Figure 2.5 under DC (steady-state) conditions. R_{ox2} and R_{ox1} are the effective thermal resistance of the overlying and underlying oxide layers respectively. This circuit can be simulated using SPICE and solved for V(t), which is equivalent to solving for temperature.

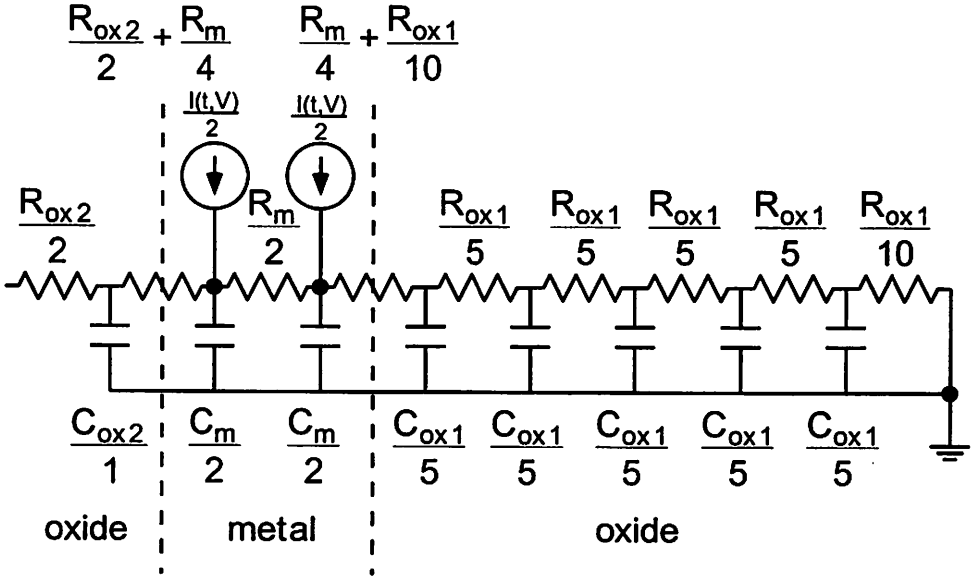


Figure 2.7 Distributed thermal circuit for the metal line of Figure 2.5 under AC (or transient) stress conditions. $R_{ox2}(C_{ox2})$ and $R_{ox1}(C_{ox1})$ are the effective thermal resistance (capacitance) of the overlying and underlying oxide layers respectively. This circuit can be simulated using SPICE and solved for V(t), which is equivalent to solving for temperature.

A simple relation for the thermal time constant of VLSI interconnects can be derived irrespective of the heat flow model employed. For the general case, the effective thermal resistance ($R_{\theta eff}$) and the effective thermal capacity ($C_{\theta eff}$) can be expressed as

$$R_{\theta eff} = \frac{t_{ox}}{K_{ox} S_{eff}}; \quad and \quad C_{\theta eff} = m_{eff} c_{ox} \tag{2.40}$$

Hence the thermal time constant (τ_θ) is given by

$$\tau_\theta = R_{\theta eff} C_{\theta eff} = \frac{t_{ox}}{K_{ox} S_{eff}} \left(\delta_{ox} S_{eff} t_{ox} c_{ox} \right) = \frac{t_{ox}^2}{\kappa} \tag{2.41}$$

Therefore the thermal time constant of interconnect lines depends only on the underlying dielectric thickness and its thermal diffusivity. Hence, for a typical level one metal line lying on an oxide layer of ~1.0 μm thickness, the thermal time constant is of the order of a few microseconds. Also, it can be easily deduced that for higher level interconnect lines the thermal time constants will be larger than those for the lower level lines.

2.5.7 Thermal Circuit for Packaged VLSI Interconnect

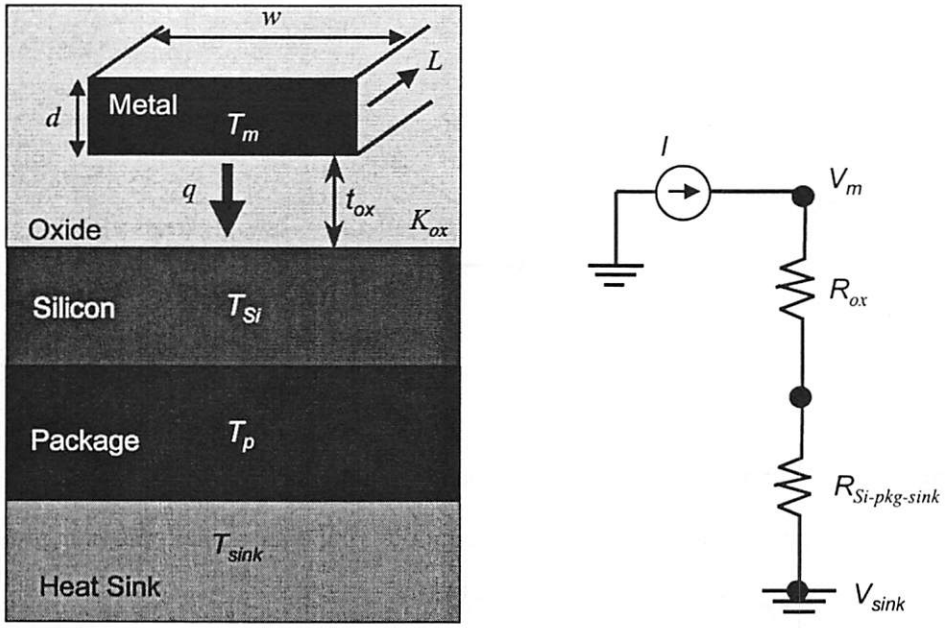
A packaged isolated interconnect line with a heat sink is shown schematically in Figure 2.8. The equivalent DC lumped thermal circuit is also shown. The voltage-current equation for the circuit is given by

$$V_m - V_{sink} = I\left(R_{ox} + R_{Si-pkg-sink} \right) \tag{2.42}$$

The thermal equation (from equation (2.36) and (2.37)) is given by

$$T_m - T_{sink} = j^2 \rho L w d \left(\frac{t_{ox}}{K_{ox} w L} + R_{Si-pkg-sink} \right) \tag{2.43}$$

Therefore by plotting ($T_m - T_{sink}$) vs t_{ox} and extrapolating to $t_{ox} = 0$, $R_{Si\text{-}pkg\text{-}sink}$ can be calculated from the intercepts to the y-axis as shown schematically in Figure 2.9.



a) b)

Figure 2.8 a) Schematic view of a packaged interconnect metal line of length L, width w, and thickness d embedded in oxide. The underlying oxide thickness is t_{ox}. Heat flow (q) is assumed to be 1-D and towards the underlying heat sink. The metal line heated by a current is at an average temperature T_m, while the heat sink is assumed to be at some temperature T_{sink}. b) Equivalent DC lumped thermal circuit. R_{ox} is the thermal resistance of the oxide layer, and $R_{Si\text{-}pkg\text{-}sink}$ is the thermal resistance from the silicon to the sink.

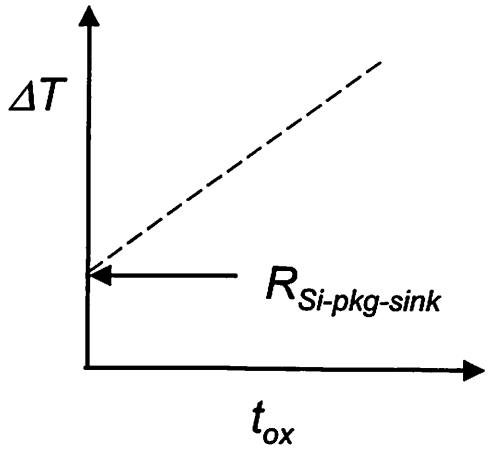


Figure 2.9 Temperature rise of the packaged interconnect line with respect to the heat sink plotted as a function of the underlying oxide thickness can provide the value of $R_{Si\text{-}pkg\text{-}sink}$, from equation (2.43).

The electrothermal analogs defined in this section are used in all the subsequent chapters. It should be kept in mind that heat conduction problems could be approached in several ways. The classical approach is described in great detail in [22], where the heat conduction equation is solved to obtain a general solution, from which a particular solution may then be evolved from the boundary conditions. Another approach proposed by Heisler uses a graphical technique [25].

In addition, the heat conduction equations can also be numerically solved using finite element method. This technique is most rigorous and suitable for analyzing 2-D or 3-D heat conduction problems in complex systems where analytical modeling is very difficult. The finite element approach has been employed to address certain problems in Chapter 3 and Chapter 10.

Chapter 3

Interconnect Heating and Failure under Short-Pulse Stress Conditions

3.1 Introduction

In this chapter self-heating and failure mechanisms of multi-level VLSI interconnect metal lines under short-pulse current stress will be analyzed in detail. The motivation behind studying interconnect heating by short-duration pulses is multifold. Firstly, short-pulse stresses are a superset of ESD type stress conditions for high pulse amplitudes (> 1 A). Secondly, as shown in this research, they can be very effectively used to study high-current and thermal failures in deep sub-micron interconnect structures. This is due to the precise and incremental transfer of thermal energy that can be achieved with short-pulses. Short-pulse stressing is a unique approach for studying the evolution of thermal damage in interconnect structures since continuous high-current stress conditions (DC, AC, or pulsed) always result in catastrophic failure modes, which convey little information on the failure process. Due to the relevance and significance of interconnect heating under short-pulse stress conditions to ESD phenomenon, a brief introduction to ESD is provided in the next section.

3.2 Electrostatic Discharge (ESD) in Integrated Circuits

ESD is a high-current ($I > 1\ A$), nanosecond time-scale ($\Delta t < 200\ ns$) phenomenon that causes irreparable damage to IC devices of all categories [10], [26]. It is now considered a major reliability threat to deep sub-micron IC technologies [27], [28]. ESD events are triggered whenever a "charged human" or a "charged machine" comes in contact with an IC resulting in a transient discharge of static charge through the IC. The equivalent circuit models used to describe these events are known as the "human body model" (HBM) ($I \sim 1\ A$, $\Delta t \sim 100\text{-}200\ ns$) and the "machine model" (MM) ($I \sim 2\text{-}3\ A$, $\Delta t < 100\ ns$) respectively. ESD events could also be triggered due to charge build-up in the chip itself followed by a discharge through internal paths upon contacting the socket. The equivalent circuit for this type of events is known as the "charged device model" (CDM) ($I > 10\ A$, $\Delta t < 1\ ns$).

However, the most common and well-known manifestation of ESD is in the form of HBM where a charge of about 0.6 μC can be induced on a body capacitance of 100 pF, leading to electrostatic potentials of 4 kV or greater. An equivalent circuit for ESD HBM is shown in Figure 3.1. The current waveform for the HBM is shown in Figure 3.2 [29]. The peak current is approximately determined by the ratio of the peak voltage to the 1.5 KΩ resistor representing the human body resistance, where the device impedance is negligible under ESD. This high current pulse generated during an HBM ESD event causes IC failures. The damage is usually thermal in nature. Protection devices at the I/O pins are usually designed to withstand a minimum of 2 kV (or 1.33 A peak current) HBM stress to meet reliability specifications.

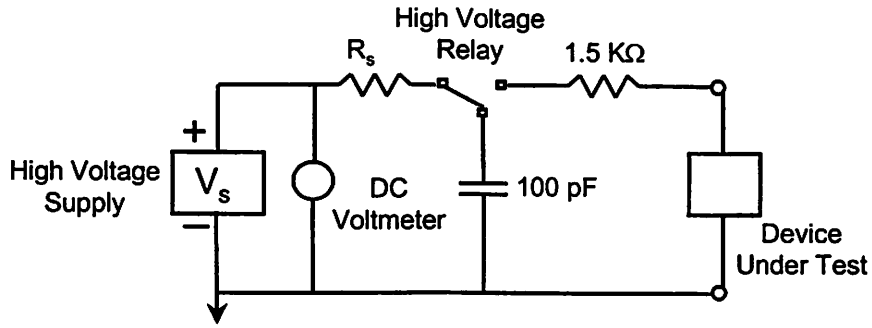


Figure 3.1 Equivalent circuit for the Human Body Model.

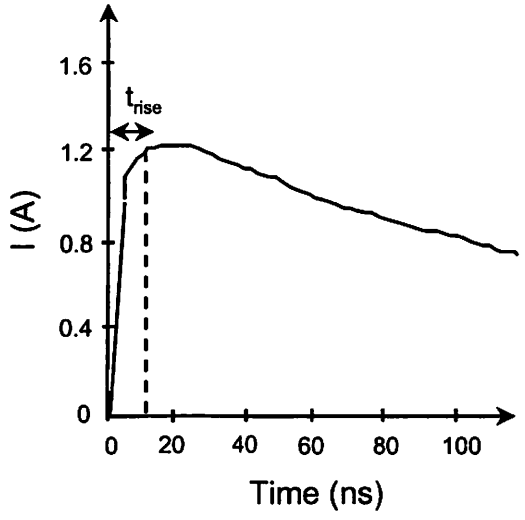


Figure 3.2 Current waveform for Human Body Model. The rise time of the waveform is < 10 ns and the decay time is ~ 150 ns.

3.3 ESD Reliability Due to Interconnect Failure

Over the years there has been much work done in the areas of ESD circuit design and the operation of protection circuit elements such as diodes, MOS and Bipolar transistors, gate oxides etc [26]. However, there is very limited information on the high current (ESD) and self-heating effects in interconnect structures including metal lines, vias, contacts, and silicides.

Protection circuits at the I/O pins are usually designed to minimize damage caused by ESD events. For improving reliability, ESD testing on these circuits is usually done by exposing them to typical discharge pulses to which the IC may be exposed during manufacturing or handling. Metal interconnections in the protection circuits are heavily stressed during these events due to excessive joule heating. The continuous drive to scale down IC devices for achieving higher circuit density and faster speed have reduced the dimensions of both the devices and the metal lines that form the interconnection system between the devices and also to the power lines in the chip. This reduction in line dimension has increased their susceptibility to ESD damage. Interfacing between multiple power supply chips, as well as between multiple power supply blocks within a chip, can also cause high current conditions at the I/O circuitry [12], **Error! Not a valid link.**. Furthermore, short time-scale high-current conditions arise in certain field

programmable gate arrays (FPGAs) devices that employ metal-to-metal voltage programmable links (VPLs) to make permanent connections between logical elements [31], [32], [33]. Thermally accelerated metal failures in ESD protection circuits have been recently reported [34]. There is an increasing need to comprehend ESD effects in metal interconnects in order to provide design guidelines for building robust interconnects and I/O buffers.

3.4 Interconnect Failures Due to High Current Pulses: Prior Literature

Kinsborn et al., [35], reported lifetime measurements of unpassivated Al conductors subjected to "continuous" high current density pulses as a function of the current density and duty cycle. They reported a sharp decrease in time-to-fail for current density above $2x10^7$ A/cm^2. They attributed the failure of the Al lines to a combination of electromigration, temperature cycling and chemical reaction between Al and SiO_2. Pierce [36], proposed a theoretical model for unpassivated metallization burnout under electrical over-stress based on the assumption that failure occurred when the metal reached its melting point due to joule heating. He also assumed adiabatic conditions for pulse duration up to $\sim 1\ \mu s$. Kim and Sachse [37], reported a systematic study of the fracture strength of unpassivated metal lines deposited on window grade quartz substrates for a single shot current pulse. Their experiments showed that the temperature to initiate fracture in Al films was ~ 300 ^{0}C. Maloney [38], [39], used passivated AlSi and AlCu lines and attributed failure to a combination of melting and evaporation. His calculations of temperature rise were again based on the adiabatic assumption. Murguia and Bernstein [40], derived a simple relationship ($j^2 t = 10^8\ A^2s/cm^4$) between the critical current density, j, to cause failure under short-pulse stress, and the pulse width, t, based on the failure temperature of 300 ^{0}C as per reference [37]. They also assumed adiabatic conditions. Detailed simulation results of passivated multilayered interconnect heating under transient conditions by Gui et al. [41], demonstrated the limitations of the model in [40], for pulse widths > $\sim 2\ ns$.

3.5 Sample Fabrication

A quadruple level AlCu(0.5%) metallization system fabricated in a state-of-the-art sub-0.5 μm industrial CMOS process flow was used in this study [42]. The stacking sequence was Si/~1.4-μm PETEOS+BPSGoxide/M1/ILD1/M2/ILD2/M3/ILD3/M4/passivation. Each metal level was isolated, i.e., they were not stacked vertically on top of each other. The interlevel dielectric (ILD) was (~ 1 μm) oxide. The passivation was composed of 1.0 μm thick oxide and nitride layers. Each metallization level was multilayered (see Figure 3.3) with the stacking sequence *TiN (50 nm)/AlCu (0.6 μm)/TiN (150 nm)* excepting metal 4 where the AlCu thickness was doubled to 1.2 μm. All the metal lines were (*W/L = 1000 μm/3 μm*) standard NIST recommended Kelvin structures. A reactive sputtering process was used to deposit the TiN and AlCu layers. The initial 50-nm layer of TiN is used to provide a controlled bottom interface for the AlCu, and it also serves as the barrier layer. The top TiN layer is used as an antireflective coating and to improve electromigration characteristics.

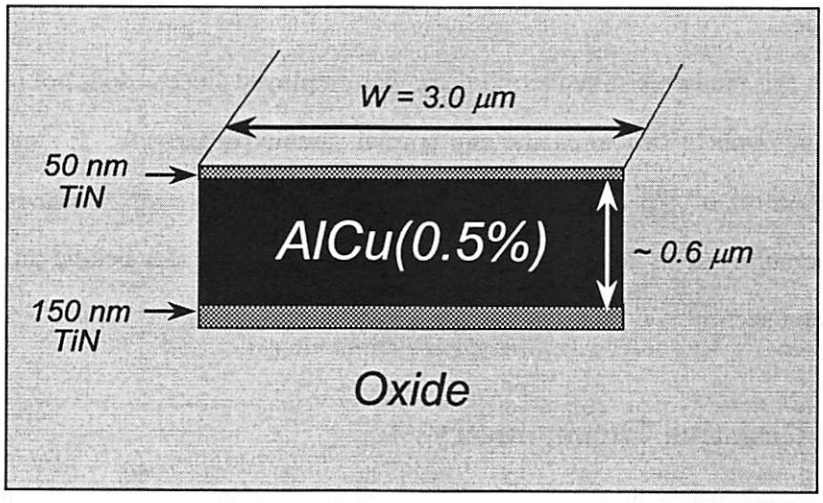


Figure 3.3 Schematic cross-sectional view of the multilayered TiN/AlCu(0.5%)/TiN metallization used in this study.

3.6 Interconnect Thermometry Techniques

In order to characterize self-heating in VLSI interconnects development of suitable thermometry techniques is essential. Thermometry techniques are generally categorized depending on whether they employ electrical or optical signals. Much of the recent work on thermometry of transistors and interconnects has been based on either high-resolution optical diagnostic techniques, or the scanning probe microscope. Optical methods use radiation interaction with the surface of the interconnect and offer good temporal and spatial temperature resolution. An excellent summary of these techniques is provided in [43]. However, most of the optical techniques, such as the laser-reflectance thermometry [44], [45], infrared thermography [46] , [47], fluorescence thermal imaging [48], and liquid-crystal thermometry [49] - [51], are not well suited for VLSI interconnect thermometry. This is due to the fact that VLSI interconnects are buried under layers of dielectric (such as oxide) and passivation materials (such as silicon nitride), which does not allow direct interaction of the radiation with the interconnects without altering the structure or special sample preparation. Electrical thermometry techniques known as scanning thermal microscopy [52] - [54], which are either based on the scanning tunneling microscope (STM) [55], or the atomic force microscope (AFM) [56], also suffer from similar problems. Presence of passivation layers or stacks of dielectric materials that isolate and protect interconnect structures prevent any direct contact between STM or AFM tip and the interconnect surface.

In this study resistive thermometry was employed since it does not require thermal contact to the interconnect structures and any special sample preparation. It simply uses the temperature dependence of the resistivity of the interconnect material to provide a spatially averaged temperature rise in the interconnect [42], [57]. The basic idea behind this technique is outlined in the next section.

3.7 Resistive Thermometry

The electrical conductivity of most metals is dominated at room temperature (300 K) by collisions of the conduction electrons with lattice phonons and at liquid helium temperature (4 K)

by collisions with impurity atoms and mechanical imperfections in the lattice. The net resistivity (ρ) is temperature dependent [21], and is given by the Matthiessen's rule:

$$\rho(T) = \rho_L(T) + \rho_i \tag{3.1}$$

where ρ_L is the lattice resistivity caused by the thermal phonons and is temperature dependent, and ρ_i is the intrinsic resistivity caused by scattering of the electron waves by static defects that disturb the periodicity of the lattice, and is often independent of temperature. At T = 0 K, ρ_L vanishes and $\rho_i(0)$, known as the residual resistivity can be extrapolated. The lattice resistivity $\rho_L(T) = \rho(T) - \rho_i(0)$, is the same for different specimens of a metal, even though $\rho_i(0)$ may itself vary widely. A linear relation gives the temperature dependence of the net resistivity:

$$\rho(T) = \rho_0(1 + \alpha\{T - T_0\}) \tag{3.2}$$

where ρ_0 is the resistivity at some reference temperature T_0. The parameter α is known as the temperature coefficient of resistance (*TCR*) defined by,

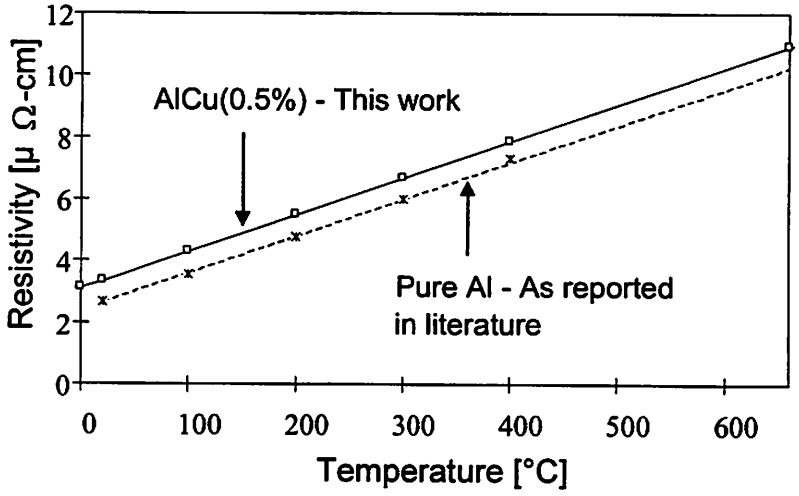


Figure 3.4 Temperature dependence of resistivity remains linear up to the melt temperatures for pure Al and for AlCu(0.5%).

$$\alpha = \frac{\rho(T) - \rho_0}{\rho_0} \frac{1}{(T - T_0)} \tag{3.3}$$

The linear dependence of resistivity on temperature is shown in Figure 3.4 for pure Al [58], and for AlCu(0.5%) based on this work. Substituting equation (3.1) in (3.3) it follows that

$$\alpha = \left[\frac{\rho_L(T) - \rho_L(T_0)}{\rho_L(T_0) + \rho_i(0)} \right] \frac{1}{(T - T_0)} \tag{3.4}$$

The presence of the residual resistivity $\rho_i(0)$ in the denominator of equation (3.4) causes slight variation in the extracted *TCR* values for different specimens of the same metal. The temperature dependence of the resistivity causes the resistance of metal interconnects to be temperature dependent. This is the basis for resistive thermometry.

3.8 Transient Resistive Thermometry (TRT) Technique

In order to characterize self-heating of the interconnect metal lines under short-pulse stress conditions, transient temperature measurements in nanosecond timescale is necessary. The transient resistive thermometry has been developed for this purpose [42], [57].

3.8.1 Basic Concept

The basic idea behind this technique comes from resistive thermometry. Consider a current pulse of duration (Δt) and amplitude I going through a metal line as shown in Figure 3.5. If this current pulse causes self-heating of the line resulting in a temperature rise (ΔT) above the ambient (reference) temperature T_0, then the resistance of the line will increase because of the temperature sensitivity of the metal resistivity as explained in the last section. Also, the resistance rise will be linear with respect to temperature due to the linear dependence of resistivity on temperature as shown in equation 3.2. The line resistance can be calculated from the instantaneous voltage pulse that develops across the line during the pulsing event. Since the magnitude of the current pulse remains constant and the pulse duration is very small, the voltage across the line rises linearly with time, as discussed in section 3.9.3. The idea is illustrated in Figure 3.5. If R_0 and R_f are the

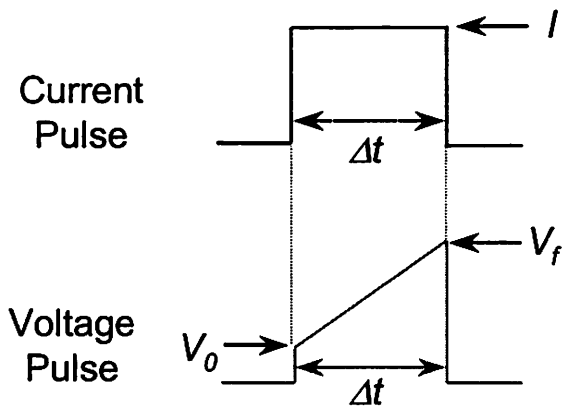


Figure 3.5 The current and the instantaneous voltage pulse across the metal line.

initial and final resistances of the metal line at the beginning and at the end of the pulse respectively then,

$$R_0 = \frac{V_0}{I} \quad and \quad R_f = \frac{V_f}{I} \tag{3.5}$$

and since the TCR for this line can be extracted from a separate experiment that involves measurement of the temperature dependence of the line resistance using a small DC current (or resistance thermometry), the transient temperature rise can be calculated using

$$\Delta T = (T - T_0) = \frac{1}{TCR}\left(\frac{R_f - R_0}{R_0}\right) \tag{3.6}$$

3.8.2 System Design

In this section the experimental set up needed for transient resistive thermometry (TRT) of interconnects is explained in detail. A block diagram in Figure 3.6 shows the overall system. A temperature controlled probe station is required to carry out all the temperature dependent measurements and the subsequent short-pulse measurements. A high frequency probe card and coaxial cables are used to minimize electrical noise. A transmission line is employed to generated constant current pulses of high magnitudes (> 2 A) [23]. This is necessary because it is

very difficult for commercially available pulse generators to provide controlled square pulses into low resistance (~ 20 Ω) devices. The transmission line pulsing scheme is very useful for generating ESD type current pulses. The high voltage power supply is capable of charging the transmission line up to 4000 volts, provided the 1 MΩ resistor is robust enough. Voltage used to charge the transmission line can be varied to generate current pulses with different amplitudes. Two switch boxes are employed to facilitate charging and discharging of the transmission line.

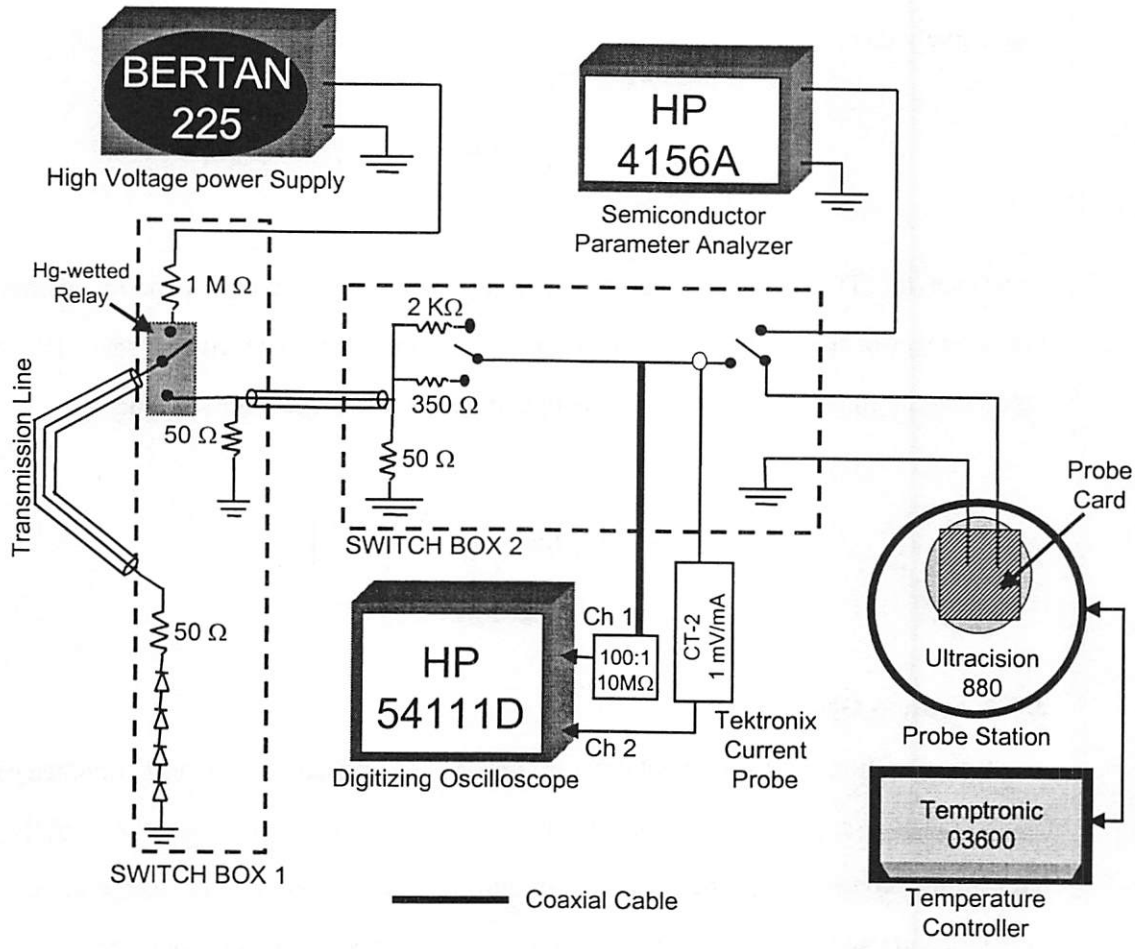


Figure 3.6 Block diagram illustrating the transient resistive thermometry system.

The resistors in series with the device under test (DUT) (2 KΩ and 350 Ω inside switch box 2) can be chosen depending on the resistance of the line being tested, to maintain constant current pulse. A mercury-wetted relay is housed inside switch box 1 to prevent voltage ringing. The probe head (or switch box 2 in Figure 3.6), which houses the unit for the voltage to current conversion, also includes a Tektronix CT-2 current probe and a contact for a oscilloscope probe for voltage measurement. The CT-2 current probe consists of a current transformer and an interconnecting cable.

The current transformers have a small hole through which a current carrying conductor is passed during circuit assembly. The current probe converts the current pulse to a voltage pulse (1mA $\leftrightarrow$ 1mV) and can handle up to 36 A of maximum pulse current. It has a bandwidth of 1 GHz, and hence for pulses with rise times > 5 ns, this is more than adequate for accurate waveform measurements. Inside the probe head an additional contact is also available for a semiconductor parameter analyzer (or curve tracer), to enable frequent monitoring of degradation characteristics of the DUT after each pulse stress. The probe head allows great flexibility to the system since it can be easily connected to any wafer-level probe card (or even Micromanipulators), and can be placed within inches from the DUT. Voltage pulse measurements are more difficult due to the inductances in the ground loop of the measuring circuit. The probe head has about 5-10 cms in the ground loop, which is an effective maximum inductance of 50 nH (since L = 5 nH/cm). This results in a voltage drop across the oscilloscope probe during the fast dI/dt of the current pulse. Hence, there is some voltage ringing measured across a short circuit, during the initial 5 ns of the pulses. The voltage drop due to these stray inductances can be estimated from the relation

$$V_L = L\frac{dI}{dt} \tag{3.7}$$

where L is the inductance in the ground loop (~ 50 nH). Therefore, for a 100 ns current pulse of amplitude 1 A, and a rise time of ~ 10 ns, V_L ~ 5V. However, since final voltage rise at the end of the current pulse is used for the temperature calculations, this inductive voltage drop during the first few nanoseconds does not influence the measurements. To capture these extremely fast

pulses a digital storage oscilloscope (HP 54111D) with a bandwidth of 250 MHz and a sampling rate of 1 GS/s is used, which is adequate for the intended measurements.

Figure 3.7 shows a simplified schematic circuit of the TRT system. The transmission line is charged by the high voltage supply. The distributed *L-C* elements of the transmission line produce a pulsed voltage source with impedance equal to the line impedance ($= \sqrt{L/C}$) resulting in a simple square waveform [23]. In order to suppress any reflections from low resistance devices (< 50 Ω) a polarized matched load (*R1 + Diodes*) is connected to one end of the transmission line. The front end of the transmission line is also terminated with a ($R2 = 50\ \Omega$) matched load to yield a rectangular pulse and to prevent unwanted reflections. The resistor *R3* is the sensing resistor, and it converts the discharge voltage to a current stress across the interconnect, and since the value of *R3* is chosen such that it is much greater than the resistance of the device under test (R_{DUT}), a constant current pulse is ensured. The relation between the pulse duration (Δt) and transmission line length (l) is given by

$$\Delta t = \frac{2l}{c} \tag{3.8}$$

where c is the propagation velocity of electromagnetic waves (= 20 cm/nsec). Hence in order to generate a 100 ns pulse the transmission line must be 10 m in length. Hence, different cable lengths can be used for generating various pulse widths. Note that the idea of charging and discharging the transmission line is similar to that of charging and discharging a capacitor. However, a charged capacitor and the device under test produce exponential R-C waveforms that are often disturbed by parasitic elements (especially if R_{DUT} is low).

In the TRT system the voltage from the transmission line is delivered to the DUT via another coaxial cable (see Figure 3.6) which is around 1 m long. This coaxial cable only introduces a small delay for the transmission line pulse. Voltage to current conversion through *R3* takes place very close to the DUT, thereby preventing capacitance loading effects of the coaxial system from degrading the current pulse.

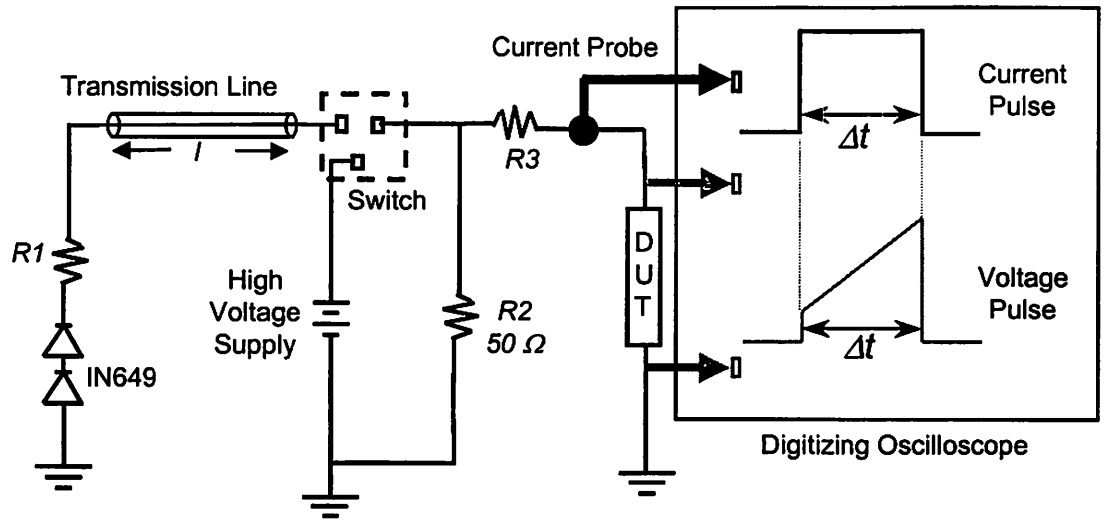


Figure 3.7 A simplified schematic circuit diagram illustrating the transient resistive thermometry (TRT) technique employing a transmission line to generate short-duration constant current pulses.

Finally, since resistance is being calculated based on very short duration current pulses it is important to check if the extracted line resistances using pulsed measurements match up to the DC line resistances in order to ensure that skin effect [59], is not influencing the data. Skin effect is normally observed at high frequencies at which the current gets confined almost entirely to a very thin sheet at the surface of the conductor. The thickness of this sheet, known as the skin depth, determines the effective cross sectional area of the conductor and its resistance. The resistance values were found to be identical under pulsed and DC stress and also theoretically checked using the formula for skin depth (δ) given by,

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{3.9}$$

Here the frequency $f = \omega/2\pi$, and μ and σ are the permeability and conductivity of the interconnect material respectively. It was found that for skin effect to start impacting the line resistance, the thickness of aluminum lines must be ≥ 3 μm for 1 GHz (1 ns) pulses. Since the dimensions of the AlCu lines used in this study were ≤ 3 μm, and the smallest pulse duration was 50 ns, skin effect is not expected to impact the measurements.

3.8.3 Equivalence of the Constant Current Square Pulse and ESD

Typical HBM and MM ESD events actually involve a double exponential or a sinusoidal waveform respectively [29]. Since the constant current square pulse is used in this study to propose design guidelines for ESD protection circuit and I/O buffer interconnects, it is instructive to note their equivalence to actual ESD waveforms, as shown in [60]. Hence, the failure thresholds, failure mechanisms, and the electrical response of metal interconnects under high constant current pulse stress can be correlated to the ESD stress conditions. Furthermore, the constant current pulse technique increases the reproducibility of the ESD stress conditions.

3.9 Interconnect Heating and Failure under Short-Pulse Stress

3.9.1 Comparison with Electromigration

Before proceeding with the characterization results of interconnect heating and failure under short-pulse stress conditions, it is important to comprehend the fundamental differences between interconnect failures under these stress conditions and those due to electromigration. Table 3.1 lists the typical stress conditions and the mechanisms responsible for interconnect failures under electromigration and short-pulse events [61]. Interconnect failures due to electromigration are always diffusion driven, even under highly accelerated wafer-level electromigration test conditions. Interconnect temperatures and current densities needed for electromigration are much lower than those that fail under short-pulse stress conditions, as will be shown in this section. Also, under short-pulse stress the failure mechanism is simply metal fusion, which is thermally driven.

Furthermore, for electromigration failures, stress has to be applied over hundreds of minutes or hundreds of hours depending on the magnitude of the stress conditions. In comparison, short-pulse failures are almost instantaneous, with failures occurring within a time scale, that is much smaller than the thermal time constant of the interconnects which is of the order of a few microseconds.

	J (A/cm^2)	Temperature (^0C)	Mechanism	Time scale
EM Failure				
Field Conditions:	$4 - 6 \times 10^5$	~ 85 - 100	diffusion	$t >> \tau_0$ steady state
Package Level:	$1 - 3 \times 10^6$	~ 100 - 200	diffusion	
Wafer Level:	$0.5 - 1 \times 10^7$	~ 150 - 300	diffusion	
High-Current Short-Pulse Failure	$> 10^7$	~ 1000	fusing	$t << \tau_0$ non-steady state

τ_0: thermal time constant (several μs)

Table 3.1 Typical stress conditions and failure mechanisms for electromigration and short-pulse (ESD) driven failures in AlCu.

3.9.2 Temperature Calibration and Steady-State Measurements

Initially the temperature dependence of resistance of AlCu lines was experimentally measured using a small DC current. This data was used to extract the value of TCR for the AlCu lines using equation 3.3. The TCR was found to be $0.0031 \pm 0.0001\ ^0C^{-1}$. For self-heating extraction, lines were stressed with increasing amounts of current (DC) at room temperature and the rise in resistance due to self-heating was converted to a spatially averaged temperature rise using the temperature dependent resistance measurements. While calculating this temperature rise, the entire length of the line is assumed to be uniformly heated due to fact that its length is much larger than the characteristic diffusion length (~ 25-50 μm for Al [62]) for heat flow along the length of the line. Furthermore, the temperature gradient (ΔT) across the thickness (d_{Al}) of the metal lead, which can be expressed as

$$\Delta T = \frac{j^2 \rho_{Al} d_{Al}^2}{2 K_{Al}} \tag{3.10}$$

is negligible due to the high thermal conductivity (K_{Al}) of Al. The DC self-heating results for the quadruple level (TiN/AlCu(0.5%)/TiN) interconnects under study are shown in Figure 3.8 [42]. The thermal impedance (R_θ) increased from 150 ^{0}C/W for metal 1 to about 250 ^{0}C/W for metal 4.

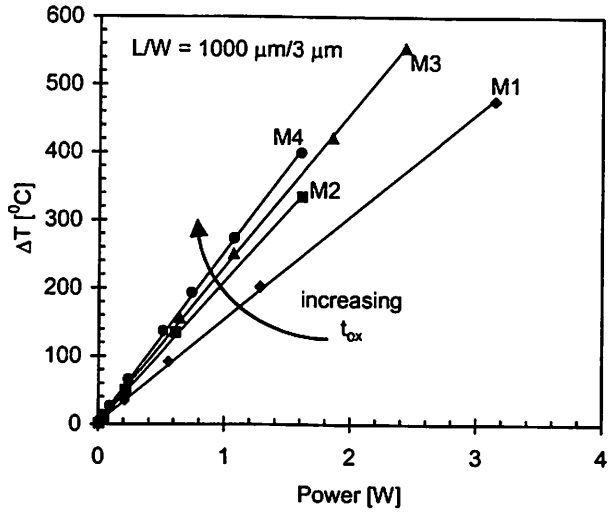


Figure 3.8 Self-heating of the quadruple level metal system under DC (steady-state) conditions. The effective thermal impedance of the interconnect to the Si substrate can be extracted from equation 2.37 ($\Delta T = P \cdot R_\theta$).

This is due to the increasing underlying oxide thickness (t_{ox}) for higher metal levels. Thus it can be observed that under DC (steady-state) conditions the net underlying oxide thickness has a strong influence on the temperature rise as expected from equation 2.37.

3.9.3 Pulsed Measurements

Identical structures from each metal level were subjected to current pulses of different widths (50 ns - 1 µs) and amplitudes using the TRT system described in Figure 3.6. The thermal time constants for the metal lines can be estimated using equation 2.41 and are of the order of several microseconds. The time dependence of the resistance and temperature of a level 1 line during a 200 ns constant current pulse stress is shown in Figure 3.9. It can be observed that the TRT technique offers excellent temporal resolution of temperature rise due to self-heating under short-pulse stress conditions. The resistance values are obtained from the voltage pulse across the interconnect for a constant current pulse stress. Their time dependence can be observed to be linear. Using the temperature dependence of resistance calibrated earlier using DC measurements (see Figure 3.8), the time dependence of the spatially averaged temperature of the metal line during the pulse can be calculated and is shown along the second y-axis.

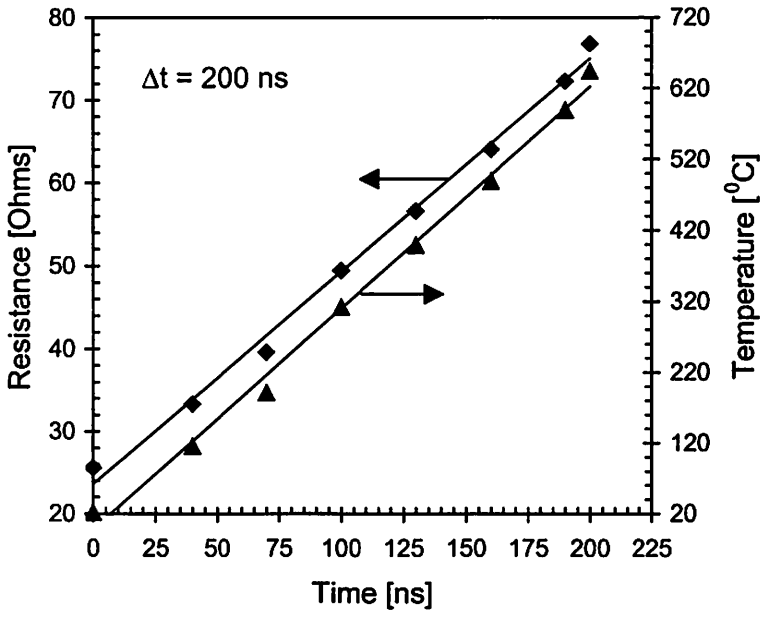


Figure 3.9 The time dependence of resistance and temperature of a 3 μm X 1000 μm AlCu line during a 200 ns constant current pulse stress obtained using the TRT technique.

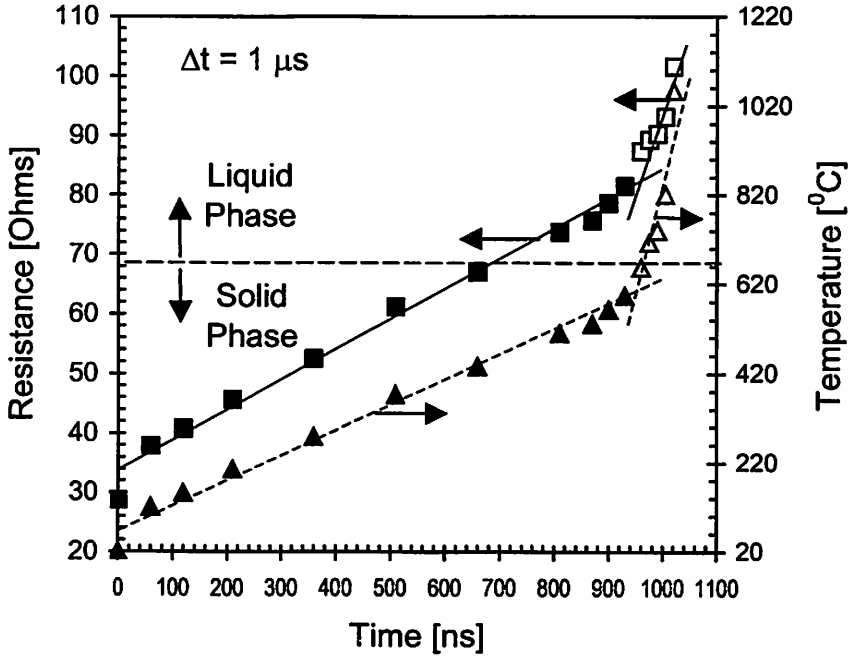


Figure 3.10 The time dependence of resistance and temperature of a 3 μm X 1000 μm AlCu line during a 1 μs constant current pulse stress obtained using the TRT technique. The sharp rise in resistance beyond ~ 950 ns indicates a phase change.

For current pulses that heat the metal lines past their melting point, the resistance rises very rapidly with time, due to the molten AlCu. This can be seen in Figure 3.10 for pulse duration of 1 μs that caused temperature rise of ~ 1000 ^{0}C. The temperature rise after the melting temperature (~ 660 ^{0}C) was calculated using the temperature dependence of resistivity of liquid Al as shown in Figure 3.11 [58]. It can be observed that the resistivity in the liquid phase also varies linearly with temperature like the resistivity of solid Al. However the resistivity of liquid Al is an order of magnitude larger than that of solid Al. As a result of this the resistance of the line increases very sharply after ~ 660 ^{0}C as seen in Figure 3.10.

In Figure 3.12 the maximum resistance and temperature rise above initial room temperature value, ($\Delta R_{max}/R_0 = \gamma$, where, $\Delta R_{max} = R_f - R_0$), for each metal level has been plotted against the current density, for a 200 ns pulse [57]. It can be observed that metal 1, 2, and 3, lines with equal dimensions (and therefore equal thermal capacity) have almost identical resistance (or temperature) rise unlike the DC case shown in Figure 3.8. This indicates that under short-pulse stress, the total underlying thickness of the insulator does not influence the temperature rise of the metal lines.

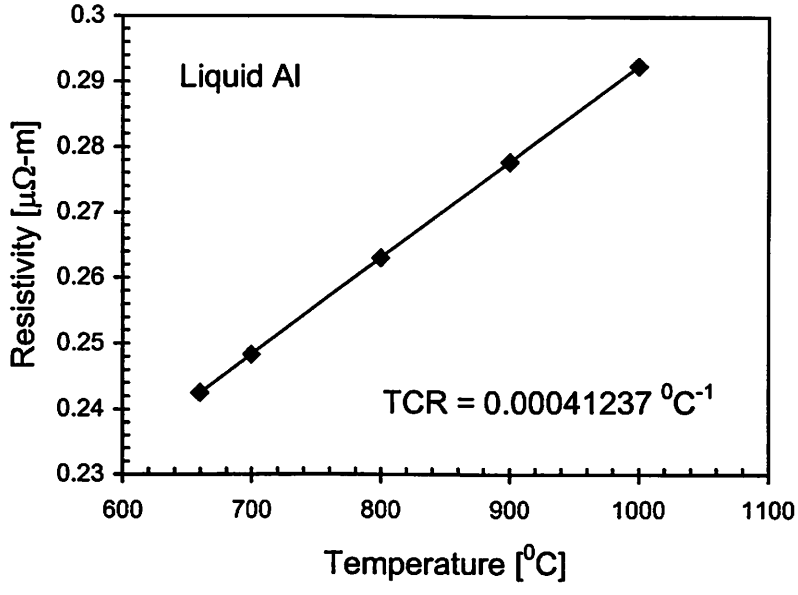


Figure 3.11 Temperature dependence of resistivity of liquid aluminum showing a linear behavior.

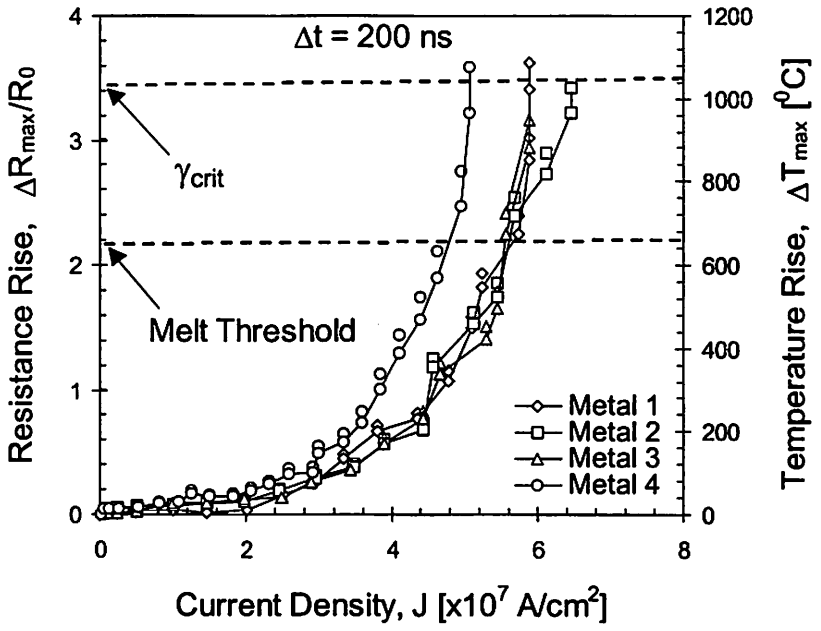


Figure 3.12 Maximum instantaneous resistance and temperature rise of the metal lines stressed by a 200 ns pulse, plotted against the current density *J*. The last data point on each curve indicates the final open circuit failure.

This can be understood from the fact that the metal lines do not reach a steady state under there short-pulse stress conditions since the heating time is much smaller than the thermal time constants of the metal lines. Metal 4, with a thicker AlCu shows higher self-heating due to its smaller surface area to volume ratio. Also, for all the metal levels γ rises superlinearly with *J*, the current density, and they all fail when γ goes beyond a critical value (γ_{crit}). Similar joule heating behavior was observed for the other pulse widths. The values of γ_{crit} were independent of the pulse width and the metal level.

3.10 Finite Element Simulations

Finite element (FE) simulations have also been conducted to determine the temperature rise of the metal lines during the short-pulse stress conditions [63]. The finite element technique numerically solves the 2-D (or 3-D) transient heat equation. The simulations are also intended to check whether or not the metal lines may actually melt during these pulses since this information

is important to understand the failure mechanisms of the metal lines, which are outlined in the last section.

3.10.1 Simulator Calibration

A 2-D FE model was set up to describe the cross section in detail as shown in Figure 3.13. In the FE model it is possible to select any metallization level to be included in the actual simulations. The test lines, surrounding dielectric material, and their thickness were selected to match the experimental structure. Hence, the metal lines were 3 μm/1000 μm and multilayered with the stacking sequence TiN/AlCu(0.5%)/TiN. The material data for the models were mostly taken from the literature. The measured resistance rise under various DC loads (Figure 3.8) was used to calibrate the simulator. One data point per level 1 and level 4 was chosen for this calibration of the finite element models with respect to thermal conductivity of the surrounding dielectric and the heat film coefficient between passivation and air. Figure 3.14 depicts the excellent agreement between the experimental and simulation results. Deviations are randomly distributed and are within 5 %.

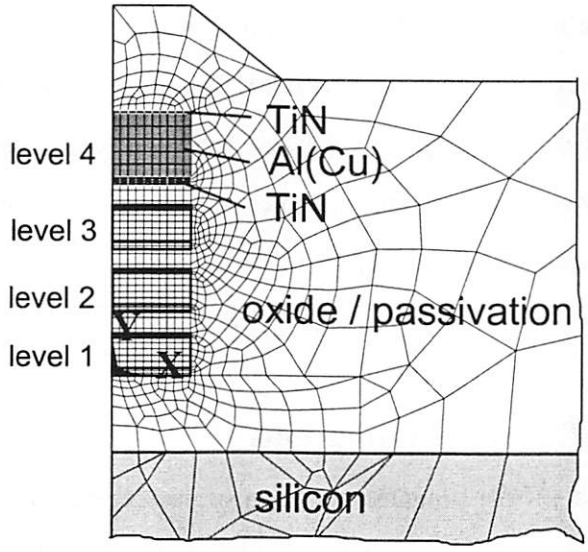


Figure 3.13 2-D finite element model showing the generated mesh for the quadruple level test structure used in this study. W = 3 μm for all levels, example shown for metal 4.

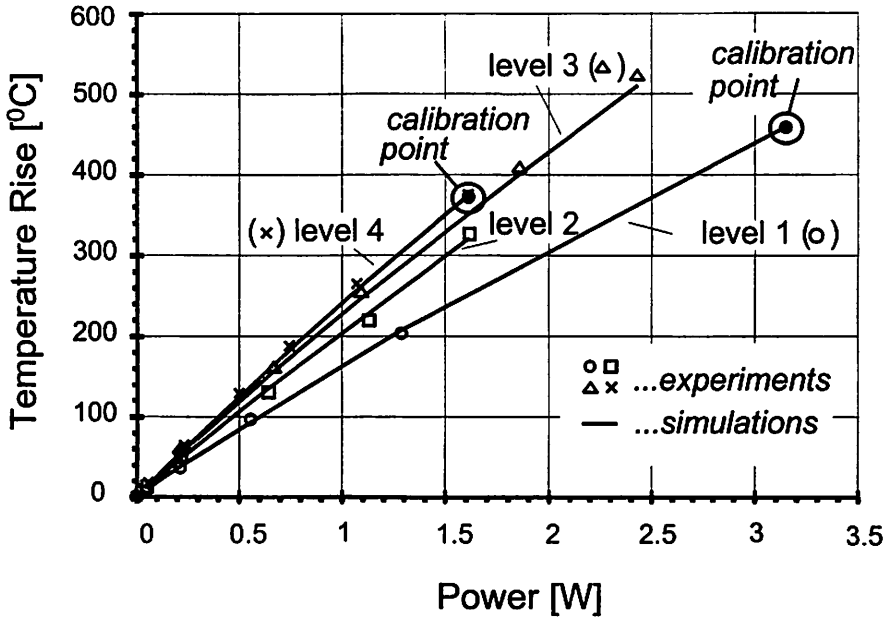


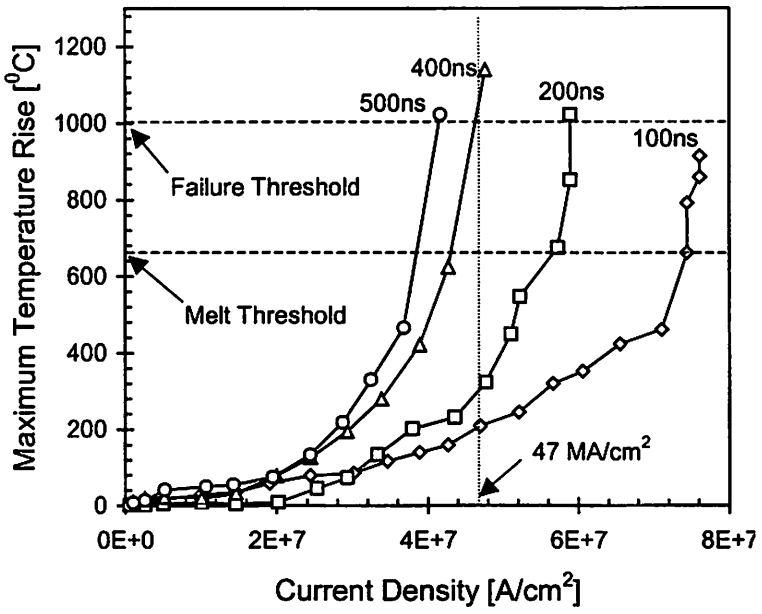
Figure 3.14 Calibration of the finite element model using DC self-heating data for the test structures. Excellent agreement between all experimental data and simulation results proves the accuracy of the model.

3.10.2 Simulation of Transient Interconnect Heating

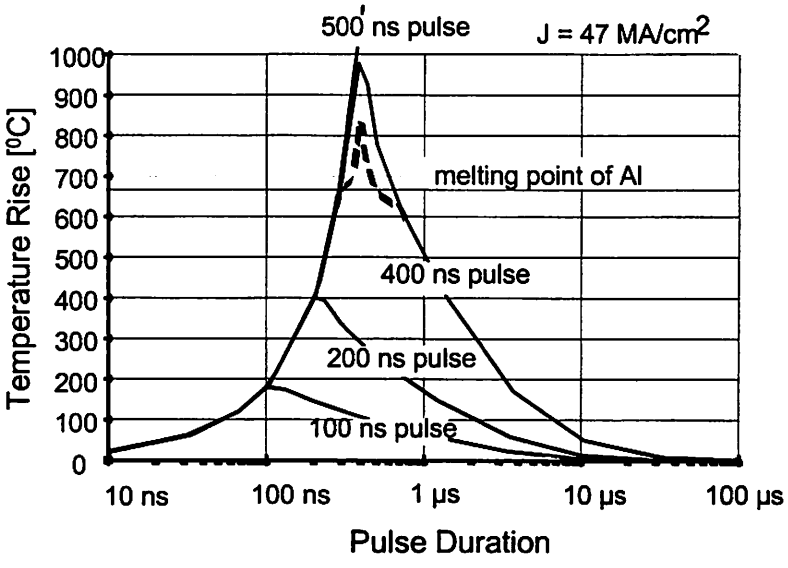
Applying the calibrated models, a level 1 metal line was selected for transient simulations. For comparison with experiments, the effect of pulse width on the self-heating behavior in metal 1 was obtained as shown in Figure 3.15a. It can be observed that the critical current density J_{crit} decreases with increasing pulse duration due to the increasing pulse energy with time. Furthermore, it can be observed that for a current density of *47 MA/cm²*, the maximum temperature rise is ~ 200 °C and ~ 1000 °C, for 100 ns and 400 ns pulses respectively.

The simulated temperatures during 100/200/400/500 ns pulses of amplitude 850 mA (J = *47 MA/cm²*) is shown in Figure 3.15b. The simulation results show good agreement with the experimental results of Figure 3.15a. They support the temperature measurements by the TRT technique, and most importantly shows that metal lines can indeed be heated beyond their melting point under short-pulse stress conditions. Furthermore, the simulation results confirm the

experimentally obtained failure temperature of 1000 ^{0}C as shown by the 400 ns and 500 ns curves. Even when the rapid increase in the line's electrical resistivity during melting was not considered (dotted curve in Figure 3.15b), the temperature increased well beyond 660 ^{0}C. In this case the latent heat of fusion caused delays in temperature evolution at the melting point. The 500 ns pulse (shown only up to 1000 ^{0}C in Figure 3.15b) led to temperatures above 1500 ^{0}C for the chosen current density. This also is consistent with the experimental self-heating curve for 500 ns shown in Figure 3.15a, where it never crosses the dotted line marking the chosen current density for simulations. The simulated value of the maximum temperature rise simply indicates that the line would have been heated to 1500 ^{0}C at this current density. But in reality it would suffer open circuit failure once the temperature reached 1000 ^{0}C as seen from the experimental curve. A recent study performed by a group of industrial researchers have also confirmed the critical temperature rise during transient heating of passivated TiN/AlCu/TiN lines by a square current pulse using RC network simulations [64].



a)



b)

Figure 3.15 a) Experimental data summarizing the effect of pulse width on heating characteristics. The vertical dotted line shows the current density value ($47MA/cm^2$) at which the transient heating during short pulses were simulated. b) Simulated temperature rise under transient currents.

3.11 Interconnect Heating and Failure Model under Short-Pulse Stress Conditions

In this section a general analytical model for interconnect heating under short-pulse stress conditions will be formulated. The model uses the open circuit failure temperature rise in passivated TiN/AlCu/TiN mealization obtained from experimental data, but the concept behind the model can also be applied to Cu interconnects.

Since it was experimentally observed that the resistance of the metal lines increases linearly with time during all the pulsing events (also see Appendix A for analytical reasoning), the input pulse energy can be calculated using the equation [42], [57],

$$E(t) = \int_0^{\Delta t} I \cdot V dt = I^2 \int_0^{\Delta t} \left(R_0 + t \frac{R_f - R_0}{\Delta t} \right) dt = \frac{1}{2} I^2 \cdot \Delta t \cdot \left(R_0 + R_f \right) \tag{3.11}$$

Here, I = current through the metal line, V = voltage developed across the line, Δt = pulse width, R_0 = initial resistance of the unstressed line and R_f = maximum resistance of line reached during the pulsing event. Since the linear temperature dependence of resistance is given by

$$R_f = R_0 (1 + TCR\, \Delta T) \tag{3.12}$$

where ΔT is the temperature rise, and the resistance rise is defined as,

$$\gamma = \frac{\left(R_f - R_0 \right)}{R_0} = TCR\, \Delta T \tag{3.13}$$

it follows that,

$$R_0 + R_f = R_0 (2 + \gamma) \tag{3.14}$$

The thermal capacity (C_θ) of the heated line is defined by

$$E(t) = C_\theta \cdot \Delta T = C_\theta \cdot \frac{\gamma}{TCR} \tag{3.15}$$

can be obtained from the self-heating data. By plotting γ vs pulse energy, the thermal capacity can be extracted from the slopes of γ vs energy curves shown in Figure 3.16 for a 200 ns pulse.

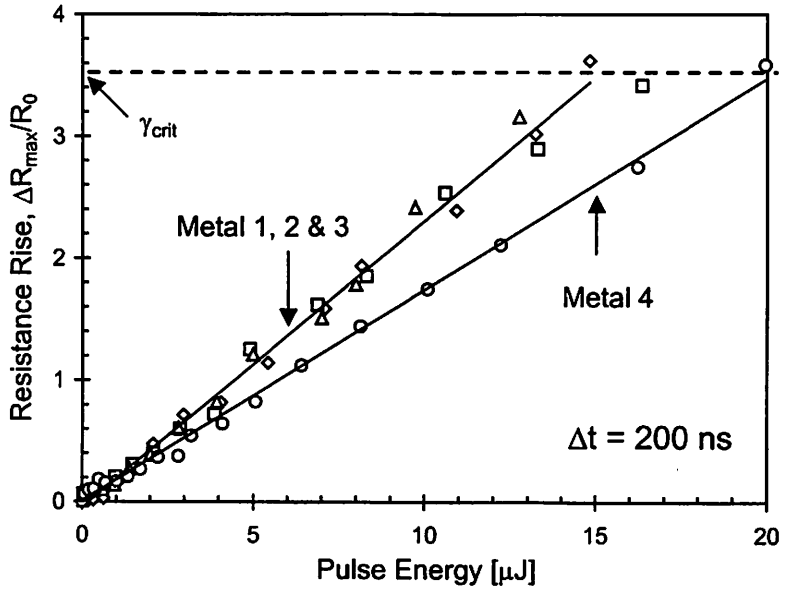


Figure 3.16 Resistance rise (γ) vs pulse energy for various metal lines stressed by a 200 ns pulse.

Difference in slopes between metal 1, 2, 3, and metal 4 is due to the bigger thermal capacity of metal 4. The thermal capacity of metal 1, 2 and 3 from Figure 3.16 is 15.09×10^{-9} J/^{0}C. This is larger than the thermal capacity of the TiN/AlCu/TiN stack ($C_m = 8.25 \times 10^{-9}$ J/^{0}C). The difference between the theoretically and experimentally determined values of the thermal capacity suggests that a thin sheath of oxide around the metal is heated during the short pulse stress.

The thermal capacity of the oxide sheath ($C_{oxide\text{-}sheath}$) can be calculated from

$$C_\theta - C_m = C_{oxide-sheath} \tag{3.16}$$

where,

$$C_{oxide-sheath} = m_{ox}\, c_{ox} = \delta_{ox}\, c_{ox}\, V_{ox} \tag{3.17}$$

Here m_{ox} and V_{ox} are the mass and volume of the heated oxide-sheath respectively, and c_{ox} and δ_{ox} are the heat capacity and density of the oxide respectively. Assuming a rectangular oxide sheath as shown in Figure 3.17, the volume of the oxide can be expressed as,

$$V_{ox} = \left[4Ld_{ox}^2 + 2L(W + d_s)d_{ox}\right] \tag{3.18}$$

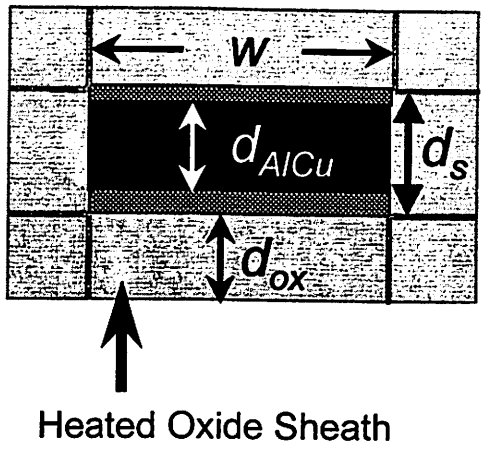


Figure 3.17 Schematic cross section of Metal line and the rectangular heated oxide sheath.

where L, W, and d_s are the length, width and thickness of the metal stack, and d_{ox} is the thickness of the heated oxide sheath. Hence, differences between C_θ and C_m can be used to calculate the thickness of the oxide sheath. From heat diffusion theory the thickness of the oxide sheath is expected to be proportional to the square root of the pulse width [22], [65], that is

$$d_{ox} = [a_d \, \Delta t]^{1/2} \tag{3.19}$$

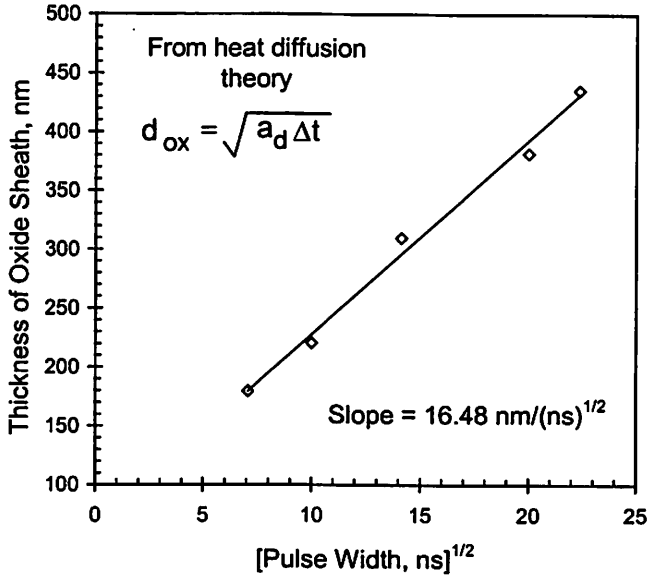


Figure 3.18 Pulse width dependence of the oxide sheath thickness.

This is confirmed by the linear dependence of d_{ox} with $\Delta t^{1/2}$, as shown in Figure 3.18. It can be observed that the thickness of the heated oxide sheath around the metal lines increases with increasing pulse width (Δt). The effective heat diffusivity, a_d, can be extracted from this graph, and was found to be ~ 2.73 x10^{-3} cm^2/s. This value is in rough agreement with those reported for deposited thin silicon dioxide films at high temperatures [66] - [68]. The TRT technique can therefore be employed to measure the effective thermal conductivity of various inter-layer dielectric films used in DSM interconnect processes.

A simple expression relating the critical current density to cause open circuit failure with the pulse duration can be derived as follows [57]. Consider a critical pulse of magnitude I_{crit}. Let R_{crit} be the resistance of the line at the end of the critical pulse and by assuming a linear dependence of temperature from R_0 to R_{crit}, the critical resistance rise γ_{crit}, is defined as,

$$\gamma_{crit} = \frac{(R_{crit} - R_0)}{R_0} = TCR\,\Delta T_{crit} \tag{3.20}$$

Then the critical energy E_{crit}, causing open circuit failure can be approximately expressed as,

$$E_{crit} \approx \frac{1}{2} I_{crit}^2\,\Delta t\,(R_0 + R_{crit}) \approx \frac{1}{2} I_{crit}^2\,\Delta t\,R_0(2 + \gamma_{crit}) \tag{3.21}$$

Equation (3.21) slightly overestimates the critical energy but it makes the analysis simpler and provides a conservative estimate for the maximum allowable current (I_{crit}), leading to safe design guidelines (see Appendix A).

During the critical pulse, the energy absorbed by the aluminum metal, barrier/capping TiN layers and the surrounding oxide sheath, E_s, can be expressed as,

$$E_s = \left[\sum_{i,j,k} mc\right]\Delta T_{crit} + m_i L_{fi} \tag{3.22}$$

Here, m is mass, c is specific heat, and L_{fi} is the latent heat of fusion of aluminum. Subscripts i, j, and k are for AlCu, TiN, and oxide respectively. While using equation (3.22) the specific heats of AlCu, TiN and oxide are taken to be independent of temperature due to their small increase in value with temperature [69]. A mean value was therefore used in all the calculations.

Furthermore, since the thermal diffusion lengths in metals and oxide are much larger than the thickness of these materials being heated, they can be assumed to be at the temperature ΔT_{crit} above the reference temperature at the end of the pulse. Equation (3.21) can be expressed as

$$E_{crit} \approx \frac{1}{2} I_{crit}^2 \, \Delta t \frac{\rho_i}{L_i W_i} (2 + \gamma_{crit}) \tag{3.23}$$

where subscript i is used for AlCu. Furthermore, equation (3.22) can be written as,

$$E_s = \left[m_i c_i + m_j c_j + m_k c_k \right] \gamma_{crit} (TCR)^{-1} + m_i L_{fi} \tag{3.24}$$

Since the Joule heat energy generated by the critical current pulse must equal the energy absorbed by the interconnect structure, the right hand sides of equation (3.23) and (3.24) must be equal. Hence,

$$J_{crit}^2 = \phi_1 \, \Delta t^{-1} + \phi_2 \, \Delta t^{-\frac{1}{2}} + \phi_3 \tag{3.25}$$

where ϕ_1 is a constant depending on metal thickness, material constants and γ_{crit}. The constants ϕ_2 and ϕ_3 have dependence on oxide properties too, including a diffusion constant (a_d) that can be extracted from Figure 3.18. ϕ_1, ϕ_2 and ϕ_3 are independent of metal length, and given by,

$$\phi_1 = \lambda_0 \left[\left(m_i c_i + m_j c_j \right) \gamma_{crit} (TCR)^{-1} + m_i L_{fi} \right] \quad A^2 \cdot s / cm^4$$

$$\phi_2 = \lambda_0 \left[\delta_k c_k \, \gamma_{crit} (TCR)^{-1} \, 2L(W + d_s) \sqrt{a_d} \right] \quad A^2 \cdot s^{\frac{1}{2}} / cm^4$$

$$\phi_3 = \lambda_0 \left[\delta_k c_k \, \gamma_{crit} (TCR)^{-1} \, 4L a_d \right] \quad A^2 / cm^4 \tag{3.26}$$

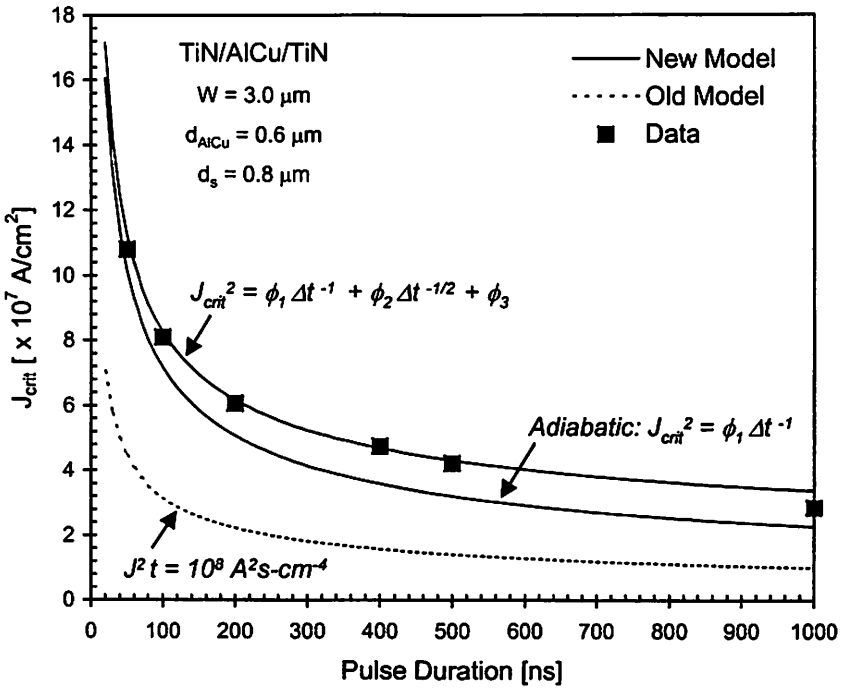


Figure 3.19 Pulse width dependence of J_{crit}.

where

$$\lambda_0 = \frac{2}{\rho_i L W d_i (2 + \gamma_{crit})} \tag{3.27}$$

Here d_i is the thickness of AlCu. For AlCu line with dimensions {W = 3 μm, and d_i = 0.6 μm, and d_s = 0.2 μm} ϕ_1, ϕ_2, and ϕ_3 are ~ 5 X 10^8, ~ 5 X 10^{11}, and ~ 1 X 10^{14} respectively. For pulse widths < 10 ns, near adiabatic condition is reached and the first term dominates. For pulse widths between 10 ns and ~ 1 μs, both the first and the second terms are significant. The ϕ_3 term which arises from the DC thermal conductance will become significant for longer pulse widths (≥ 1 μs). Further, ϕ_2 and ϕ_3 will become increasingly important for narrower line widths.

The critical current density values obtained using this relation are in excellent agreement with the experimentally observed values as shown in Figure 3.19. This model can be used to provide interconnect design guidelines for high current robustness. The adiabatic approximation of equation (3.25), i.e., $J_{crit}^2 = \phi_1(\Delta t)^{-1}$ is also plotted. It is clear that heat dissipation into the surrounding oxide is significant under pulsed conditions of approximately 100 ns and hence the

ϕ_2 term must be included in the model. In Figure 3.19 we also compare an earlier work [40], that proposed a theoretical relationship $J_{crit}^2 \Delta t = 10^8$ and it is shown to be lower than the experimental values. The model in [40] did not take into account any heat dissipation into the surrounding oxide, nor the latent heat of fusion of aluminum and was formulated using a critical temperature value of 300 ^{0}C based on work in [37], where unpassivated metal lines deposited on window grade quartz substrates were studied under pulsed stress. This is much lower than the 1000 ^{0}C which is extracted from the present experimental work. Detailed thermal simulations using transmission line models of joule heating under a high current pulse, by Gui et al [41], had also indicated that the model from [40] has limitations for pulse widths bigger than ~ 2 ns.

It should be noted that in the present work the heated oxide sheath is assumed to have square edges, but to be more realistic the edges should be cylindrical, since heat dissipation from a point source is circular. This will slightly reduce the volume of the heated oxide sheath, which will be given by

$$V_{ox} = \left[4L\left(\frac{1}{4}\pi d_{ox}^2\right) + 2L(W + d_s)d_{ox} \right] = \left[\pi L d_{ox}^2 + 2L(W + d_s)d_{ox} \right] \qquad (3.28)$$

By comparing equation (3.18) and (3.28) it can be observed that the volume contributed by the four corners of the heated oxide sheath gives rise to ϕ_3 in equation (3.26). Hence, the cylindrical corners will only reduce the ϕ_3 term in equation (3.26) by around 20%. However, in order for this to impact the J_{crit} values, pulse widths much longer than 1 μs must be involved which are comparable to the thermal time constant, and the concept of heated oxide sheath will not be valid for such pulse widths.

3.12 Interconnect Design Rules for High-Current Robustness

From equations (3.21) and (3.22), the critical value of the current pulse can be written as [57],

$$I_{crit} = \sqrt{\frac{2\left[\left(\sum_{i,j,k} mc\right)\Delta T_{crit} + m_i L_{fi}\right]}{\Delta t\, R_0 \left(2 + \gamma_{crit}\right)}} \tag{3.29}$$

which can be expressed as,

$$I_{crit} = \left(\frac{W(Wa_1 + a_2 V_{ox})}{\Delta t}\right)^{\frac{1}{2}} \tag{3.30}$$

Here, a_1 and a_2 are constants that depend only on the thickness and material properties of the metal and the oxide and γ_{crit}. They are independent of the length of the line. This result is valid for line lengths much greater than the thermal diffusion length in Al (25-50 μm), as discussed earlier.

Finally, the model developed in the form of equation (3.30) can be used to determine the critical current for open circuit metal failure in terms of the pulse width and the line width as shown in Figure 3.20. These curves provide design guidelines for ESD/EOS protection circuit interconnects. For example, a typical Human Body Model (HBM) ESD pulse, which can be described as a ~100 ns/1.3 A event [18], would require the width of the metal line to be greater than ~2.5 μm. Similarly for an electrical overstress (EOS) event of 1 μs the minimum line width should be more than ~7 μm. Hence a safe design guideline for these stress conditions would be ~10 μm. A recent study from IBM has shown that the model can also be applied to design damascence Cu interconnects [20].

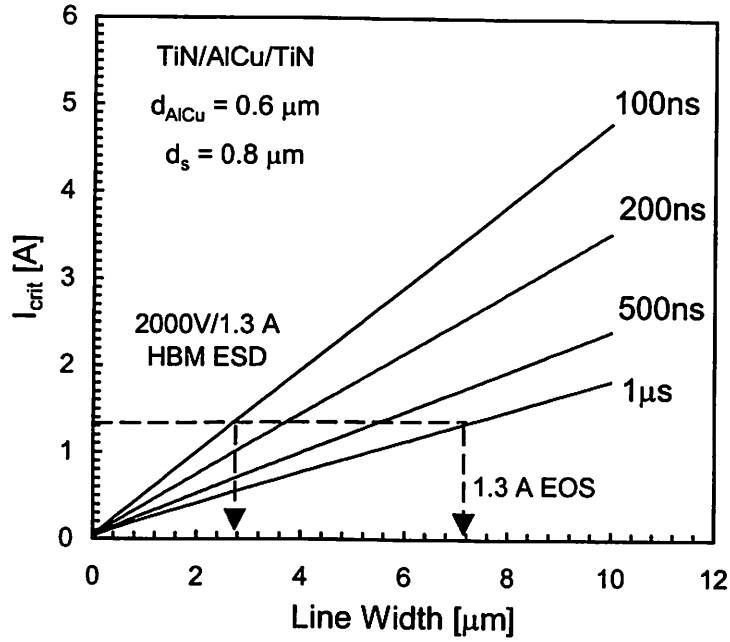
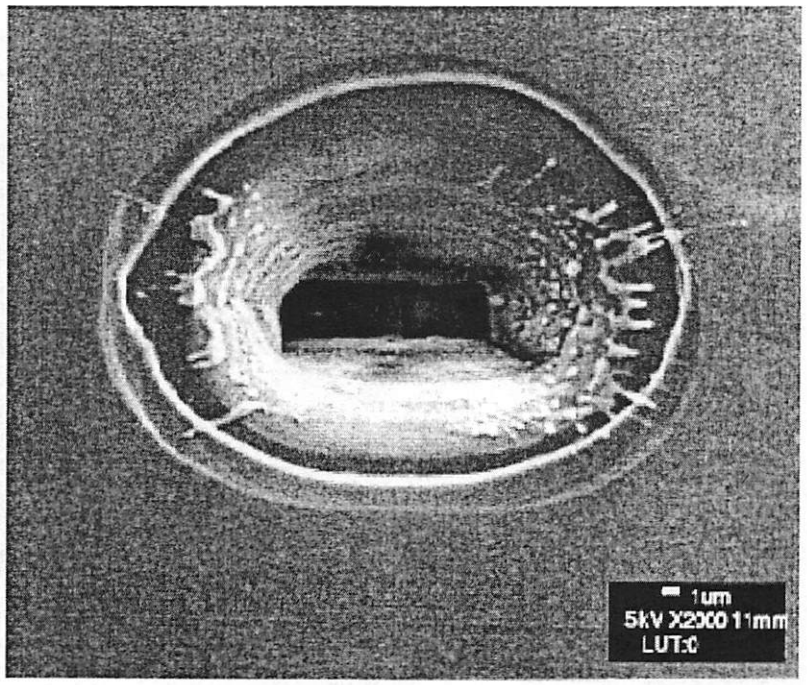


Figure 3.20 High-current short-pulse and EOS/ESD design guidelines for AlCu.

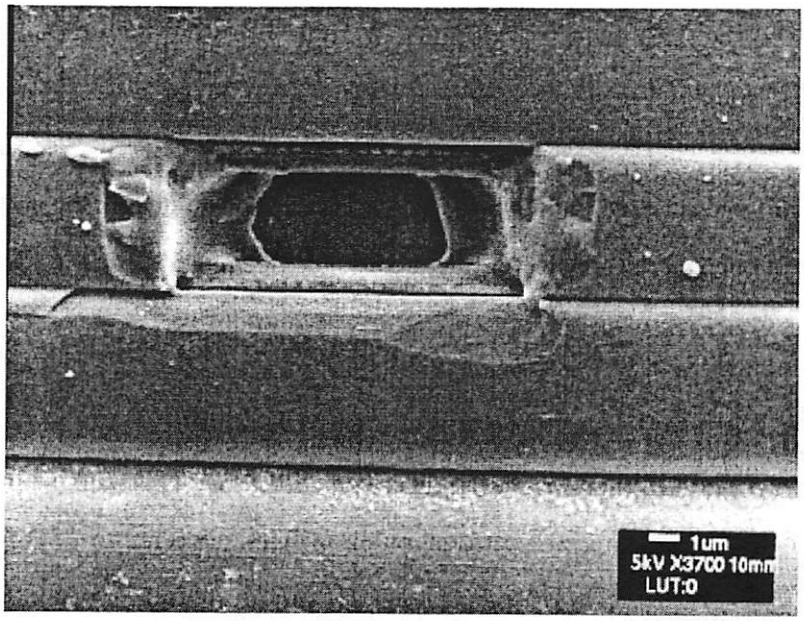
3.13 Failure Modes under High-Current Short-Pulse Stress

3.13.1 Open Circuit Failure Mechanism

As illustrated in section 3.9 high-current short-duration or ESD type pulses can heat up the metal lines well beyond their melting points and a maximum temperature rise value of 1000 ^{0}C (see Figure 3.11) is observed for all the metal lines at various levels. This critical value for temperature rise agrees with those obtained using finite element simulations, as shown in section 3.10. The open circuit failure of the lines was usually accompanied by cracking of the protective oxide/nitride layers. The fracture strength of silicon nitride is around 1 GPa [70]. At around the failure temperature the thermo-mechanical stress generated by the expansion of the molten AlCu exceeds the fracture strength of the oxide/nitride layers. Therefore all the lines fail when the temperature rises beyond that critical value. SEM micrographs of metal 1 and metal 4 lines (W = 3 μm) showing passivation fracture is shown in Figure 3.21. It is interesting to note that metal 1, which has a thicker overlying dielectric than that of metal 4, failed at the same temperature. This is due to the fact that fracture strength is independent of the thickness of the thin film



Metal 1



Metal 4

Figure 3.21 SEM micrograph showing open circuit failure mode of passivated TiN/AlCu/TiN metallization. The molten AlCu can be seen to have broken through the oxide/nitride passivation layer.

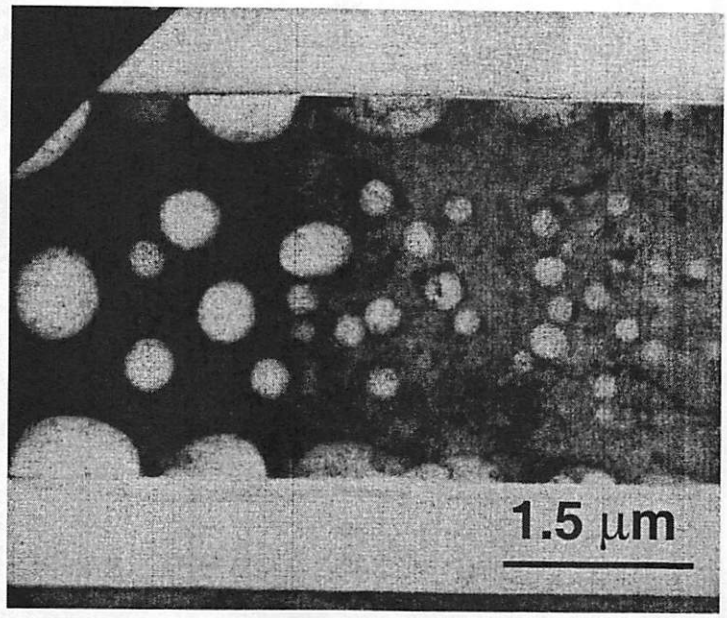


Figure 3.22 TEM micrograph of an AlCu line that has failed due to passivation fracture showing voids along the line.

material [71]. Similar failures were observed for metal lines at other levels. Furthermore, no preferential failure site is observed, as expected. Figure 3.22 shows a TEM micrograph of a section of a metal line that failed due to passivation fracture. It can be observed that there are several voids created in the line, which are likely to result from the sudden loss of material at the time of open circuit failure via fracture of the oxide/nitride layers.

3.13.2 Latent Damage: Impact on EM Reliability

A potential reliability hazard has been identified that significantly degrades interconnect lifetime due to enhanced electromigration in lines subjected to ESD type stress conditions [42]. This *"latent damage"* is introduced by high-current short-duration pulses that heat the metal lines past their highest equilibrium solid-solution temperature. The metal lines offer no visible clue of any damage. Furthermore, no change in line resistance can be observed, even with a Kelvin measurement after the pulse stressing. However, the electromigartion lifetime of these stressed lines were found be significantly lower that of the unstressed ones. This reliability degradation

effect was verified by comparing electromigration lifetimes of pulsed and unstressed AlCu metal lines.

Accelerated wafer level EM tests were performed on unstressed AlCu lines and then on two groups of pulsed lines. One group of lines were stressed by a pulse that caused a temperature rise of ~ 950 ^{0}C, which was just below $\Delta T_{crit.}$ The other group was stressed by a pulse that caused a temperature rise of ~ 700 ^{0}C. All the lines from both these groups showed no visible damage or any significant change in resistance after the pulsing. However the EM lifetime data revealed that the pulsed lines from the first group had lifetimes reduced by a factor of 3. The EM lifetime degradation was even more for the second group of pulsed lines and the MTTF (mean time to fail) was reduced by a factor of > 4.

In order to identify the mechanism responsible for this reliability degradation, detailed microstructure analysis were performed. Figure 3.23 shows a TEM micrograph of an unstressed 3.0 μm AlCu line. It can be observed that the grain size is ~ 1.5 μm. A TEM diffraction analysis is also shown below. The diffraction "spots" confirm the information from the TEM micrograph, that the unstressed lines have small number of large grains. Similar analysis was performed on AlCu(0.5%) lines stressed above their equilibrium solid solution temperature (~ 660 ^{0}C) [72], but well below the critical temperature of 1000 ^{0}C. It can be observed from Figure 3.23 that these lines have very different microstructure. A large number of small grains can be seen. This is well supported by the TEM diffraction rings, which arise due to large number of diffraction spots resulting from the random orientation of a large number of grains.

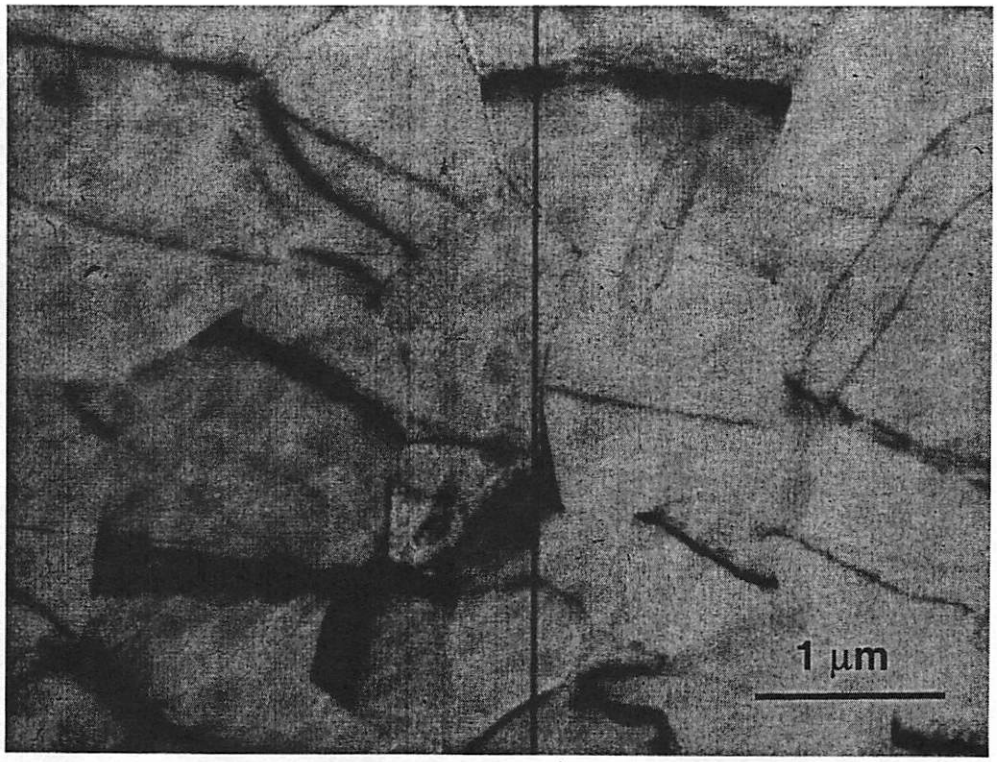
The change in microstructure of the stressed lines without any noticeable change in their resistance can be explained using solidification physics [72]. The heating during the ESD event is followed by quenching, since the thermal time constant is usually of the order of a few microseconds. This rapid cooling from a molten or partially molten state after pulsing, results in the formation of a large number of small grains throughout the line or in localized segments. This happens due to the combined effect of heterogeneous nucleation (which introduces a high nucleation rate) and extremely small growth time. Large number of small grains result in the formation of a increased number of grain boundaries throughout the line, which aids the diffusion of metal atoms under an applied electric field, causing enhanced EM.

In addition to grain size reduction, the distribution of Cu within the grains will also affect the EM failure times. After resolidifying, the AlCu lines are likely to have a gradient in Cu concentration due to constitutional supercooling [72], which will force the grains to contain increasingly higher concentration of Cu during the progress of the resolidification process, with the last few grains containing the highest Cu concentration.

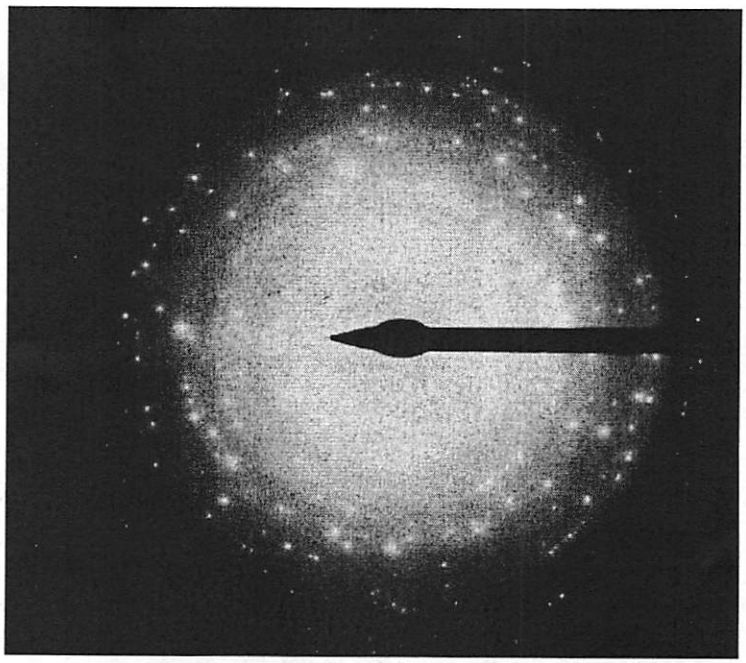
The difference in the observed MTTF of the two groups of pulsed lines can be attributed to localized melting caused by the pulses that heated the lines to ~ 700 ^{0}C, since the spatially averaged temperature rise was just above the melting point. The pulses that heated the lines more strongly to ~ 950 ^{0}C, likely caused more uniform melting of the line which resulted in a more uniform grain size distribution in the line after resolidification. In the case of localized melting, segments with smaller grain sizes are created in the line. These introduce a gradient in the microstructure and the effect of electromigration under such microstructural gradients is expected to be even more severe [73].

A recent study done in the industry has confirmed this new interconnect damage mechanism introduced by short-duration high-current pulses [64]. Therefore, this EM lifetime degradation effect on reliability needs careful consideration.

It should be noted that another report [74], concludes that there is no EM lifetime degradation in passivated multilayered AlCu lines after a short-pulse stress during which the temperature rise was well below an arbitrarily chosen (300 ^{0}C) value for the failure temperature. This value for the failure temperature is inapplicable for an actual IC interconnect system since it is based on [37], where unpassivated single layer Al lines deposited on window grade quartz were tested.

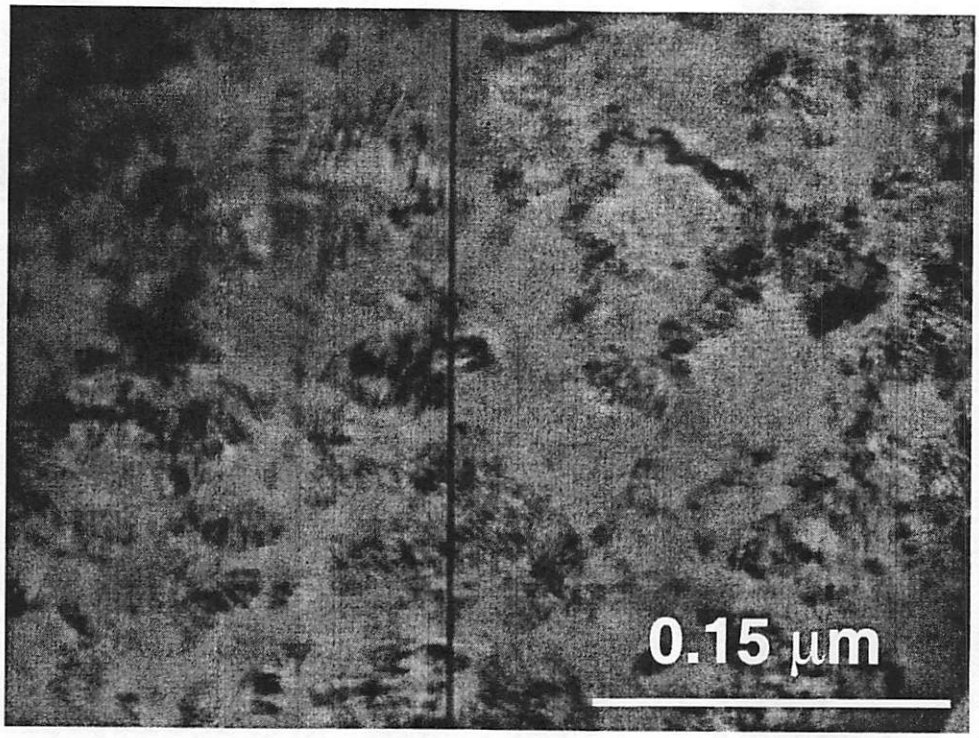


a)

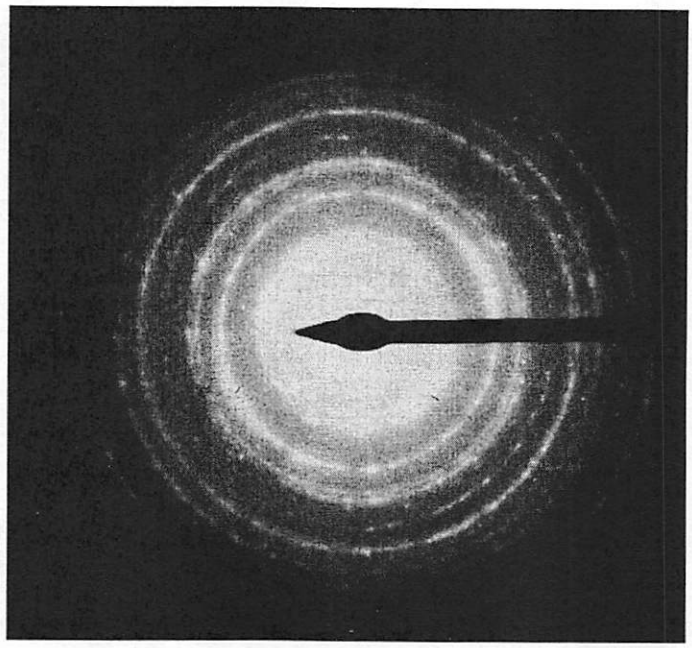


b)

Figure 3.23 a) TEM micrograph showing the microstructure of unstressed AlCu line. b) The corresponding TEM diffraction pattern showing scattered spots arising due to small number of large grains.



a)



b)

Figure 3.24 a) TEM micrograph showing the microstructure of a stressed AlCu line. b) The corresponding TEM diffraction pattern showing rings arising due to large number of small grains.

3.14 Summary

In this chapter self-heating behavior of passivated TiN/AlCu/TiN lines under short-pulse stress conditions have been presented in detail. A transient resistive thermometry technique has been developed for the fast temperature measurements. The TRT technique has been applied to obtain excellent temporal resolution of temperature rise in interconnects under short-pulse stress. It has been demonstrated that self-heating under short-pulse (ESD) conditions is not completely adiabatic and heat diffusion into the surrounding oxide is important. It has also been shown that under these conditions, interconnect heating is independent of total underlying dielectric thickness and only a thin sheath of surrounding oxide gets heated. Our characterization of passivated TiN/AlCu/TiN metal system under short-pulse stress conditions have demonstrated that, at the point of open circuit failure, the temperature rises to ~ 1000 ^{0}C. This temperature rise has been confirmed by solving the transient heat equation for the interconnect system using finite element analysis. An analytical failure model under ESD type stress conditions have been formulated that incorporates heat dissipation into the surrounding oxide sheath and the latent heat of fusion of the Al metal. The model has shown excellent agreement with experimental data and has been applied to generate design guidelines for high current robustness including ESD/EOS and I/O buffer interconnects.

The open circuit failure has been shown to be due to the fracture of the overlying passivation layers and independent of the overlying dielectric thickness. A latent interconnect failure mechanism has been identified for the first time. This failure mode have been shown to occur due to microstructural changes introduced in the lines that are stressed beyond the liquidus point of AlCu (0.5%) but below the open circuit failure temperature. The microstructural change has been shown to be undetectable by resistance measurements or physical inspection, but has been shown to cause significant electromigration lifetime degradation, which is a new IC reliability hazard.

Chapter 4

Impact of Scaling and Low-k Dielectric on Thermal Characteristics of Interconnects

4.1 Introduction

As discussed in Chapter 1, aggressive scaling of Si based IC devices motivated by the desire for faster circuit speed and higher packing density has increased the functional complexity of VLSI circuits. This has in turn, reduced the interconnect metal pitch and increased the number of metallization levels. Reduction in metal pitch however degrades interconnect RC delay which tends to curtail the benefits of interconnect scaling [1]. Low dielectric constant (Low-k) materials have been introduced [9], [75] - [80], as an alternative intra-level or interlayer insulator to reduce interconnect capacitance (therefore delay) and cross-talk noise to enhance circuit performance. Recently it has been demonstrated that thermal effects, instead of electromigration itself, will start to dominate interconnect design guidelines for advanced high performance interconnects [81], [82]. Furthermore, as discussed in Chapter 3, metal lines can also suffer from thermal damage under short duration high-current stress conditions such as during an ESD event. In Chapter 3, a model for interconnect heating and failure under these conditions was presented.

Characterization of the effect of interconnect scaling and low-k dielectric material on the thermal behavior of the IC metal is desirable to provide thermal design guidelines in the near future. The purpose of this chapter is to analyze the thermal implications of interconnect scaling using low-k dielectric structures and to demonstrate the usefulness of the TRT technique developed in Chapter 3, in comparing the thermal characteristics of different dielectric materials.

4.2 Minimization of Interconnect (RC) Delay using Low-k Dielectric

The implications of technology scaling on technology performance was discussed in Chapter 1, and it was illustrated why interconnect delays are beginning to dominate technology performance. The minimization of the total interconnect delay given by $0.5rcl^2$ in equation (1.1) is driving the introduction of low-k materials. It is instructive to examine a simple first order interconnect RC delay model proposed in [1]. The interconnect line resistance R (= rl) can be expressed as

$$R = 2\rho L / P H \tag{4.1}$$

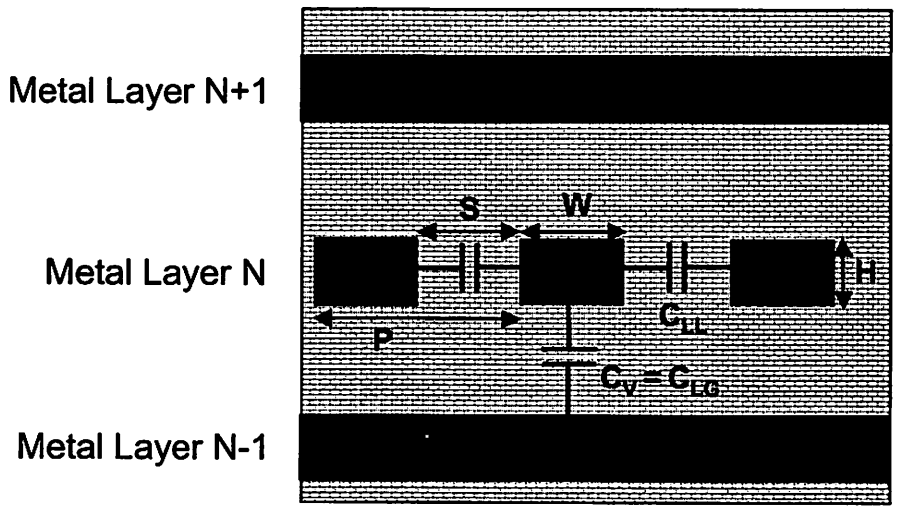


Figure 4.1 Schematic cross-sectional diagram of a multilevel interconnect system showing line-to-line (C_{LL}), the vertical layer-to-layer (C_V), and line-to-ground (C_{LG}) capacitances.

where ρ is the resistivity of the interconnect material, L is the interconnect length, H is interconnect height and also inter layer dielectric thickness, P is interconnect pitch ($=W+S$), and $W = S$, as illustrated in Figure 4.1. C is the total parallel plate capacitance comprising of the lateral line-to-line capacitance, C_{LL}, and the vertical layer-to-layer capacitance, C_V. Assuming that the line-to-ground capacitance (C_{LG}) equals C_V, and ignoring the effects of fringing fields for simplicity, C can be expressed as

$$C=2\left(C_{LL}+C_V\right)=2k\varepsilon_0\left(2LH/P + LP/2H\right) \tag{4.2}$$

where k is the relative dielectric constant and ε_0 is the permittivity of free space. Hence the product of R and C is given by [1], [76],

$$RC=2\rho k\varepsilon_0\left(4L^2/P^2 + L^2/H^2\right) \tag{4.3}$$

It can be observed from equation (4.3) that as P decreases the RC delay will increase and by introducing materials with lower values of k, the RC delay can be lowered. For example, the substitution of SiO_2 (k = 4) with air (k = 1) will lower the RC delay by about 75% [76]. Furthermore, for interconnect spacing S < 0.3 μm, C is dominated by C_{LL}. In fact, almost 90% of C is expected to be contributed by C_{LL} at sub-0.25 μm feature sizes [83]. Hence, reduction of C_{LL} is more important for reducing RC delay. Therefore, interconnect processes with low-k as intra layer dielectric only (gap-fill), have been introduced [84].

4.3 Sample Fabrication and Experiments

A double level metallization system fabricated using a state-of-the-art 0.25 μm CMOS process flow, which employed two different intra-level dielectric processes was used in this study. The standard dielectric process had SiO_2 (k ~ 4) as the insulating material everywhere, while for the low-k process, a dielectric (HSQ) with k ~ 3 was used as the gap-fill insulator for the level 1 metallization only [85]. The schematic cross section is shown in Figure 4.2. The metal system was multilayered with the stacking sequence of TiN/AlCu/TiN, and with all layer

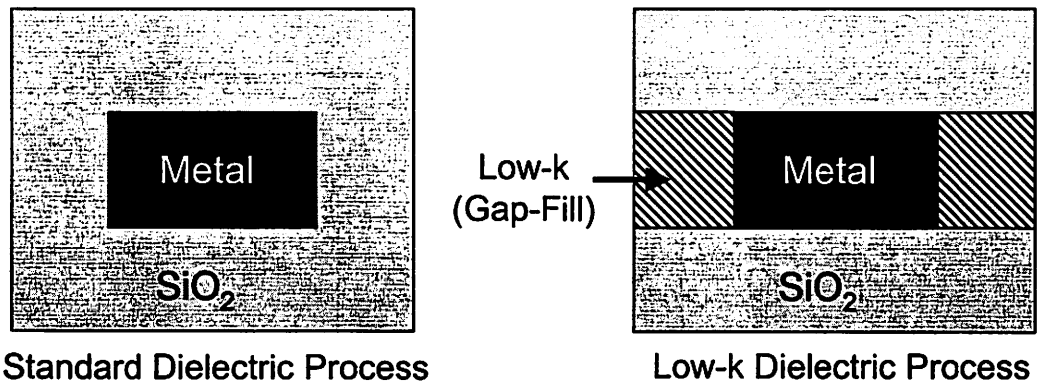
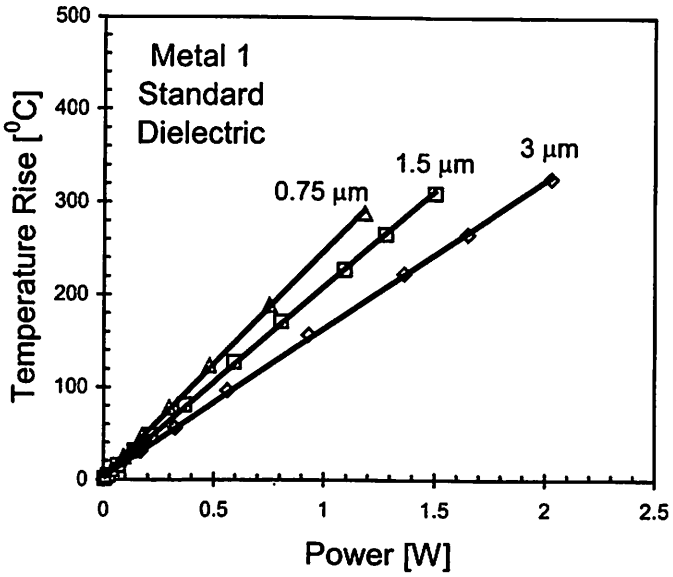


Figure 4.2 Schematic cross section of the two intra-layer dielectric processes used in this study.

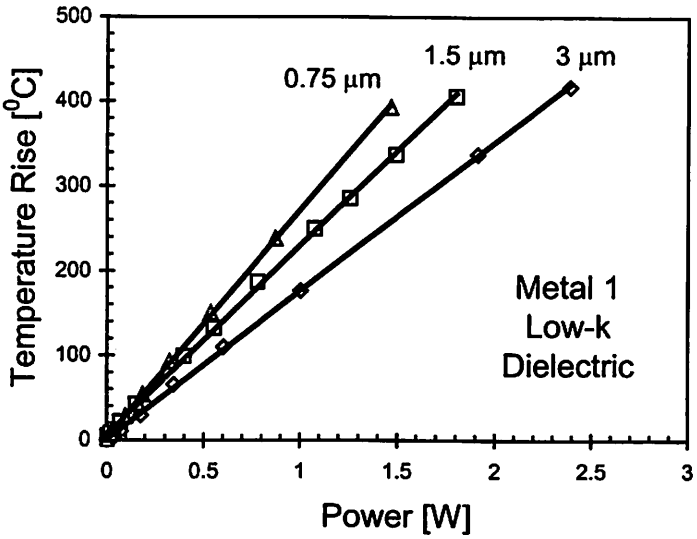
thickness equal to that of the quadruple level metal system described in Chapter 3. All the metal lines were standard NIST recommended 1000 μm long test structures with varying line width (3.0 μm, 1.5 μm and 0.75 μm). The TRT technique introduced in Chapter 3 was used to generate constant current pulses of varying widths (Δt=100ns, 200ns and 500ns) and magnitudes. The voltage, and hence the resistance of the metal lines were observed to increase linearly with time during all the pulsing events in agreement with experimental observations in Chapter 3.

4.4 Thermal Characterization under DC Stress Conditions

Initially the DC joule heating was measured for all the structures. Figure 4.3(a) shows the effect of interconnect scaling on the self heating of metal 1 lines. The self-heating gets more severe for smaller line widths for a given input power. This result is due to the fact that under DC conditions most of the joule heat generated is dissipated through the underlying oxide layer to the Si substrate which acts like a heat sink. Scaling reduces the effective surface area in contact with this underlying oxide. Figure 4.3(b) shows the same effect with low-k dielectric. Apart from the increase in self-heating with decreasing line width the temperature rise (ΔT) is higher for a given input power (P). These results indicate that the thermal impedance (R_θ) defined in equation (2.36) increases as line width decreases and that interconnect structures with low-k materials have even higher thermal impedance that gets worse with scaling as shown in Figure 4.4.



a)



b)

Figure 4.3 Effect of interconnect scaling on the self-heating of metal 1 lines under DC stress conditions for the a) standard dielectric process, and b) the low-k dielectric process.

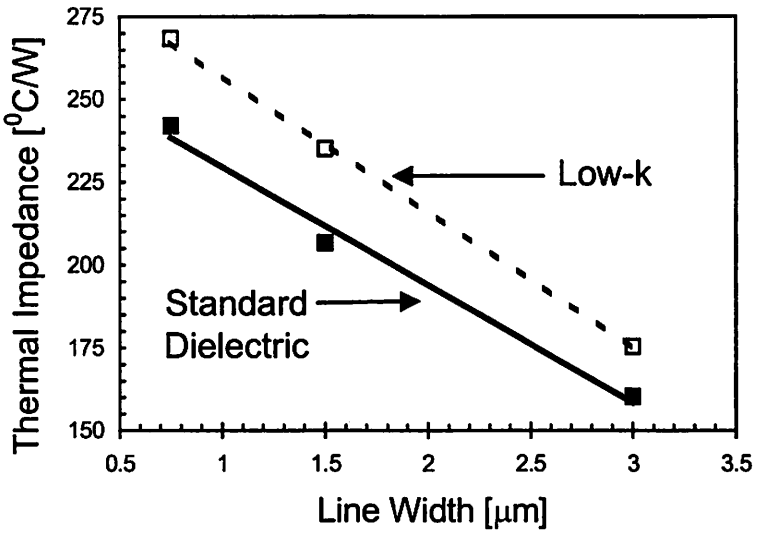
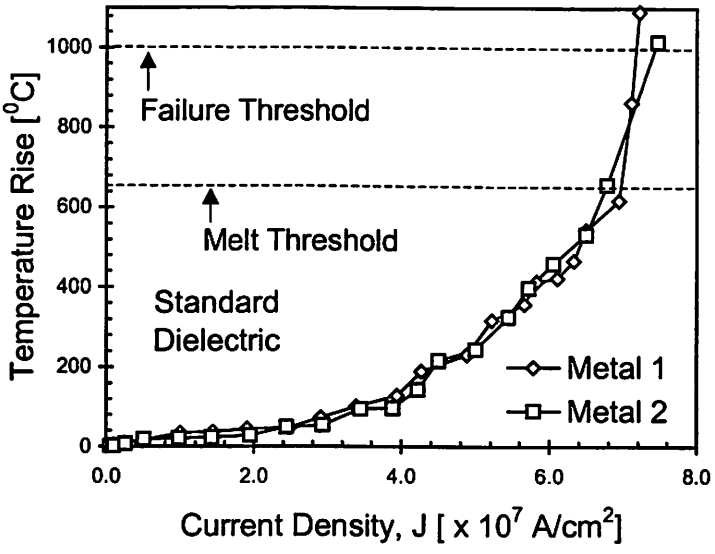


Figure 4.4 Effect of scaling and low-k dielectric on the thermal impedance of interconnects.

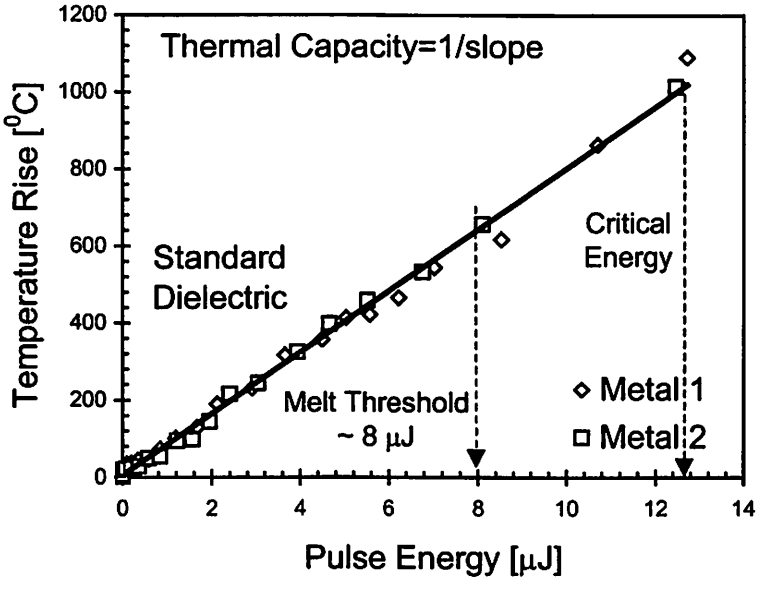
4.5 Thermal Characterization under Pulsed Stress Conditions

4.5.1 Standard Dielectric Process

The TRT technique developed in Chapter 3 is used to analyze the self-heating characteristics of the metal lines under short pulse stress conditions. Figure 4.5(a) shows the self-heating of 3μm metal lines for a 100 ns pulse for the standard dielectric process. It is observed that the metal lines are heated well beyond their melting point. The open circuit failure occurs at a ΔT of 1000 ^{0}C as observed in Chapter 3. Similar heating behavior is observed for other pulse widths. In agreement with experiments in Chapter 3, ΔT is identical for both metal 1 and 2. This shows that the underlying oxide thickness has no impact on the ΔT (contrary to self-heating under DC conditions) and that the lines have identical thermal capacities. Such behavior arises due to the fact that the pulse widths are much smaller than the thermal time constants of the interconnect structures, as explained in Chapter 3. As a result, the total underlying oxide thickness doesn't influence the self-heating of metal lines under these short time current pulses. Figure 4.5(b) demonstrates this effect, where the input pulse energy defined in equation (3.11) is plotted against ΔT. Since the inverse of the slope of this line gives the thermal capacity of the metal system, which is defined in equation (3.15), it can be observed that the thermal capacities of metal 1 and 2 are identical.



a)



b)

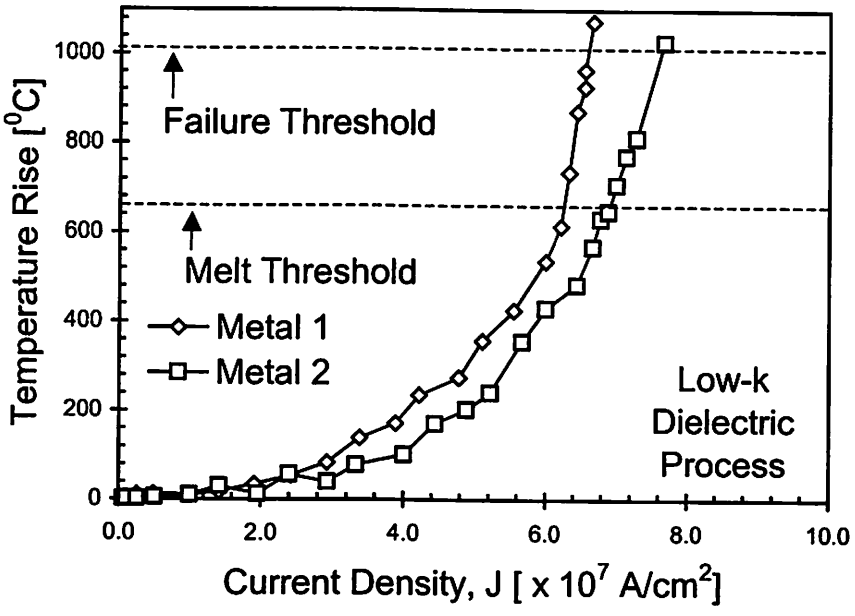
Figure 4.5 Self-heating behavior of 3.0 μm metal 1 and metal 2 lines for the standard dielectric process under a 100 ns pulse stress showing identical a) temperature rise, and b) thermal capacity.

The minimum energy required to melt the given volume of the 3 μm wide AlCu line was calculated to be 8 μJ, including the latent heat of fusion. However, it can be observed from Figure 4.5(b) that the energy dissipation capability of the interconnect structure is much larger. This suggests that a thin sheath of surrounding insulator also gets heated during the pulse and the failure temperature is much higher than the melting temperature, in agreement with the results in Chapter 3.

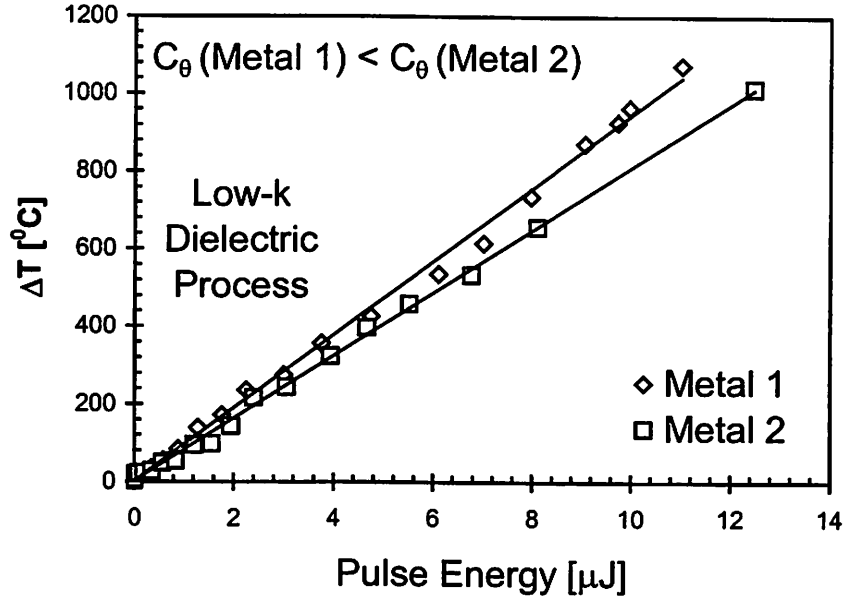
4.5.2 Low-k Dielectric Process

Figure 4.6(a) shows similar self heating effect for the low-k dielectric process. Contrary to the standard dielectric process it can be observed that metal 1 heats up and fails more quickly than metal 2. To comprehend this observation the ΔT is plotted as a function of pulse energy in Figure 4.6(b). It can be observed that metal 1 requires less energy to heat up to a given temperature. This is due to the dissipation of a smaller fraction of heat energy into the surrounding dielectric. The different thermal response of the low-k interconnect structure is explained below.

The thermal capacity of metal 1 as extracted from Figure 4.6(b) is smaller than the thermal capacity of metal 2. This is expected since the low-k material has a lower thermal conductivity and therefore the thermal diffusion length into the low-k material is shorter. Due to this smaller thermal capacity effect in the low-k dielectric structures the fraction of the pulse energy that goes into the dielectric is lower and thus the metal fails at a lower current density (or pulse energy). This was observed for all the other pulse widths as well and the results are summarized in Figure 4.7. Note that the critical current densities decrease with increasing pulse widths, as observed in Chapter 3. It should also be noted that the open circuit failure temperature is nearly constant (~ 1000 ^{0}C) for all the pulse widths and for both processes in agreement with the results in Chapter 3. Furthermore, the effect of the low-k dielectric gets stronger for longer pulse widths. This is due to the fact that as heat diffusion time increases, the effect of the lower thermal conductivity of the low-k material becomes more prominent. Figure 4.8 clearly show that the increase in thermal capacity with pulse width is lower for the line with low-k intra-level dielectric.



a)



b)

Figure 4.6 Self-heating behavior of 3.0 μm metal 1 and metal 2 lines for the low-k dielectric process under a 100 ns pulse stress showing different a) temperature rise, and b) thermal capacity.

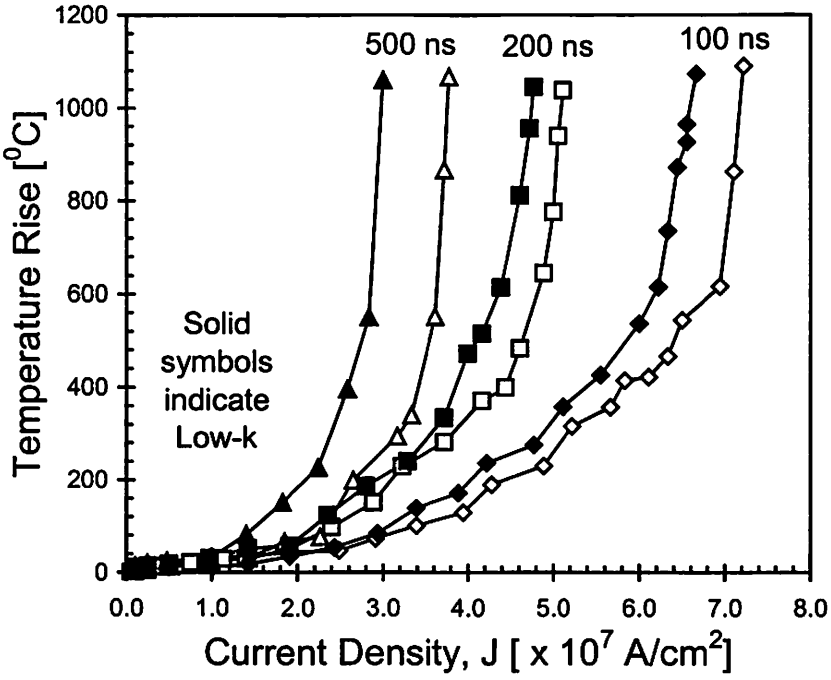


Figure 4.7 Effect of low-k dielectric on the self-heating behavior of 3 μm wide lines under different pulse durations.

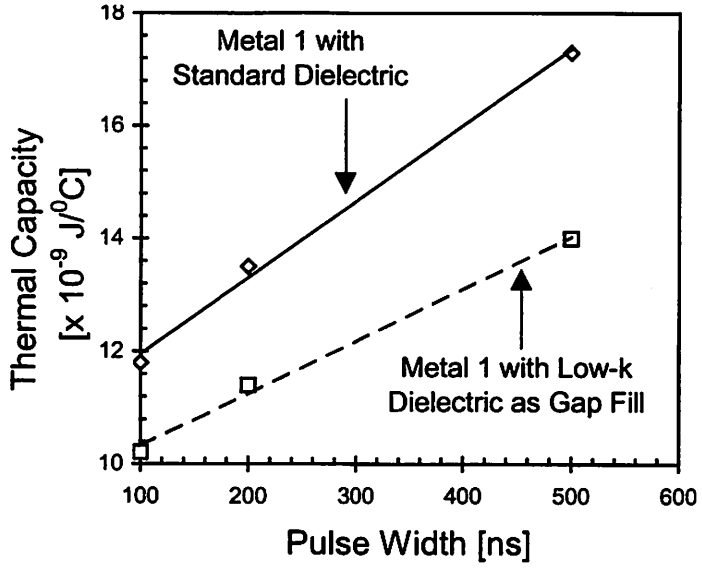


Figure 4.8 Increase in thermal capacity with pulse width is lower for a 3.0 μm metal line with low-k intra-layer dielectric process.

The lower thermal capacity of the line with low-k is due to the smaller thermal diffusion length in the lateral direction ($L_{d\text{-}lateral}$), which (from heat diffusion theory) can be expressed as

$$L_{d-lateral} \propto \sqrt{\frac{\kappa_{low-k}\,\Delta t}{\rho c}} \tag{4.4}$$

Where $\kappa_{low\text{-}k}$ is the effective thermal conductivity, ρ is the density, and c is the specific heat of the low-k dielectric material.

Therefore, the effective thickness of the heated insulator sheath around the metal line with the low-k dielectric is smaller, as shown in Figure 4.9. It can also be observed that the standard dielectric provides a larger sheath thickness, and therefore a bigger extra thermal capacity. As a result of the smaller thermal capacities of the low-k process the critical current densities for failure are also relatively smaller for all pulse widths as shown in Figure 4.10. The J_{crit} values are ~ 10% lower for the low-k structures and gets further reduced for longer pulses which carry more energy.

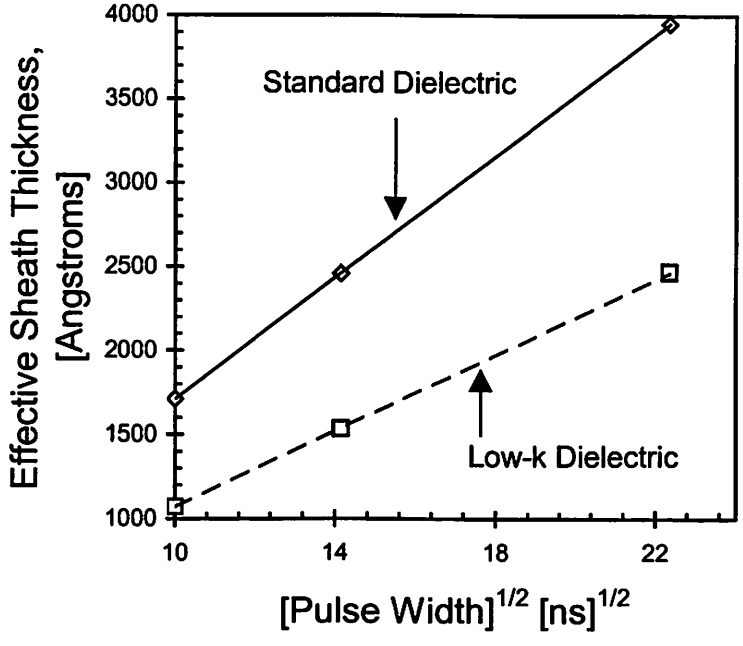


Figure 4.9 Effective thickness of the heated dielectric sheath around the 3.0 μm metal lines are lower for the low-k process.

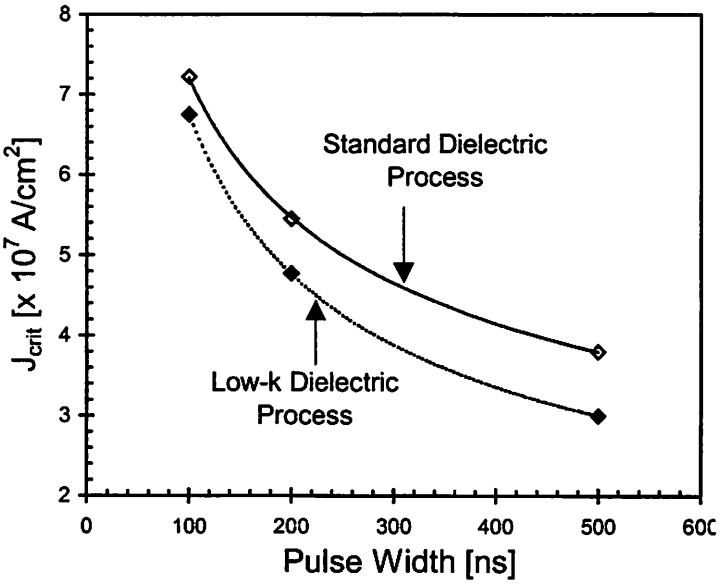


Figure 4.10 The critical current density in a 3.0 μm wide metal lead with the low-k process is smaller than that with a standard dielectric process.

4.6 Implications of Interconnect Scaling

The effect of interconnect scaling on the heating of metal lines with short duration pulsed current is shown in Figure 4.11. It can be observed that the wider metal lead ($W = 3$ μm) heats up more quickly as compared to the narrower lines. This is due to the smaller surface area to volume ratio for the wider lead. Hence smaller J is needed to reach the same ΔT. Similar effects were observed for the other pulse widths.

Furthermore, as shown in Figure 4.12(a) the thermal capacities decrease with narrower line widths as expected and the low-k structures have even smaller thermal capacities as discussed earlier. Figure 12(b) plots the ratio of the experimentally extracted total thermal capacity to the thermal capacity of the metal against various line widths. It can be observed that this ratio increases with scaling which explains the results of Figure 4.11.

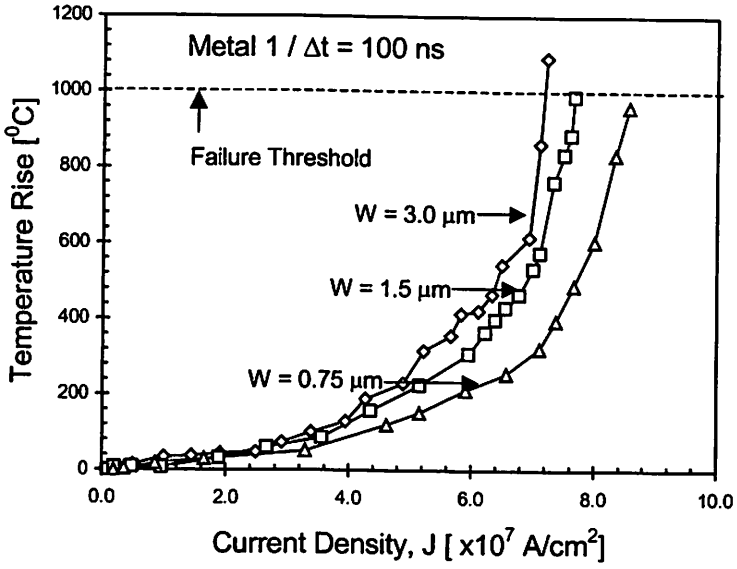
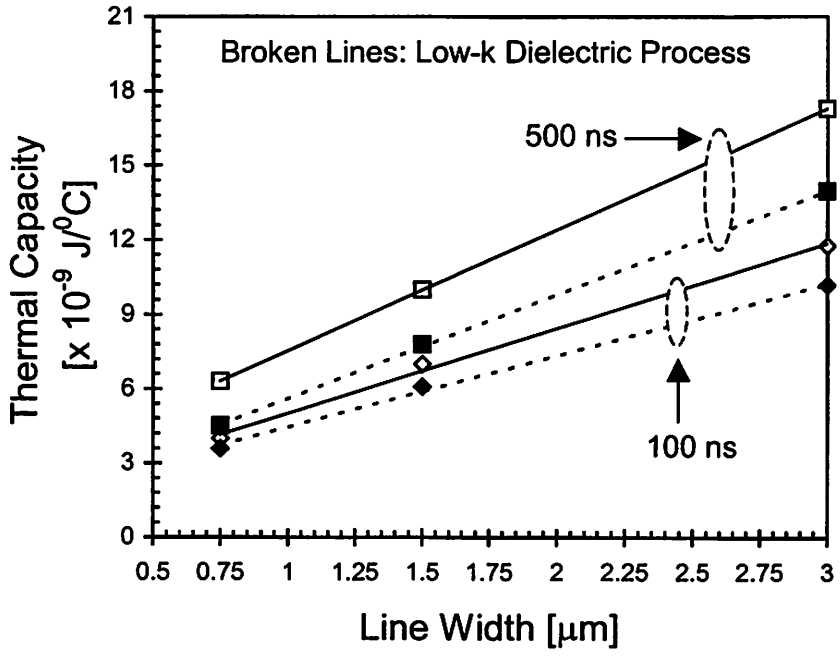
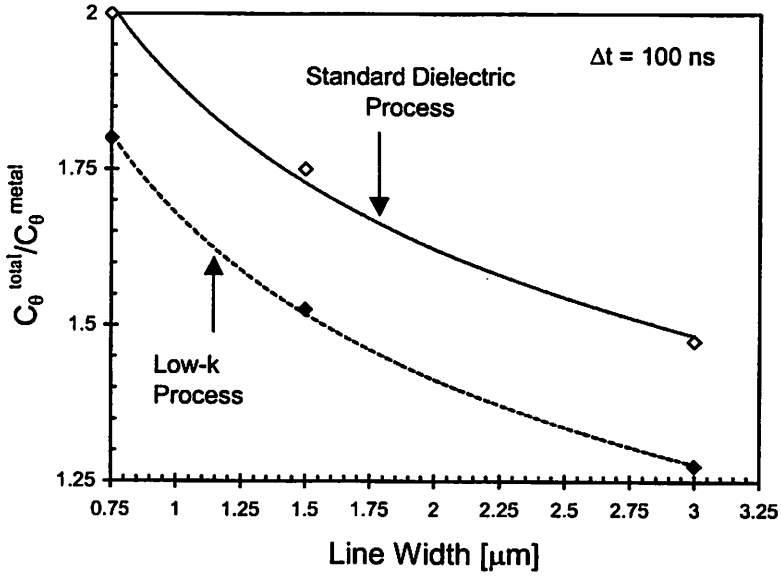


Figure 4.11 Effect of interconnect scaling on their self-heating characteristics under short pulse stress conditions.

Finally, in Figure 4.13, the effect of interconnect scaling using low-k dielectric on pulsed failure thresholds is illustrated. The J_{crit} values are ~10-30% lower for the low-k structures and show a tendency to be getting even lower for sub-micron lines with longer pulse widths.



a)



b)

Figure 4.12 Effect of interconnect scaling using low-k dielectric material on a) the thermal capacity of the metal lines under two different pulse widths, and b) the ratio of the total thermal capacity to the metal thermal capacity shown for 100 ns pulses.

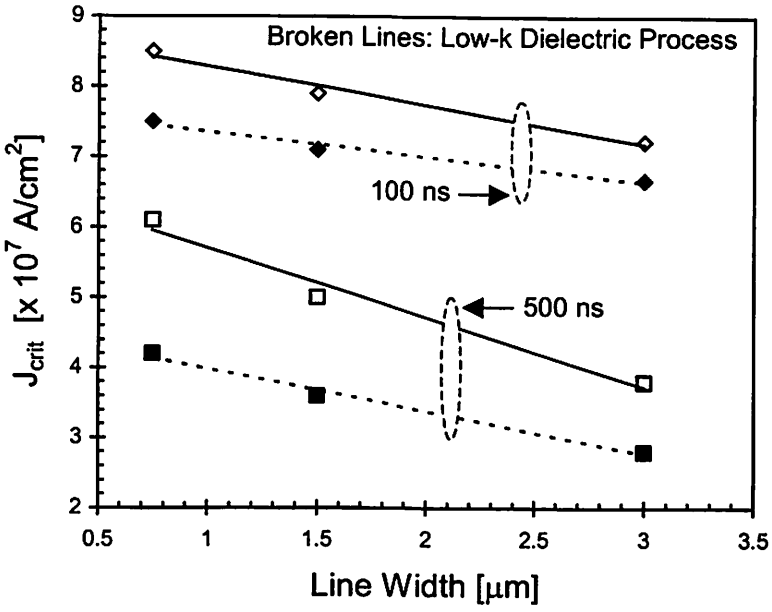


Figure 4.13 Effect of interconnect scaling using low-k dielectric material showing their increasing impact on J_{crit} with pulse duration.

4.7 Summary

In this chapter the effects of interconnect scaling using low-k dielectric material has been analyzed under DC and pulsed stress conditions. It has been shown that a low-k dielectric process, which uses the low-k dielectric for gap fill only, raises the DC thermal impedance of metal lines by about 10%. The TRT technique developed in Chapter 3 has been employed to study the dynamic self-heating behavior of metal lines with low-k dielectric process. The thermal impedance under short current pulse increases even more and the critical failure current density is reduced by 10 to 30% depending on line and pulse widths. The TRT technique is particularly useful for studying the thermal characteristics of intra-layer and inter-layer dielectric materials, since heat diffusion during short duration stress is limited to small distances from the metal, which helps avoid the influence of other underlying dielectric materials on the thermal characteristics of the metal line. These thermal characteristics will have significant implications on the design of deep sub-micron VLSI interconnects that employ low-k dielectrics and must be considered while developing electromigration, ESD/EOS and I/O buffer interconnect design rules.

Chapter 5

Failure Mechanisms of Contacts and Vias under High Current Stress Conditions

5.1 Introduction

As discussed in Chapter 1, increasing complexity of VLSI circuits has reduced the interconnect metal pitch and the width of diffusion regions. The reduction in the width of diffusion regions and metal pitch has consequently resulted in smaller contact and via critical dimensions as illustrated in Figure 1.5. Therefore the current density is increasing in these structures. Recently it has been demonstrated that thermal effects, instead of electromigration itself, will start to dominate interconnect design guidelines for advanced high performance interconnects [81], [82]. Furthermore, metal lines have been reported to thermally breakdown under ESD events [19].

Characterization of Interconnect heating and failure, under high-current (ESD) conditions have been presented in Chapter 3 and the effects of these high pulsed currents on interconnect scaling using low-k dielectric have been analyzed in Chapter 4. Aggressive scaling can cause considerable self-heating in contacts and vias due to increasing current densities, which can lead

to either a breakdown of the structure or a degradation of performance. Furthermore, failure of contacts and vias under high-current (ESD) conditions can also become a reliability hazard. Characterization of high-current and ESD effects on contact and via structures is desirable to achieve improved understanding of their failure mechanisms and to provide thermal design guidelines in the near future.

In order to characterize these high-current effects on contact and via structures, it is necessary to develop suitable thermometry and diagnostic techniques. High-current characterization under DC conditions is not very useful in identifying the failure modes since failures are usually catastrophic. The TRT method (described in Chapter 3) is an attractive alternative. However, due to the vertical geometry of the contact and via structures, the precise measurement of their resistance must involve a four-point (Kelvin) measurement [86]. Kelvin measurements under DC stress are typical, but transient measurements during a single short-duration pulse are relatively difficult and require careful calibration.

5.2 Prior Work

In general there is little information on the high current behavior of contacts and vias and their failure mechanisms under these conditions. One study by Fu and Pyle [87], on $TiN/TiSi_2$ barrier contact failures under DC stress reports a critical temperature of 465 ^{0}C and attributes the failure to the maximum current carrying capacity of the underlying diffusion region, which became intrinsic past this critical temperature. The increase of intrinsic carriers for a current bigger than the critical value was believed to cause rapid heating of the diffusion region that eventually lead to a failure of the $TiN/TiSi_2$ diffusion barrier. No physical evidence was provided to support the proposed failure mechanisms. The critical temperature estimated in [87] was found to be independent of the type of impurity dopant in the diffusion region. Another study on metal/Si contact failure under DC stress proposed localized melting near the metal-silicon interface and subsequent metal migration along Si (electro-thermomigration) close to the $Si-SiO_2$ interface to be the cause of failure [88]. But it was also pointed out that the presence of a p-n

junction is necessary to initiate metal migration between contacts. Furthermore, this study did not involve silicided diffusion or contacts.

The diffusion regions involved in the present work always have a much bigger cross sectional area as compared to the contact itself. Therefore the current density in a single contact structure increases more rapidly than in the diffusion region causing greater self-heating in the contact. This is especially important for self-heating during short pulses where there is minimal heat dissipation due to the non-equilibrium nature of the heating. Hence an estimation of the upper limit of temperature rise in the contact along with an understanding of degradation mechanism is necessary for providing design guidelines for advanced high performance interconnects.

5.3 Sample Fabrication

A number of different contact structures made of CVD-W and Force-Fill (FF)-Al in a state-of-the-art sub-0.25 μm CMOS process were analyzed in this study. The contact sizes were 0.3, 0.35 and 0.4 μm in diameter. Contacts were patterned on pre-silicided diffusion and poly-Si, using deep ultra-violet (DUV) lithography and the oxide was etched in a high-density plasma (HDP) etcher. Prior to contact liner/barrier metallization, the wafers were cleaned in dilute HF. Following the liner deposition a thin layer (20 or 40 nm) of TiN barrier was sputter deposited before the contact plug formation. W-plugs were formed by CVD tungsten deposition and etchback. Al plugs were formed via force-fill process [170]. The different types of diffusion layers underneath the silicide included n+/p+ Si and poly-Si of 3.0 μm width. Furthermore, structures with multiple contacts were also examined. Figure 5.1 shows the layout of a single contact structure on Si. Figure 5.2 shows the schematic cross section of a single contact structure on Si and poly-Si.

The vias were made of CVD-W with a thin TiN barrier. A schematic cross section of a W-plug via is shown in Figure 5.3. Structures with variation in the number of vias (1 through 9) were used to analyze their high current behavior.

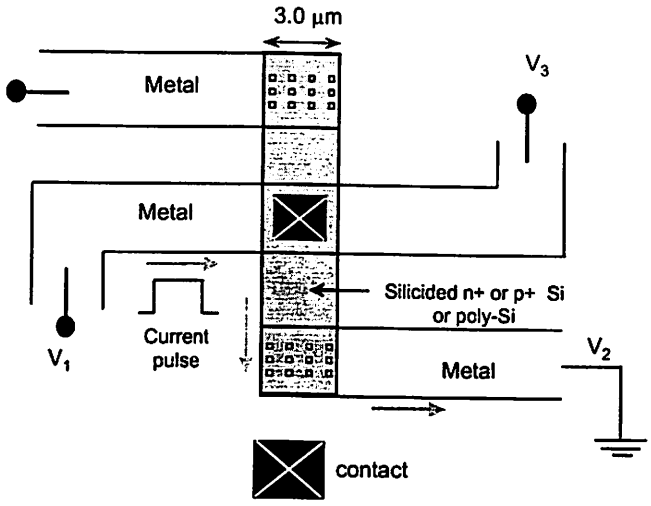


Figure 5.1 The layout of a Kelvin contact structure under study.

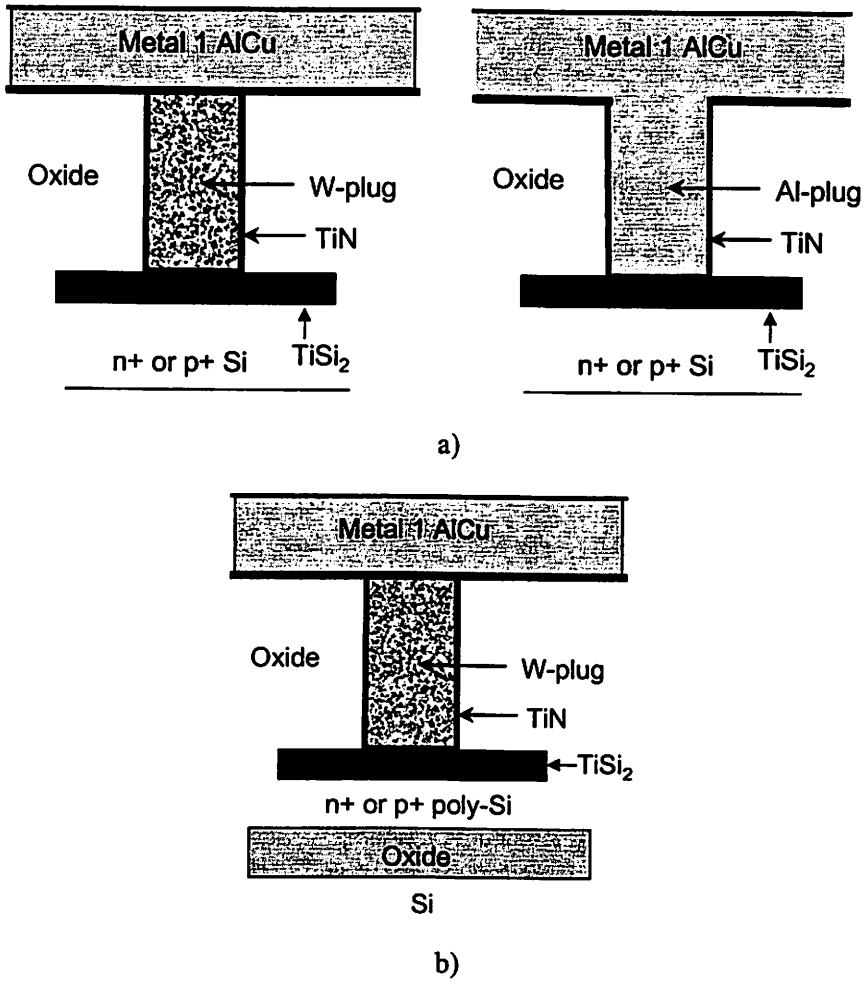


Figure 5.2 The schematic cross sectional view of the contact structures on a) Si and b) poly-Si.

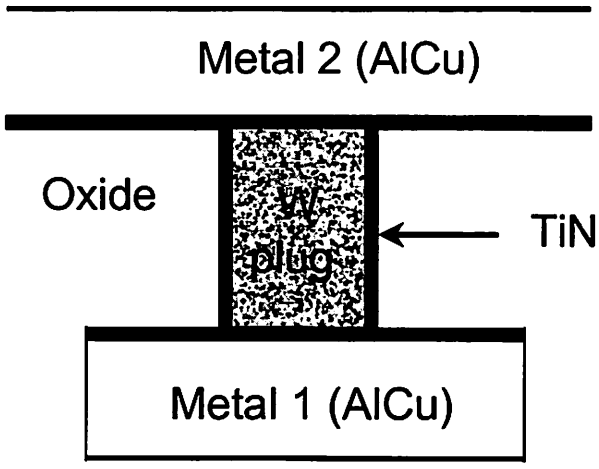


Figure 5.3 The schematic cross sectional view of a via structure used in this study.

5.4 Transient Kelvin Thermometry (TKT) Technique

The transient Kelvin Thermometry technique has been developed to characterize high-current effects in contacts and vias [89]. The system design for the TKT measurements is very similar to the design of the TRT technique described in Chapter 3 (see Figure 3.6). The HP 54111D in the TRT system is replaced by a Tektronix TDS 684B, which is equipped with four channels. Figure 5.4 shows the simplified TKT system. A Kelvin type set up was employed to capture the voltage pulse across the contacts/vias using a digitizing oscilloscope. The voltage across these structures was observed to vary linearly with time during all the pulsing events in agreement with previous work on metal lines [42], [57]. The instantaneous voltage pulse across the contacts and vias were captured using a difference math function channel on the TDS 684B as shown in Figure 5.5.

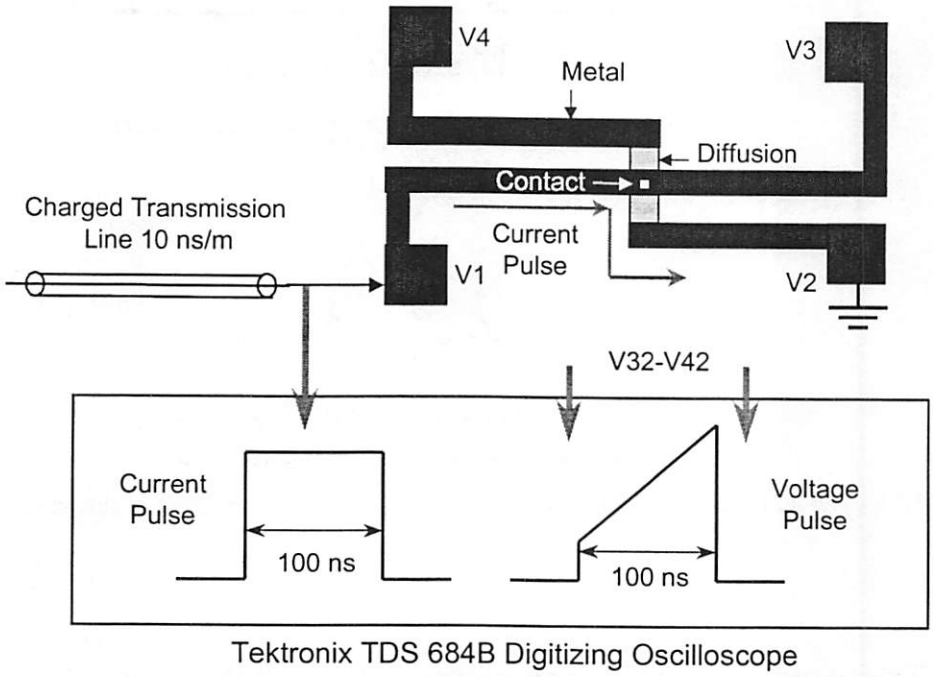
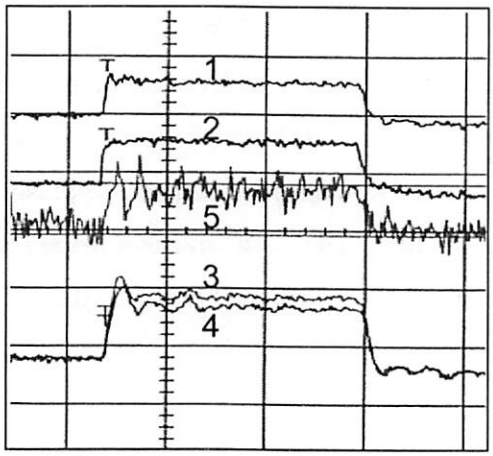


Figure 5.4 The schematic view of the simplified TKT system.



1: Voltage pulse across pads 1 and 2 (V12)
2: Current Pulse
3: Voltage pulse across pads 3 and 2 (V32)
4: Voltage pulse across pads 4 and 2 (V42)
5: Voltage pulse across contact (V5 = V32-V42)

Figure 5.5 Voltage pulse across contacts/vias captured using a difference math function in a digitizing oscilloscope. The difference between pulse 3 and 4 is electronically calculated and displayed as pulse 5.

5.5 Characterization of Contact Structures

The TKT technique described above is employed to study the high current behavior of contacts. For all contact sizes as pulse width increases the critical current density J_{crit} decreases as shown in Figure 5.6. These critical currents were determined at that current level after which there was an irreversible increase in the resistance of the contact as measured with a parameter analyzer so as to avoid joule heating. No visual damage was observed since resistance increase is used as failure criteria. At the critical point the resistance rises past 1.6 times the initial resistance corresponding to a temperature rise past ~ 800 ^{0}C, which was computed using the measured temperature dependence of the contact plug resistance using DC current.

At this critical point the resistance of the contact structure suddenly increases by more than 100%. As shown later by TEM analysis this resistance rise is caused by the interface degradation.

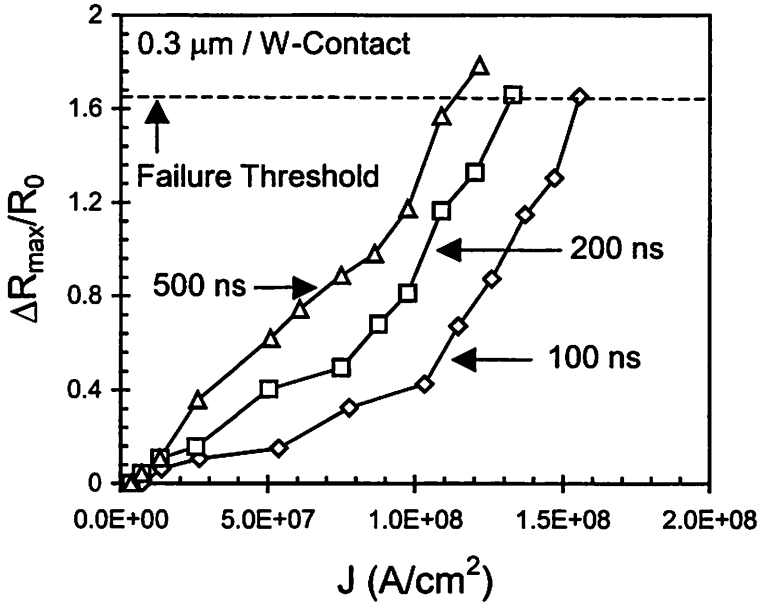


Figure 5.6 Resistance rise factor as a function of current density for single 0.3 µm W-contacts shown for various pulse widths.

Furthermore, for a given metal width W, as the contact area increases, I_{crit} increases but J_{crit} decreases as shown in Figure 5.7. This is due to decreasing heat dissipation ($\propto$ surface area) with increasing thermal capacity ($\propto$ volume) of the contact metal, in agreement with our observations on scaled interconnect lines in Chapter 4. It can also be observed that for 3 x 0.3 μm multiple contact structures J_{crit} is lower than for any of the single contact structures. This is expected, since for a multiple contact structure the heat dissipating capacity is reduced due to thermal coupling between adjacent columns and hence I_{crit} per contact is lower causing lowering of J_{crit}.

It was also observed that contacts to poly silicon substrates had a lower J_{crit} than those to Si substrates as shown in Figure 5.8. This is due to the poly Si being isolated from the Si substrate by an oxide layer, which causes lower heat dissipation into the substrate during the pulsing events

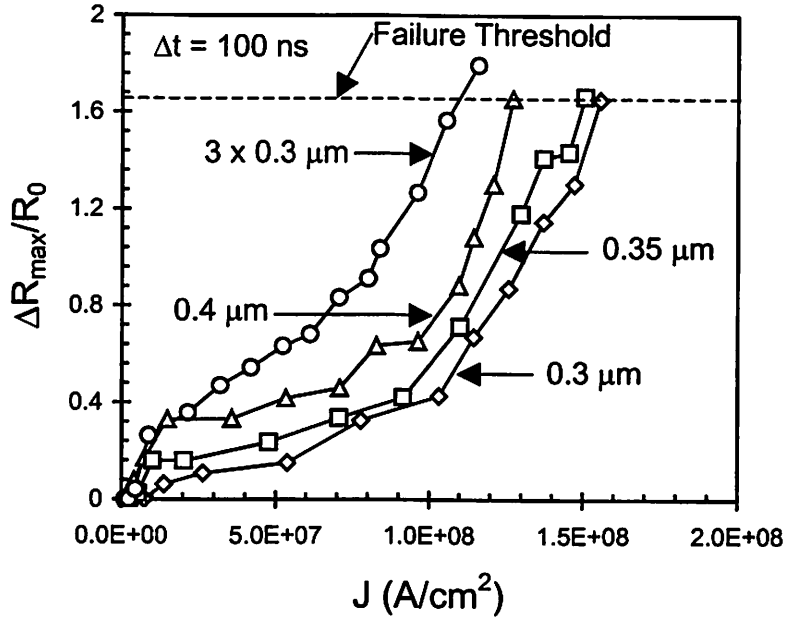


Figure 5.7 Effect of contact size and number on the resistance rise factor for 100 ns pulses (results shown for W).

Contacts to n and p type diffusion regions had similar J_{crit}, which indicated that the contact degradation is insensitive to the sheet resistivity of the underlying diffusion region for silicide strapped diffusions. Also, the type of dopant impurity has little effect on the thermal conductivity of Si. This result is in agreement with previously published work on $TiN/TiSi_2$ barrier contact failure under DC stress [87]. In addition, the rise in contact resistance arises primarily due to the failure of the $TiN/TiSi_2$ interface as against contact spiking proposed in [90].

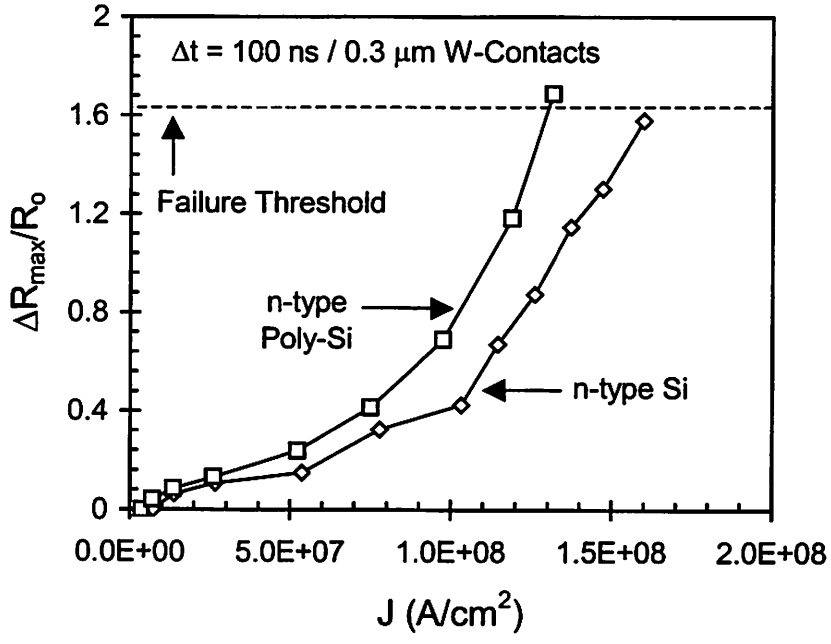


Figure 5.8 Heating and failure of W-contacts on n type Si and Poly-Si.

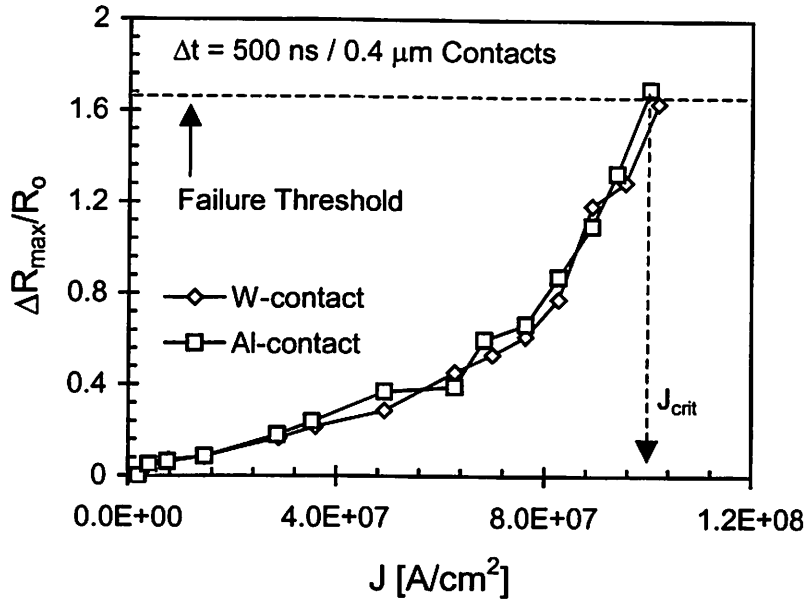


Figure 5.9 Heating and failure of single 0.4 μm, W and Al contacts under 500 ns pulsed stress.

Analysis of contact structures made of Al showed similar current carrying capability and the results are shown for 0.4 μm single contact structures in Figure 5.9. Equal failure thresholds of W and Al contacts indicate that the W/TiN or Al/TiN interfaces are not the weakest link amongst the three interfaces involved at the bottom of the contacts. Again, this result also supports the conclusion that contact degradation is not dependent on the material used to fill the contact plugs. Instead, it is the nitride/silicide interface that suffers degradation because the TiN/$TiSi_2$ interface resistances dominate the resistance of the contact.

Furthermore, for all contact structures the failure thresholds were also found to be independent of the direction of the current (electron) flow because the failure mechanism is thermally driven. In order to verify this breakdown mechanism, contact structures made of Al but with a thicker TiN liner were pulsed under similar stress conditions. The result in Figure 5.10 shows that the contact with the thicker TiN liner heats up and fails more rapidly compared to the contact with thinner TiN liner. This is due to the thicker TiN film in contact with the silicide layer, which forms a larger quantity of high resistivity interface material. This confirms that the TiN has an important role in the failure mechanism.

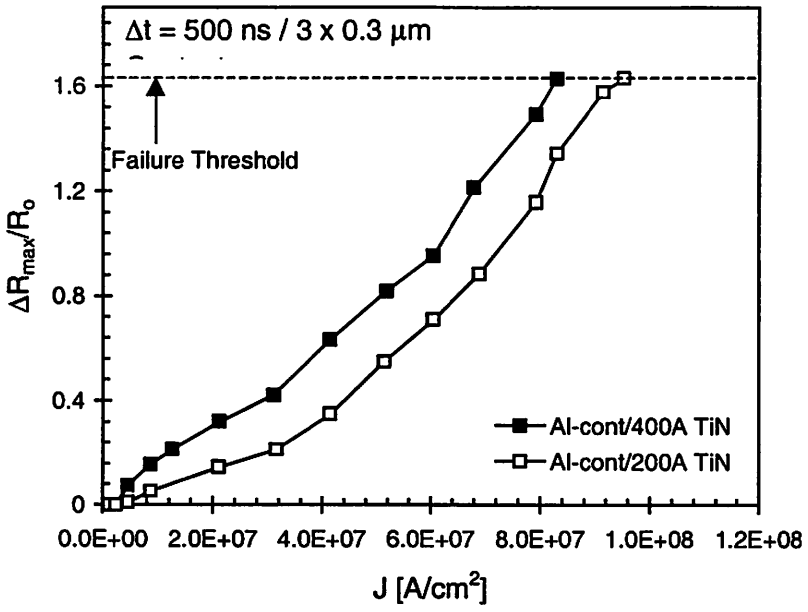
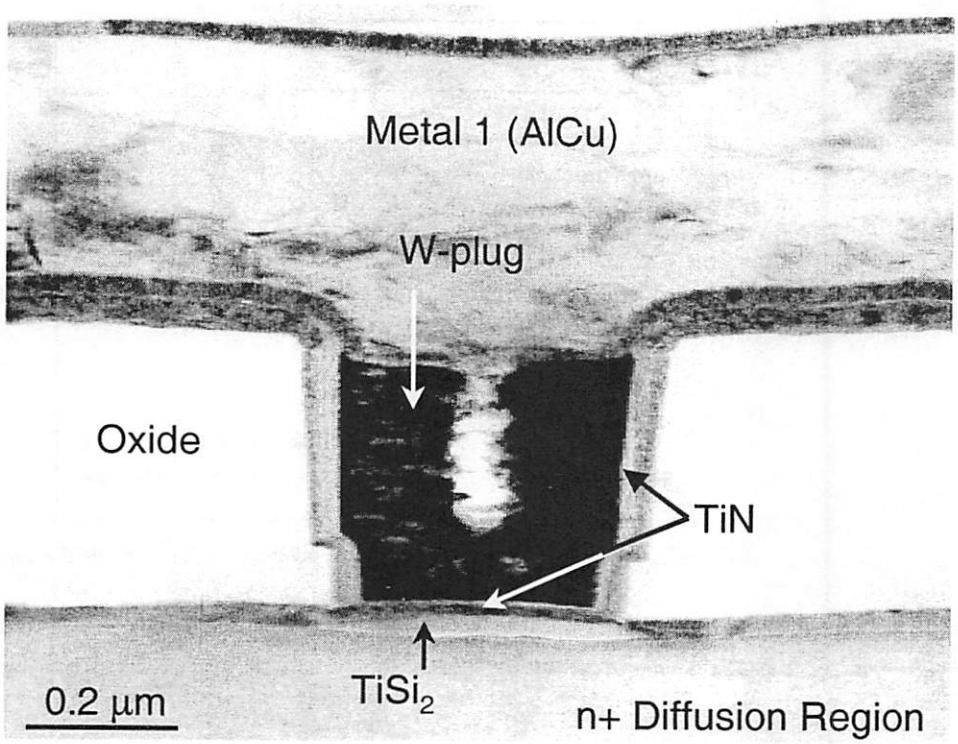


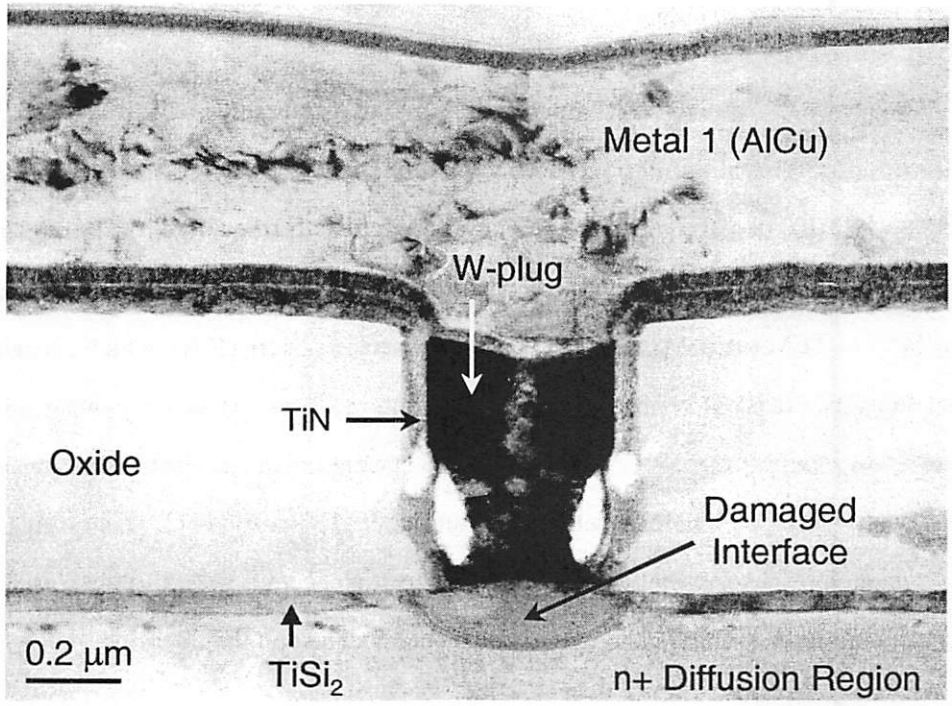
Figure 5.10 Effect of TiN liner thickness on the robustness of 3 x 0.3μm Al contacts under 500 ns pulsed stress.

5.6 Contact Failure Mechanism

Figure 5.11 shows TEM micrographs of an unstressed and stressed W-contact structure showing the TiN/$TiSi_2$ interface. Both structures had a 200 Å TiN liner. It can be observed from Figure 5.11(b) that an interfacial reaction product has formed underneath the contact area, which is likely the cause for the rise of the contact resistance. The voids seen in the contacts are simply artifacts of TEM sample preparation using focused ion beam (FIB), which caused overthinning of the samples. It is also observed that the failure mode observed for contact degradation in this work was independent of the direction of the current pulse (or electrons), contrary to previously published work on silicided contact wearout under DC stress [91]. This result is expected, since Si depletion or accumulation under contacts subjected to DC stress is mainly driven by electromigration, which is known to have a dependence on the direction of electron flow [92].

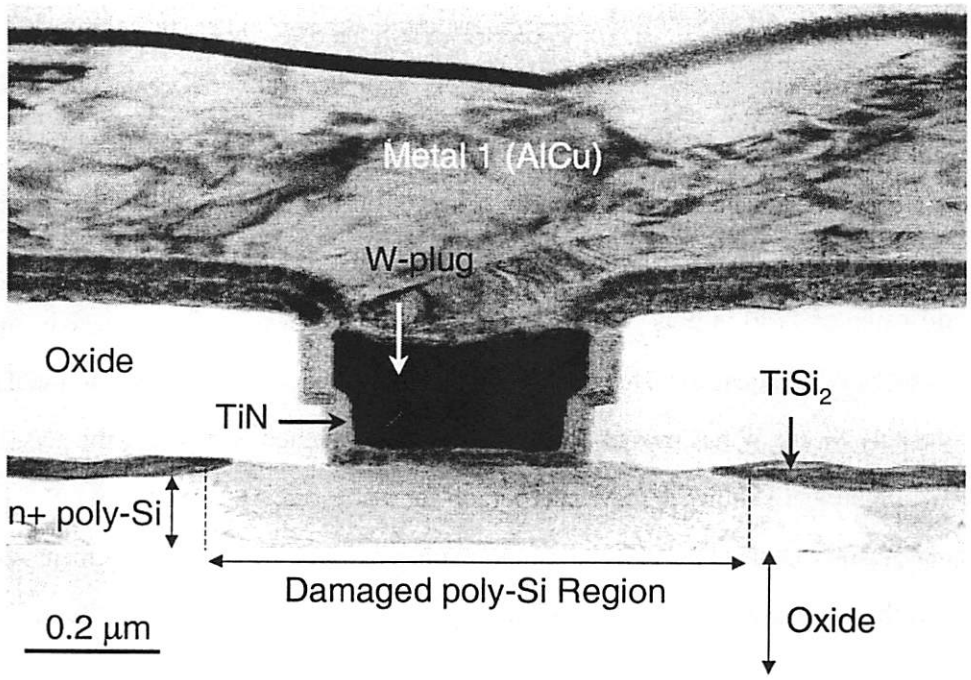


a)

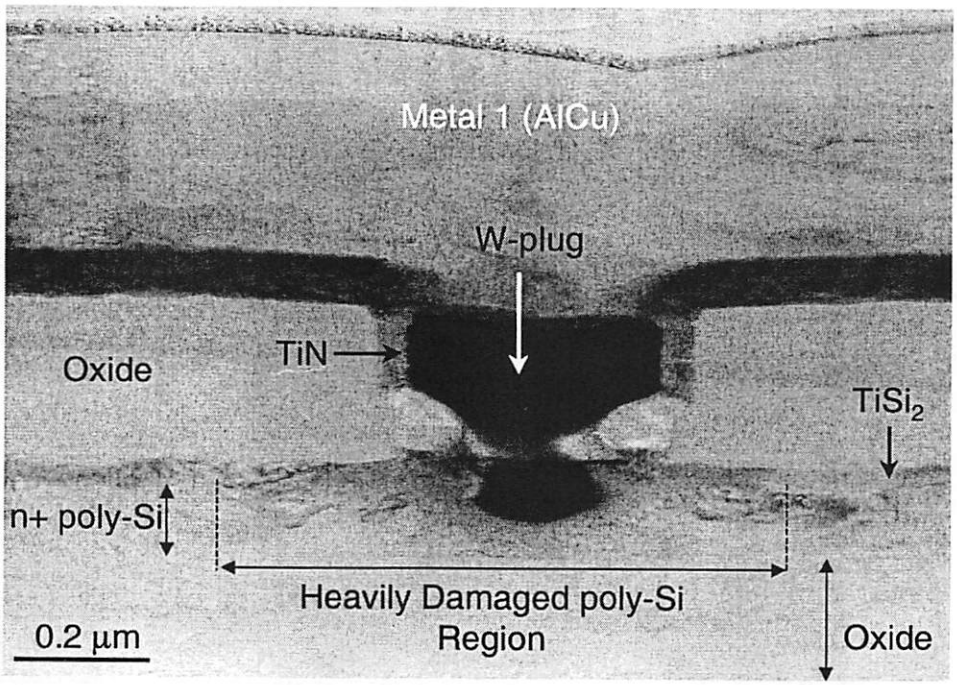


b)

Figure 5.11 TEM micrographs showing 0.3 µm W-contact structures to n+ Si with (a) unstressed $TiN/TiSi_2$ interface and (b) stressed $TiN/TiSi_2$ interface.



a)



b)

Figure 5.12 TEM micrographs showing 0.3 μm W-contact structures to n+ poly-Si with stressed $TiN/TiSi_2$ interfaces with (a) 100 ns and (b) 500 ns pulses.

In Figure 5.12 the TEM micrographs show damage under pulsed stress for contacts to n+ poly-Si. It can be observed that the poly-Si underneath the contact also suffers significant damage. In Figure 5.12(a) the poly underneath the contact has a different contrast and the $TiSi_2$ has been consumed. The differences observed in the failure mechanisms of contacts to Si and poly-Si is due to the poly-crystalline microstructure of the poly-Si, which causes enhanced diffusion of $TiSi_2$ in poly-Si through the grain boundaries. In Figure 5.12(b) the substrate seems to be heavily damaged. Due to more severe joule heating, in addition to the $TiSi_2$ diffusion into the poly-Si, the W has moved down through the weakened barrier into the poly-Si.

Figure 5.13 shows a TEM micrograph for a stressed Al contact to n+ Si. It can be observed once again that the TiN/$TiSi_2$ interface area has reacted. This is in agreement with our conclusion that the breakdown initiates at the interface.

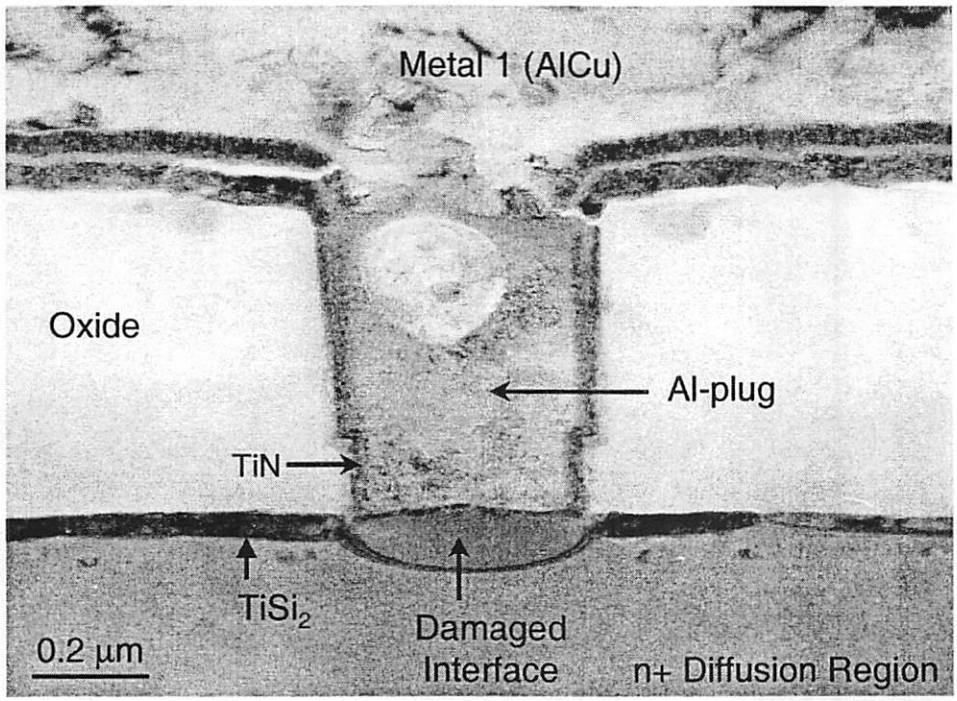


Figure 5.13 TEM micrograph showing 0.4 µm Al-contact structure to n+ Si stressed by a 100 ns pulse.

5.7 Characterization of Via Structures

Thermal breakdown of vias under high-current stress conditions has also been investigated. A via structure connects two levels of metal. It was observed that, for all via sizes, as pulse width increases J_{crit} decreases. This is due to the higher energy in longer pulse widths that result in a higher temperature rise. The critical temperature rise is more than 1000 ^{0}C for via failure as compared to ~ 800 ^{0}C for contacts. In comparison to contacts, where failure is always characterized by a degradation of the TiN/$TiSi_2$ interface, vias exhibit catastrophic damage once the critical value of the current is exceeded. Figure 5.14 shows a SEM micrograph of a stressed single W-via structure. It can be observed that unlike the contact structures the via has undergone severe damage as a result of excessive joule heating.

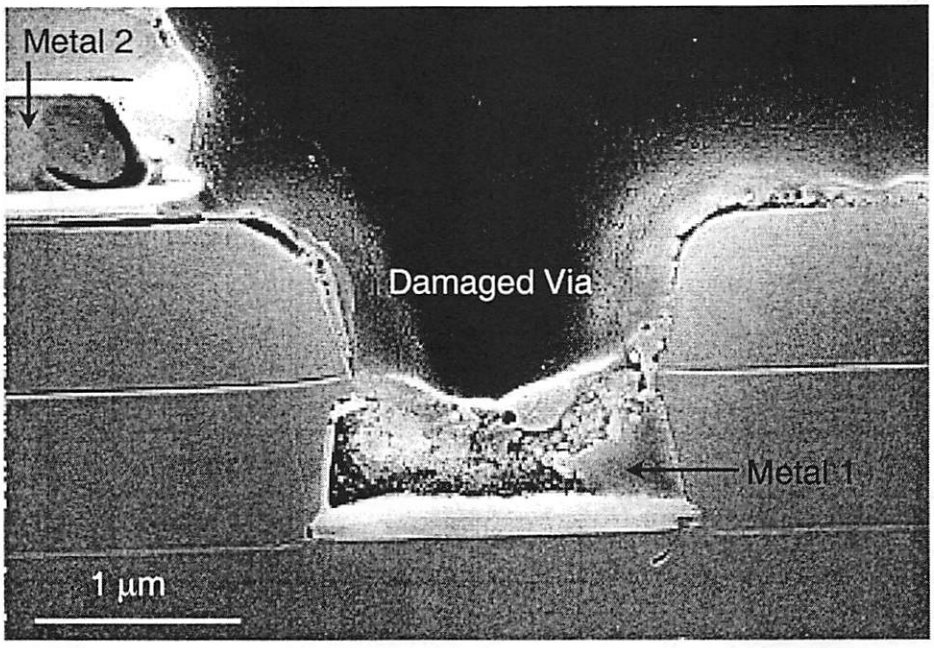


Figure 5.14 SEM micrograph showing a 0.4 μm single W-via damaged under high pulsed current stress.

In the case of interfacial reactions, the change of via resistance due to the formation of any metal-metal compounds is expected to be much less than that of metal-Si compounds in contacts. This can explain the absence of a resistance rise in the case of vias. Instead, failure occurs due to an open circuit, which is caused by the eventual thermal runaway and catastrophic damage of the structures. Again, via failure under short duration pulsed stress exhibited no dependence on the stress polarity, indicating that this is not an electromigration induced failure that have been reported under continuos stress conditions [93], [94].

Furthermore, as shown in Figure 5.15, the critical current density decreases as number of via is increased. This is due to incresed thermal coupling between the neighboring via columns. Figure 5.16 shows that vias display greater robustness to high current degradation and failure than contacts. This is due to their lower resistance path as compared to the contacts, which have the higher resistance TiN/$TiSi_2$ interface in the current path. The vias also have metal layers at the top and bottom, which helps them to dissipate heat more quickly. Hence, the via structures have lower self-heating and better heat conduction compared to the contact structures.

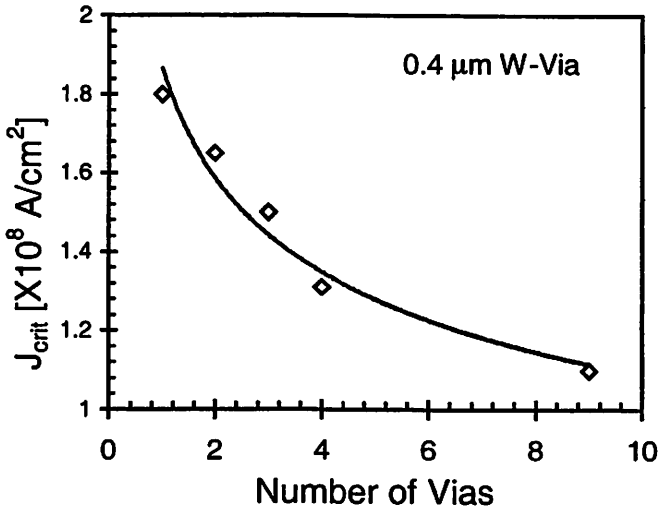


Figure 5.15 Critical current density of 0.4 μm W via structures under 100 ns pulsed stress.

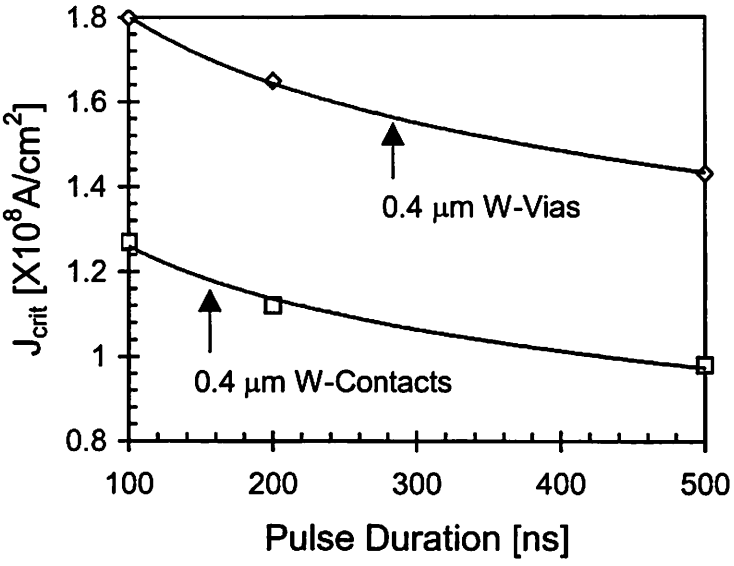


Figure 5.16 Critical current density to cause degradation/failure is higher for vias as compared to contacts.

5.8 Summary

In this chapter, a new failure mode of silicided barrier contact structures made of W and Al, under high-current stress has been identified using the transient Kelvin thermometry technique. It has been shown that the critical current is independent of the plug material but strongly dependent on the pulse width, and the cross sectional area of the contact. It is also shown that for contact structures, the critical current is influenced by the thermal conductivity of the underlying material, and is independent of the electrical properties of the underlying diffusion, such as sheet resistivity. Using extensive microstructure analysis it has been shown that contact failure under high-current stress conditions is characterized by a breakdown of the TiN/$TiSi_2$ interface due to interfacial reaction, which results in the formation of a new phase with a higher resistivity. This interface degradation, being thermally driven, occurs past a constant critical resistance rise, which is independent of contact size and number. The corresponding critical temperature rise is ~ 800 ^{0}C. Furthermore, thermal coupling between adjacent contacts/vias in multiple contact/via structures has been shown to strongly influence the self-heating behavior and failure thresholds.

In addition, contact and via failures under short time joule heating has been found to be independent of the direction of the current flow. Vias have been observed to fail catastrophically, once the threshold current limit is exceeded. TEM and SEM micrographs detailing the degradation of contacts/vias have been provided. Hence, an understanding of the failure mechanisms has been developed that can be used for further analysis and for designing interconnect structures for high current robustness.

Chapter 6

Thermal Effects in Thin Silicide Films

6.1 Introduction

Continuous scaling of VLSI devices into the deep sub- micron region has led to the increased use of silicided metalization schemes for low-resistivity gates, interconnections and contacts between the metal and Si [95], [96]. Currently, self-aligned silicide (salicide) processes are widely used in advanced CMOS technologies [97], [98], as shown in Figure 6.1. In addition to lowering the gate sheet resistance (and therefore RC delay), they also reduce the source/drain parasitic resistance, by forming ohmic contacts in the source/drain regions of MOS transistors, thereby increasing the drive current of the transistors. Furthermore, silicided diffusion structures are frequently used as resistors in I/O buffers and ESD protection circuits. These silicide films are often subjected to high current stress in MOS devices and as well as during electrostatic discharge (ESD) and electrical overstress (EOS) events. The thickness of the silicide is known to significantly impact ESD performance [99]. In Chapter 5, failure of CVD-W and force-fill Al contacts to silicided diffusions under high current stress conditions was shown to initiate due to

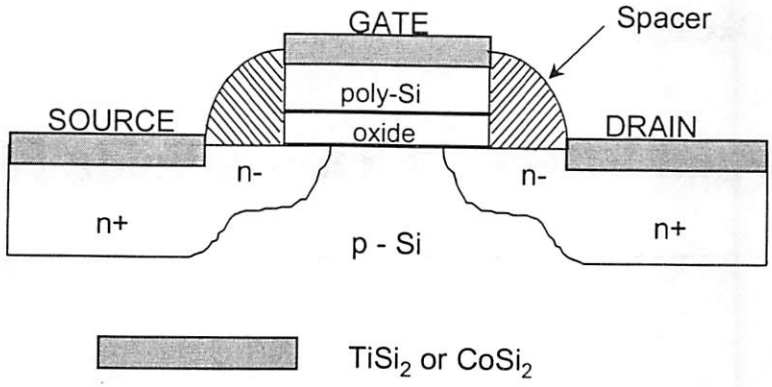


Figure 6.1 Sample cross-section of a salicided LDD NMOSFET.

the degradation of the silicide under the contact plug. The silicide thickness also affects the contact resistance sensitivity to temperature and current as will be shown in the next chapter.

In this chapter, characterization and modeling of high current effects in the two most commonly used silicides namely, $TiSi_2$, and $CoSi_2$, under DC and short-pulse stress conditions are presented. The transient resistive thermometry technique developed in Chapter 3 is employed for the short-pulse characterization. This technique is shown to be an effective tool to evaluate the electrical and thermal stability of these structures. Currently, there is limited information available on the high current and self-heating effects [95], [96], [100], in these silicides.

This chapter will identify important parameters related to the high current behavior in silicides and present analytical models. Furthermore, failure mechanisms of the silicide films will be identified. An understanding of these high current effects will enable the impact of technology scaling of silicide films to be defined for the development of deep sub-micron technologies.

6.2 Sample Fabrication

$TiSi_2$ and $CoSi_2$ films of different geometries and 50 nm thickness were formed using the salicide process for a state-of-the-art, 0.18 μm, 1.5 V, CMOS technology on n+ Si and n+ poly-Si. The various technology and process splits used in this work are summarized in Table 1. The sheet resistance values were measured with a four-point probe test structure. The $TiSi_2$ films were formed after the gate-etch, followed by either a Ge pre-amorphization implant (PAI) or a Mo implant. The Ge PAI process, which requires a high-energy dose of Ge to amorphize the

substrate, is used to increase nucleation density and achieve improved sheet resistivity [101], [102]. In case of the $TiSi_2$ process with Mo implant (pre-gate), the implant energy and dose are lower and hence there is no amorphization of the substrate, but the presence of Mo itself influences the silicide formation, resulting in low sheet resistance [103]. The Ge PAI or Mo implant is followed by a deposition of ~ 30 nm of Ti, which forms the $TiSi_2$ {C49 phase [104] - [106]} upon RTP heat treatment in N_2 at 750 ^{0}C for ~ 30 seconds. A thin layer of TiN also forms on top of the silicide layer during this step, which is stripped. This is followed by an anneal at 900 ^{0}C for ~ 10 seconds to form the low resistivity C54 phase of $TiSi_2$ [104]- [106].

For the $CoSi_2$ process [96], [107], [108], there is no implantation step after the gate-etch. Instead, an 8 nm layer of Co and 50 nm of TiN was deposited and then annealed in N_2 at 575 ^{0}C for ~ 30 seconds to form the silicide and then annealed at 850 ^{0}C for another 30 seconds.

Figure 6.2(a) shows the schematic cross sectional view of the silicide film on n+ Si [109]. A lumped circuit representation of this silicide structure is shown in Figure 6.2(b). The temperature coefficient of resistance (TCR) were measured to be 0.0029 ± 0.0001 and 0.0031 ± 0.0001 $^0C^{-1}$ for $TiSi_2$ and $CoSi_2$ films (formed on n+ Si) respectively at low DC current. The films were then subjected to high DC and pulse stress. The TRT technique was used for high-current short-pulse characterization.

Silicide	Thickness	Si-Type	Implant Species	Sheet Resistance
$TiSi_2$	50 nm	n+ Si	Ge (PAI)	~ 3.0 Ω/square
		n+ polySi		~ 3.2 Ω/square
$TiSi_2$	50 nm	n+ Si	Mo	~ 3.2 Ω/square
$CoSi_2$	50 nm	n+ Si	N/A	~ 5.2 Ω/square
		n+ polySi	N/A	~ 5.5 Ω/square

L/W = 25/5, 50/5, and 10/2

Table 6.1 Silicide technologies, processes and sample geometry used in this work [109].

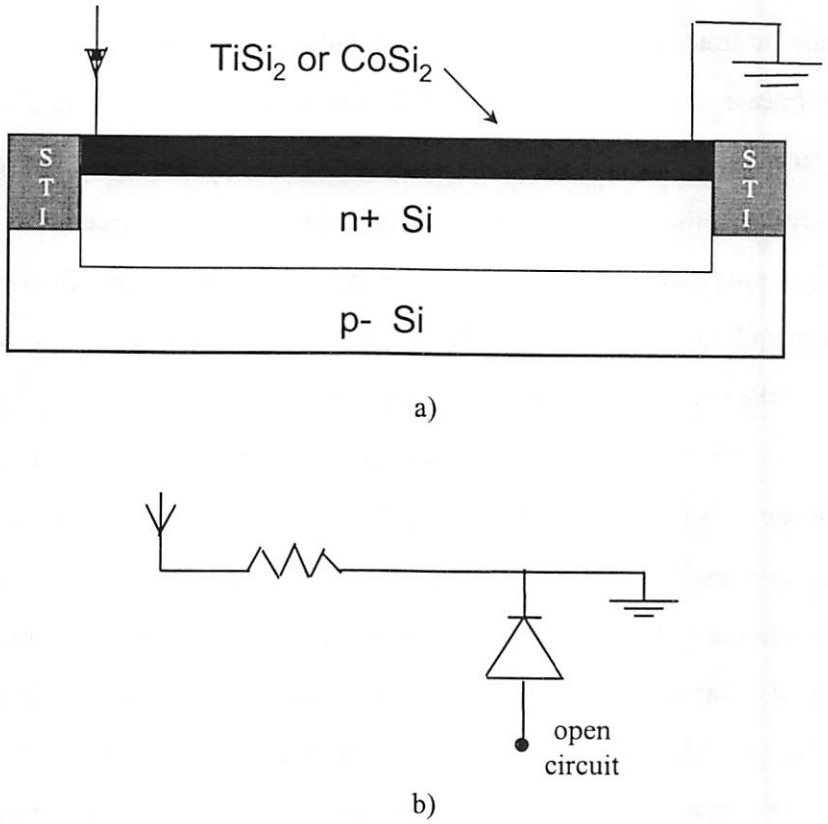


Figure 6.2 a) Schematic cross section of silicide structures on n+ Si used in this study. b) A lumped equivalent circuit for the silicide structures.

6.3 Characterization and Modeling of High Current Effects in $TiSi_2$ Films

6.3.1 Characterization of High DC (steady state) Conduction

Figure 6.3 shows the low DC current I-V characteristics for the $TiSi_2$ (Ge-PAI) on n+ Si structures with different *L/W* ratios. The curves indicate that under low current conditions the silicide films display ohmic behavior.

The high current I-V curves for these structures under DC stress conditions is shown in Figure 6.4. It can be observed that the I-V curves become non-linear in the high current regime due to self-heating.

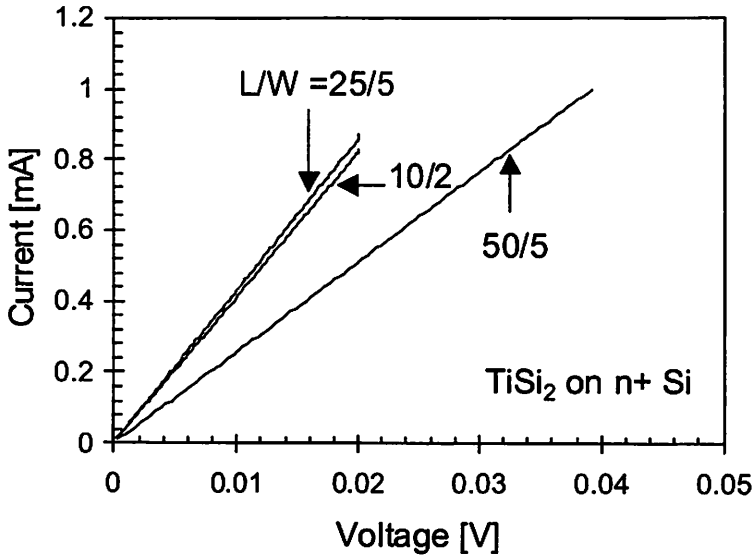


Figure 6.3 Low current I-V characteristics for the $TiSi_2$ films on n+ Si showing ohmic behavior.

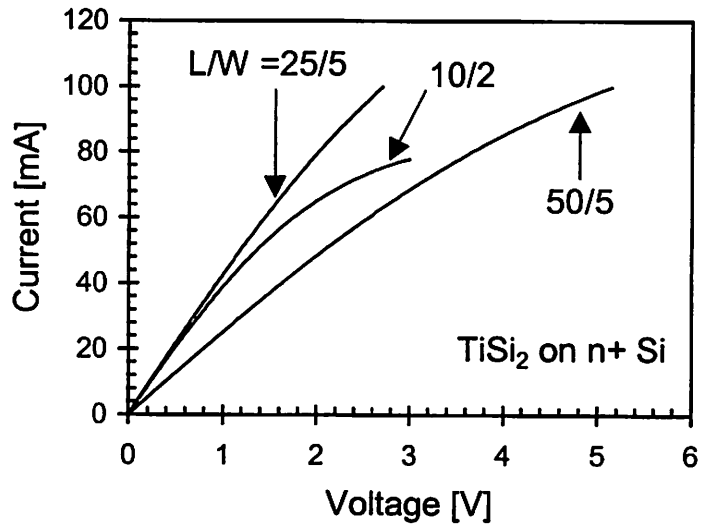


Figure 6.4 High current I-V characteristics for the $TiSi_2$ films on n+ Si showing non-linear behavior.

6.3.2 Model for Resistance Variation under High DC Stress

Current conduction in thin silicide films under high DC stress conditions can be modeled as following. The voltage, V, across the film under high current, I, can be expressed as,

$$V = I \cdot R = I \cdot R_0(1 + TCR\,\Delta T) = I[R_0 + B(I \cdot V)] \qquad (6.1)$$

where R_0 is the initial resistance of the film under low current stress, and $\Delta T = P.R_\theta$, as defined in equation (2.36). B is introduced as a parameter that depends on the sheet resistance, geometry, TCR, and thermal impedance of the structures and has the unit of Ω/Watt. Here, temperature rise is assumed to be proportional to the power dissipation, $I \cdot V$. Equation (6.1) can be rearranged to give [109],

$$R = \frac{R_0}{1 - BI^2} \tag{6.2}$$

The analytical model developed as equation (6.2) has been used to explain resistance rise with current and is shown in Figure 6.5. The values of B were found to be 16.5, 55, and 23 Ω/Watt for L/W = 25/5, 10/2 and the 50/5 respectively. Note that, B is linearly dependent on the geometry (L/W) and thermal impedance of these structures. Although, the R_0 is identical for the 10/2 and 25/5 structures, thermal impedance and hence B for the 10/2 structure is higher due to the smaller surface area in contact with Si. In order to separate the design geometry effects from materials issues, B/R_0 for identical L/W, may be used while comparing different technologies and process effects.

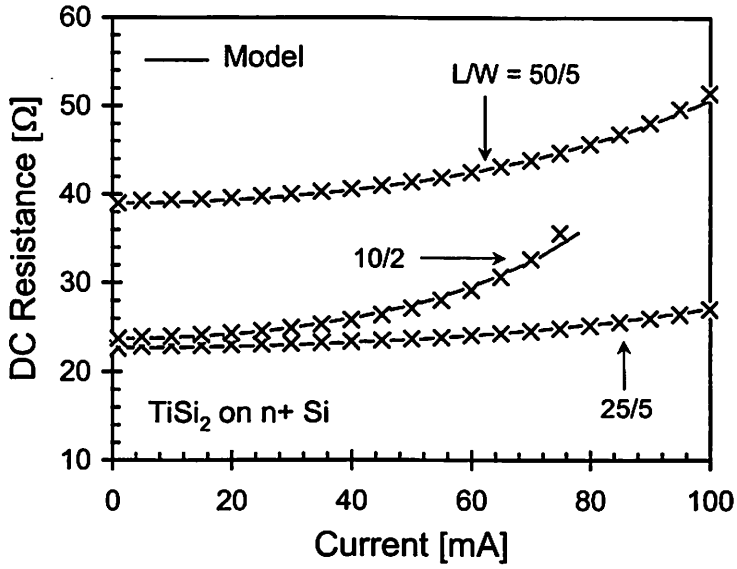


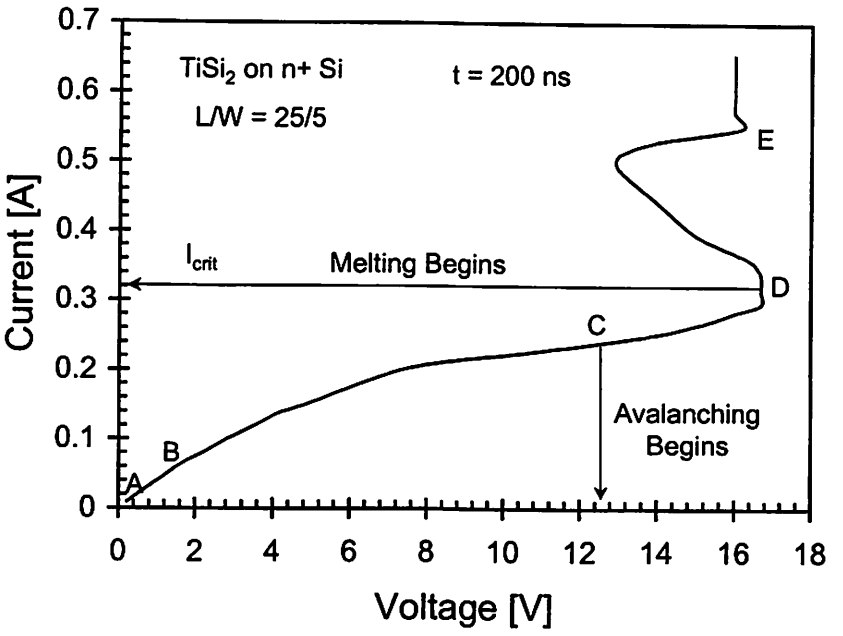
Figure 6.5 The model for resistance as a function of current through the silicide is shown to be in excellent agreement with data for various geometry.

6.3.3 Characterization of High Pulsed Current Conduction

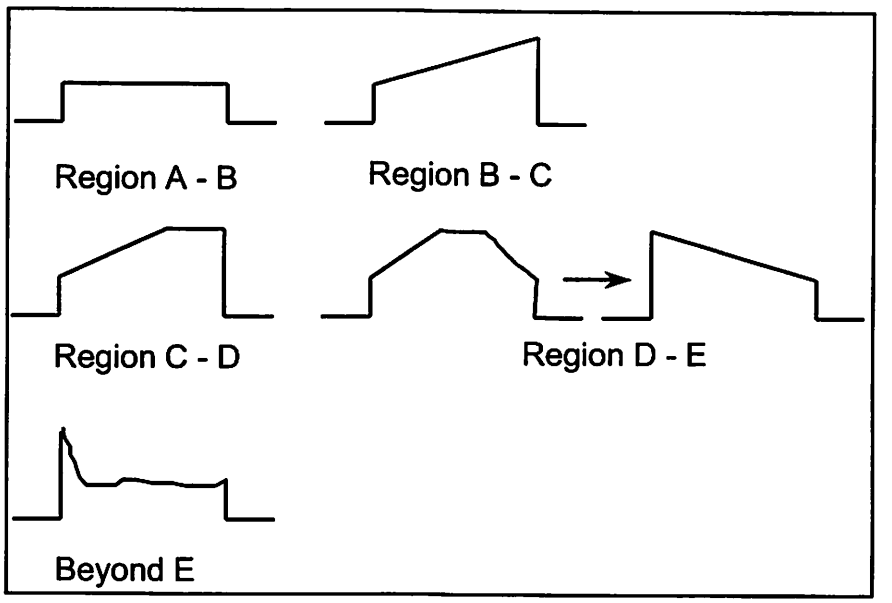
Figure 6.6(a) shows the high current I-V characteristics for a $TiSi_2$ film (L/W=25/5) on n+ Si measured with 200 ns pulses. The high current curve displays several characteristic regions as explained below. The instantaneous voltage pulse shapes in the various regions are shown in Figure 6.6(b). The voltage was always measured at the end of the pulse. The post-stress resistance of the film was also monitored after each pulse with a small DC current. Figure 6.7 shows the instantaneous resistance as a function of the pulse current corresponding to Figure 6.6(a) with the different regions.

In the region from A to B {in Figure 6.6(a)} the curve is almost linear, as expected for a constant resistance. Beyond point B the I-V curve becomes nonlinear due to self-heating of the silicide film and the voltage pulse can be observed to rise linearly with time, similar to results obtained in Chapter 3 for metal lines. At point C the underlying p-n junction begins to avalanche (at ~ 12.5 V determined experimentally in Figure 6.20) and the voltage pulse saturates. As the magnitude of the current pulse is increased further, beyond point D, the voltage pulse begins to fall with time after rising sharply. At this point the junction begins to melt and as a result the silicide-Si interface begins to degrade causing an irreversible increase in the post-stress resistance of the film.

This threshold point of damage is defined as I_{crit}. At point E, the junction fails completely due to spiking and the voltage level remains at a constant high value until the structure fails like a fuse upon further increase in current. The effects beyond point C leading to failure are discussed in detail in section 6.5.



a)



b)

Figure 6.6 a) High current I-V characteristics for a $TiSi_2$ film under a 200 ns pulsed current stress. Voltage is measured at the end of the pulse. b) The voltage pulse shapes in the different regions of the I-V curve.

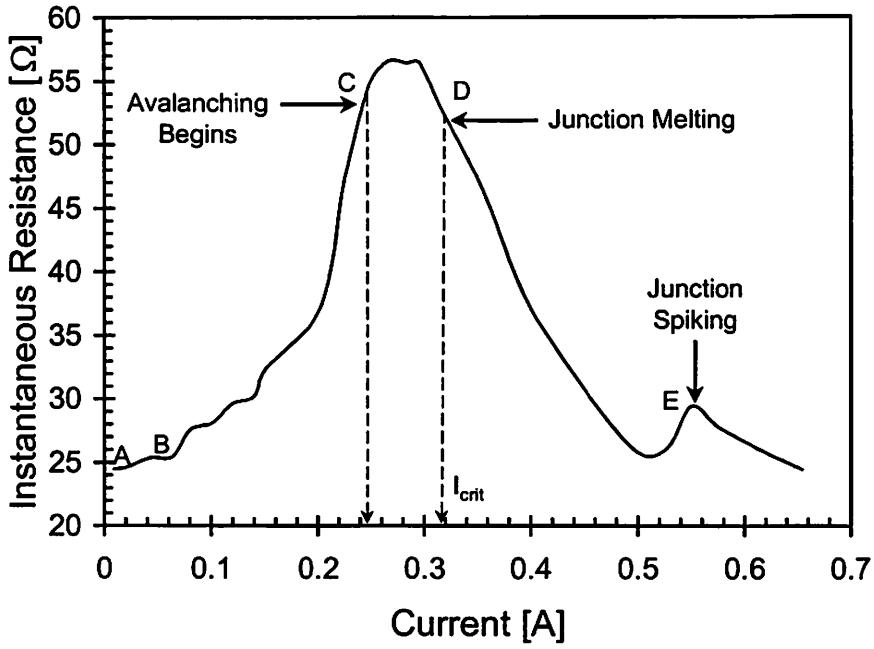


Figure 6.7 Instantaneous resistance of the silicide film as a function of the current amplitude for 200 ns pulses.

6.3.4 Model for Resistance Variation Under Pulsed Current

The model for high current conduction under DC stress is now extended to include time dependence. This model can be used to describe the I-V characteristics up or close to point C, i.e., the onset of avalanching in the underlying p-n junction. Rewriting (6.1) as,

$$V = I \cdot R = I[R_0 + F \cdot \Delta T] \tag{6.3}$$

where F is a parameter that depends on the sheet resistance, geometry and TCR. The pulse energy can be expressed as,

$$E = \frac{1}{2} I^2 \cdot t \cdot (R_0 + R) = C_\theta \cdot \Delta T \tag{6.4}$$

where R_0 is the initial resistance at the beginning of the pulse (t = 0) and R is the resistance at the end of the pulse. C_θ is the effective thermal capacity of the structure. Substituting ΔT from equation (6.4) in equation (6.3) gives [109],

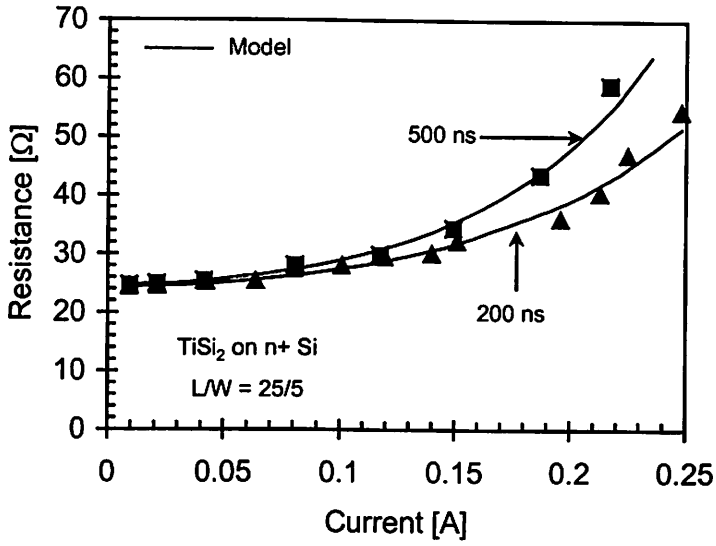


Figure 6.8 High current conduction model for $TiSi_2$ films under pulsed stress conditions.

$$R = R_0\left(\frac{1+\lambda I^2}{1-\lambda I^2}\right) \tag{6.5}$$

where $\lambda = (Ft)/(2C_\theta)$ in units of Ω/Watt or A^{-2}. The model developed in the form of equation (6.5) has been used to simulate high current conduction in the $TiSi_2$ film for two different pulse widths and the results are shown to be in good agreement with data in Figure 6.8. For the $TiSi_2$ film with $L/W = 25/5$, the values of λ were found to be 5.8 and 8.0 Ω/Watt for 200 ns and 500 ns pulse duration respectively. Increase in λ with t is not linear since the effective thermal capacity also increases with t, due to increasing heat diffusion into the surrounding.

6.4 Technology and Process Dependence of High Current Conduction

In this section, effects of technology and processes on the high current conduction under DC and pulsed current stress conditions will be examined. First a $CoSi_2$ technology will be compared with the $TiSi_2$ technology discussed in the last section. Secondly, results of using Mo implant on the high current effects in $TiSi_2$ will be compared with that for the Ge-PAI $TiSi_2$ process discussed in the last section. Finally, high current effects in $TiSi_2$ and $CoSi_2$ films formed on n+ poly-Si will be analyzed.

6.4.1 High Current Conduction in $CoSi_2$ Films

High current tests under DC and pulsed stress conditions were carried out on the $CoSi_2$ films formed on n+ Si by high temperature rapid thermal process (RTP) (see Table 6.1). Comparison will only be made to the $TiSi_2$ (Ge-PAI) technology discussed in the last section.

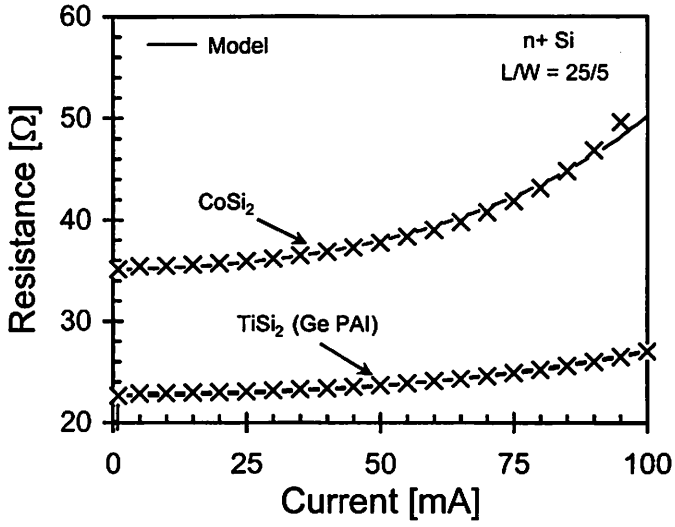


Figure 6.9 High current behavior of $CoSi_2$ and $TiSi_2$ films under DC stress conditions along with the model developed in section 6.3.

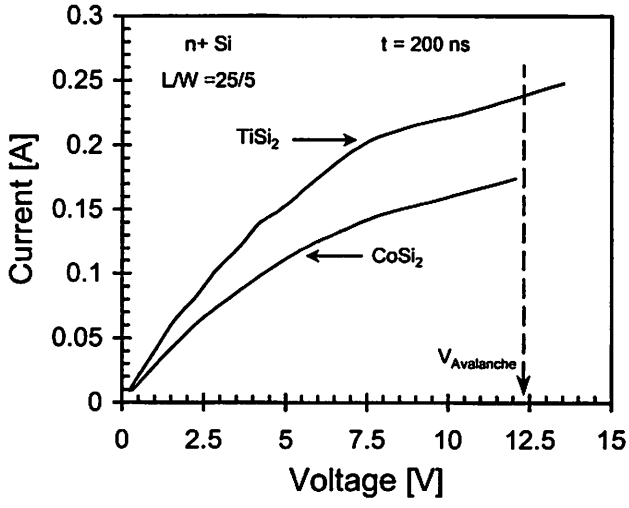


Figure 6.10 Comparison of high current conduction between $CoSi_2$ and $TiSi_2$ films under a 200 ns pulsed stress condition.

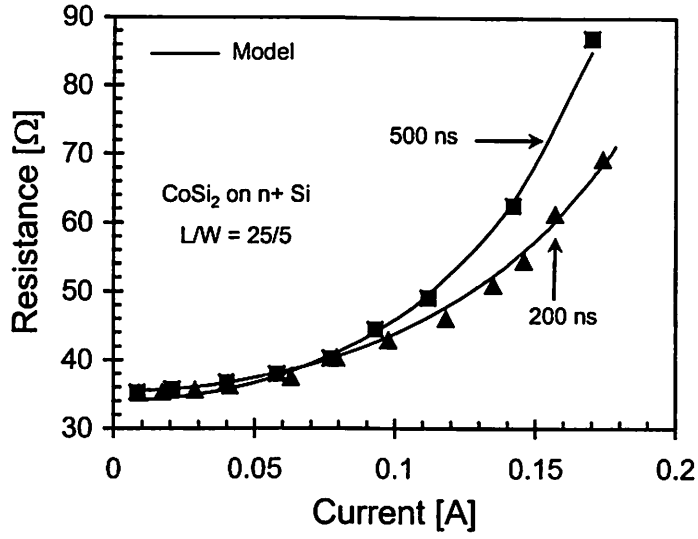


Figure 6.11 High current conduction model for $CoSi_2$ films under pulsed stress conditions.

Figure 6.9 provides a comparison of high current conduction under DC stress conditions between $CoSi_2$ and $TiSi_2$ with L/W = 25/5. The high current model developed in the last section as equation (6.2) is shown to hold for the $CoSi_2$ film as well, though with a higher value of B (= 30 Ω/Watt) as compared to 16.5 Ω/Watt for the $TiSi_2$. B/R_0 is slightly bigger for $CoSi_2$ than $TiSi_2$ (0.853 per Watt Vs 0.728 per Watt). This is due to the slightly higher TCR of $CoSi_2$.

Figure 6.10 shows the high current I-V curves (up to the avalanche voltage) for the $CoSi_2$ and $TiSi_2$ films under a 200 ns pulsed stress condition. In Figure 6.11 equation (6.5) has been used to simulate high current conduction in the $CoSi_2$ film for two different pulse widths and the results are shown in good agreement with data. For the $CoSi_2$ film with L/W = 25/5, the values of λ were found to be 10.5 and 14.8 Ω/Watt for 200 ns and 500 ns pulse duration respectively.

6.4.2 Comparison of Mo Implant and Ge Pre-Amorphization Implant (PAI)

Figure 6.12 compares the effect of using Mo implant with Ge-PAI before forming $TiSi_2$ films, on high current conduction under DC stress conditions. It can be observed that Mo implant makes the silicide resistance more strongly dependent on the current. This is verified by applying the model in equation (6.2) for L/W=10/2, which gives a slightly higher value for B/R_0 (= 2.485 W^{-1}) as compared to 2.317 W^{-1} for the $TiSi_2$ film with Ge-PAI. This result indicates that $TiSi_2$ films with Mo implant have a slightly higher TCR.

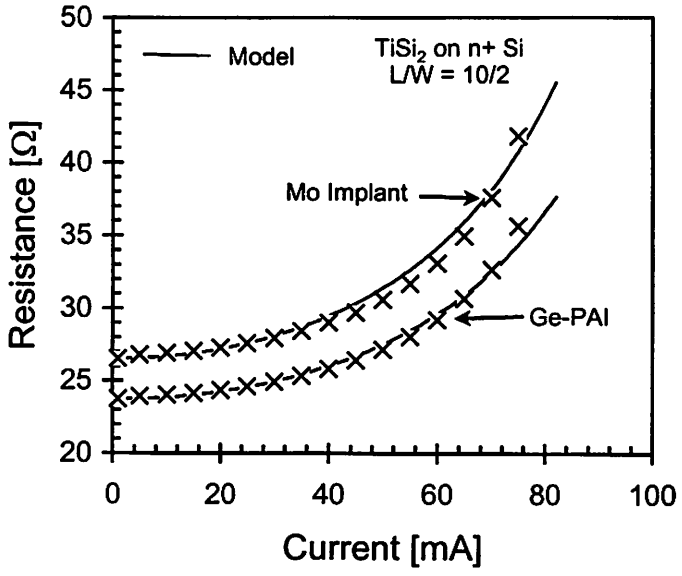


Figure 6.12 High current conduction in $TiSi_2$ film under DC stress condition showing the effect of Mo implant.

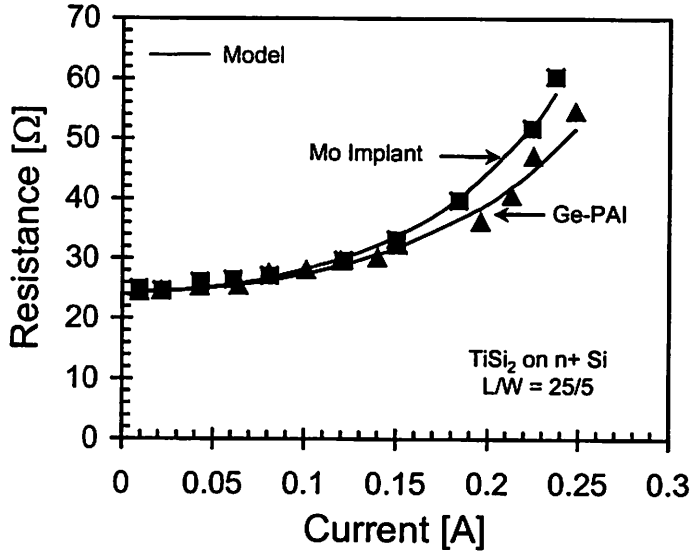


Figure 6.13 Effect of Mo implant on the high current conduction under a 200 ns pulse showing a stronger dependence of film resistance on the current.

The effect of Mo implant on the high current behavior under pulsed condition is also shown in Figure 6.13, for *L/W*=25/5. The Mo implanted $TiSi_2$ film's resistance increases more rapidly with current giving a higher value for λ (=7.2 Ω/Watt), the parameter from the model in equation (6.5),

as compared to 5.8 Ω/Watt for the Ge-PAI $TiSi_2$ film. Here R_0 is identical for the two cases. Again, the values of λ indicate that Mo implant results in a higher TCR.

The higher TCR resulting in higher self-heating and the fact that Mo implanted $TiSi_2$ films have lower thermal stability [102], would make them more susceptible to high current degradation. This was found to be consistent with measured values of I_{crit} that were lower by up to ~ 25%.

6.4.3 High Current Effects in $TiSi_2$ and $CoSi_2$ Films on Poly-Si

The schematic representation of the test structure is shown in Figure 6.14. The temperature coefficient of resistance (TCR) were measured to be 0.0025 ± 0.0002 and 0.0029 ± 0.0001 $^0C^{-1}$ for $TiSi_2$ and $CoSi_2$ films (formed on n+ poly-Si) respectively using low DC current. Figure 6.15 shows the high current curves under DC stress conditions for $TiSi_2$ films on n+ poly-Si for two different geometries. Curves for $TiSi_2$ on n+ Si are also included for comparison. It can be observed that the films on poly-Si display larger sensitivity to current. This is due to the higher thermal impedance of the oxide layer.

Figure 6.16 shows the high current effects under DC stress conditions in $CoSi_2$ films on n+ poly-Si. Curves for the $TiSi_2$ films are also included for comparison. The $CoSi_2$ films display even larger sensitivity of resistance to current, as expected. The model from equation (6.2) is again shown to be valid and the values of B were found to be 85 and 345 Ω/Watt for the 25/5 and 10/2 $TiSi_2$ films respectively, and 148 and 610 Ω/Watt for the 25/5 and 10/2 $CoSi_2$ films respectively.

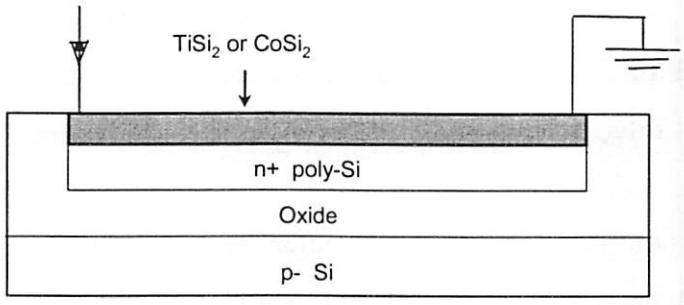


Figure 6.14 Silicide structures on n+ poly-Si used in this study.

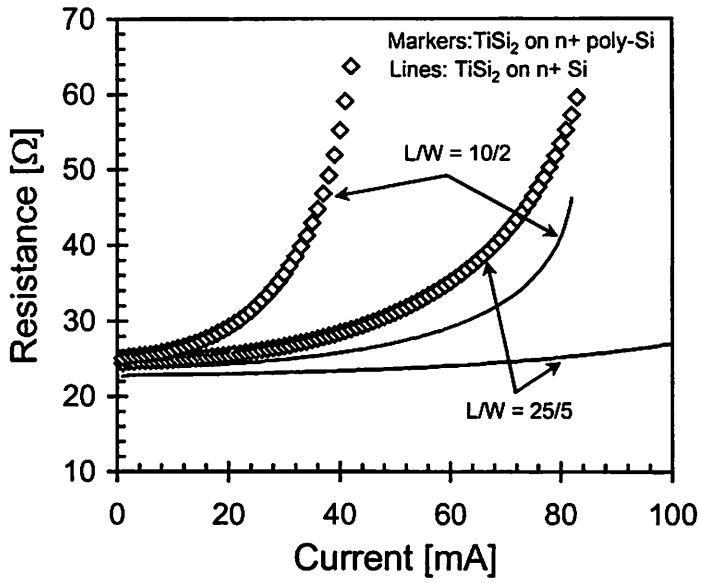


Figure 6.15 High current conduction in $TiSi_2$ films on n+ poly-Si under DC stress displaying larger current sensitivity of resistance.

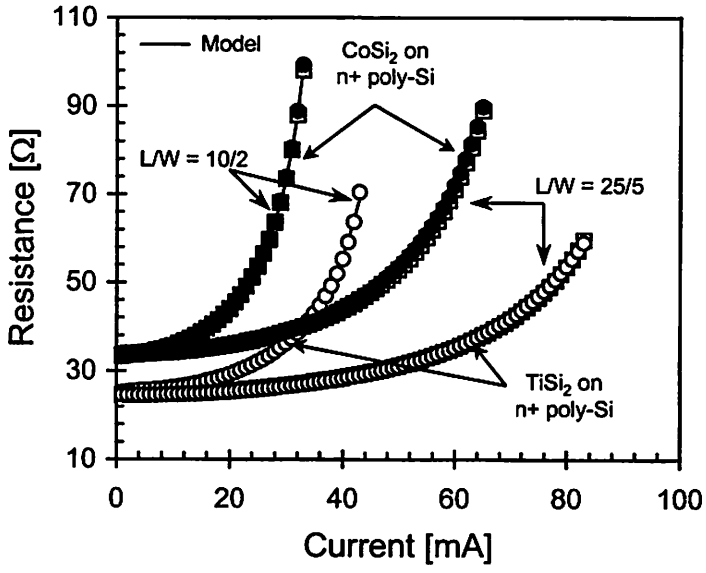


Figure 6.16 Resistance sensitivity to current under DC stress is larger for $CoSi_2$ films, compared to that of $TiSi_2$, formed on n+ poly-Si.

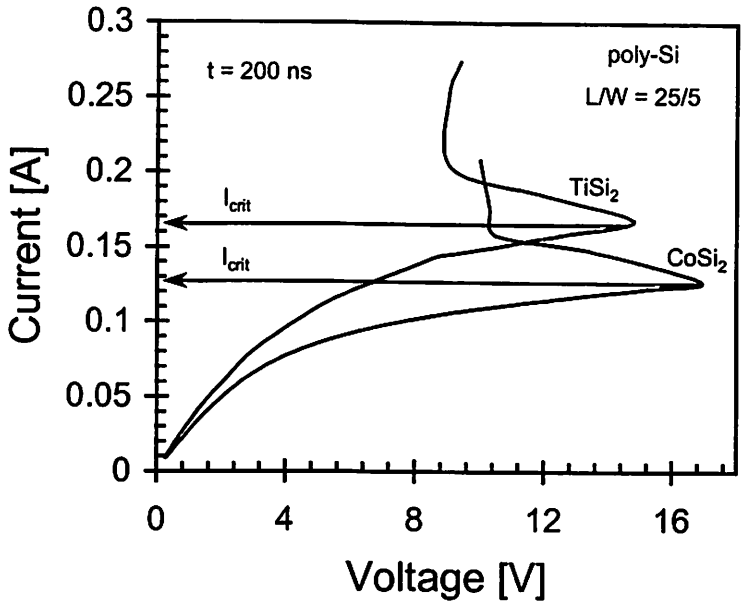


Figure 6.17 The high current I-V curves for $TiSi_2$ and $CoSi_2$ films formed on n+ poly-Si under a 200 ns pulsed stress condition.

The high current curves under pulsed stress conditions are shown in Figure 6.17 for $TiSi_2$ and $CoSi_2$ films on n+ poly-Si. The curves display characteristics similar to that of a silicide film on n+ Si {Figure 6.6(a)}. However, since there is no underlying junction, the resistance of the films continues to increase until the critical points. The snapback is believed to occur when the structure begins to melt. The failure currents are lower than that of silicides on n+ Si. These failure mechanisms will be discussed in detail in the next section. The high current model under pulsed current is shown to be in excellent agreement with the data in Figure 6.18, all the way till snapback. The λ values, for a L/W=25/5 and pulse width of 200 ns, were found to be 18 Ω/Watt (or A^{-2}) and 38 Ω/Watt for the $TiSi_2$ and $CoSi_2$ films respectively. The higher λ compared to the n+ diffusion structures is due to the higher sheet resistance and lower thermal capacity of the poly-Si structures.

Finally, the effects of different salicide technology and processes on the high current behavior of thin silicide films are summarized in Figure 6.19 for L/W=25/5. The parameter B, from the high current model under DC stress, in equation (6.2) and the parameter λ, from the high current

model under pulsed stress, in equation (6.5) are shown to be good monitors of the impact of technology and process variations on the current dependence of resistance. The difference in the values of B for silicide films on n+ Si are mainly due to different sheet resistance and TCR. For silicide films on poly-Si the difference is mainly due to higher thermal impedance. Likewise, for silicide films on n+ Si, the difference in the values of λ are due to different sheet resistance and TCR, while for silicide films on poly-Si the difference is attributed to lower thermal capacity.

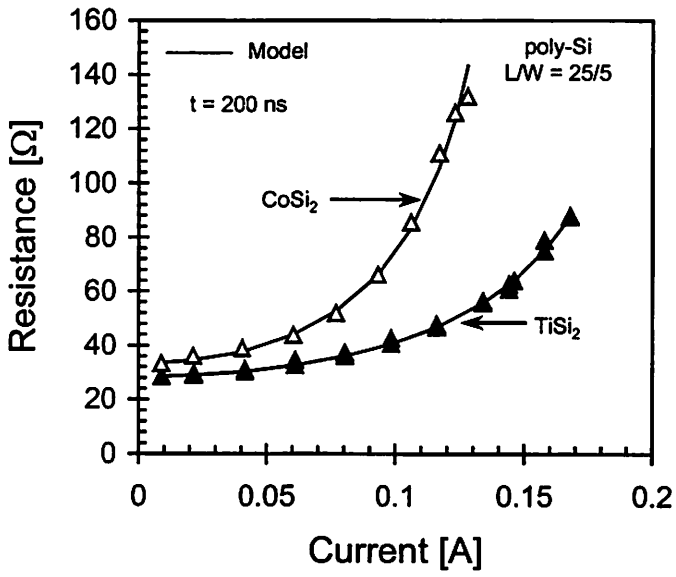


Figure 6.18 High current conduction model under pulsed stress condition shown for $TiSi_2$ and $CoSi_2$ films on n+ poly-Si.

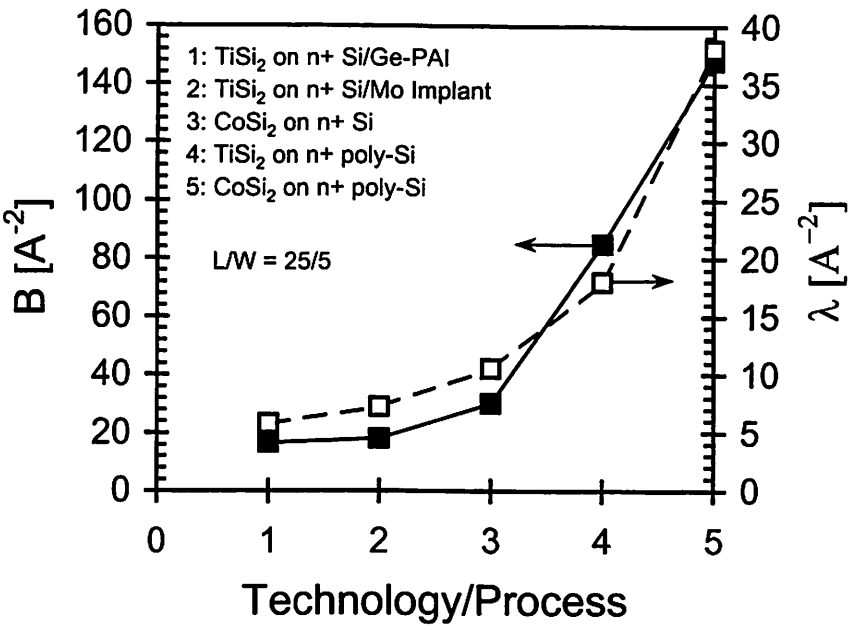


Figure 6.19 Impact of salicide technology and process variations on the current sensitivity of resistance [109].

6.5 Failure Mechanisms

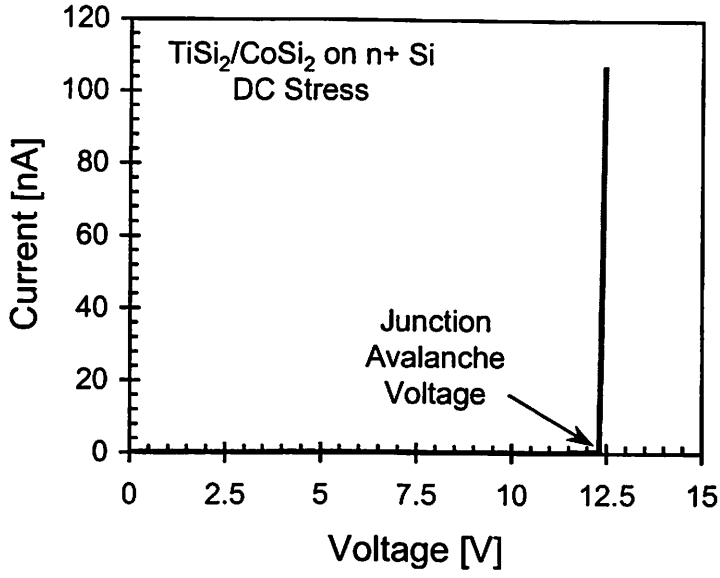
6.5.1 Failure Mechanisms of $TiSi_2$ and $CoSi_2$ films on n+ Si

For studying the failure mechanisms of silicide films under high current stress conditions, the TRT technique provides more insight into the physics of degradation. The high current DC stress causes catastrophic failures, after which it becomes difficult to separate out one contributing factor from another. This section will therefore focus on deciphering the causes of silicide film degradation and failure under pulsed conditions.

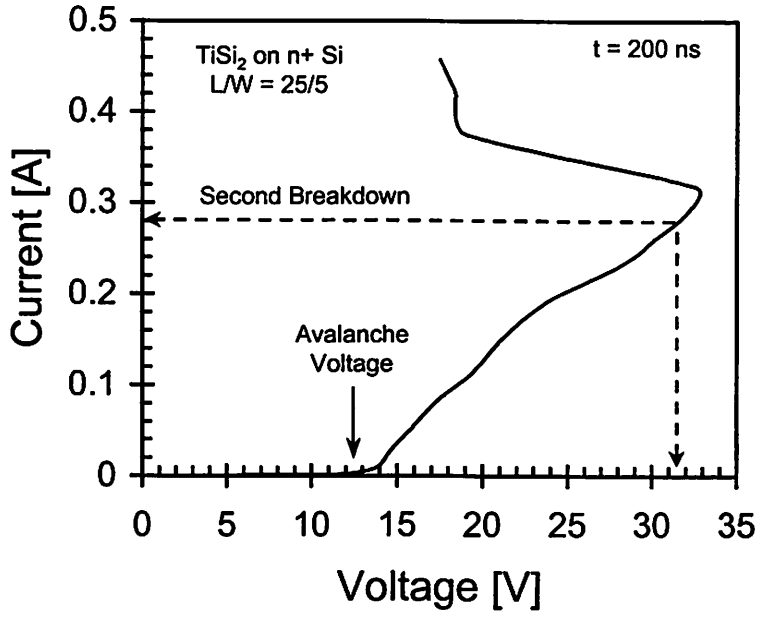
It was shown in Figure 6.6(a) that in region CD, the voltage pulse increased linearly with time, up to a point at which the underlying p-n junction carried significant current by avalanche and the voltage pulse saturated. This is verified by stressing the p-n junction under the silicide structure with a DC voltage as shown in Figure 6.20(a). Figure 6.20(b) shows the I-V characteristics of the junction under a 200 ns pulse stress. The junction avalanche voltage is in good agreement with that under DC stress {Figure 6.20(a)}. The voltage at the second breakdown point in Figure 6.20(b) is ~ 32 V which is nearly twice the second breakdown voltage

in Figure 6.6(a). This is due to the voltage drop across the high resistance due to the current flow through the substrate.

Now, as the magnitude of the current pulse was increased beyond point D {in Figure 6.6(a)}, the instantaneous voltage pulse started to fall with time after rising sharply and the resistance of the post-stress silicide film increased irreversibly. This threshold point in the I-V curve was used to define I_{crit}. No electrical or physical degradation (from TEMs) was observed until this point which indicates that the non-linearity in the I-V characteristics under DC and pulsed stress conditions is due to self-heating of the silicide films.



a)



b)

Figure 6.20 a) The reverse-bias DC I-V characteristics of the p-n junction showing avalanche voltage of ~ 12.5 V. b) The I-V characteristics of the p-n junction under a 200 ns pulse stress.

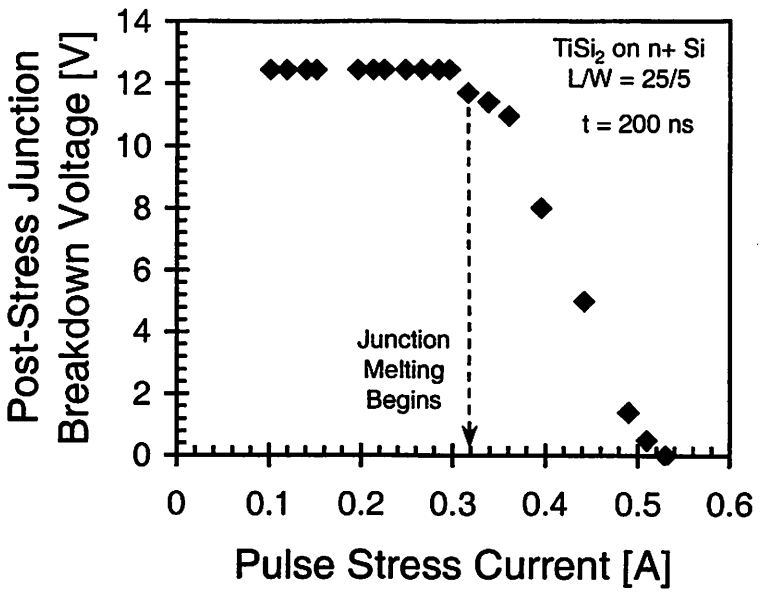
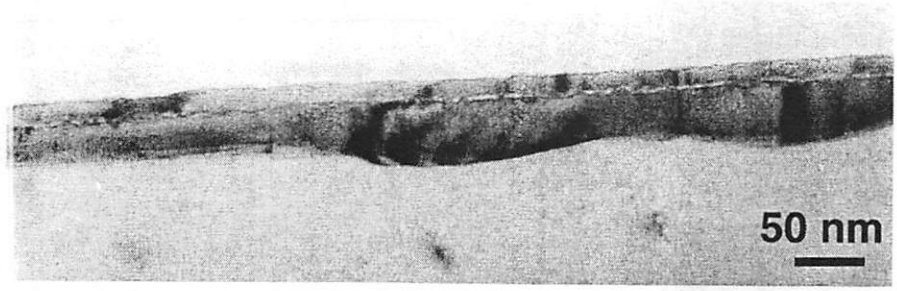


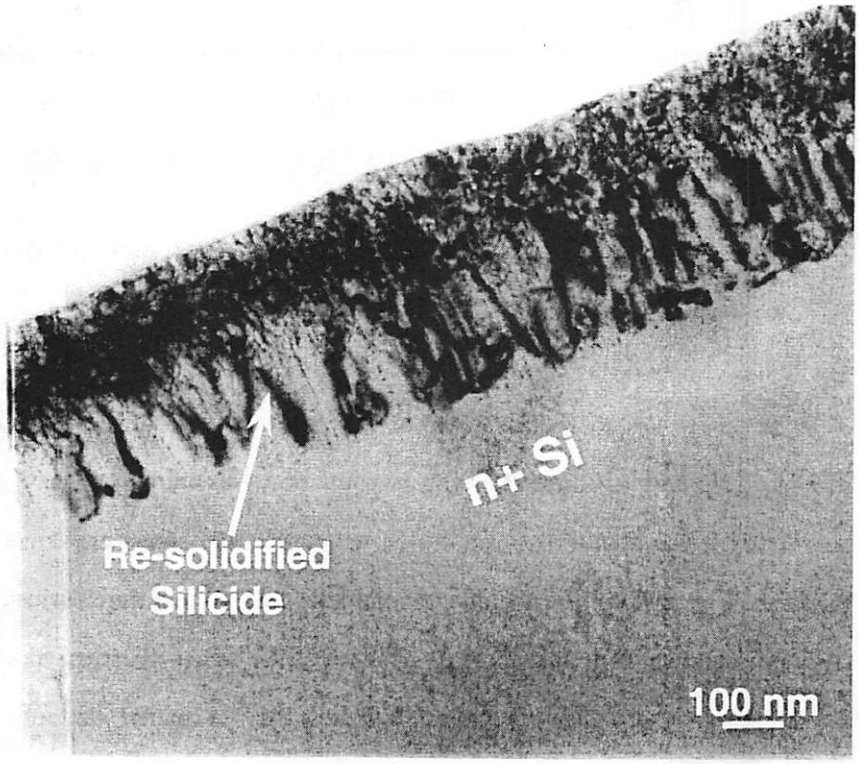
Figure 6.21 Junction avalanche voltage measured after each pulse of increasing amplitude decreases rapidly.

In order to comprehend the real cause for the observed degradation, the post-stress breakdown voltage of the underlying p-n junction was monitored with each pulse of increasing magnitude. It was found that at point D, where the post-stress resistance of the $TiSi_2$ film begins to increase irreversibly, the avalanche voltage of the p-n junction begins to decrease slowly.

This is shown in Figure 6.21 for a $TiSi_2$ film on n+ Si stressed with 200 ns pulses. The decrease in the breakdown voltage indicates that Si in the junction region has melted and polycrystalline formation is responsible for the lower breakdown voltage. This point can also be termed as the "second breakdown" point, since there is an irreversible change in the junction characteristics. The junction heating now causes the silicide to begin to melt and degrade due to morphological changes after re-solidification, causing an irreversible increase in the resistance of the structure.



a)



b)

Figure 6.22 TEM micrograph showing a) unstressed silicide film, and b) morphological change in a silicide film upon re-solidification after being stressed past the critical point by a 200 ns pulse.

Figure 6.22(a) shows a TEM micrograph of an unstressed silicide film and Figure 6.22(b) shows a degraded silicide layer with distinct morphological changes. Figure 6.23 shows the post- stress resistance rise of the *L*/*W*=25/5, $TiSi_2$ and $CoSi_2$ films monitored with a low stress DC measurement after each pulse of increasing magnitude. The critical current for both types of

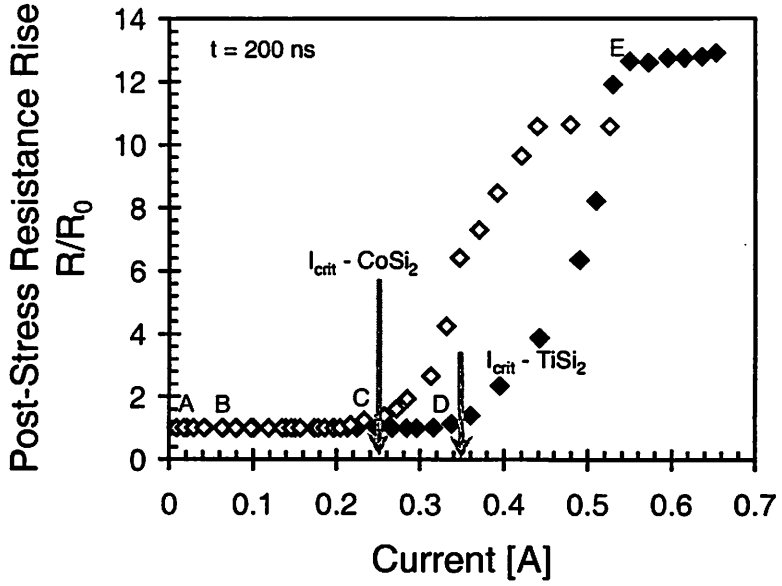
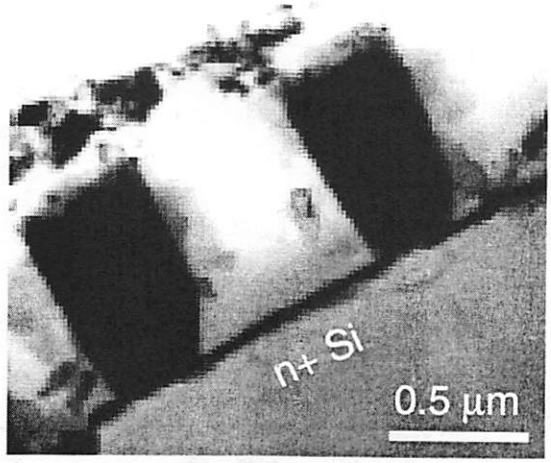
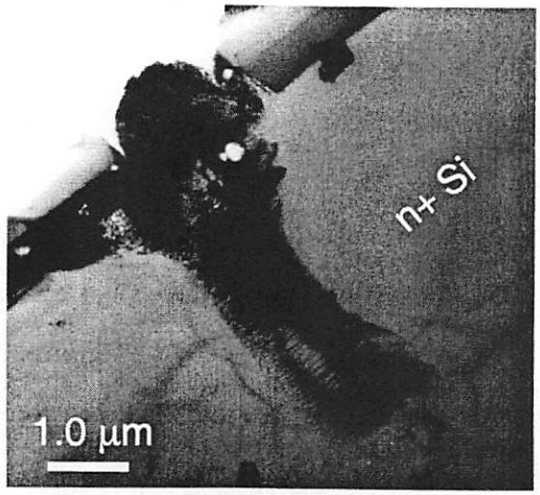


Figure 6.23 The post-pulse resistance rise for $TiSi_2$ (solid markers) and $CoSi_2$ (empty markers) films formed on n+ Si under a 200 ns pulsed stress condition. The letters indicate the various regions in the high current curve from Figure 6.6(a) for the $TiSi_2$ film.

silicide are indicated with vertical arrows. These are points at which the resistance of the films begins to increase. The irreversible change in the silicide resistance is introduced due to thermal effects resulting from second breakdown [100], [110], in the underlying diffusion region. This effect is introduced by current localization due to the negative resistance coefficient of Si beyond a critical temperature. This causes the silicide film to melt at hot spots. Since the melting point of $CoSi_2$ (~ 1326 ^{0}C) is lower than that of $TiSi_2$ (~ 1500 ^{0}C) [111], and that $CoSi_2$ has higher resistivity than $TiSi_2$, the $CoSi_2$ film degrades at a lower current. Beyond point E, the junction is spiked by the contact as shown by the TEM in Figure 6.24. The saturation point in Figure 6.23 (for $TiSi_2$) agrees well with the junction short point in Figure 6.21. When the contact spike reaches the junction, R/R_0 saturates as shown in Figure 6.23. The current now flows from one contact to another through the p substrate in parallel with the degraded silicide. The resistance in that region is greater than 10(12) times the initial resistance of the $CoSi_2$($TiSi_2$) film. This indicates that the sheet resistance of the new material is ~ 10(12) times higher than that of the silicides and is in rough agreement with that of a poly-Si film. This poly-Si film eventually fails like a fuse as current is increased further.



a)



b)

Figure 6.24 TEM micrographs showing a) unstressed contacts and b) damaged contact spike into the substrate.

6.5.2 Failure Mechanisms of $TiSi_2$ and $CoSi_2$ films on n+ poly-Si

Silicide films formed on n+ poly-Si degrade at lower current levels as pointed out in section 6.4.3, due to the higher thermal resistance of these structures. The post-stress resistance rise in these films after the critical (snapback) points in the high current I-V characteristics (Figure 6.17), follow the behavior of silicide films on n+ Si, shown in Figure 6.23. That is, the film resistance

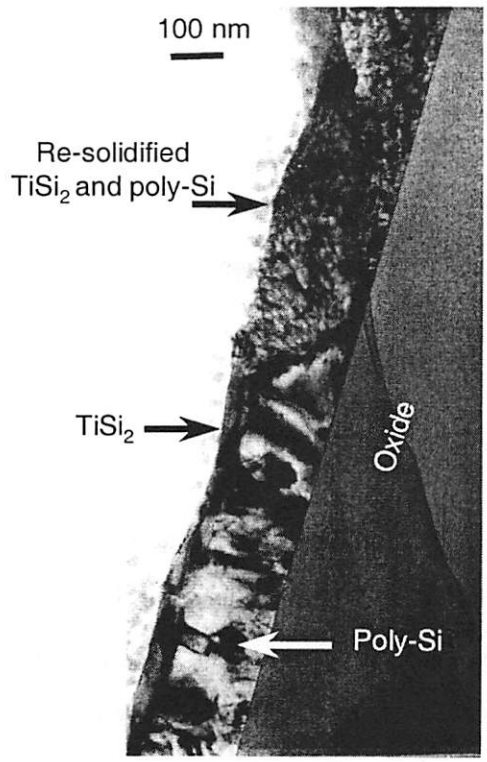


Figure 6.25 TEM micrograph showing a $TiSi_2$ film on n+ poly-Si showing a section that has undergone morphological changes upon re-solidification.

saturates to a value (~12 times that of the unstressed film) before failing like a fuse. The resistance rise is again associated with morphological changes in the film upon re-solidification after melting during the high current pulsing. This is shown by a TEM micrograph in Figure 6.25. The micrograph clearly shows a section of the film that is still intact and a section that has undergone morphological changes.

6.5.3 Geometry and Pulse Width Dependence of I_{crit}

The dependence of the critical current on the geometry of the films and pulse width can be modeled as follows. Since, it has been shown that the silicide films reach melt temperatures, the energy, E, required to melt a film of length L and width W can be expressed as,

$$E = a(W \cdot L) \tag{6.6}$$

where a is a proportionality constant. The energy for a pulse of amplitude I_{crit}, and duration t, can be expressed as,

$$E_{crit} = I_{crit}^2 \cdot R \cdot t = b \cdot \left(\frac{L}{W}\right) \cdot I_{crit}^2 \cdot t \tag{6.7}$$

Here b is proportionality constant. From equation (6.6) and (6.7), we get, under adiabatic assumption,

$$I_{crit} = K \cdot \frac{W}{\sqrt{t}} \tag{6.8}$$

where K depends on the material properties and the thickness of the silicide. I_{crit} is shown to be linearly dependent on W for different silicides in Figure 6.26. It can also be observed that the $CoSi_2$ films have lower I_{crit} compared to $TiSi_2$ films and poly-Si lowers these values even further. The pulse width dependence of I_{crit} is shown in Figure 6.27 for $TiSi_2$ and $CoSi_2$ films on poly-Si. The adiabatic model from equation (6.8) is shown to be in good agreement with the data. Note that the model in equation (6.8) will be less accurate for silicide films on Si, due to heat diffusion into the substrate.

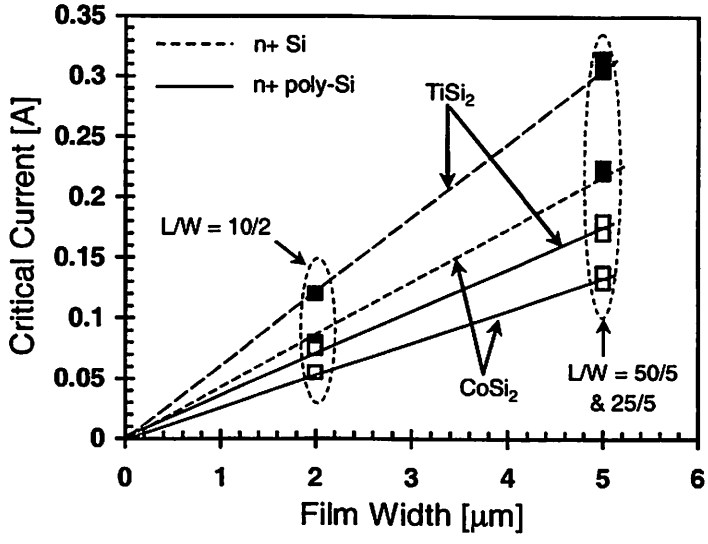


Figure 6.26 Proportionality of I_{crit} to the film width, W, shown for $TiSi_2$ and $CoSi_2$ films on n+ Si and n+ poly-Si. The pulse duration was 200 ns.

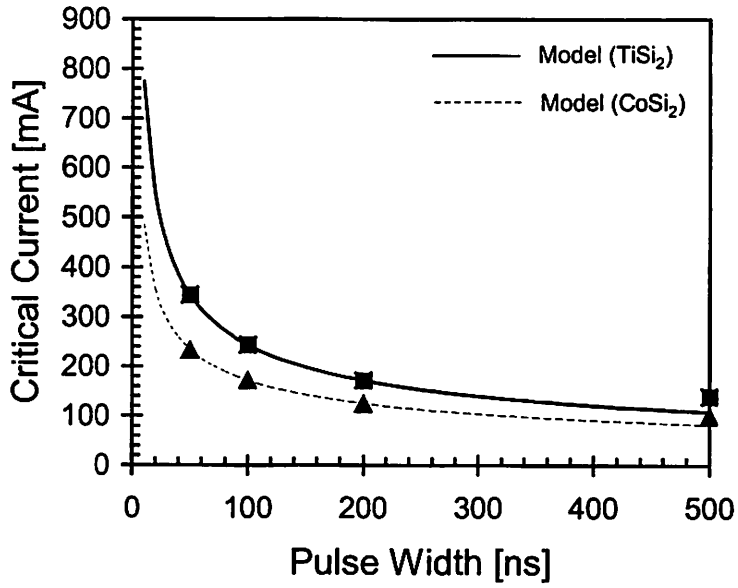


Figure 6.27 Pulse width dependence of I_{crit} for L/W = 25/5, $TiSi_2$ and $CoSi_2$ films on n+ poly-Si. The model is based on equation (6.8).

6.6 Summary

In this chapter the high current behavior of thin $TiSi_2$ and $CoSi_2$ films used in advanced CMOS technologies have been characterized and modeled under DC and pulsed stress conditions. High current conduction in silicides has been shown to be strongly affected by the technology and process conditions. The non-linear resistance rise of silicide films under DC and pulsed high current stress has been shown to be due to self-heating. Two physical parameters, B and λ, associated with DC and pulsed current stress, have been shown to be able to describe the sensitivity of the films to high current conduction. At high current an abrupt lowering of the resistance of silicided diffusions has been observed that can be important in the operation of advanced ESD/EOS and I/O buffer circuits. The sudden resistance drop in these structures under high current stress can be related to melting of the silicide structures. The critical current for this irreversible degradation has been shown to be determined by the silicide film width and the time duration of the pulse. Furthermore, it has been experimentally demonstrated that the $CoSi_2$ films

and silicides on poly-Si have a lower failure threshold under high current stress conditions. Extensive microstructure characterization of the silicide films using TEM has been performed to comprehend the evolution of silicide film degradation under high current stress. Thus, an understanding of these high current effects has been developed along with the physics of the failure mechanisms, which will enable the impact of technology design and scaling of silicide films to be defined for the reliability of deep sub-micron CMOS technologies.

Chapter 7

Impact of Thermal Effects on Small-Geometry Titanium Silicide-Si Contact Resistance

7.1 Introduction

As MOS transistors continue to be scaled in the deep sub-micron regime, the intrinsic parasitic series resistance is known to become increasingly important in limiting device performance [112]. Contact resistance *(R_c)* between the silicide and doped Si source/drain regions of advanced MOSFETs constitutes a significant fraction of this parasitic series resistance in the path of the drain current, thereby contributing to a loss in circuit performance. This parasitic source/drain resistance is known to affect the MOS I-V characteristics [113]- [116]. As discussed in Chapter 1, aggressive scaling of IC devices has resulted in smaller contact sizes and higher current densities. Characterization and modeling of temperature and current effects on the behavior of small-geometry-contact resistance is necessary to comprehend the full impact of their parasitic behavior on transistor and circuit performance. This is also desirable in order to comprehend self-heating effects on R_c. Presently, R_c values are empirically determined for a

specific technology. There is an increasing need for a quantitative model, which relates R_c to temperature, and technology variables such as silicide thickness, and source/drain doping concentration. Understanding the influence of thermal effects on the silicide/Si contact resistance is also very important for high density DRAM cells that employ high aspect ratio silicided contacts, and also for Schottky contact MOSFETs [117]. In this chapter a simple quantitative model is presented for R_c, which accounts for the effect of temperature and silicide thickness on current transport across the silicide/Si interface, and directly correlates R_c to temperature and interface doping concentration.

In addition, high-current effects on R_c have been analyzed and thermal parameters have been extracted. Furthermore, failure modes of the contact structures have been studied using the Transient Resistive Thermometry (TKT) technique described in Chapter 5.

7.2 Contact Resistance in Deep Sub-Micron MOSFETs

The contact resistance between the silicide and n+/p+ source-drain (S/D) in deep submicron CMOS transistors significantly contributes to the total parasitic S/D resistance. The various resistive components contributing to this parasitic resistance are schematically shown in Figure 7.1 for an NMOSFET.

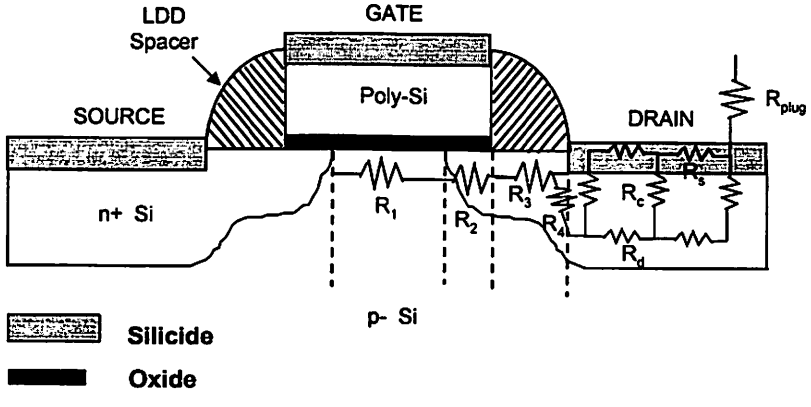


Figure 7.1 Schematic cross-section of a salicided LDD NMOSFET illustrating the various resistances in the path of the drain current.

R_1 is the channel resistance, R_2 is the accumulation layer resistance, R_3 is the resistance under the sidewall spacer, R_4 is the spreading resistance, and R_c is the contact resistance between the silicide and the n+ Si forming the drain and source. R_s is the resistance of the silicide layer and R_d is the resistance of the diffusion region underneath the silicide. All these resistances are dependent on the technology and process conditions. However, for a given technology and process, R_1, R_2, R_3, R_4, R_s, and R_d must decrease with scaling in order to maintain device performance. R_c depends on the interface properties of the metallurgical junction between the silicide and the Si and increases with scaling. Therefore, its contribution to the overall parasitic resistance increases with scaling.

The contact resistance contribution to the parasitic series S/D resistance (R_{sc} and R_{dc}) can be obtained using the transmission line model [118], and can be expressed as [115],

$$R_{sc} + R_{dc} = \frac{2\sqrt{\rho_c \rho_d^s}}{W \tanh\left(L/L_c\right)} \tag{7.1}$$

where ρ_c is the normalized contact resistance known as the specific contact resistance, $\rho_d{}^s$ is the sheet resistance of the diffusion under the silicide and W is the channel width. L is spacing between the spacer edge and the drain edge as shown in Figure 7.2. L_c is the transfer contact length [119]. It characterizes the distance over which the current (I_d) moves from the silicide into the diffusion.

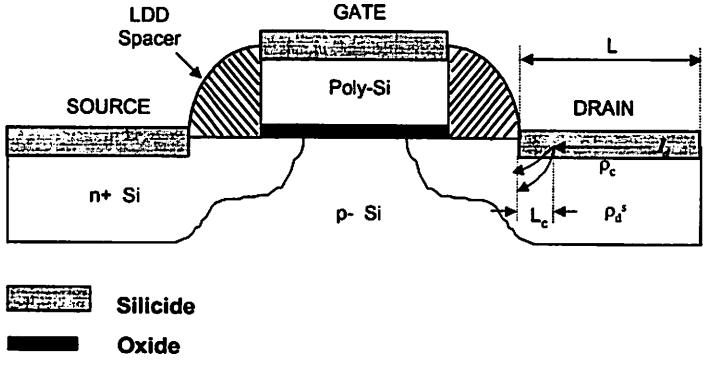


Figure 7.2 Schematic cross-section of a salicided LDD NMOSFET illustrating the contact transfer length over which the current enters the diffusion from the silicide layer.

L_c is given by

$$L_c = \sqrt{\frac{\rho_c}{\rho_d^{\ s}}} \tag{7.2}$$

Therefore, high values of L_c result from high specific contact resistivities. At $L = L_c$, from the silicide edge, the series resistance due to the diffusion $(\rho_d{}^s L_c)/W$ exactly equals the contact resistance due to the interface, $\rho_c/(L_c W)$.

It can be observed from equation (7.1) that for large values of L_c, the series S/D resistance will depend on the layout of the diffusion. The following two limiting cases in equation (7.1) are useful to consider: For L >> L_c, equation (7.1) reduces to

$$R_{sc} + R_{dc} = \frac{2\sqrt{\rho_c \rho_d^{\ s}}}{W} \tag{7.3}$$

which is the ideal case for the contact resistance contribution to the total series S/D resistance. In this case, the series S/D resistance is independent of the field edge spacing to the gate and is inversely proportional to transistor width. For L << L_c, equation (7.1) reduces to

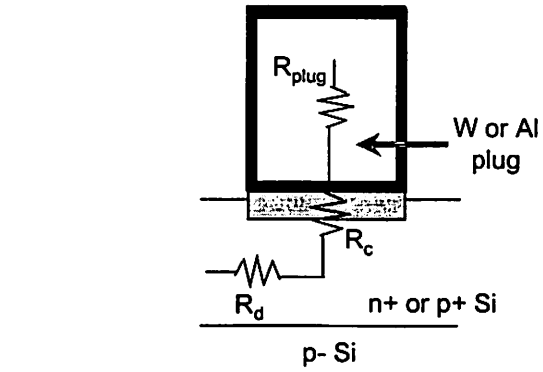
$$R_{sc} + R_{dc} = \frac{2\rho_c}{LW} \tag{7.4}$$

which is the contact resistance due to the entire source and drain contact area.

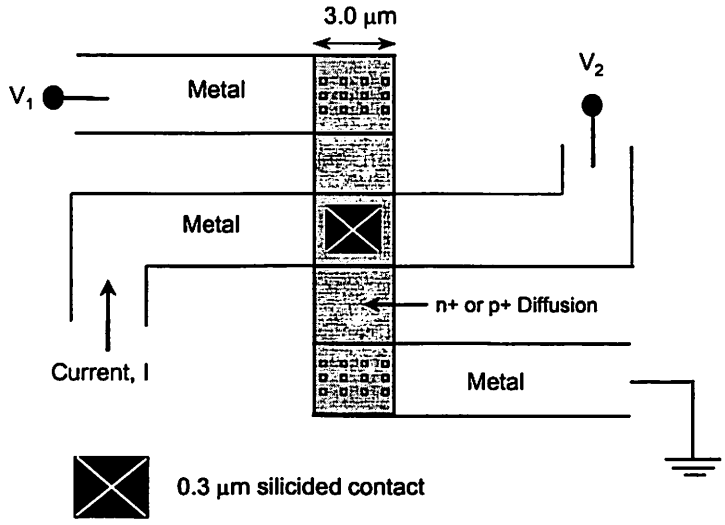
The goal of salicide technology is to achieve a low enough value of ρ_c such that equation (7.3) applies for all device geometries. However, in either case, since ρ_c remains nearly invariant, and L decreases with scaling, it can be deduced from equation (7.3) and (7.4) that the contact resistance contribution to the total parasitic series S/D resistance will increase with scaling.

7.3 Test Structure

In order to separate the effect of R_c from the other parasitic resistances shown in Figure 7.1, small-geometry vertical contact structures were used in this study as shown schematically in Figure 7.3(a). Kelvin-type test structures were employed to measure the R_c between the silicide and the n+ or p+ Si. A schematic of the test structure is shown in Figure 7.3(b).



a)



b)

Figure 7.3 a) Schematic cross-section of the contact structure used in this study, b) Schematic view of the Kelvin structure used for contact resistance measurements.

R_c was obtained by forcing a current (I) through two terminals of the Kelvin structure, and measuring the voltage at the other two terminals [120]- [122]. That is,

$$R_c = \frac{(V_2 - V_1)}{I} \tag{7.5}$$

Equation (7.5) assumes 1-D current flow through the contact and across the interface [123]-[125].

However, the contact resistance obtained using the Kelvin test structure involves some parasitic resistance due to two-dimensional current crowding effects [126], [127], which results in overestimation of R_c and also in a sublinear dependence of R_c on contact area [128]- [130].

Hence, the Kelvin measurement provides a lumped value, which consists of the resistance of the contact plug (R_{plug}), the interface contact resistance (R_c), and the parasitic resistance due to 2-D current crowding ($R_{2\text{-}D}$). However, since $R_{plug} << R_c$, it can be assumed that the Kelvin measurement provides a lumped value of R_c and $R_{2\text{-}D}$. A two-dimensional analytical model to extract $R_{2\text{-}D}$ for the Kelvin structure is given in [131]. The Kelvin structures used in this study were designed to minimize this 2-D effect, and $R_{2\text{-}D}$ constituted a small fraction of the total measured contact resistance.

7.4 Device Fabrication and Experiments

Silicon wafers were processed through a single-level-metal, salicided, 0.25-μm industrial CMOS flow. Chemical mechanical planarization (CMP) was used for the polysilicon-metal-dielectric (PMD). Contacts were patterned using deep ultra-violet (DUV) lithography and the oxide was etched in a high-density plasma (HDP) etcher. Prior to contact liner/barrier metallization, the wafers were cleaned in dilute HF. Titanium liner films were deposited via HDP physical vapor deposition (PVD) or $TiCl_4$ based chemical vapor deposition (CVD), which formed the silicide. Following the liner deposition the wafers were annealed followed by CVD TiN barrier deposition. W-plugs were formed by CVD tungsten deposition and etchback. Al plugs were formed via force-fill process. Al-Cu was then deposited, patterned, and etched prior to the oxide/ nitride passivation deposition. A TEM micrograph showing the cross section of a 0.3-μm W-plug contact is shown in Figure 7.4.

The different contact technologies that were studied are listed in Table 7.1. In all the measurements n+ Si was negatively biased relative to $TiSi_2$ and p+ Si was positively biased relative to $TiSi_2$. This was done to get majority carrier injection from the Si into the silicide. The $TiSi_2$ thickness was measured using TEM. The interface doping concentration at the silicide-Si interface was obtained using the SIMS measurements done prior to silicidation and their variation with depth is shown in Figure 7.5.

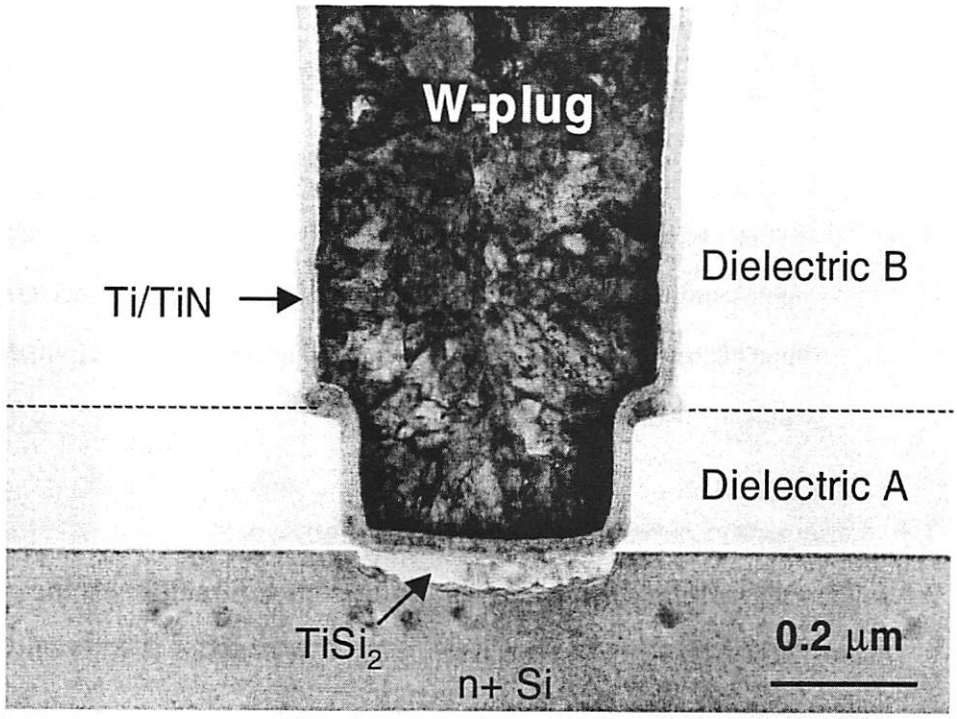


Figure 7.4 TEM micrograph of a 0.3 μm silicided (~ 35nm) W-plug contact.

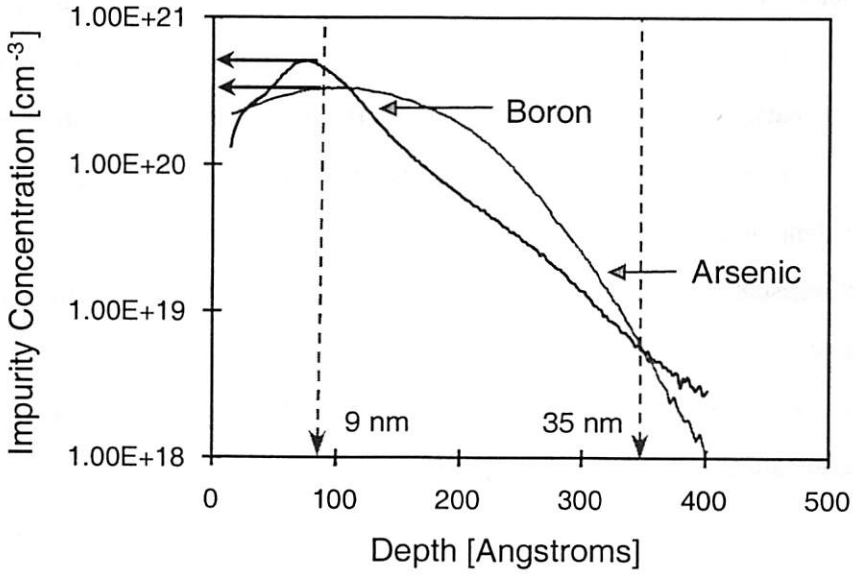


Figure 7.5 The impurity doping concentration variation with depth obtained using SIMS.

Contact Plug Material	Contact Size	Substrate	TiN liner Thickness	$TiSi_2$ Thickness
CVD-W	0.30 μm	n+/p+ Si	20 nm	35 nm
CVD-W	0.30 μm	n+/p+ Si	20 nm	9.0 nm
FF-Al	0.30 μm	n+/p+ Si	20 nm	9.0 nm

Table 7.1 Contact technologies evaluated in this study. The contact plug processes are chemical vapor deposited (CVD) W and force-fill (FF) Al. Silicide was formed from physical vapor deposited (PVD) Ti in all cases. The sheet resistance of the n+/p+ diffusion region is ~ 50 Ω/square.

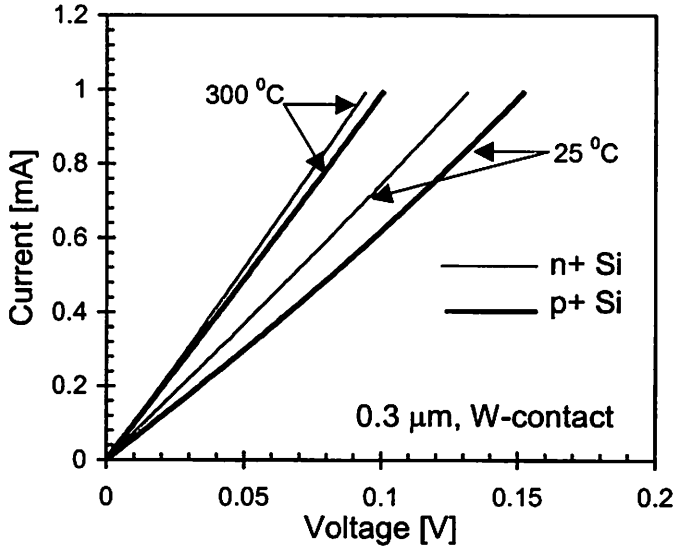
7.5 Characterization of Thermal Effects on R_c

7.5.1 Temperature and Low Current Effects on Contact Resistance

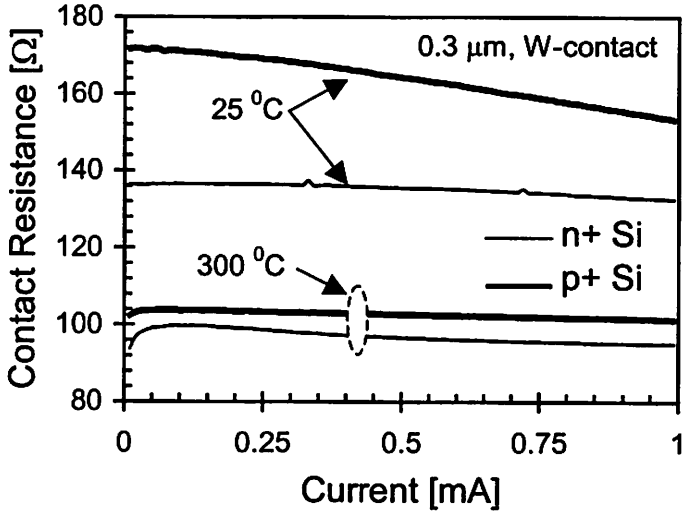
Figure 7.6(a) shows the I-V characteristics of W-contacts with 35 nm silicide. It can be observed that the contacts to p+ Si have higher resistance and greater nonlinearity in the I-V characteristics. Figure 7.6(b) clearly shows that the resistance of contact to p+ Si is more sensitive to current and temperature.

Figure 7.7 depicts the behavior of the contact resistance with ambient temperature for the W-contact with thick silicide (35 nm). Kelvin structures were used for the contact resistance measurements. A small current corresponding to a voltage < 0.02 V was used for these measurements. Also, since the resistance of a 0.3 μm W or Al plug is only ~ 2-3 Ω, most of the contact resistance comes from that of the $TiSi_2$/Si interface barrier. Thus R_c values are independent of the contact plug material. Under such small current (voltage) conditions the contact resistance values were found to be weakly dependent on the polarity of the bias. Therefore, all the contact resistance measurements were carried out under a single polarity condition as stated earlier. It can be observed from Figure 7.7 that contact resistance decreases with temperature and that contact to p+ Si shows bigger temperature sensitivity.

This explains the larger current sensitivity of the contact to p+ in Figure 7.6b. Being more sensitive to temperature, it is more sensitive to self-heating and therefore more sensitive to current.



a)



b)

Figure 7.6 a) I-V characteristics of W-contact structures with 35 nm silicide at two different temperatures. b) Contact resistance sensitivity with current (low current regime) at two different temperatures for W-contacts to n+ and p+ Si.

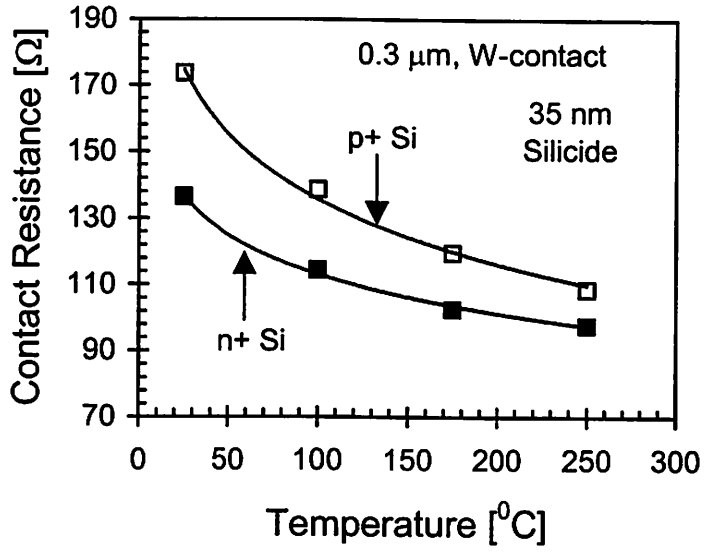


Figure 7.7 Contact resistance variation with temperature for contacts to n+ and p+ Si. The p+ contact shows larger temperature sensitivity.

Figure 7.8 shows the I-V characteristics of both W and Al plug contacts to n+ Si where the silicide thickness is only ~ 9 nm. It is observed that unlike the W-contact with the thicker silicide, these contacts are more ohmic in character and shows very little sensitivity to temperature, and therefore to current.

As the silicide thickness increases more dopants may segregate into the silicide and the interface moves down into the Si where the doping concentration may be lower. Thinner silicide therefore results in a silicide-Si interface with a higher impurity doping concentration. Higher doping concentration with thin silicide results in a narrower depletion region (smaller barrier width), enhanced tunneling, and lower contact resistance. Figure 7.9 shows similar results for contacts with thin silicide to p+ Si. It can also be noticed from Figure 7.8 and Figure 7.9 that for the thin silicide case the contact resistance to n+ and p+ Si is nearly identical unlike the thick silicide case. Figure 7.10 shows the contact resistance sensitivity to current for the thin silicide. It can be observed that the contact resistance remains almost invariant even at 300 ^{0}C as compared to that of the thick silicide case shown in Figure 7.6(b).

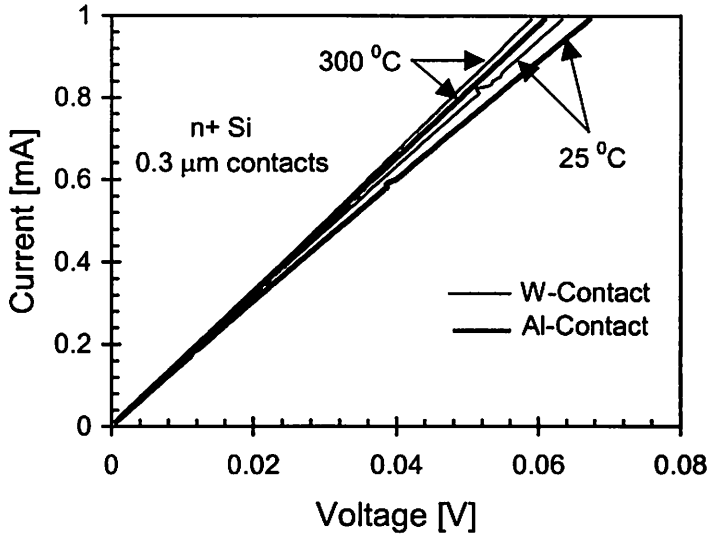


Figure 7.8 I-V characteristics of W and Al plug contacts to n+ Si with 9-nm silicide gives nearly equal contact resistance and are only slightly temperature sensitive.

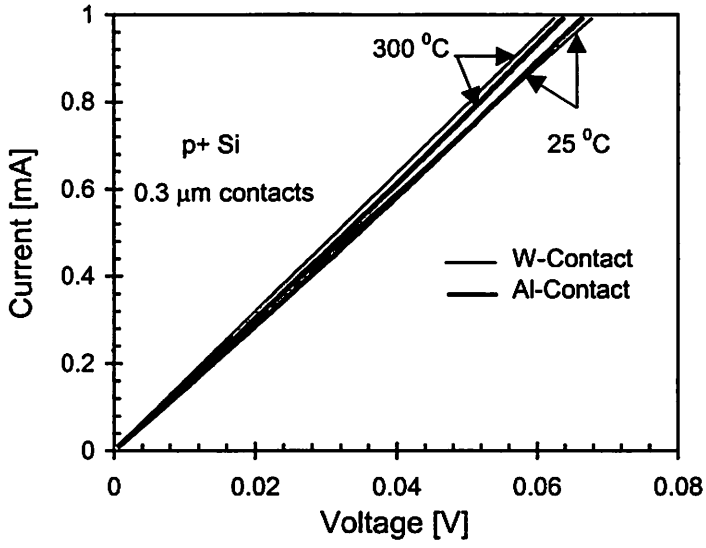


Figure 7.9 I-V characteristics of W and Al plug contacts to p+ Si with 9 nm silicide gives nearly identical contact resistance values as compared to those of n+ Si and are also only slightly temperature sensitive.

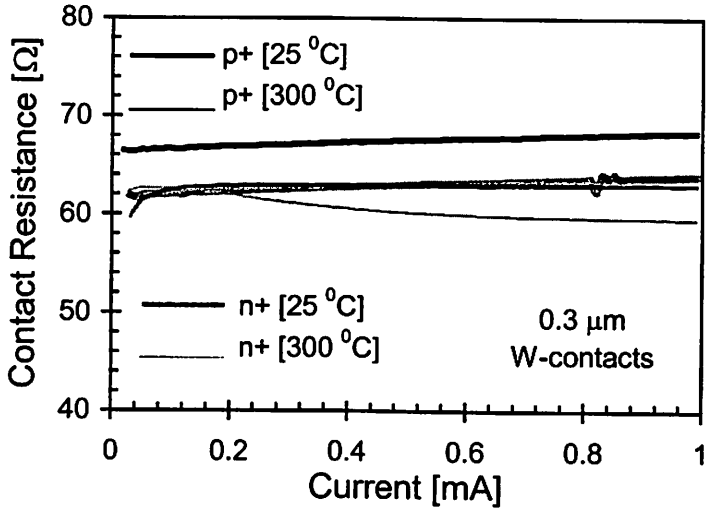


Figure 7.10 Contact resistance sensitivity with current (low current regime) at two different temperatures for W-contacts to n+ and p+ Si with 9 nm silicide.

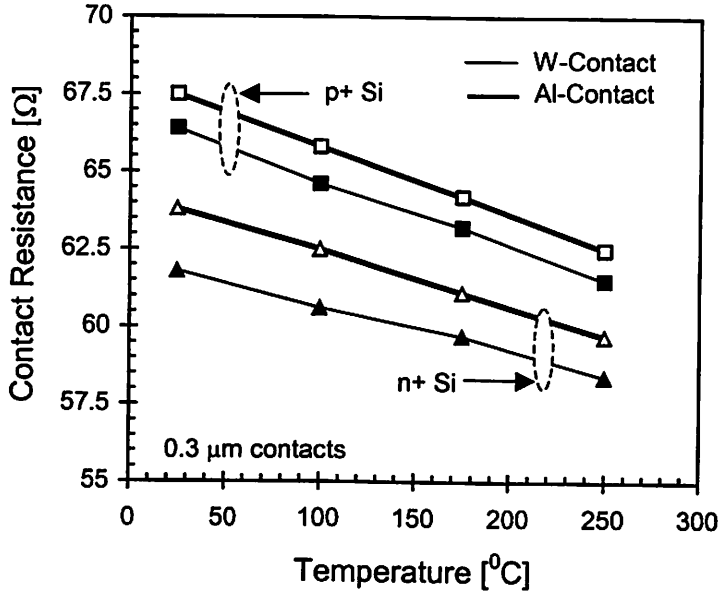


Figure 7.11 Temperature sensitivity of the contact resistance of W and Al plug contacts with 9 nm silicide.

Figure 7.11 shows the temperature sensitivity of the contacts to n+ and p+ Si for both the W-plug and the Al plug processes with 9 nm thick silicide. It is observed that these contacts have smaller contact resistance, and decrease very slowly with temperature. All these observed

differences in the temperature and silicide thickness sensitivity of R_c will be explained with a model presented in the next section.

7.6 Quantitative Model for $TiSi_2$-Si Contact Resistance

7.6.1 Carrier Transport Mechanisms

In order to analyze the effect of temperature and current on the contact resistance between the silicide and Si, it is important to understand the mechanisms responsible for carrier transport at the metal-semiconductor interface. The key mechanisms responsible for carrier transport across the metal-Si interface are the thermionic emission (TE) [132] - [134], field emission (FE) [134], [135], and thermionic field emission (TFE) [135], [136].

Current conduction due to TE results from carriers being thermally excited over the metal-semiconductor barrier, whereas for FE carriers simply tunnel through the barrier, as illustrated in Figure 7.12. TE tends to dominate in lowly doped semiconductors, where the barrier thickness (width of space charge region) is large enough to prevent direct tunneling. For heavily doped semiconductors, the barrier is thin and carriers simply tunnel through. In the intermediate doping range, the carrier transport is known to be dominated by TFE. In this mechanism, thermally excited carriers reach an energy level where the barrier is narrow enough for tunneling to occur

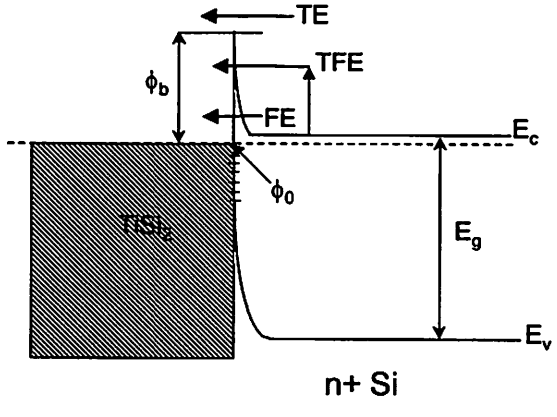


Figure 7.12 Metal-Si energy band diagram illustrating various transport mechanisms. ϕ_b is the silicide-Si barrier height, and ϕ_0 denotes the energy level at which the Fermi level is pinned.

The dominant carrier transport mechanism across the metal-semiconductor interface at a given temperature is determined by the ratio kT/E_{00} [137]. E_{00} represents a characteristic energy that is related to the tunneling probability, and is defined as

$$E_{00} = \frac{q\eta}{2}\sqrt{\frac{N}{\varepsilon_s m_t^*}} \tag{7.6}$$

where N is the doping concentration at the interface, ε_s, is the semiconductor permittivity, and m_t^* is the tunneling effective mass. In equation (7.6), the increase in E_{00} with N reflects the barrier getting thinner at higher doping levels, leading to increased tunneling. The ratio kT/E_{00} is a measure of the relative importance of the thermionic process with respect to the tunneling process. For low interface doping concentration, $kT/E_{00} >> 1$ and TE dominates. When interface doping concentration is high $kT/E_{00} << 1$, and FE dominates. In the intermediate doping range $kT/E_{00} \sim 1$ and TFE dominates.

For deep submicron CMOS technologies, $N > 10^{18}$ cm^{-3}. Therefore, the ratio of kT/E_{00} is either much less than one or around one, depending on the temperature. This implies that either FE or TFE will dominate the carrier transport mechanism, depending on the temperature.

7.6.2 Specific Contact Resistance under FE and TFE

The specific contact resistance (ρ_c) is defined as [132],

$$\rho_c = \left(\frac{dJ}{dV}\right)^{-1}_{V=0} \tag{7.7}$$

where J is the current density, and V is the applied voltage across the metal-semiconductor contact. ρ_c can be determined from equation (7.6), once the theoretical relation for the current-voltage characteristics accounting for the appropriate physical mechanisms responsible for carrier transport across the metal–semiconductor interface is known. Also, ρ_c is related to the contact resistance, R_c through the contact area, A, as [123],

$$R_c = \frac{\rho_c}{A} \tag{7.8}$$

The expression for the specific contact resistance, which is valid for both FE and TFE, has been shown to be given by [138],

$$\rho_c = \frac{k}{qA^*T} c \exp\left(\frac{q\phi_b}{E_0}\right) \tag{7.9}$$

where c = 0.425 for n-type Si and c = 0.355 for p-type Si, for a wide range of doping concentrations ($10^{16} \leq N \leq 10^{21}$) [138], A^* is the Richardson's constant [132], k is Boltzmann's constant, and E_0 is given by

$$E_0 = E_{00} \coth\left(\frac{E_{00}}{kT}\right) \tag{7.10}$$

For $E_{00} >> kT$, $E_0 = E_{00}$, and equation (7.9) reduces to

$$\rho_c = \frac{k}{qA^*T} c \exp\left(\frac{q\phi_b}{E_{00}}\right) \tag{7.11}$$

which is the specific contact resistance under FE [138].

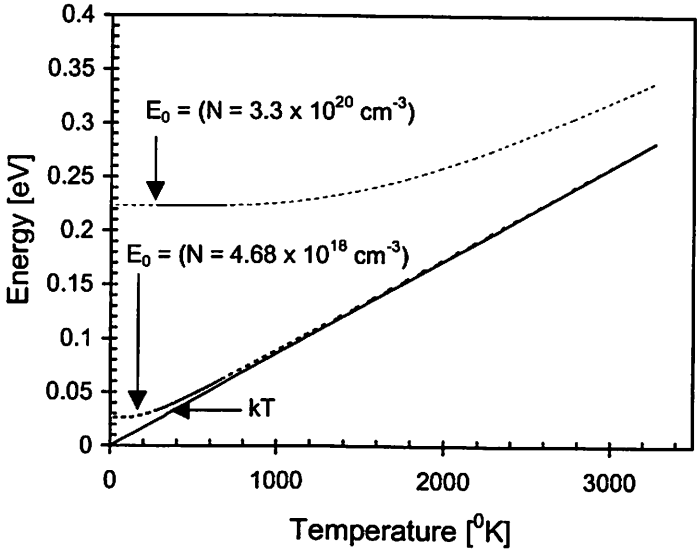
7.6.3 Temperature Dependence of Specific Contact Resistance

From equation (7.9) it can be observed that the temperature variation of the specific contact resistance has a strong dependence on the temperature dependence of E_0. Figure 7.13 shows a plot of E_0 with temperature for two different values of E_{00} (or N) corresponding to the interface doping concentrations obtained for the 9 nm and the 35 nm silicided contacts using SIMS. To simplify the illustration, E_{00} is assumed to be independent of temperature, although m_t^* has a weak positive temperature dependence [139]. It can be observed that for the smaller doping concentration E_0 varies almost linearly with temperature and is nearly equal to kT, within the temperature range of interest as shown in Figure 7.13(b). Whereas, for the higher doping concentration, E_0 remains invariant with temperature up to ~ 1000 ^{0}K and is equal to E_{00}.

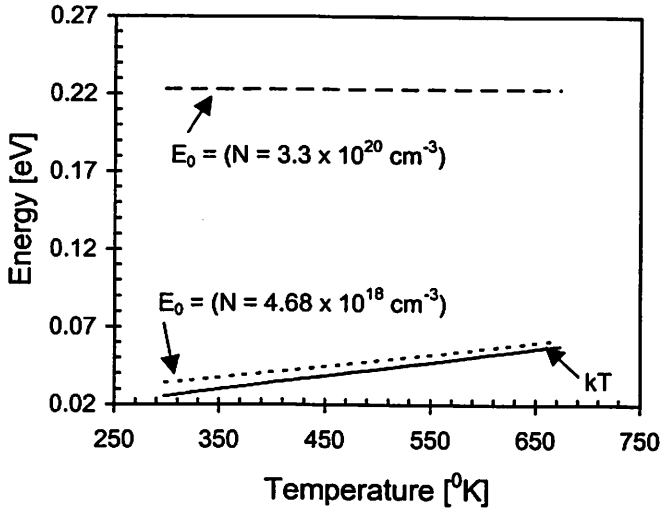
Apart from E_0, equation (7.9) contains other temperature dependent parameters. The silicide-Si barrier height (ϕ_b) in equation (7.9) which can be expressed as [140],

$$\phi_b = E_g - \phi_0 \tag{7.12}$$

is also temperature dependent [141]. Here ϕ_0 is the energy level at which the Fermi level is pinned due to the presence of high density of surface states [140], [142]. The value of ϕ_0, is known to be roughly one third of the band gap, E_g [140]. In equation (7.12), at least, E_g, is known to be temperature dependent [132], and is given by



a)



b)

Figure 7.13 Variation of E_0 with temperature shown for two different values of N for a) large temperature range and b) in the experimental temperature range. E_{00} is assumed to be independent of temperature.

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{(T+\beta)} \tag{7.13}$$

where $E_g(0)$ = 1.519 eV, α = 4.73 x 10^{-4}, and β = 636, for Si. Furthermore, as mentioned earlier, the tunneling effective mass is also known to have small positive temperature dependence. It should also be kept in mind that the band gap of Si is also known to decrease with the doping concentration if it is heavily doped [143].

7.6.4 Unified Quantitative Model for Contact Resistance

As discussed in the last section, several parameters influencing the specific contact resistance under FE or TFE, have complex temperature dependencies, which are difficult to measure or calculate. Furthermore, since it was observed in section 7.5 that apart from their temperature variations, R_c and its sensitivity to temperature also depends on the silicide thickness and on the type of dopants in the diffusion region, a unified quantitative model has been formulated which can account for all the different dependencies of R_c. The model takes into account both FE and TFE transport mechanisms by expressing the measured contact resistance as [144], [145],

$$R_c = H\, exp\left(\frac{B_n(T)}{\sqrt{N_n}}\right) \tag{7.14}$$

$$R_c = H\, exp\left(\frac{B_p(T)}{\sqrt{N_p}}\right) \tag{7.15}$$

for contacts to n+ and p+ Si respectively. Here H is introduced as a temperature independent parameter. Suffixes n and p are used to distinguish between contacts to n+ and p+ Si. The entire temperature dependence of R_c has been lumped into the parameter B. Furthermore, distinction has been made between B_n and B_p, since $\phi_{bn} \neq \phi_{bp}$, (see Figure 7.14) and the tunneling effective mass is different for electrons and holes (see equation 7.6 and 7.9). Differences in B_n and B_p accounts for differences in measured R_c for a given silicide thickness. Observed dependence of R_c on silicide thickness is incorporated through the effective impurity doping concentration, N_n and N_p, at the silicide-Si interface. For thin silicides, $N_{n,p}$ is large, which results in narrower barrier

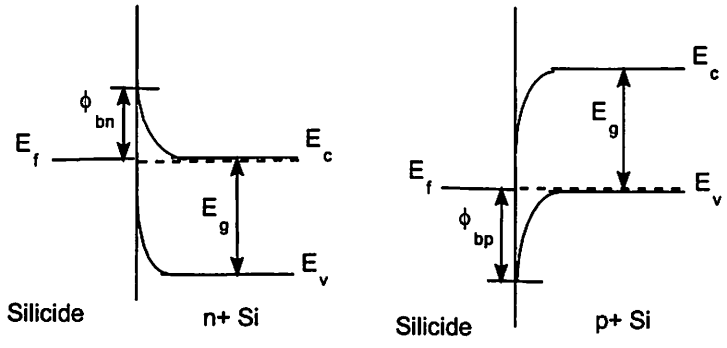


Figure 7.14 Silicide-Si band diagram illustrating Fermi level pinning in contacts to n/p type Si.

causing enhanced tunneling and therefore a lower value of R_c. Note that the temperature sensitivity of R_c is incorporated in a unified way for both FE and TFE. As discussed earlier TFE will tend to dominate for lower values of $N_{n,p}$ (for thicker silicide) causing stronger temperature variations of R_c. However, for thinner silicides FE will be the dominant transport mechanism even at higher temperatures. In the unified model, this is implicitly incorporated, since thinner silicide with higher values of $N_{n,p}$ will lower the effect of the temperature dependency of B_n and B_p on R_c.

Using equation (7.14) and (7.15) and the ratios of temperature dependent contact resistance values in Figure 7.7 and Figure 7.11, values of H, $B_n(T)$, $B_p(T)$, N_n *and* N_p have been determined that can be used to calculate the doping concentration from a knowledge of the contact resistance and results are shown in Table 7.2. The doping concentration measured by SIMS at 9 nm depth was used as the only input to the model. It can be observed that the model is in excellent agreement with measured values of R_c and SIMS data.

Figure 7.15 shows the extracted temperature dependence of the parameter B for n and p type Si. The negative temperature dependence of the parameter B introduced in equation (7.14) and (7.15) is expected due to the incorporation of the TFE carrier transport mechanism and also the temperature dependency of the barrier height. It should be also noted that both ϕ_{bn} and ϕ_{bp} (shown in Figure 7.14) can decrease with increasing temperature, since the Fermi level E_f is generally not pinned at the same point in n+ and p+ semiconductor.

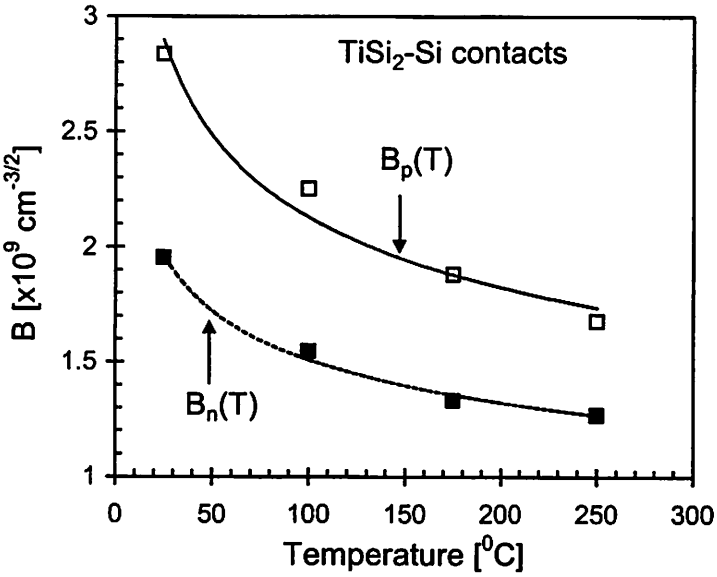


Figure 7.15 Temperature dependence of B for n and p type Si.

H = 55.845 Ω

$TiSi_2$ Thickness		Measured R_c [Ω] T=25°C	T=175°C	B(T) x10^9 $cm^{-3/2}$ T=25°C	T=175°C	Doping Concentration N(SIMS)	N(Model)	Calculated R_c [Ω] T=25°C	T=175°C
35 nm	n+ :	136.5	102.5	1.954	1.334	$4.68x10^{18}cm^{-3}$	$4.78x10^{18}cm^{-3}$	136.5	102.79
	p+ :	173.77	119.6	2.836	1.881	$6.24x10^{18}cm^{-3}$	$6.54x10^{18}cm^{-3}$	171.07	116.53
9 nm	n+ :	61.8	59.7	1.954	1.334	$3.28x10^{20}cm^{-3}$	$3.28x10^{20}cm^{-3}$	62.20	60.11
	p+ :	66.4	63.2	2.836	1.881	$3.74x10^{20}cm^{-3}$	$3.74x10^{20}cm^{-3}$	64.67	61.55

Table 7.2 Comparison of measured and calculated contact resistance values using the model. Here, only the doping concentration for 9 nm $TiSi_2$ contact (from SIMS data) was used as input to the model. All other parameters like $B_{n,p}$, H and $N_{n,p}$ (for 35 nm $TiSi_2$) were calculated using the model.

7.7 High Current Effects on Contact Resistance

7.7.1 Self-Heating Effects on Contact Resistance

The high current behavior of these contact structures has also been analyzed to understand their characteristics under high stress conditions and determine an upper limit of their current carrying capacity along with the failure mechanisms. Figure 7.16 shows the high current (DC) behavior of the contact structures (under forward bias: n+ Si biased negatively w.r.t silicide, and p+ Si biased positively w.r.t silicide) for the 35 nm silicide. It can be observed that resistance decreases with increasing current. This is due to self-heating as explained earlier. The high current I-V characteristics shown in Figure 7.16 is used along with Figure 7.7 to plot the input power, P vs. temperature rise, ΔT in Figure 7.17. It can be observed that the thermal impedance for both the n+ and p+ structures are nearly identical as expected. The slope of the best-fit line gives the thermal impedance R_θ of the contact system defined in equation (2.36), and is extracted to be ~ 7.5×10^4 ^{0}C/W. The self-heating for the present design rule (~ 1 mA) for contacts is shown to be ~ 10 ^{0}C. The thermal impedance extracted from Figure 7.17 is used to estimate a maximum allowable critical temperature for these contacts.

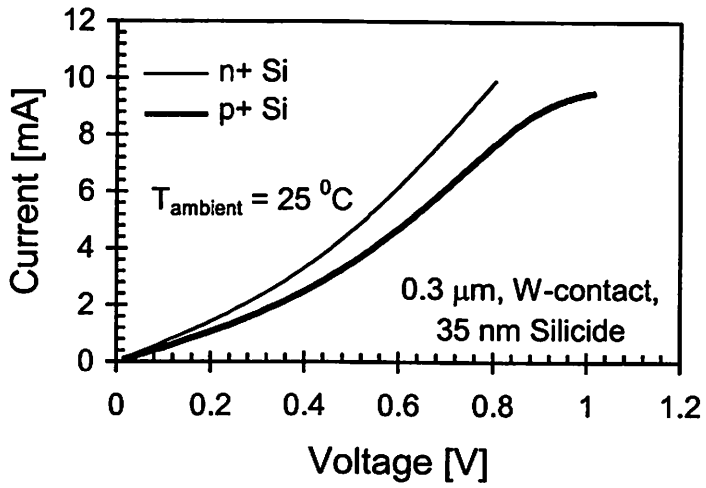


Figure 7.16 Forward bias I-V characteristics of n+ and p+ contacts in the high current regime becomes non linear due to severe self-heating.

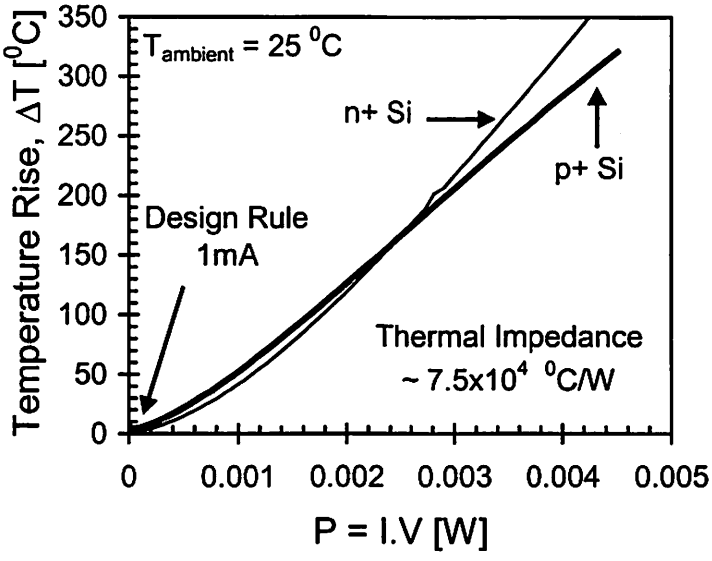


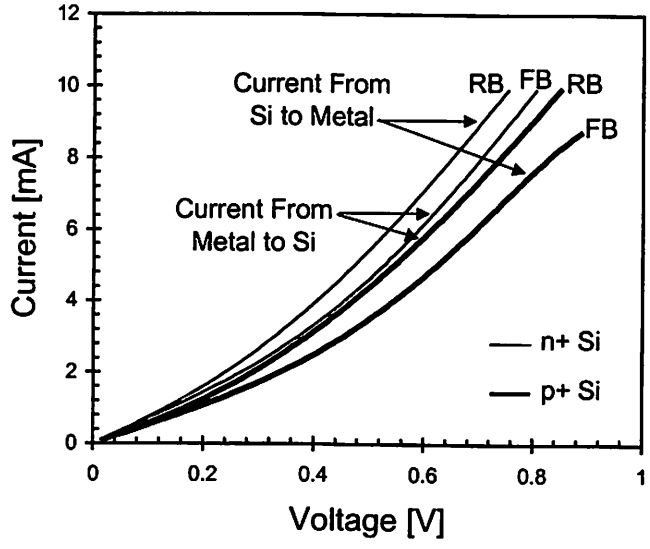
Figure 7.17 Thermal impedance of n+ and p+ contacts extracted from Figure 7.7 and Figure 7.16 at an ambient temperature of 25 ^{0}C are equal.

7.7.2 Polarity Dependence of High Current Behavior

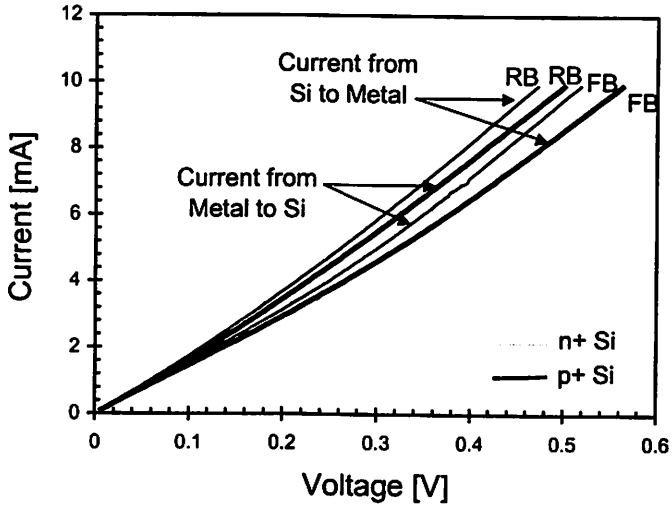
As mentioned earlier, under low-current stress negligible dependence of R_c on the polarity of the biasing current was observed. However, under high current stress a significant dependence on the polarity of the bias is observed as shown in Figure 7.18 for the 35 nm and 9 nm silicided contacts. Forward bias (FB) means n+ Si biased negatively w.r.t silicide, and p+ Si biased positively w.r.t silicide. Reverse bias (RB) means n+ Si biased positively w.r.t silicide, and p+ Si biased negatively w.r.t silicide. Note that the barrier height for carriers in the Si will be reduced under FB and increased under RB. The barrier height for carriers in the metal will remain unchanged irrespective of the polarity of the bias. Also for contacts to n+ Si, FB will cause current flow from silicide to Si, and RB will cause current flow from Si to silicide. For contacts to p+ Si the directions of current flow will be reversed under FB and RB conditions.

It can be observed from Figure 7.18 that RB always gives rise to higher current (or lower resistance) as compared to FB for a given type of diffusion (n+ or p+). This is expected, since reverse biasing the silicide-Si junction will cause a narrowing of the barrier. It can also be observed that the contacts with 35 nm silicide shows greater sensitivity to high current stress, since they are more sensitive to temperature as observed earlier in Figure 7.7. The contacts with

9 nm silicide shows lower sensitivity to high current stress due to their smaller sensitivity to temperature.



a)



b)

Figure 7.18 Biasing polarity dependence of Silicide-Si contacts under high current stress conditions for a) 35 nm silicide and b) 9 nm silicide.

7.8 Failure Mechanisms

The critical temperature for failure can be calculated using the thermal impedance obtained from Figure 7.17 and is shown to be ~ 1400 ^{0}C in Figure 7.19. This suggests that the silicon near the interface reaches melt temperature and the $TiSi_2$ gets dissolved in it forming some polycrystalline Si rich compound. Also, the failure current was found to be independent of the contact plug material for a given bias polarity, and displayed a small dependence on the siliside thickness, which is expected. The actual failure site is difficult to capture under DC stress, since the contact is usually completely destroyed. Since, the barrier resistance dominates the net contact resistance, maximum heat dissipation also takes place near the interface. In order to capture the degradation front, the Transient Kelvin Thermometry technique described in Chapter 5 was employed and the contacts were stressed in small increments of current by a single 500 ns pulse. Figure 7.20 (a) shows a TEM micrograph of a contact stressed just into degradation. The failure location clearly shows that the damage initiates at the interface. A more severe degradation state is shown in Figure 7.20 (b). In this case the barrier has undergone complete breakdown and molten W has reacted with the underlying Si forming some new compound (most likely WSi_2).

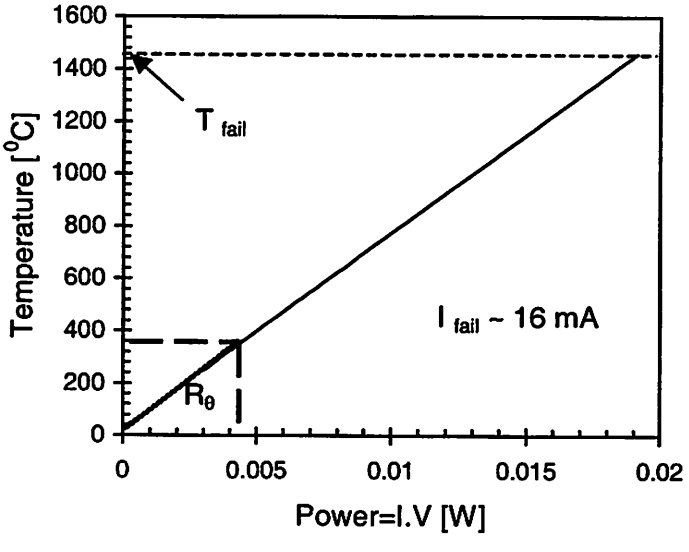
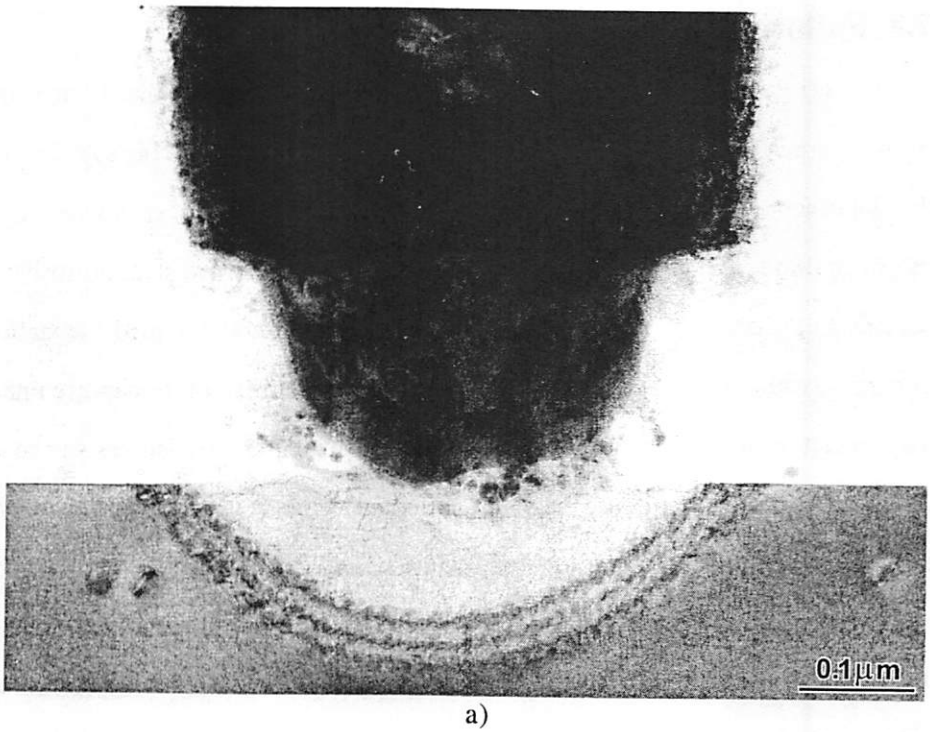
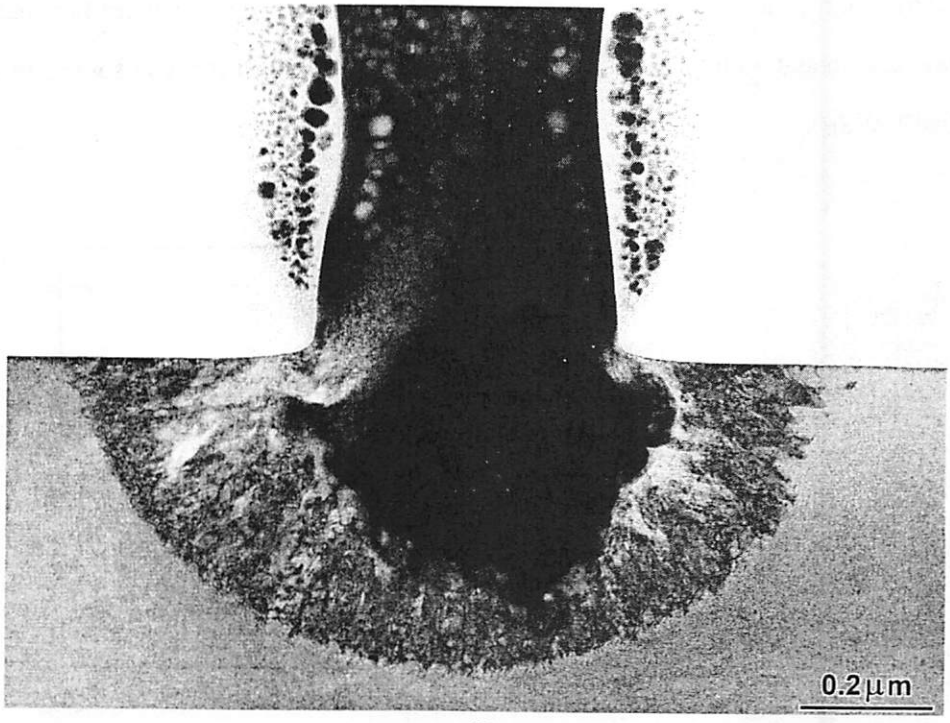


Figure 7.19 Temperature rise at the point of failure under high current stress conditions is calculated from the thermal impedance in Figure 7.17.



a)



b)

Figure 7.20 TEM micrographs of 0.3 μm W-contacts showing progression of failure at the silicide-Si interface showing a) initiation of the failure and b) severe degradation.

7.9 Summary

In this chapter the effects of temperature and current on the contact resistance of small geometry $TiSi_2$-Si contact structures have been characterized and modeled. Both, temperature and high current induced self heating have been shown to cause contact resistance lowering which can be significant in the performance of advanced ICs and I/O protection circuits. It is demonstrated that contact-resistance sensitivity to temperature and current can be controlled by the silicide thickness, which influences the interface doping concentration. Behavior of W-plug and force-fill (FF) Al plug contacts has been investigated in detail.

A simple model that comprehends field emission and thermionic field emission carrier transport mechanism has been formulated which directly correlates contact resistance to temperature and interface doping concentration. Furthermore, the high current behavior of these contact structures has been characterized. It has been shown that under high current stress conditions self-heating can cause significant variations in the I-V characteristics and that these variations are strongly influenced by the silicide thickness and polarity of the biasing current. Finally, the thermal impedance of these contact structures has been extracted and a critical failure temperature has been determined that can be used to design robust contact structures. In addition, the TKT technique developed in Chapter 5 has been successfully applied to analyze the failure modes of these contact structures.

Chapter 8

Thermal Characterization of Small Geometry Vias using Scanning Joule Expansion Microscopy

8.1 Introduction

As outlined in Chapter 1 the continuous scaling of VLSI circuits has resulted in an increase in the aspect ratio of the vias (connection between adjacent metallization levels) and increases in the current density and associated thermal effects, namely self-heating. Current crowding and localized heating [146], [147], [148], in deep sub-micrometer vias are known to strongly impact reliability of VLSI interconnects. The magnitude and spatial distribution of the temperature rise in the via are important to accurately estimate interconnect lifetime under electromigration (EM), which is temperature dependent. Localized temperature rise can also cause stress gradients inside the via structures and can also lead to melting under short-duration high-current stress conditions [89], such as electrostatic discharge (ESD) events as discussed in Chapter 5. Hence, measurements of the magnitude and spatial distribution of the temperature rise in deep sub-micrometer vias are important to accurately model their reliability and provide thermal design guidelines for various via technologies.

As illustrated in Chapter 3, interconnect thermometry based on temperature-dependent electrical resistivity of the interconnect metal can provide a spatially averaged temperature rise along the interconnects [57]. However, this does not provide spatial distribution of temperature across the interconnect structures.

The spatial resolution of far-field optical techniques, such as scanning thermoreflectance thermometry [45], infrared thermography [47], and liquid crystal thermography [51], is diffraction-limited to about 1 μm. This is insufficient to probe deep sub-micrometer vias in the size range of 0.1-0.5 μm. Near-field optics can be employed to overcome the diffraction limit [149], but it is still under development and hence cannot provide accurate measurements of all desired parameters are concerned. The SJEM has recently been developed with spatial resolution in the sub-0.1 μm range [150]. In this chapter, SJEM is used to study the thermal characteristics of small geometry W-plug vias [151].

8.2 SJEM Experimental Setup

Figure 8.1 shows the schematic diagram of the scanning Joule expansion microscope (SJEM) system used in this study. An atomic force microscope (AFM) is used to bring a sharp tip into force-controlled contact with the sample surface and perform a raster scan.

A sinusoidal or pulsed voltage is applied to the electrically conducting sample (W-Via) which produces sample Joule heating and temperature rise, resulting in sample thermal expansion. A low-power laser beam incident on the AFM tip changes its location on the photodiode due to deflection of the tip. The location of the laser beam on the photodiode determines the output signal of the photodiode. The AFM photodiode detects the cantilever deflection due to both expansion and sample topography. Since the feedback controller of the AFM has a bandwidth of 5 KHz, the photodiode signal below 5 KHz is processed for feedback control of the *z*-piezo to image surface topography under constant tip-sample force or cantilever deflection. The Joule heating frequency is kept above 5 KHz to avoid feedback response. The lock-in amplifier is tuned to the Joule heating frequency, which detects only the expansion signal and provides this to an auxiliary AFM channel to form the expansion image. The system can also be operated without

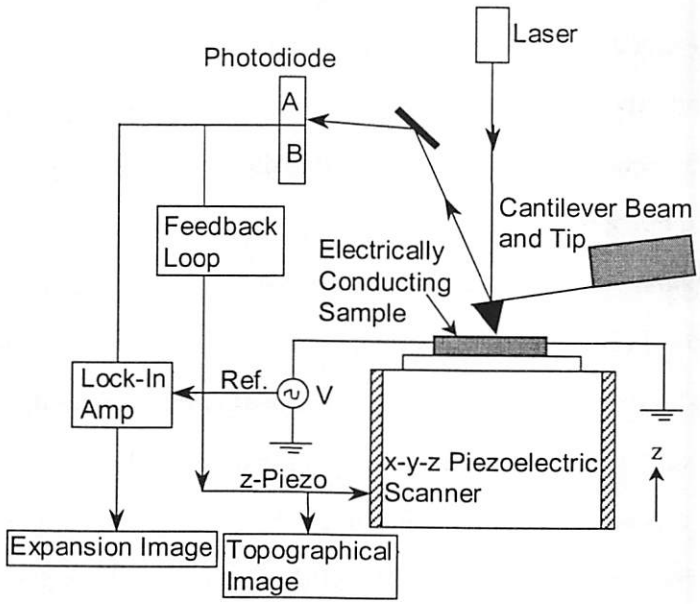


Figure 8.1 Schematic diagram of the experimental setup used for the scanning Joule expansion microscopy (SJEM).

feedback in which case the heating frequency can be below the controller bandwidth [150], [152].

8.3 Via Fabrication and Sample Preparation

The W-plug via structures were fabricated in a state-of-the-art 0.25 μm industrial CMOS process flow. The samples used in the experiments contained 0.6 μm-thick Al-Cu interconnects (with top and bottom layers of 0.05 and 0.15 μm-thick TiN) at two levels of metallization that were separated by a layer of ~ 0.9 μm-thick silicon dioxide.

These interconnects crossed each other forming an overlapping region and were bridged in this region by a single W-plug via with 0.4 μm diameter as shown schematically in Figure 8.2. Both levels of AlCu lines were 1.6 μm wide. The samples were coated with a standard passivation layer of ~1.0 μm thick silicon dioxide followed by a capping layer of 0.3 μm thick silicon nitride. A Chemical Mechanical Polishing (CMP) process was used to planarize the inter-layer dielectric and the overlying passivation layers. Since the thermal expansion coefficient of the oxide and nitride is low, a 0.28 μm thick film of polymethyl methacrylate (PMMA) was spin coated on the sample to amplify the expansion signal.

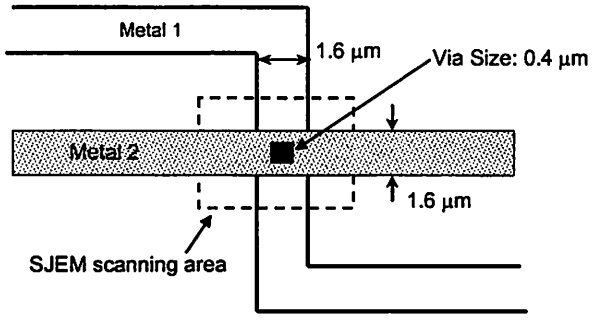


Figure 8.2 Top view of the via structure showing the scanning area (15 μm x 15 μm) of the AFM tip in the SJEM experiments.

Figure 8.3 shows the schematic cross section of the via sample including the PMMA coating. The PMMA film can be easily stripped off using acetone after all the measurements. This procedure is feasible for routine investigations since it is easy and safe to be accomplished. The thermal expansion coefficient of PMMA is typically about 7×10^{-5} K^{-1}, which results in a sensitivity of 2.0×10^{-11} m/K and a temperature resolution of about 0.2 K.

The temperature drop across the thickness of the passivation layer and the PMMA film can be expected to be quite small as illustrated in Figure 8.4. This is because the temperature rise exponentially decays from the heat source with the decay length given by $\sqrt{\kappa/\pi f}$, where κ is the thermal diffusivity of the material and f is the applied frequency [152]. The thermal diffusivity of PMMA is the lowest of the three passivation materials, and is ~ 10^{-7} m^2/s. Hence, if the modulation frequencies are kept below 50 KHz, the penetration depth in PMMA should be larger than 0.8 μm, which is much larger than the PMMA film thickness (0.28 μm).

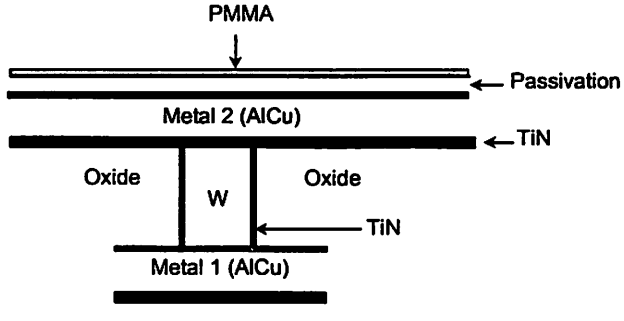
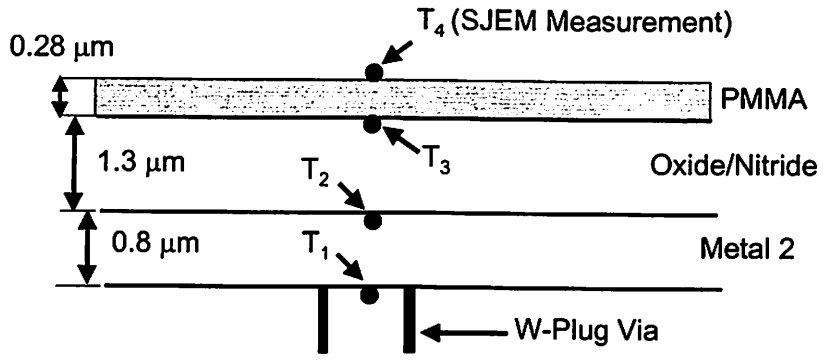


Figure 8.3 Via cross section showing the thin layer of PMMA film coated over the passivation layers to enhance the expansion signal.



$$\text{Penetration Depth, } \delta = \begin{cases} 8\ \mu\text{m for oxide} \\ 2.6\ \mu\text{m for PMMA} \end{cases}$$

$$\Rightarrow T_1 \approx T_2 \approx T_3 \approx T_4$$

Figure 8.4 A schematic diagram illustrating the equivalence of actual via temperature and measured temperature on PMMA film by SJEM. The penetration depths in the oxide and the PMMA film shown above were calculated at f = 5KHz. For the metal film on top of the W-plug via, the thermal diffusivity (κ) is even higher resulting in a penetration depth >> metal film thickness. Hence, the temperature measured on the surface of PMMA film (T_4) is expected to equal that on top of the via (T_1).

8.4 Temperature Calibration

Since the expansion signal of the PMMA film dominated over other layers on the sample, it was necessary to accurately determine the expansion coefficient of the PMMA film in order to convert expansion signals to temperature rise. This was achieved by first measuring the electrical resistance of an AlCu interconnect as a function of temperature and extracting the temperature coefficient of resistance (β) of AlCu. The expansion signal of the PMMA film over the AlCu interconnect was then measured using an AC bias current given by,

$$I = I_0 \cos \omega t \tag{8.1}$$

The temperature rise due to self-heating is given by,

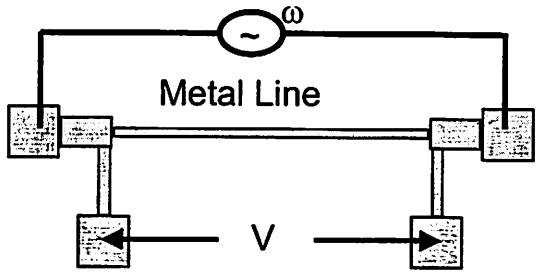


Figure 8.5 Schematic of the interconnect line heated by a sinusoidal current and the corresponding voltage across it.

$$T = T_{DC} + T_{AC}\cos 2\omega t \tag{8.2}$$

The resistance of the AlCu line under such bias is then given by,

$$R = R_0\{1 + \beta(T_{DC} + T_{AC}\cos(2\omega t + \varphi))\} \tag{8.3}$$

Where T_{DC} and T_{AC} are the DC and peak AC temperature rises respectively. ϕ is the phase lag between the current and the temperature and $\omega\ (=2\pi f)$ is the angular frequency. Therefore, the voltage across the line (illustrated in Figure 8.5) is given by,

$$V = I_0 R_0\left[\begin{array}{l}\left\{(1+\beta T_{DC})\cos\omega t + \frac{\beta T_{AC}}{2}\cos(\omega t+\varphi)\right\} \\ + \frac{\beta T_{AC}}{2}\cos(3\omega t+\varphi)\end{array}\right] \tag{8.4}$$

The lock-in amplifier locks in to the 3ω component of the voltage and measures the amplitude $(\frac{I_0 R_0 \beta T_{AC}}{2})$ of the 3ω component. We can therefore calculate T_{AC} from the measured amplitude of the 3ω component of the voltage signal.

Since the expansion signal frequency is directly proportional to the AC temperature frequency (2ω), the locally measured 2ω component of the expansion signal of the PMMA film, δL, can be expressed as,

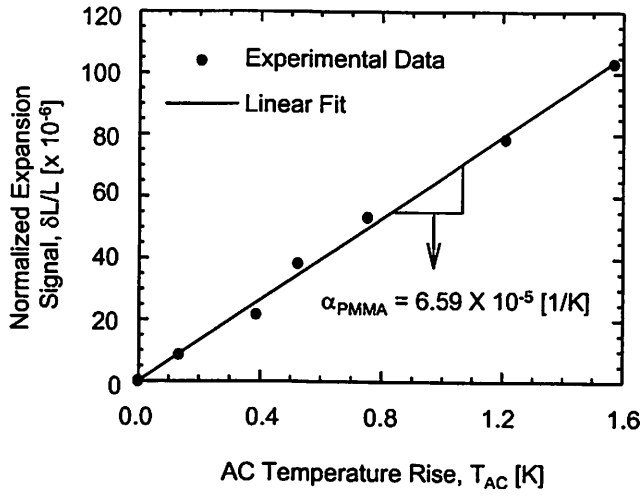


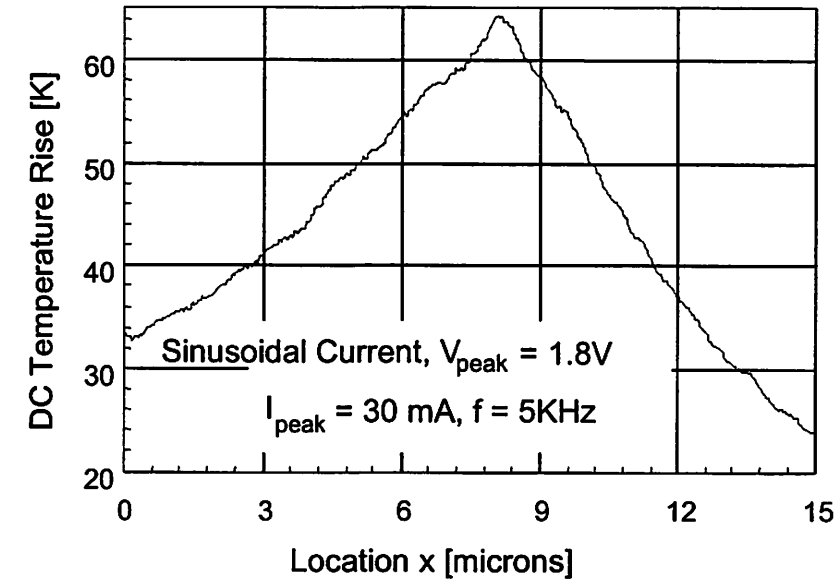
Figure 8.6 The thermal expansion coefficient of PMMA determined by measuring the expansion signal and calculating the corresponding AC temperature rise of an Al-Cu line from the measured amplitude of the 3ω component of the voltage signal using the lock-in amplifier.

$$\delta L = \alpha \cdot L \cdot T_{AC} \tag{8.5}$$

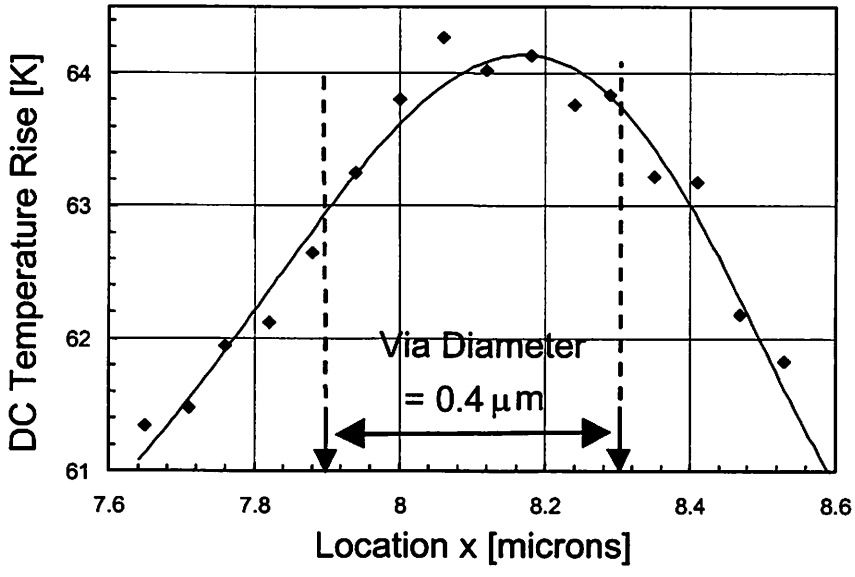
Here α is the expansion coefficient of the PMMA film, L is the PMMA film thickness. It should be noted that the expansion signal is mainly due to the PMMA film since $(\alpha L)_{PMMA} >> (\alpha L)_{oxide}$. The thermal expansion coefficient of PMMA can be experimentally obtained from equation (8.5) and was found to be $65.9 \pm 3.3 \times 10^{-6}$ K^{-1}as illustrated in Figure 8.6. Once α has been calibrated locally, T_{AC} can be calculated directly from the measured expansion signal using equation (8.5).

8.5 Self-Heating Analysis under Sinusoidal Bias

Figure 8.7 (a) shows the spatial temperature profile around the via sample using SJEM. The magnified spatial temperature distribution across the via is also shown in Figure 8.7(b). The temperature contour image of Figure 8.8 shows the hot region on top of the via. Although the vias were 0.4 μm in diameter, diffusion in the passivation layers spread the temperature peak to about 10 μm at full-width half-maximum. It must be noted that despite this lateral spread, the temperature drop across the thickness of the passivation layer and the PMMA film can be expected to be quite small as explained in section 8.3.



(a)



(b)

Figure 8.7 a) Temperature profile around the 0.4 μm via sample measured by SJEM and b) magnified spatial temperature distribution along a micron length on top of the via. The corresponding peak current densities are: $j_{peak(line)}$ = 0.63 MA/cm^2, and $j_{peak(via)}$ = 24 MA/cm^2. The spatial resolution of temperature rise is ~ 0.06 μm. The temperature gradient across the via is ~ 6 ^{0}K/μm.

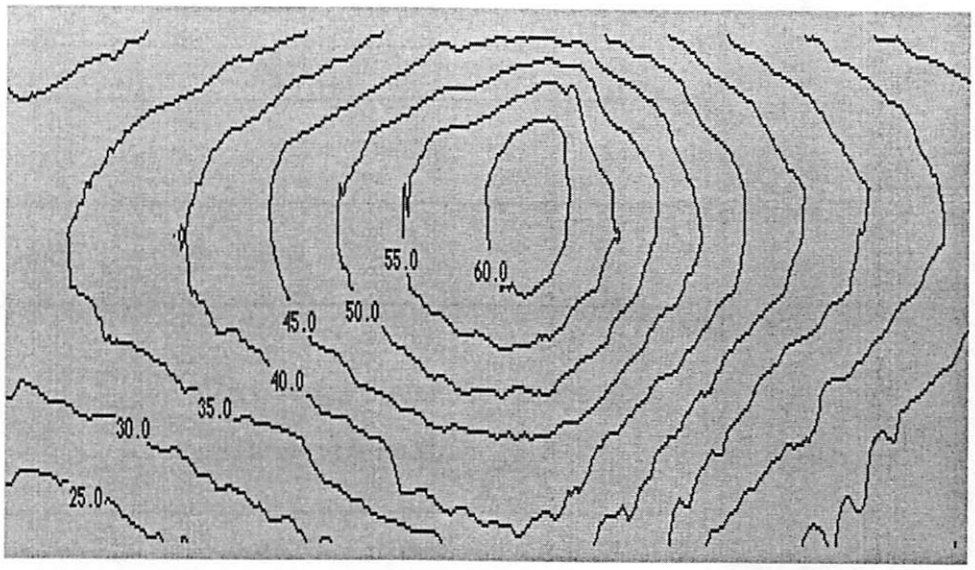


Figure 8.8 The temperature contour map across the 0.4 μm via sample measured by SJEM.

In addition to an AC temperature rise under a sinusoidal bias, there is also a DC temperature rise due to the value of RMS power dissipation. The DC temperature rise is often much higher than the AC component and is, therefore, of interest. Since SJEM measures only the AC component, estimation of the DC component requires analysis.

The thermal behavior follows a first-order system and is characterized by a time constant τ. The governing heat equation [22] is given by

$$mc\frac{\partial T}{\partial t} = -(h_c A)(T - T_0) + I^2 R \tag{8.6}$$

where m is the sample mass, c is specific heat, R is the resistance of sample, h_c is the effective heat transfer coefficient averaged over the via surface area A. The first term on the right hand side of equation (8.6) represents an effective heat dissipation term under the assumption that there are no temperature gradients within the sample (Via). The validity of this assumption is based on the fact that the Biot number, Bi is < 0.1. Bi is defined as [153],

$$\begin{aligned} Bi &= \frac{Internal - conduction - resistance}{External - heat - transfer - resistance} \\ &= \frac{L/k_s A}{1/h_c A} = \frac{h_c L}{k_s} \end{aligned} \tag{8.7}$$

where L is the characteristic length of the solid sample (W-Via), and k_s is the thermal conductivity of sample material (W). Now under an AC bias the current is given by equation (8.1). Substituting this expression for current in equation (8.6) we get,

$$mc\frac{\partial\theta}{\partial t} = -h_c A\theta + \frac{I_0^2 R}{2}\left[1 + \cos 2\omega t\right] \tag{8.8}$$

where $\theta = T\text{-}T_0$. The dependence of R on θ has been neglected since $R = R_0(1+\beta\theta)$ and $\beta\theta << 1$. Equation (8.8) can be rewritten as,

$$\frac{\partial\theta}{\partial t} = -\frac{\theta}{\tau} + \psi\left[1 + \cos 2\omega t\right] \tag{8.9}$$

where
$$\tau = \frac{mc}{h_c A} \text{ and } \psi = \frac{I_0^2 R}{2mc} \tag{8.10}$$

Thus the heat generation term in equation (8.9) has two components:

(i) a DC component = ψ, (ii) an AC component = $\psi \cos 2\omega t$. Since this is a linear equation in θ, we can write $\theta = T_{DC} + T_{AC}$. Note that T_{DC} is the DC temperature rise due to an AC current. For T_{DC} the governing equation is,

$$0 = \frac{-T_{DC}}{\tau} + \psi \Rightarrow T_{DC} = \tau\psi = \frac{I_0^2 R}{2h_c A} \tag{8.11}$$

The governing equation for T_{AC} is,

$$\frac{\partial T_{AC}}{\partial t} = -\frac{T_{AC}}{\tau} + \psi \cos 2\omega t \tag{8.12}$$

Now let $T_{AC} = B\cos 2\omega t + C\sin 2\omega t$. Substituting in equation (8.12) we can show that

$$B = \left[\frac{\psi\tau}{1+(2\omega\tau)^2}\right] \text{ and } C = \left[\frac{\psi 2\omega\tau^2}{1+(2\omega\tau)^2}\right]$$

Thus T_{AC} can be written as,

$$T_{AC} = \left[\frac{\psi\tau}{1+(2\omega\tau)^2}\right]\cos 2\omega t + \left[\frac{\psi 2\omega\tau^2}{1+(2\omega\tau)^2}\right]\sin 2\omega t \tag{8.13}$$

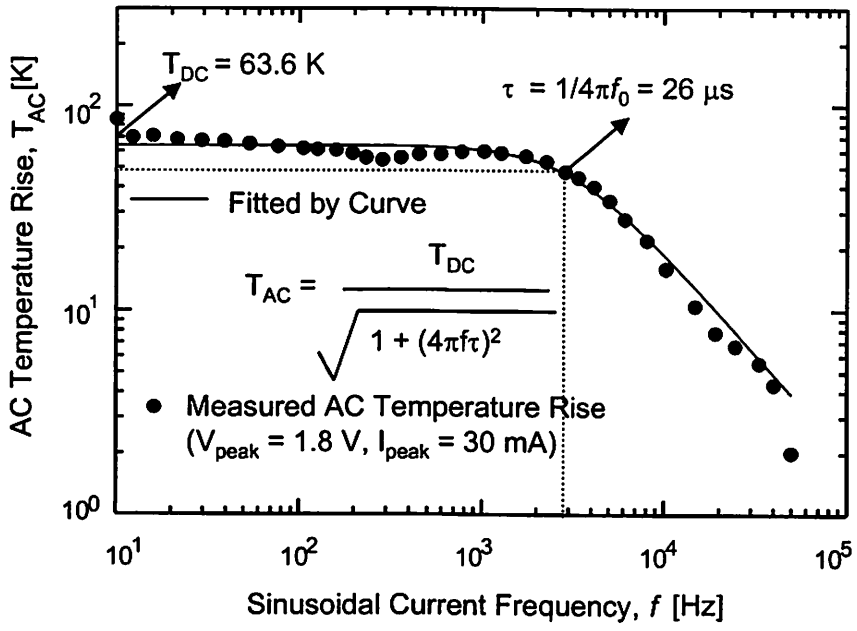


Figure 8.9 Experimental data showing extraction of a time constant τ = 26 μs for the via structure. Note that τ is a characteristic of the thermal system and is independent of the waveform.

Now the amplitude or the maximum value of T_{AC} is given by,

$$|T_{AC}| = \frac{\psi \tau}{\left[1+(2\omega\tau)^2\right]}\sqrt{1+(2\omega\tau)^2} \tag{8.14}$$

From equation (8.10) and (8.11) we see that $T_{DC} = \psi\tau$. Hence, the peak AC and DC temperature rises can be related as

$$T_{AC} = \frac{T_{DC}}{\sqrt{1+(4\pi f\tau)^2}} \tag{8.15}$$

where $\omega = 2\pi f$. Figure 8.9 plots experimental data indicating this behavior and yields a time constant τ = 26 μs. The experiments were performed in vacuum in order to get rid of the effect on the deflection of the cantilever due to the heating by the surrounding gas [154]. Since the internal time constant of the W-plug via, $\tau_{via} = d^2/\alpha_W$ (here d is the via diameter, and α_w is the thermal diffusivity of W), is of the order of 1 ns, the observation of τ = 26 μs suggests that the surrounding materials, composed mainly of oxide and nitride, must be involved in heat spreading.

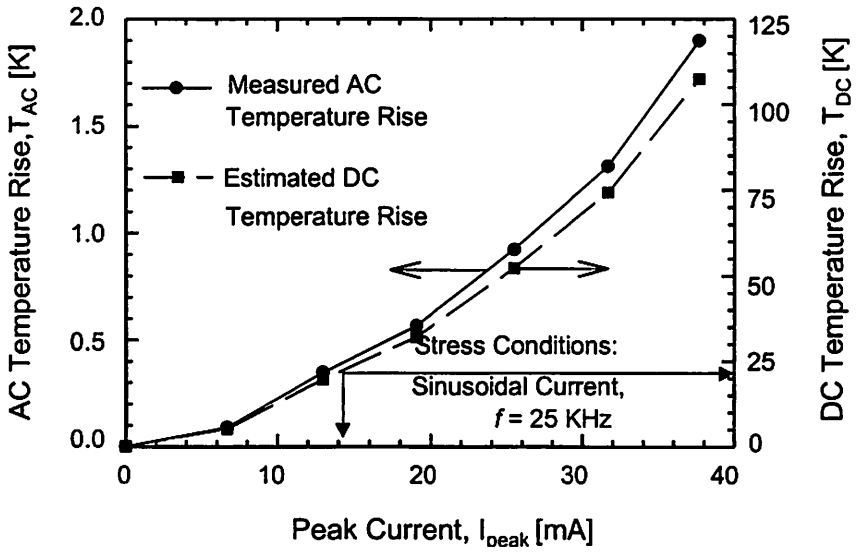


Figure 8.10 The maximum AC temperature rise and the corresponding DC temperature rise under different sinusoidal current stress. The time to get one data point typically is around 15 minutes. The DC temperature rise was estimated from AC temperature rise using equation (8.15) and the analysis is described in the text.

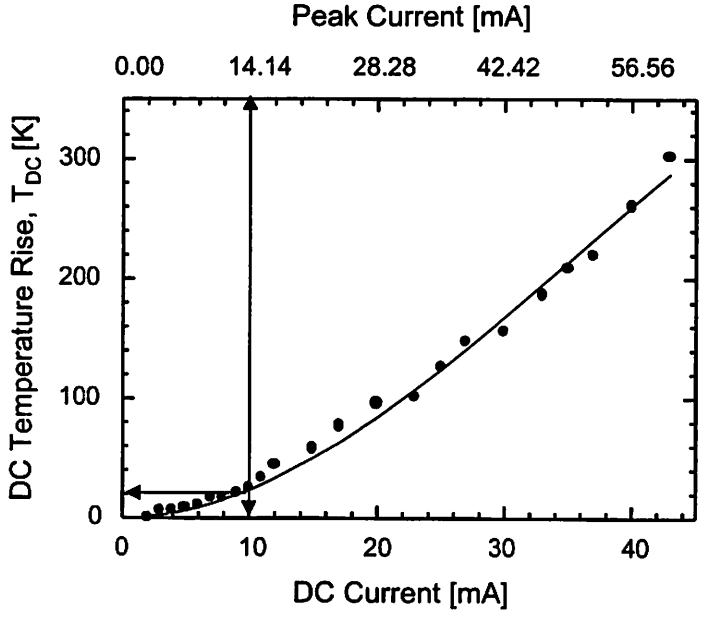


Figure 8.11 Average temperature rise of the via sample determined using resistive thermometry. The corresponding peak current values for the sinusoidal bias is also shown along the second x-axis based on $I_{peak} = \sqrt{2} I_{DC}$.

Using this first-order analysis, T_{DC} was found as a function of sinusoidal current stress and plotted in Figure 8.10, where the $T_{DC} \propto I_{peak}^2$ behavior can be seen.

We also measured the self-heating of the W-via under a DC current stress using the temperature dependence of the resistivity of the W-plug as shown in Figure 8.11. For this, the temperature coefficient of resistance (TCR) of the W-plug was experimentally determined to be $\sim 1.01 \times 10^{-3}$ K^{-1} from Kelvin measurements. We can observe from Figure 8.11 that a DC current of (say 10 mA) results in a temperature rise of about 25 ^{0}C. The corresponding peak current for the sinusoidal bias is 14.14 mA, which gives a T_{DC} (from Figure 8.10) of around 20 ^{0}C, which is lower than the temperature rise under DC current stress. This is expected since the DC temperature rise under sinusoidal bias is estimated from the measured AC temperature rise using SJEM experiment, where heat diffusion into the dielectric surrounding the via, and into the metal at the top and bottom of the via, results in a slightly lower measured temperature.

8.5 Self-Heating Analysis under Pulsed Bias

To study the thermal behavior of the W-via under a pulsed bias, experiments were conducted at f = 25 KHz, pulse amplitude of 2 V, and pulse widths, t_p, from 50 ns to 500 ns. The schematic representation of the pulsed current and the corresponding temperature rise is shown in Figure 8.12.

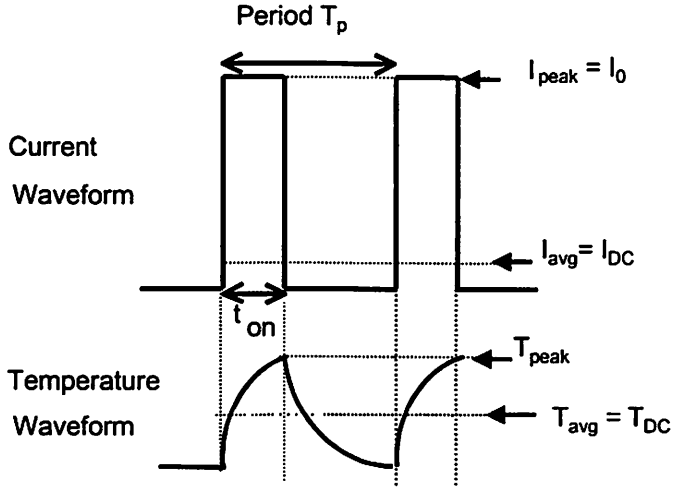


Figure 8.12 The schematic waveforms of the pulsed bias and the corresponding temperature rise.

The current amplitude for the pulsed bias can be expressed as,

$$I = \begin{Bmatrix} I_0, & for \;\; nT < t < nT + t_p \\ 0, & for \;\; nT + t_p < t < (n+1)T \end{Bmatrix} \tag{8.16}$$

where n is some integer, t_p is the on time of the pulses and T is the time period. The governing heat equation can be written as,

$$\frac{\partial \theta}{\partial t} = -\frac{\theta}{\tau} + \frac{I^2 R}{mc} \tag{8.17}$$

Here τ is the characteristic time constant of the first order equation. The Fourier analysis of the temperature rise yields,

$$\theta = \theta_0 + \sum_{n=1}^{\infty} \theta_n e^{in\omega t} \tag{8.18}$$

where

$$\theta_n = \frac{1}{T}\int_0^T \theta e^{-in\omega t} dt \tag{8.19}$$

and

$$\theta_0 = \frac{1}{T}\int_0^T \theta dt \tag{8.20}$$

The first harmonic (25 KHz) of the via temperature rise, θ_1, can be calculated from equation (8.19) by multiplying equation (8.17) with $e^{-i\omega t}$ and integrating each term from 0 to T. This yields

$$\theta_1 = \frac{1}{T}\int_0^T \theta\, e^{-i\omega t} dt = \frac{{I_0}^2 R}{mcT} \frac{(1 - e^{-i\omega t_p})}{i\omega(i\omega + \frac{1}{\tau})} \tag{8.21}$$

For $t_p << T$, $1 - e^{-i\omega t_p} \approx i\omega t_p$ and thus θ_1 is given by,

$$\theta_1 = \frac{{I_0}^2 R}{mcT} \frac{t_p}{(i\omega + \frac{1}{\tau})} \tag{8.22}$$

and thus the amplitude of θ_1 is given by,

$$|\theta_1| = \sqrt{\theta_1 \overline{\theta_1}} = \frac{{I_0}^2 R t_p}{mcT} \frac{1}{\sqrt{\omega^2 + \frac{1}{\tau^2}}} \tag{8.23}$$

Similarly, the average temperature rise θ_0 can be calculated from equation (8.20) by simply integrating equation (8.17) from 0 to T. This yields,

$$\theta_0 = \frac{1}{T}\int_0^T \theta\, dt = \frac{I_0^2 R t_p}{mcT}\tau \tag{8.24}$$

In our experiments $T_{AC} = \theta_1$ was measured by measuring the amplitude of the first harmonics of the expansion signal. Figure 8.13 shows that T_{AC} varies linearly with pulse widths, in accordance with equation (8.23). Here I_0 is the peak current, R is the via electrical resistance, and (mc) is the thermal capacity of W-via. The average temperature $T_{avg} = T_{DC} = \theta_0$ can be estimated by this analysis from the ratio of equation (8.23) to equation (8.24) which gives,

$$\frac{T_{AC}}{T_{DC}} = \frac{1}{\sqrt{1+(2\pi f\tau)^2}} \tag{8.25}$$

Similarly one can obtain the peak temperature rise, T_{peak} in the following way. The heat equation during the "on" and "off" times during the pulsed stress can be written as,

$$\frac{\partial\theta_{on}}{\partial t} = -\frac{\theta_{on}}{\tau} + \frac{I_0^2 R}{mc} \quad \text{for} \quad nT < t < nT + t_p \tag{8.26}$$

and

$$\frac{\partial\theta_{off}}{\partial t} = -\frac{\theta_{off}}{\tau} \qquad \text{for} \quad nT + t_p < t < (n+1)T \tag{8.27}$$

The solution to equation (8.26) is given by,

$$\theta_{on} = C_1 \exp(-\frac{t}{\tau}) + \frac{I_0^2 R\tau}{mc} \tag{8.28}$$

Similarly, the solution to equation (8.27) is given by,

$$\theta_{off} = C_2 \exp(-\frac{t}{\tau}) \tag{8.29}$$

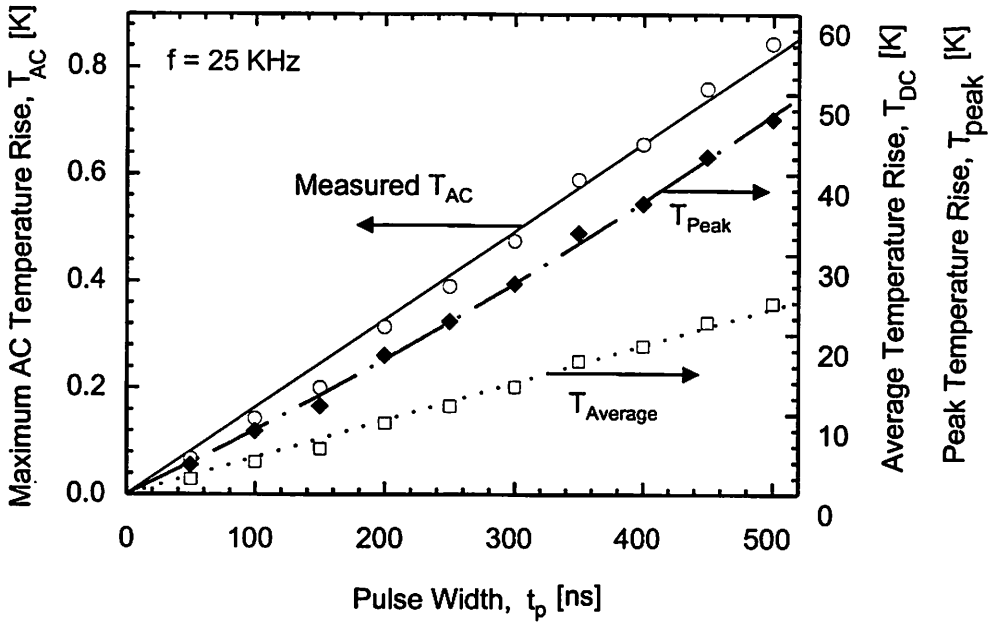


Figure 8.13 The first harmonic, average and peak temperature rise as a function of the pulsed width of applied pulsed current. The stress condition corresponds to a current density of 2.65×10^7 A/cm^2 in the via and 3.47×10^6 A/cm^2 in the metal line.

C_1 and C_2 can be obtained by applying the following two boundary conditions to equation (8.28) and equation (8.29).

(i) Steady periodic behavior: $\theta_{on}(nT) = \theta_{on}((n+1)T)$ and,

(ii) Temperature Continuity: $\theta_{on}(nT + t_p) = \theta_{off}(nT + t_p)$

and therefore we get T_{peak} from the fact that $T_{peak} = \theta_{on}(nT + t_p) = \theta_{off}(nT + t_p)$

This yields,

$$T_{peak} = \frac{I_0^2 R\tau}{mc}\left[\frac{1-exp(-\frac{t_p}{\tau})}{1-exp(-\frac{T}{\tau})}\right] \qquad (8.30)$$

Therefore, the peak temperature rise T_{peak} can be estimated by this analysis from the ratio of equation (8.24) to equation (8.30) supposing $t_p << \tau$, which gives,

$$\frac{T_{DC}}{T_{peak}} = (f\tau)[1 - exp(-1/f\tau)] \tag{8.31}$$

These values are also shown in Figure 8.13.

8.6 Summary

In this chapter it has been demonstrated that SJEM can be used to measure the spatial temperature distribution in VLSI interconnect leads and vias and to study their steady state and dynamic thermal behavior under sinusoidal and pulsed current stress. The spatial distribution of temperature across a 0.4 μm via has been obtained with a spatial resolution of 0.06 μm, and temperature resolution of ~ 0.25 ^{0}K. The temperature gradient across the via has been shown to be ~ 6 ^{0}K/μm.

The temperature contour map over a larger area around the via has also been extracted, which gives quantitative information on the spatial distribution of temperature surrounding a small geometry via. A precise knowledge of the spatial distribution of temperature rise will be useful in modeling via electromigration and other thermally induced via failure mechanisms, which are strongly influenced by temperature gradients around the via. The thermal time constant of the via structure has also been determined from the measured AC frequency dependence of the temperature rise. Furthermore, models have been developed to estimate the average (DC) and peak temperature rise under pulsed stress condition from the measured first harmonic temperature rise. Thus, the sub-100 nm resolution of SJEM can be potentially employed to design thermally robust multilevel interconnect structures, and improve reliability. It can be used to investigate current crowding effects and can help determine the impact of different via and lead designs on the thermal characteristics of multi-level deep submicron interconnect structures.

Chapter 9

Highly Accelerated Electromigration Tests for Interconnect Reliability Evaluation

9.1 Introduction

In this chapter, we have analyzed the implications of using highly accelerated stress conditions chosen in EM testing on reliability assessment of various interconnect processes [155]. The EM failure times obtained using accelerated stress conditions are strongly influenced by thermal effects [62], [156],[157]. Package level tests are generally used in IC industry to evaluate electromigration (EM) reliability of metal interconnects. For this purpose, standard test structures and test methodology has been defined by NIST [62], [156]. These tests are typically performed under moderately accelerated stress (current densities around 1-3 MA/cm^2, and stripe temperatures around 150-250 ^{0}C). Based on the lognormally distributed lifetime data, the lifetimes under field conditions are then estimated by using the Black's equation [92], for determining the acceleration factor (AF) in the EM tests. Black's equation is given by,

$$t_{50} = A \cdot j^{-n} \cdot exp(\frac{Q}{k_B T_m}) \tag{9.1}$$

where t_{50} is the time-to-fail at 50% cumulative failure fraction, or median-time-to-fail. A is a constant that is dependent on the geometry and microstructure of the interconnect, j is the DC or average current density. The exponent n equals 2 under normal use conditions, Q is the activation energy for grain-boundary diffusion and equals ~ 0.7 eV for Al-Cu, k_B is the Boltzmann's constant, and T_m is the metal temperature. The typical goal for EM reliability is to achieve less than 0.5% cumulative failures in 10 years of chip operation, at a temperature of ~ 100 ^{0}C (= T_{ref}). The acceleration factor is defined as,

$$AF = \frac{t_{50}(field)}{t_{50}(stress)} = \left(\frac{j_{stress}}{j_{field}}\right)^n exp\left[\frac{Q}{k_B}\left(\frac{1}{T_{ref}} - \frac{1}{T_{stress}}\right)\right] \tag{9.2}$$

with t_{50} chosen as the mean time that the metal conductors take to attain a certain failure criterion (usually a certain percentage rise in resistance). The "field" and "stress" in equation (9.2) denote field and stress conditions for EM. The lifetime under field conditions, corresponding to 0.5% cumulative failure is given by

$$t_{field} = \frac{AF\, t_{50}}{exp(2.58 \cdot \sigma)} \tag{9.3}$$

where σ is the dispersion of time to failure, usually defined as

$$\sigma = ln\left(t_{50}/t_{16}\right) \tag{9.4}$$

As IC technology evolves, interconnect systems becomes more complicated, which demands an increasing number of process modules to be evaluated for reliability assurance. To allow quick turn around and reduce test cycle time, highly accelerated stress at wafer level is needed in EM testing. As illustrated in Figure 9.1, large stress around 10 MA/cm^2 is required to reduce the test cycle time to the desired level below 10 minutes. Due to inadequate understanding of EM kinetics and lack of a valid lifetime model under large stress, the lifetime data taken under large stress can not be used for field lifetime determination. Furthermore, there are concerns regarding the measurement sensitivity for reliability ranking for EM tests using large stress.

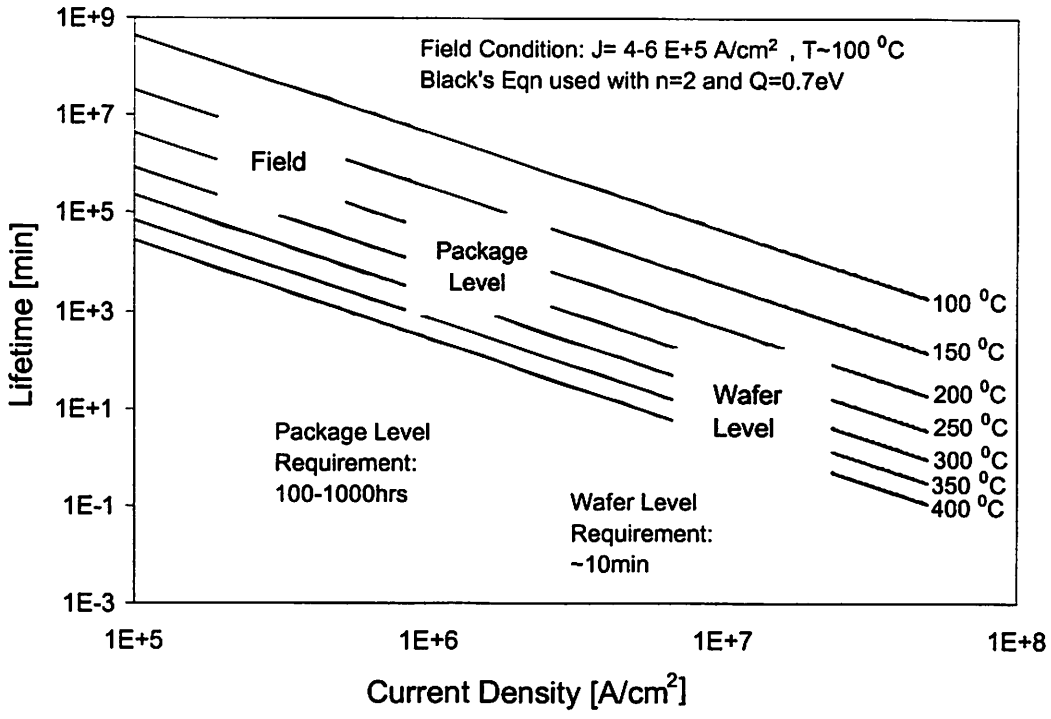


Figure 9.1 Typical lifetimes under field, package and wafer level stress conditions.

9.2 Sample Fabrication

NIST recommended standard test structures with stripe dimensions of W/L = 3/1000 μm/μm were used in this study. Multilayered AlCu(0.5%) metallization system with three splits in barrier metal design, shown in Figure 9.2, were tested under moderate stress conditions (j = 3 MA/cm^2, T_m = 245 °C) at the package level. The results are shown in Table 9.1.

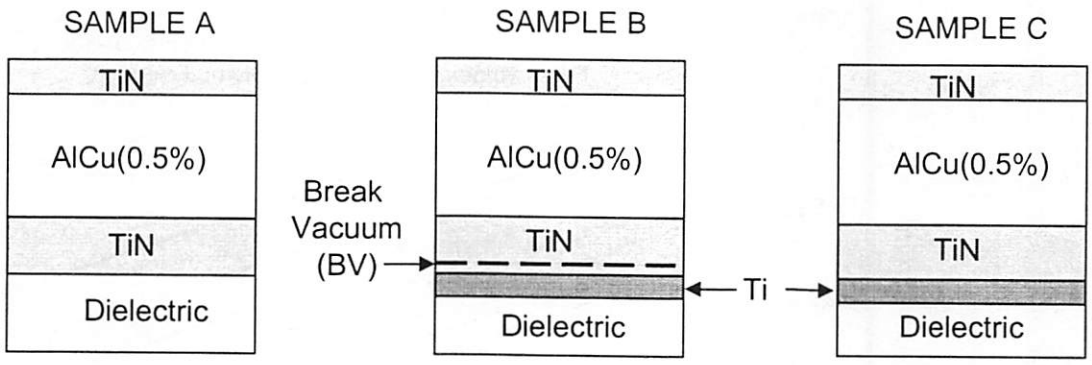


Figure 9.2 Schematic cross-sectional view of the three samples used in this study showing variation in barrier metal design. The dashed line in sample B indicates a break in vacuum.

• Package level lifetimes (J = 3MA/cm^2, T_m = 245 ^{0}C)

Barrier Metal Design		Actual t50 (hours)	Normalized t50
TiN	(A)	53.7	1
Ti/TiN/BV/TiN	(B)	116.8	2
Ti/TiN (in-situ)	(C)	222.1-696	>4

Table 9.1 Lifetimes obtained by package level EM tests on three splits of a multilayered Al Cu(0.5%) metallization system with varying barrier metal design.

Prior to the wafer level testing the joule-heating at various current density and two different chuck temperatures was measured for each split using NIST recommended methodology. Figure 9.3 shows the joule heating results for the three splits at two different chuck temperatures (20 ^{0}C and 200 ^{0}C). As seen from this figure, there is significant variation in Joule heating among the three splits due to process variation. Also, as we move towards the accelerated testing zone the difference in the temperature rise (ΔT) increases. Wafer level tests at current density of 10 MA/cm^2 were then performed while varying the chuck temperature from 49 ^{0}C to 185 ^{0}C. This introduced a variation in the stripe temperature ranging from 225 ^{0}C to 390 ^{0}C.

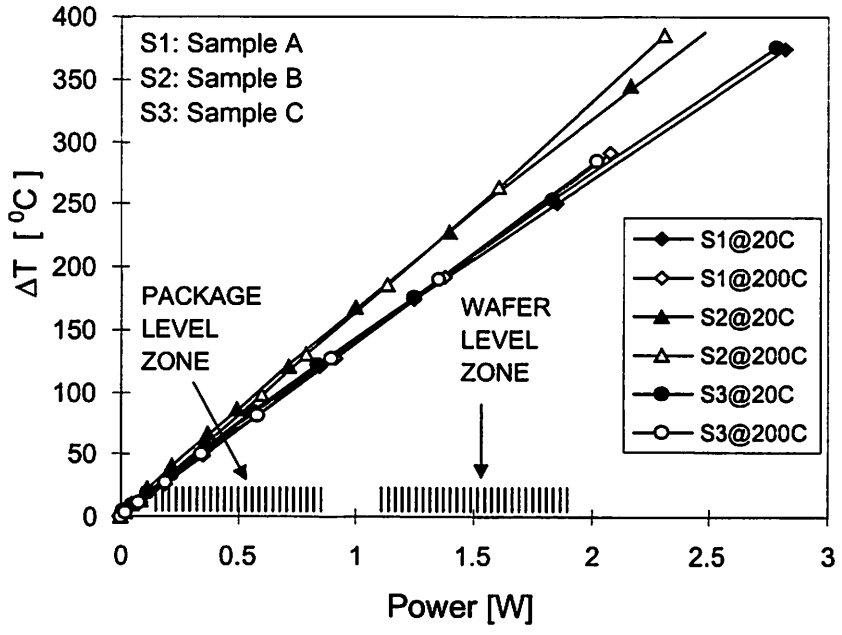


Figure 9.3 DC Joule heating results for the three samples at two different chuck temperatures.

9.3 Results and Discussion

The normalized values [t_{50}/t_{50}(sample A)] of the t_{50}s were relatively in good agreement with the normalized package level values (Table 9.1) up to 390 °C as shown in Figure 9.4. From this figure we can also observe that as the temperature of the stripe increases the normalized lifetimes for sample C show a tendency to converge. This is expected under high current and temperature stressing, and the effect is more prominent for metal lines with robust microstructure like that of sample C. We can also observe that, eventually with increasing temperature and current the failure lifetimes will become independent of the microstructure of the metal lines. The failure will be governed by thermally activated breakdown instead of voiding caused by metal migration under the influence of an electron wind force.

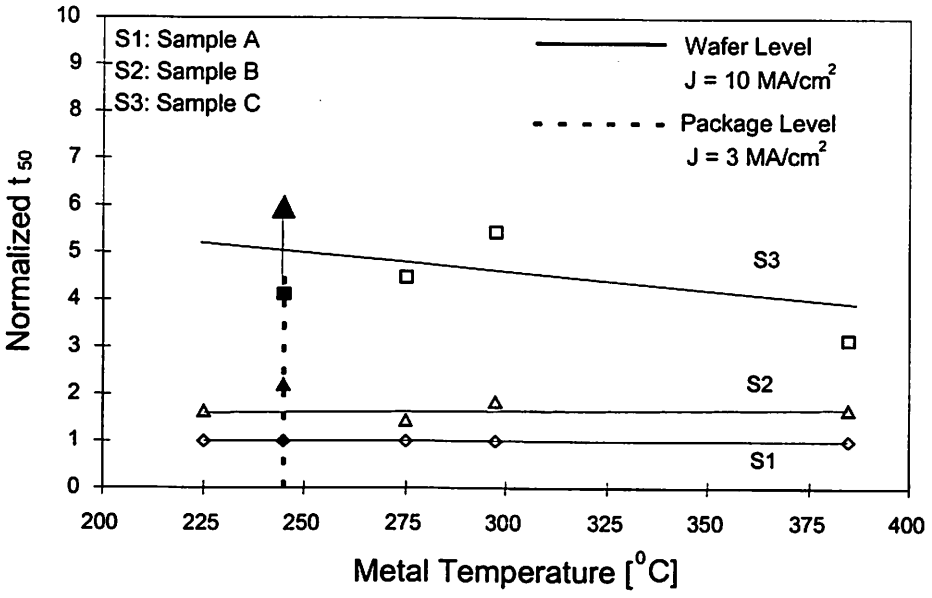


Figure 9.4 Accelerated stress EM t_{50} sensitivity with metal temperature. The solid symbols represent the normalized t_{50}s at package level stress where T_m = 245 ^{0}C and j = 3 MA/cm^2. The arrow is shown to indicate a greater than 4 normalized t_{50} value.

In Figure 9.5 the activation energies extracted from package and wafer level metal lifetimes are plotted. The activation energy values from the wafer level data are in the range of 0.63-0.73 eV. These values are close to the value of 0.78 eV obtained from package level test data.

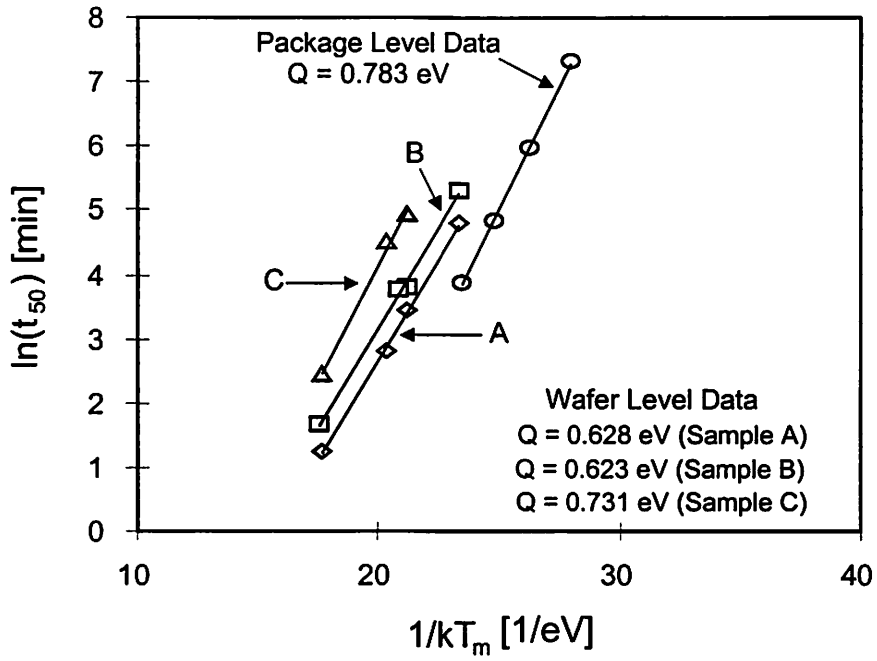


Figure 9.5 Activation energy (Q) for electromigration under package level and wafer level stress conditions.

These results indicate that for EM reliability assessment using large stress as typically done in wafer level testing, the metal joule heating needs to be accurately determined. This metal temperature variation can seriously affect EM lifetime evaluation. Also, the failure modes of the interconnect structures stressed under high current density conditions are similar to those stressed under package level (moderate stress) conditions. Optical micrographs clearly showed that the voids are created uniformly and randomly over the entire length of the line. However, their degree of occurrence is higher under high current density stress conditions due to the increased metal ion diffusivity at microstructure divergence sites.

These observations indicate that the physics of the failure mechanism is not completely changed by using up to 10 MA/cm^2 of current density, provided the metal stripe temperature is kept below 390 ^{0}C. We have been able to achieve a reduction in the EM testing time from greater than 50 hours at package level to less than 10 minutes at the wafer level, i.e., by a factor of more than 300. These results lend support to the possibility of extending the metal EM test current density to 10 MA/cm^2. This would significantly reduce EM testing time in a high volume production environment.

9.4 Summary

In this chapter we have briefly analyzed the effects of using highly accelerated stress conditions on the EM lifetimes of AlCu(0.5%) interconnects and on the sensitivity of the lifetime data for reliability ranking of various process designs. We have demonstrated that EM lifetimes under high current density stress conditions are in relatively good agreement with those under package level moderately accelerated stress conditions. Reduced testing times of less than ten minutes has been achieved using high current stressing as compared to greater than fifty hours for package level tests, while retaining excellent lifetime sensitivity. Also, no significant change in the failure mode is observed. These results support the possibility of extending the wafer level EM test current density to 10 MA/cm^2.

Chapter 10

Implications of Thermal Effects on Interconnect Reliability and Performance

10.1 Introduction

As discussed in Chapter 1, thermal effects in deep-submicron (DSM) interconnects are increasing with VLSI scaling. This is due to a dramatic reduction of the interconnect metal pitch and increasing number of metallization levels, which causes increasing current densities [8], [81], and associated thermal effects.

As mentioned earlier thermal effects are an inseparable aspect of electrical power distribution and signal transmission through the interconnects due to self-heating (or Joule heating) caused by the flow of current. Thermal effects impact interconnect design and reliability in the following ways. Firstly, they limit the maximum allowable RMS current density, $j_{RMS\text{-}max}$ (since the RMS value of the current density is responsible for heat generation) in the interconnects, in order to limit the temperature increase. Secondly, interconnect lifetime (reliability) which is limited by electromigration (EM, has an exponential dependence on the

inverse metal temperature [17]. Hence, temperature rise of metal interconnects due to self-heating phenomenon can also limit the maximum allowed average current density, $j_{avg\text{-}max}$, since EM capability is dependent on the average current density [18]. Thirdly, thermally induced open circuit metal failure under short-duration high peak currents including electrostatic discharge (ESD) is also a reliability concern [57], and as shown in Chapter 3, can introduce latent EM damage [42], that has important reliability implications.

Chatterjee et al., [81], have argued that thermal effects will increasingly dominate interconnect design rules that specify maximum current densities for circuit designers. Recently, Hunter [158], solved the EM lifetime equation for Al-Cu, and the 1-D heat equation, in a self-consistent manner. In this approach, both EM and self-heating can be comprehended simultaneously. Rzepka et al., [159], have carried out detailed simulations of self-heating in multilevel interconnects using finite element analysis. Their study concluded that in the near future, 3-D interconnect arrays will be affected by self-heating more severely than those of today's.

Furthermore, as discussed in Chapter 1 and 4, low dielectric constant (Low-k) materials are being introduced as an alternative intra-level insulator to reduce interconnect capacitance (therefore delay) and cross-talk noise to enhance circuit performance [9]. These materials can further exacerbate thermal effects owing to their lower thermal conductivity than silicon dioxide (see Figure 1.6). In Chapter 4 it was shown that interconnect scaling using low-k as the intra-level (gap-fill) dielectric material can cause significant increase in thermal effects.

Presently, interconnect design rules are generated in a non self-consistent manner [160], as explained in the next section. As deep sub-micron interconnect technologies, like Cu and low-k dielectric materials are rapidly evolving, there is an increasing need to understand their effects simultaneously on the thermal characteristics of these interconnects and on their EM reliability, in order to provide robust design guidelines. Furthermore, it is not clear whether thermal constraints conflict with the performance optimization steps employed at the circuit level. Hence, a thorough analysis of thermal effects in DSM interconnects is necessary to comprehend their full impact on circuit design, accurately model their reliability, and provide thermally safe design guidelines for various technologies.

This chapter presents a comprehensive analysis of the thermal effects in advanced high performance interconnect systems arising from self-heating under various circuit conditions, including ESD [161]. The analysis examines the self-consistent solutions for allowed interconnect current density for technologies up to 0.1-μm involving Cu and various low-k materials in an eight-level metallization system as per the National Technology Roadmap for Semiconductors (NTRS) [13].

10.2 Interconnect Design Rules: Current Approach

10.2.1 Average, RMS, and Peak Current Density

Circuit designers are typically provided with the maximum allowable values for three interconnect current densities. These are the average current density, j_{avg}, the RMS current density, j_{rms}, and the peak current density, j_{peak}. These quantities are defined as follows:

The peak current density is simply the current density corresponding to the peak current of the waveform,

$$j_{peak} = \frac{I_{peak}}{A} \tag{10.1}$$

where A is the cross-sectional area of the interconnect.

The average current density is defined as,

$$j_{avg} = \frac{1}{T}\int_0^T j(t)dt \tag{10.2}$$

where T is the time period of the current waveform. The RMS current density is defined as,

$$j_{rms} = \sqrt{\frac{1}{T}\int_0^T j^2(t)dt} \tag{10.3}$$

For a fixed temperature, EM lifetime of interconnects is known to be determined by j_{avg} [18]. Self-heating is determined by j_{rms}. Presently, high performance interconnect design is based on the specified limits for the maximum values of the average, RMS, and peak current densities [160].

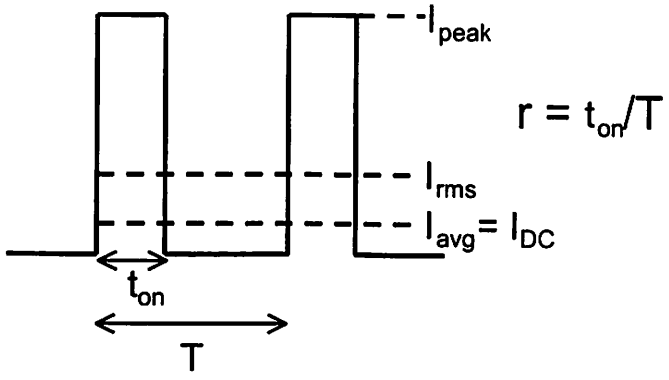


Figure 10.1 A unipolar pulsed waveform illustrating various current definitions.

They are however not self-consistent, i.e., these values do not simultaneously comprehend the two temperature dependent mechanisms: EM and self-heating.

Interconnects can be broadly classified into two categories: signal lines and power lines. They differ in that currents in signal lines are bi-directional (or bipolar) [18], while those in power lines are usually uni-directional (or unipolar).

We will now consider unipolar current waveforms for some illustrative analysis. Using the above definitions for the three current densities, one can easily show that

$$j_{avg} = r \cdot j_{peak} \tag{10.4}$$

and,

$$j_{rms} = \sqrt{r} \cdot j_{peak} \tag{10.5}$$

where r is the duty cycle defined as t_{on}/T in Figure 10.1.

10.2.2 Electromigration and Self-Heating

EM is the transport of mass in metals under an applied current density and is widely regarded as a major wear out or failure mechanism of VLSI interconnects [17]. When current flows through the interconnect metal, an electronic wind is set up opposite to the direction of current flow. These electrons upon colliding with the metal ions, impart sufficient momentum, and displace the metal ions from their lattice sites creating vacancies. These vacancies condense to

form voids that result in increase of interconnect resistance or even open circuit conditions [162]. EM lifetime reliability of metal interconnects is modeled by the well known Black's equation [17], given by,

$$TTF = A^{*} \cdot j^{-n} \cdot exp(\frac{Q}{k_B T_m}) \tag{10.6}$$

where TTF is the time-to-fail (typically for 0.1% cumulative failure). A^{*} is a constant that is dependent on the geometry and microstructure of the interconnect, j is the DC or average current density. The exponent n is typically 2 under normal use conditions, Q is the activation energy for grain-boundary diffusion and equals ~ 0.7 eV for Al-Cu, k_B is the Boltzmann's constant, and T_m is the metal temperature. The typical goal is to achieve 10 year lifetime at 100 ^{0}C, for which equation (10.6) and accelerated testing data produce a design rule value for the acceptable current density at T_{ref}, j_0. However, this design rule value does not comprehend self-heating.

The effect of self-heating can be analyzed from the following: The metal temperature, T_m in equation (10.6) is given by,

$$T_m = T_{ref} + \Delta T_{self-heating} \tag{10.7}$$

and,

$$\Delta T_{self-heating} = \left(T_m - T_{ref}\right) = I^2{}_{rms} \cdot R \cdot R_\theta \tag{10.8}$$

where T_{ref} is the reference chip (silicon junction) temperature and is typically taken as 100 ^{0}C, $\Delta T_{self\text{-}heating}$ is the temperature rise of the metal interconnect due to the flow of current, R is the interconnect resistance, and R_θ is the thermal impedance of the interconnect line to the chip. Thus, both EM and self-heating are temperature dependent effects, and as self-heating increases, EM lifetime decreases exponentially according to equation (10.6).

10.3 Thermal Effects in DSM Interconnects

10.3.1 Self-Consistent Interconnect Design Analysis

In this section we will first introduce the formulation of the self-consistent solutions [158] for allowed interconnect current density, and then apply them to analyze Cu interconnects. The $\Delta T_{self\text{-}heating}$ in interconnects given by equation (10.8) can be written in terms of the RMS current density as,

$$j^2{}_{rms} = \frac{(T_m - T_{ref}) \cdot K_{ox} \cdot W_{eff}}{t_{ox} \cdot t_m \cdot W_m \cdot \rho_m(T_m)} \tag{10.9}$$

Here t_m and W_m are the thickness and width of interconnect metal line, and $\rho_m(T_m)$ is the metal resistivity at temperature T_m. Note that the thermal impedance R_θ in equation (10.8) has been expressed as,

$$R_\theta = \frac{t_{ox}}{K_{ox} \cdot L \cdot W_{eff}} \tag{10.10}$$

This expression for the thermal impedance is based on a quasi-1-D heat conduction model with $W_{eff} = W_m + 0.88t_{ox}$, valid for $W_m/t_{ox} > 0.4$ and is accurate to within 3% [163]. Here t_{ox} is the total thickness of the underlying dielectric, K_{ox} is the thermal conductivity normal to the plane of the dielectric, and L is the length of the interconnect.

Now, in order to achieve an EM reliability lifetime goal mentioned in section 10.2, we must have the lifetime at any (j_{avg}) current density and metal temperature T_m, equal to or larger than the lifetime value (eg. 10 year) under the design rule current density stress j_0, at the temperature T_{ref}. This value of j_0 is dependent on the specific interconnect metal technology. Therefore we have,

$$\frac{exp\left(\frac{Q}{k_B T_m}\right)}{j^2{}_{avg}} \geq \frac{exp\left(\frac{Q}{k_B T_{ref}}\right)}{j_0{}^2} \tag{10.11}$$

From equation (10.4) and (10.5) we have after eliminating j_{peak}

$$\frac{j^2_{avg}}{j^2_{rms}} = r \tag{10.12}$$

Substituting for j^2_{rms} from equation (10.9) and j^2_{avg} from equation (10.11) in (10.12) we get the self-consistent equation given by

$$r = j_0^{\,2}\left[\frac{exp\left(\frac{Q}{k_B T_m}\right)}{exp\left(\frac{Q}{k_B T_{ref}}\right)}\right]\frac{t_{ox}\cdot t_m \cdot W_m \cdot \rho_m(T_m)}{(T_m - T_{ref})\cdot K_{ox}\cdot W_{eff}} \tag{10.13}$$

Note that this is a single equation in the single unknown temperature T_m. Once this self-consistent temperature is obtained from equation (10.13), the corresponding maximum allowed j_{peak} and j_{rms} can be calculated from equation (10.5) and (10.9). The self-consistent equation given by equation (10.13) for unipolar pulses is also valid for more general time varying waveforms with an effective duty cycle r_{eff} [164] .

One of the consequences of the self-consistent equation is that, for a certain j_0, as r decreases the self-consistent temperature and the maximum allowed j_{peak} increases. This effect is shown in Figure 10.2 for Cu interconnects. Secondly, it can be observed that as r decreases the ratio j_{peak}*(self-consistent)/*j_{peak}*(without self-heating, i.e., the line labeled 1/r)* decreases monotonically. At $r = 10^{-2}$, the self-consistent j_{peak} is nearly 2 times smaller than the j_{peak} obtained from EM constraint only, i.e., without self-heating. From the EM lifetime relation given by equation (10.6) this implies that if a design used only the average (EM) current as the design guideline without comprehending self-heating it could have a lifetime nearly three times smaller than the reliability requirement. Thirdly, it can be observed from Figure 10.3 that, as j_0 is increased, the self-consistent interconnect metal temperature (T_m) increases as expected from equation (10.13). However, the maximum allowed j_{peak} does not increase much for values of $r < 10^{-3}$. That is, j_0 becomes increasingly ineffective in increasing j_{peak} as the duty cycle r, decreases. From the above analysis, it apparently seems that EM capability might cease to be the dominant driver of interconnect technology evolution and thermal effects will limit the maximum allowed j_{peak}.

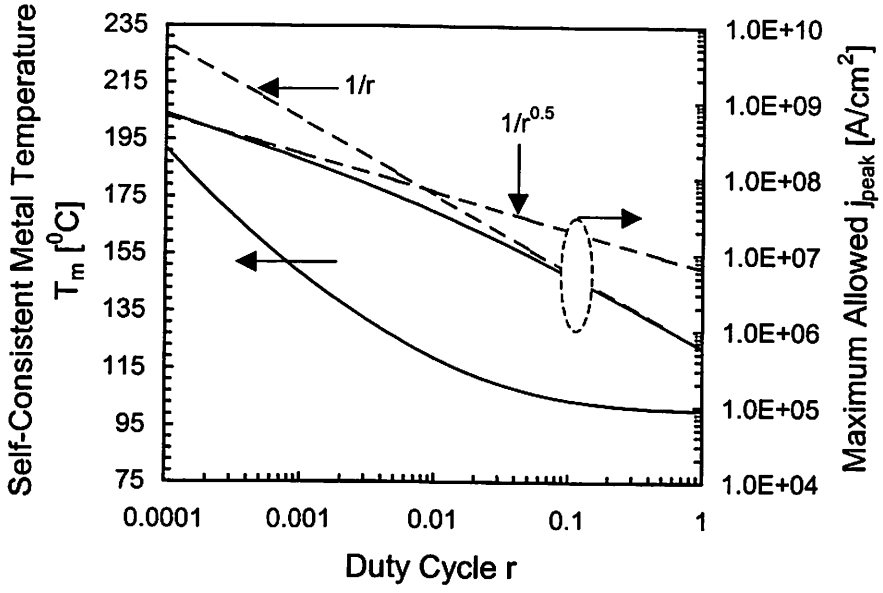


Figure 10.2 Self-consistent solutions for T_m and j_{peak}. Parameters used here are: $j_0 = 0.6$ MA/cm^2, $t_{ox} = 3$ μm, $t_m = 0.5$ μm, and $W_m = 3$ μm. Interconnect metal is Cu with $\rho_m(T_m) = 1.67 \times 10^{-6}$ Ω-cm $[1 + 6.8 \times 10^{-3}\ ^0C^{-1}\ (T_m - T_{ref})]$. The two dotted lines indicate j_{peak} values based on a) $j_{peak} = j_0/r$, and b) $j_{peak} = j_{rms}/r^{0.5}$.

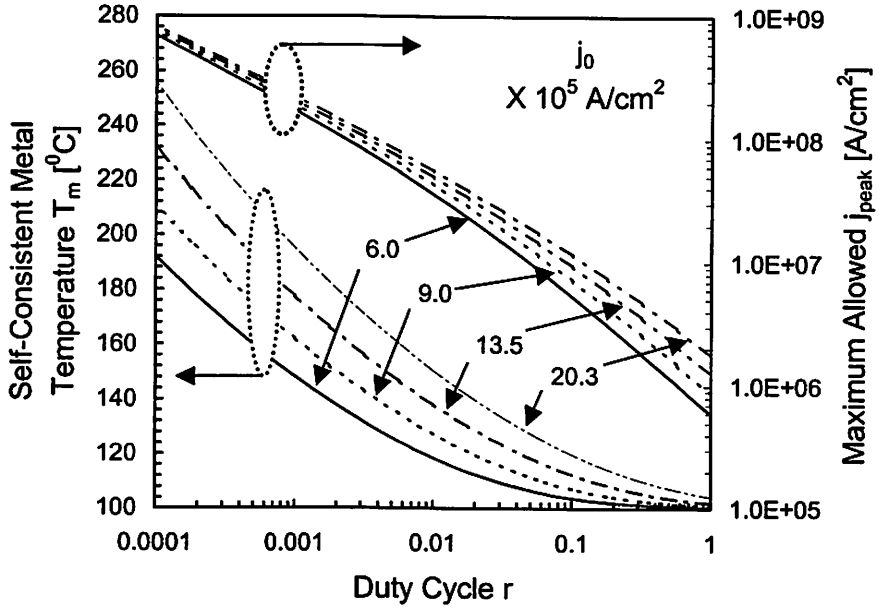


Figure 10.3 Self-consistent analysis showing dependence of T_m and j_{peak} on j_0. Parameters used here are: $t_{ox} = 3.0$ μm, $t_m = 0.5$ μm, and $W_m = 3.0$ μm. Interconnect metal is Cu.

The above analysis however motivates the following questions:

- Since the analysis assumes a quasi-1D heat conduction model, what will be the effects on the self-consistent solutions if a more realistic model is used?
- Since the analysis assumed that the intra-level and inter-level dielectric material is SiO_2, what will be the effects on the self-consistent solutions if low-k materials are considered as candidates for intra-level dielectric material?
- Since the analysis assumed, simplistic values of interconnect geometry, and underlying oxide thickness, what will be the effects on the self-consistent solutions, if technology and interconnect scaling were considered?
- Since the waveform shape or duty cycle r plays an important role in determining the maximum allowed T_m and j_{peak}, what is a realistic value for r from a circuit designer's point of view?
- For optimized driver and interconnect signal length (between repeaters) design, how do the j_{peak} and j_{rms} values (from purely electrical considerations such as signal delay) compare with those generated using the self-consistent approach?
- Since the analysis assumed an isolated interconnect line, what will be the effects on the self-consistent solutions if real 3-D interconnect arrays, with heat generation in all the lines, are considered?
- Finally, the question also arises that how do these j_{peak} values compare with peak current density values causing open circuit or latent metal damage in metal lines under ESD conditions?

10.3.2 Technology Scaling Effects on Self-Consistent Design Rules

In this section we will address the first three issues outlined above. We will first analyze the effects of line width scaling on the thermal impedance (or W_{eff}) defined by equation (10.10) which is accurate to within 3% for $W_m/t_{ox} > 0.4$ [163]. In current DSM technologies the ratio W_m/t_{ox} gets even smaller (especially for the top-level metal lines), and W_{eff}, which is meant to account for the extra heat conduction from the sides (illustrated in Figure 10.4) must be obtained for real DSM interconnect structures experimentally. We will first express W_{eff} in a more general form,

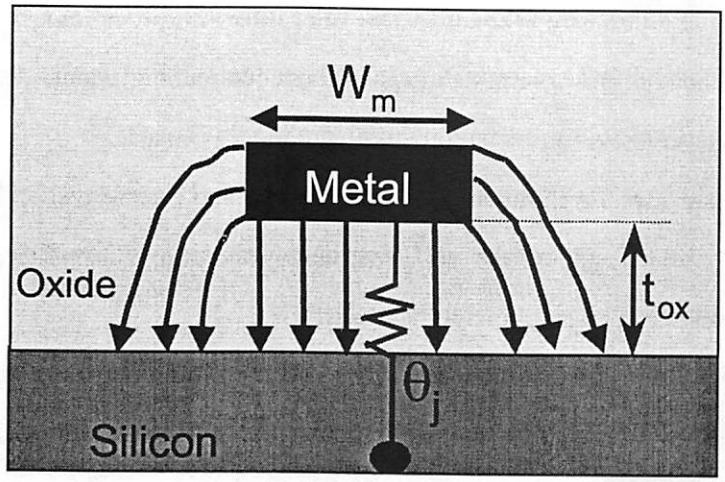


Figure 10.4 Schematic representation of quasi 2-D heat conduction scenario in DSM interconnects.

$$W_{eff} = W_m + \phi \cdot t_{ox} \tag{10.14}$$

Here ϕ is introduced as the heat spreading parameter that must be extracted from data. Furthermore, new intra-level insulating materials for DSM interconnects such as low-k dielectrics, which lower interconnect capacitance (and hence RC delay and cross-talk noise) and enhance circuit performance, are known to have poor thermal properties [165]. The thermal characteristics of these low-k dielectrics and their impact on interconnect reliability is important.

For extracting the appropriate value of ϕ, the thermal impedance was experimentally determined using equation (10.8), for AlCu interconnects fabricated in a state-of-the-art 0.25-μm industrial CMOS process flow. Results are shown in Figure 10.5 for a standard oxide process and a low-k (HSQ) intra-level (gap-fill) process. We can observe that for the narrowest line width (W_m = 0.35 μm) the thermal impedance for the low-k gap-fill process is around 20% higher than the value for standard oxide in agreement with the trends shown in [85]. Furthermore, since the ratio of W_m/t_{ox} (= 0.29) is the lowest for this line width, the heat spreading pattern for this case can be expected to be similar for top level metal lines in DSM technologies. Using the thermal impedance value for W_m = 0.35 μm and using equation (10.10) and (10.14), ϕ was extracted to be 2.45 (for the standard oxide case). The K_{ox} value used in the calculation is given in Table 10.1.

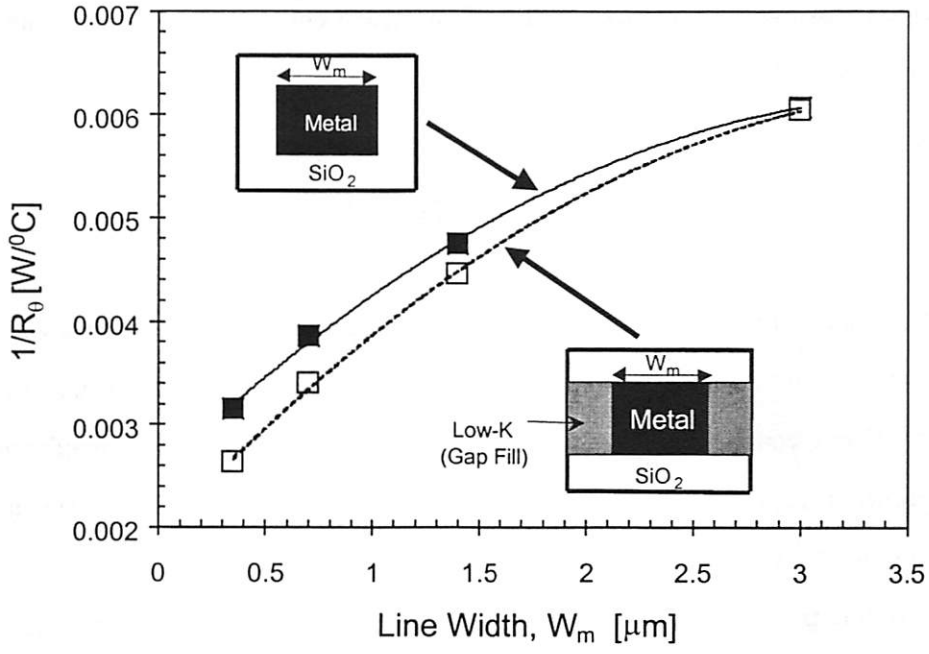


Figure 10.5 Experimentally obtained effective thermal impedance values for level 1 AlCu metal lines in a 0.25 μm CMOS process, with two different intra-level dielectrics. The t_{ox} was 1.2 μm, and interconnect length, L = 1000 μm.

Dielectric Material	Thermal Conductivity [W/m ^{0}K]
Oxide (PETEOS)	1.15*
HSQ	0.6**
Polyimide k	0.25**
Air	0.024***

Table 10.1 Thermal conductivity values (normal to the dielectric plane) for various dielectrics analyzed in this study. These values were obtained from [165] for * and [166] for ** and [22] for ***.

We now consider the technology specifications for a 0.25 μm and 0.1 μm process as per the NTRS [13] shown in Appendix B. The interconnect metal is assumed to be Cu, and the effects of using various intra-level dielectrics (listed in Table 10.1) are analyzed. We then solve for the self-consistent metal temperature, T_m and the corresponding maximum allowed j_{rms} and j_{peak}

values. However, while solving the self-consistent equation (10.13), we replace the ($t_{ox}/K_{ox}W_{eff}$) term by the following generalized term:

$$\left(\frac{t_{ox}}{K_{ox}} + \frac{t_{low-k}}{K_{low-k}}\right)\frac{1}{W_m + \phi \cdot t_{total}} \tag{10.15}$$

This is necessary to take care of the heat spreading effect in DSM structures and the use of low-k materials as intra-level gap-fill dielectric. The value of ϕ is kept unchanged even for the case of gap-fill interconnect structures, since the heat flow pattern is not significantly altered in the presence of gap-filled dielectric. Also, the duty cycle r in equation (10.13) is taken as 0.1, for all cases, due to reasons explained in the next section.

Before presenting the self-consistent solutions, we must point out an important caveat, which is the distinction between *thermally long* and *thermally short* metal lines. Since interconnect leads are typically connected to the diffusion (or another metal lead) through a contact (or via), the temperature at the end regions of these lines are in general lower than the temperature in regions far away from the ends. The solution of the differential equation for such 2-D heat conduction yields that the spatially dependent parts of the solution depend exponentially on a characteristic thermal length, λ [62], which is of the order of 25 μm - 200 μm. Metal lines that are >> λ are termed *thermally long* while those that are close to the value of λ are termed *thermally short*.

We will focus on the analysis of the thermally long lines which give rise to worst case scenario. We will not concentrate on thermal effects for small inter-block interconnects. Their lengths are usually of the same order of magnitude as the thermal characteristic length. Also these nets are typically routed on the first few levels of metal, which are closer to the silicon substrate, and hence the thermal problem is not as severe. Long interconnects which could be much larger than the thermal characteristic length will therefore form the main focus of the following analysis. These long lines usually realize the inter-block communication and hence for delay purposes, they are routed on higher, less resistive top few layers of metal. However, these metal layers are usually far away from the silicon substrate and hence can potentially experience thermal problems.

Therefore, self-consistent solutions were obtained for the top metal layers only in the 0.25-μm and 0.1-μm technologies.

We begin by analyzing the effect of introducing new interconnect (Cu) and dielectric materials on reliability. In Figure 10.6 the self-consistent values of j_{rms} and j_{peak} are plotted as a function of duty cycle (r) for different dielectrics. It can be observed that j_{rms} and j_{peak} decrease significantly as dielectrics with lower thermal conductivity are introduced. For small values of r, j_{rms} varies very slowly with r, thereby signifying the increasing importance of self-heating.

In Figure 10.7 we plot j_{peak} and T_m as a function of r for SiO_2 and air as the dielectric, and Cu metal, for different values of j_0. It can be observed that increasing j_0 does not change j_{peak} appreciably. In fact, as r reduces, the increase in j_{peak} with j_0 gets smaller. Furthermore, for a dielectric with low thermal conductivity such as air, the increase in j_{peak} with increasing j_0 is almost insignificant.

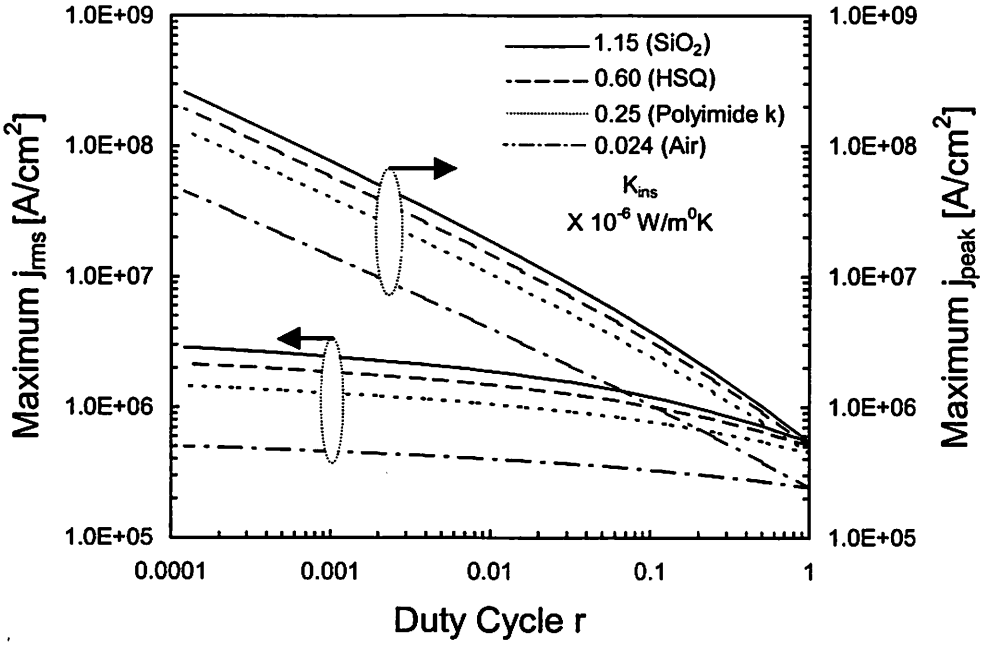


Figure 10.6 Self-consistent solutions for maximum allowed values of j_{rms} and j_{peak} for a metal 6 in 0.25 μm technology. Parameters used here are: $j_0 = 0.6$ MA/cm^2, $t_{ox} = 11.2$ μm, $t_m = 2.5$ μm, and $W_m = 2$ μm. Interconnect metal is Cu with $\rho_m(T_m) = 1.67 \times 10^{-6}$ Ω-cm [1 + 6.8×10^{-3} $^0C^{-1}$ $(T_m - T_{ref})$].

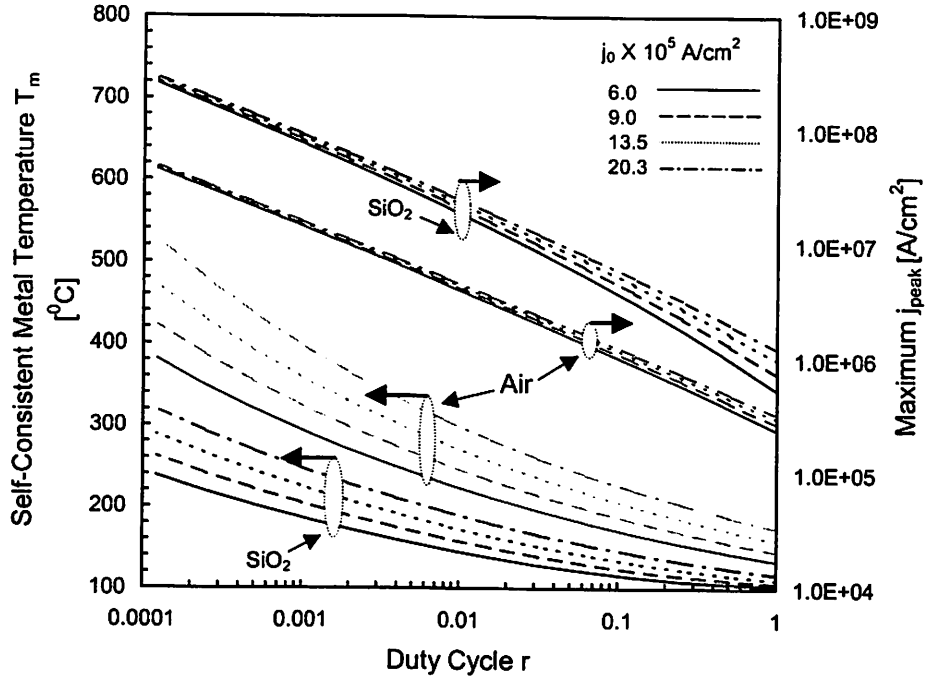
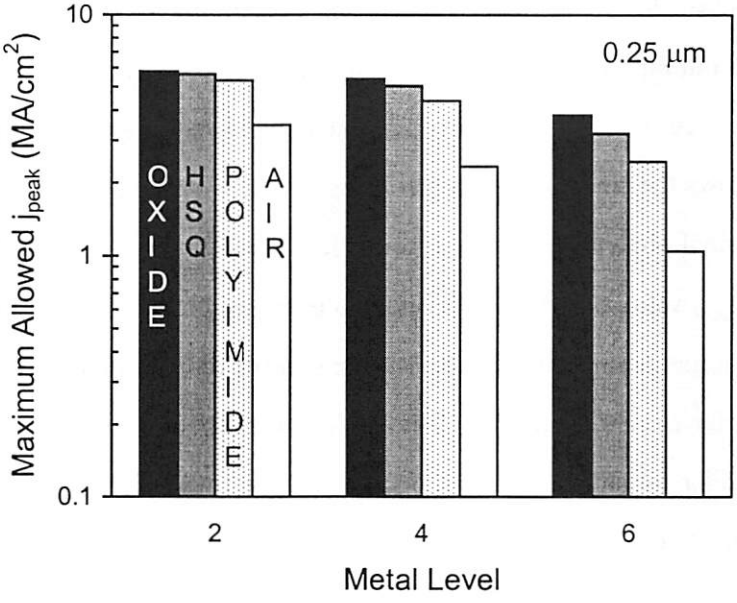


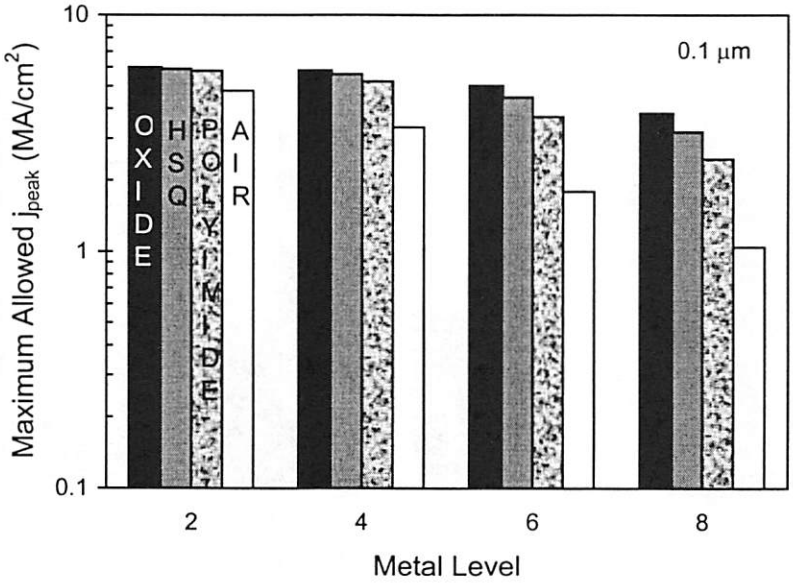
Figure 10.7 Self-consistent metal temperature and the maximum allowed j_{peak} for a metal 6 in a 0.25 μm technology, as a function of duty cycle for different values of j_0 for SiO_2 and air as the dielectric.

This indicates that introduction of better interconnect materials becomes increasingly ineffective in increasing j_{peak} for dielectrics with poor thermal properties. Another detrimental effect of using low-k dielectric materials is that the metal temperature increases which makes these lines more susceptible to high current failures, as discussed in Chapter 4.

Next the effect of increasing metal layers on interconnect reliability is examined. As mentioned in Chapter 1 thermal effects are more pronounced in higher layer metal lines since these lines are farther away from the substrate. This is observed in Figure 10.8(a) and Figure 10.8(b) where the self-consistent j_{peak} is plotted for various Cu metal layers for 0.25 μm and 0.1 μm technologies respectively. It is observed that the self-consistent j_{peak} decreases with increasing metal layers for both technology nodes and for all dielectrics. Furthermore, the reduction in j_{peak} with increasing metal layers is more pronounced for lower dielectric materials.



a)



b)

Figure 10.8 a) Effect of metallization levels and dielectric materials on maximum allowed j_{peak} for a a) 0.25 μm technology (r = 0.1) b) 0.10 μm technology (r = 0.1). $j_0 = 6 \times 10^5$ A/cm^2.

We have also compared current carrying capabilities of signal and power lines and the results are shown in Figure 10.9. Consistent with Figure 10.6, it can be observed that the j_{peak} values are about an order of magnitude lower for the power lines due to increased thermal effects. Figure 10.10 shows the diminishing returns of using Cu as the interconnect material as dielectrics with poor thermal properties are introduced. It can be observed that the difference in the maximum allowed j_{peak} values between AlCu and Cu interconnects reduces as dielectric materials with poor thermal properties are introduced. For the specific case of air as the dielectric, the j_{peak} values are very similar for AlCu and Cu. This is demonstrated for two different values of j_0 for Cu, one value similar to that of AlCu and the other three times higher than this value. This is done to accommodate process dependence of EM performance in Cu.

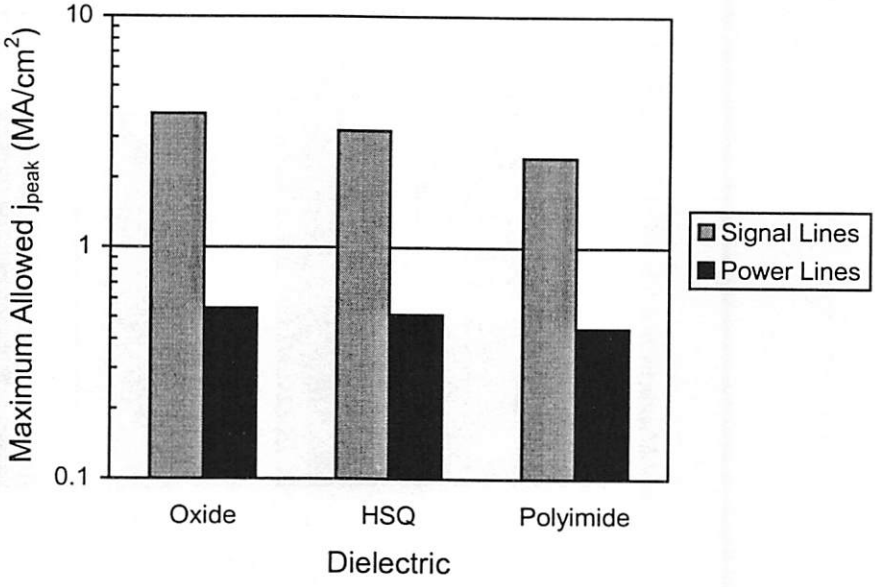


Figure 10.9 Comparison of the maximum allowed j_{peak} values for signal (r = 0.1) and power lines (r = 1.0) for metal 8 of a 0.10 μm technology shown for different dielectric materials. The lower values of j_{peak} in power lines are because of their higher temperature rise resulting from carrying DC current.

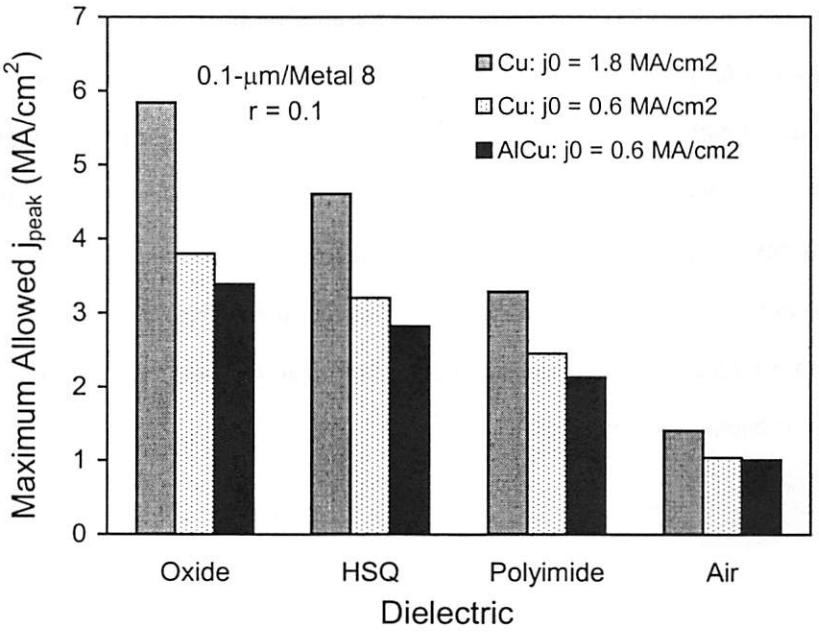


Figure 10.10 Comparison of the maximum allowed j_{peak} values for Al-Cu and Cu lines with two different j_0 values for metal 8 of a 0.10 μm technology shown for different dielectric materials.

10.4 Implications for Circuit Performance and Design

10.4.1 Effects on Signal Line Length, and Driver Size Optimization

In this section we will address the next two issues pointed out in section 10.3.1. We will compare j_{peak} values from thermal and reliability considerations with those obtained from circuit performance considerations. As we have mentioned in Section 10.3.2, thermal effects are predominant in semi-global and global interconnects which are thermally long. These are used to connect functional blocks (macros) over relatively long distances and hence utilize the upper layers of interconnect. Furthermore, long interconnect lines are split into buffered segments [167]. Buffer insertion is also used to reduce the crosstalk noise in high performance circuits [168]. It has been shown [167], that for point to point connections, the optimal length of interconnect at which to insert repeaters is given by

$$l_{opt} = 1.3\sqrt{\{r_0 \frac{(c_0 + c_p)}{r \cdot c}\}} \qquad (10.16)$$

where r_0 is the effective driver resistance for a minimum sized driver, c_0 is the input capacitance, c_p is the output parasitic capacitance, r is the interconnect resistance per unit length and c is the interconnect capacitance per unit length (see Figure 10.11). For DSM technologies, a significant fraction of c would be contributed by coupling capacitances to neighboring lines. For our simulations, we performed a full 3D-capacitance extraction using SPACE3D [169] for the signal lines to obtain the value of c for every metal layer for both technologies. The model assumes that the output V_{tr} makes a transition as soon as the input voltage crosses a threshold (say $0.5V_{dd}$). The optimum repeater size is given by,

$$s_{opt} = \sqrt{\frac{r_0 \cdot c}{c_0 \cdot r}} \tag{10.17}$$

which means that the width of the NMOS and PMOS for this inverter are obtained by scaling the width of the corresponding transistors in a minimum sized inverter by a factor s_{opt}. This also implies that the delay between any two optimally spaced and sized repeaters is independent of the layer [167]. This model also implies that it is not optimal to insert buffer in a line, which is of length less than l_{opt} for a given metal layer. Our simulations suggest that drivers and interconnects optimized with this model maintain good slew rates for rising and falling transitions for all the metal layers and across technologies. For lines of length less than l_{opt}, the buffer size can also be reduced to $\{s_{opt}(l/l_{opt})\}$ to reduce the power dissipation while still maintaining good slew rates. Also, given the distributed nature of the interconnect, the maximum RMS current occurs close to the repeater output. Hence, we need to first verify whether this maximum current, which is dictated by delay considerations also meets the thermal specifications obtained in Section 10.3.

In the following, we perform some calculations using a well-reviewed technology file (see Appendix B) based on [13] for 0.25 μm and 0.10 μm technology that employs copper and low-k interconnect structures. We use these parameters to determine the peak and RMS current (and therefore current density) by SPICE simulation where we take into account all the device parasitics. We denote the maximum current density due to electrical delay considerations as $j_{\text{-electrical}}$ and due to thermal and EM considerations as $j_{\text{-self-consistent}}$.

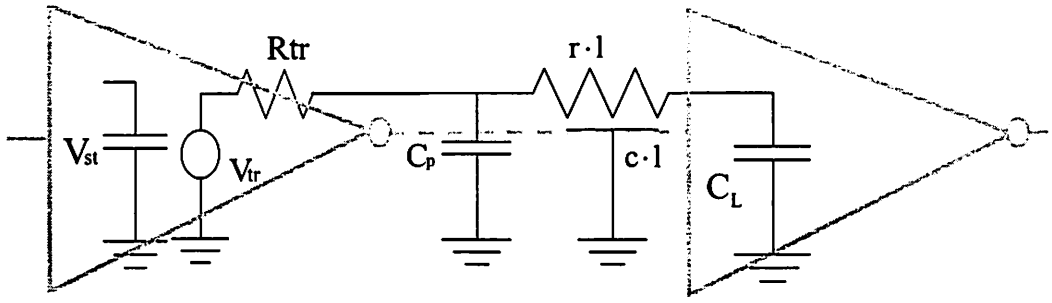


Figure 10.11 Equivalent distributed network for driver and interconnect. *Vst* is the voltage at the input capacitance that controls the voltage source. *Rtr* is the driver transistor resistance and C_L is the load capacitance of the next stage. *l* is the interconnect length.

As suggested earlier, we find that the relative slew rate defined as the 10%-90% rise time as a fraction of the clock period is almost constant across all metal layers and across technologies. This is also reflected by the fact that the effective duty cycle obtained for these simulations was 0.12±0.01 for every case considered, as shown in Table 10.2. The interconnect current waveform in the top layer metal lines for 0.25 μm and 0.1 μm technologies, as obtained from transient SPICE simulations, is plotted in Figure 10.12, which shows that the relative rise and fall skew is the same across both technologies. For lines which are not on critical path, the buffer size may be reduced to save power in which case the effective duty cycle will actually increase slightly, but, as suggested from Figures 10.2 and 10.3, this will not change $j_{\text{-self-consistent}}$ significantly. This justifies the selection of the duty cycle value as 0.1 in section 10.3.2 for the self-consistent analysis on various technologies.

Figure 10.13 shows the comparison of $j_{\text{peak-electrical}}$ with the values of $j_{\text{peak-self-consistent}}$ for 0.25 μm and 0.1 μm Cu technologies respectively. We find that $j_{\text{peak-electrical}}$ is lower than $j_{\text{peak-selfconsistent}}$ for silicon dioxide as the dielectric. However, for low-k dielectric materials, we find that the difference between $j_{\text{peak-self-consistent}}$ and $j_{\text{peak-electrical}}$ reduces due to increasing thermal effects.

Layer	L_{opt} (mm)	S_{opt}	j_{rms} (x10^{5}A/cm^2)	j_{peak} (x10^{5}A/cm^2)	r_{eff}
1	2.51	75	9.23	27.02	0.11669
2	2.41	80	8.76	26.38	0.11027
3	4.79	157	5.03	14.69	0.117244
4	5.48	148	4.16	13.48	0.100237
5	12.5	494	1.76	5.18	0.115443
6	13.9	454	1.58	4.69	0.113493

a)

Layer	L_{opt} (mm)	S_{opt}	j_{rms} (x10^{5}A/cm^2)	j_{peak} (x10^{5}A/cm^2)	r_{eff}
1	0.85	36	14.24	38.99	0.133387
2	0.9	40	14.32	39.48	0.131562
3	1.57	97	7.57	20.93	0.130814
4	1.79	85	6.63	18.27	0.131689
5	4.47	276	3.56	9.78	0.132502
6	4.68	249	3.35	9.00	0.138549
7	8.8	500	1.35	3.67	0.135312
8	9.58	522	1.26	3.5	0.1296

b)

Table 10.2 Optimized interconnect and buffer parameters, corresponding RMS and peak current densities ($j_{RMS\text{-}electrical}$ and $j_{peak\text{-}electrical}$) and effcetive duty cycle ($r_{eff} = j^2_{avg}/j^2_{rms}$) for a) 0.25-μm Cu technology with insulator dielectric constant = 3.3, and b) 0.1-μm Cu technology with insulator dielectric constant = 2.0.

Low-k dielectrics impact interconnect performance in two ways. Firstly, as suggested by equation (10.16) and (10.17) the optimum unbuffered interconnect length increases and the optimum repeater size decreases. It can be shown that s_{opt} and $c.l_{opt}$ decrease by the same factor and hence the RMS current density remains almost unchanged. Secondly, the lower thermal conductivity of these materials (see Table 10.1) causes $j_{peak\text{-}self\text{-}consistent}$ to decrease by almost a factor of two for these low-k dielectrics and the margin between $j_{peak\text{-}self\text{-}consistent}$ and $j_{peak\text{-}electrical}$ reduces. As we will see in Section 10.5, the thermal situation is exacerbated with the presence of

other active metal lines at the same metal layer as well as other layers. This further reduces $j_{peak\text{-}self\text{-}consistent}$ for these configurations. Hence, self-heating needs to be considered in high performance deep sub-micron interconnect design that employs low-k dielectrics.

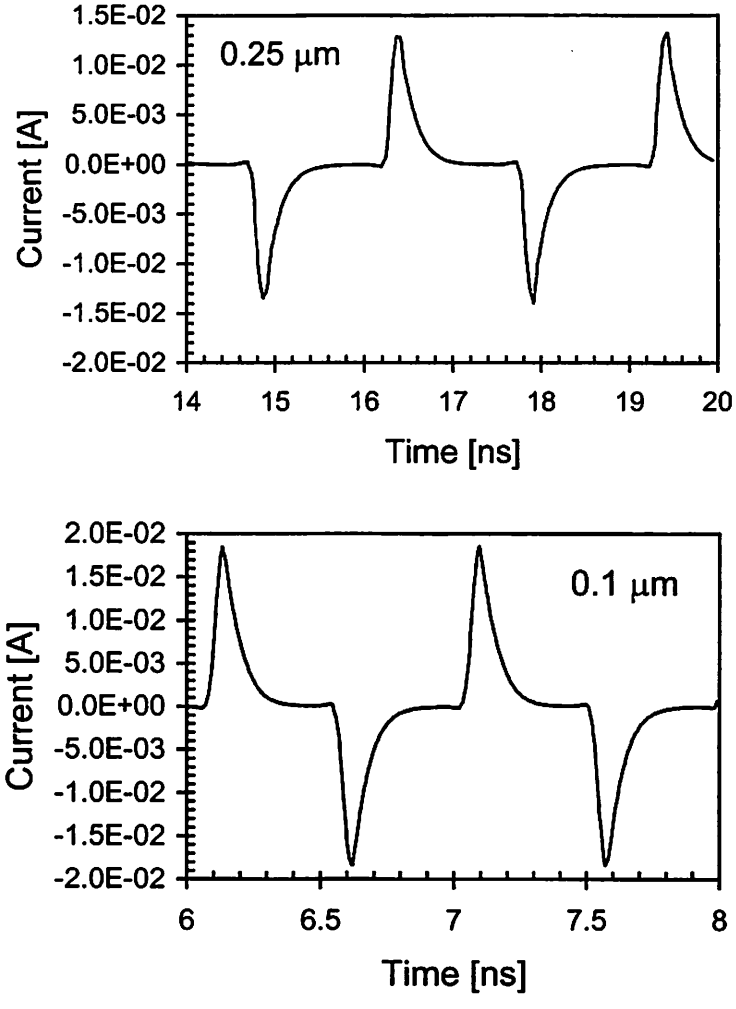
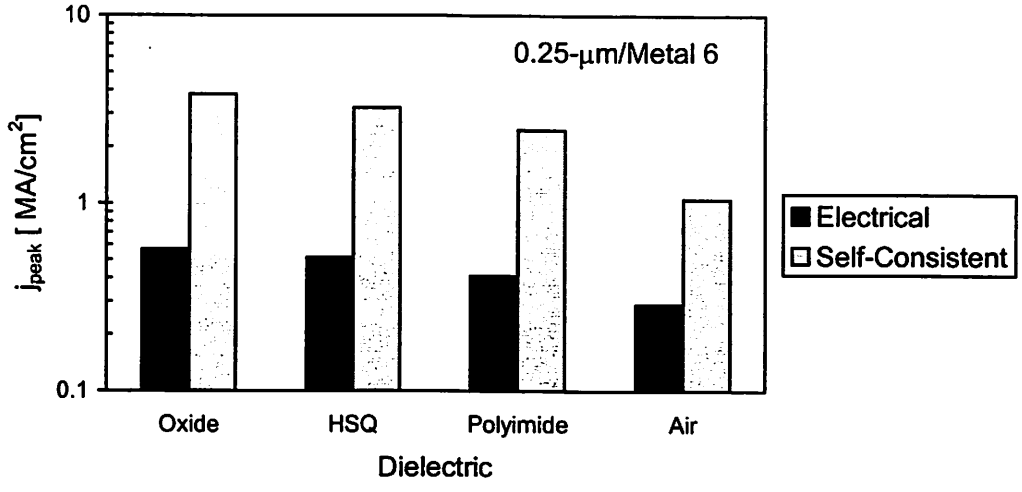
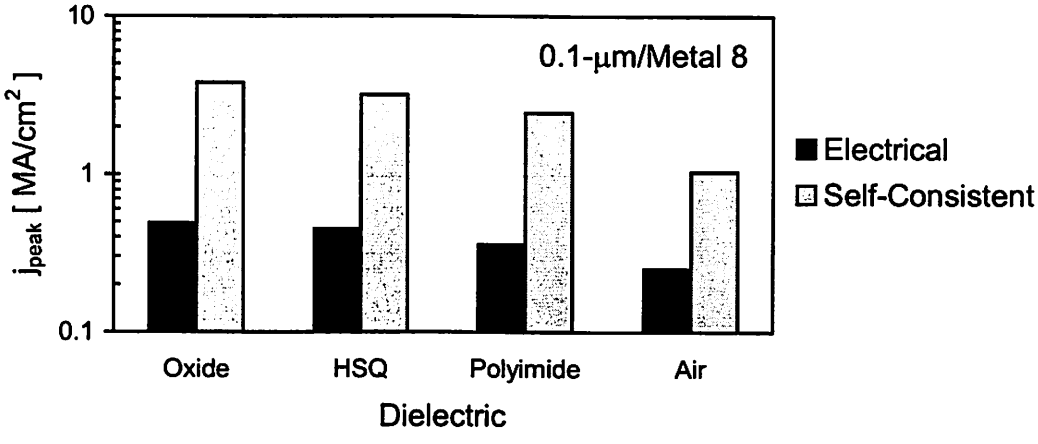


Figure 10.12 Current waveforms in the top layer metal lines obtained from SPICE simulations for the 0.25 µm and the 0.1 µm technologies.



a)



b)

Figure 10.13 Comparison of j_{peak} values obtained from reliability and electrical performance considerations for a) metal 6 of a 0.25 μm technology, and b) metal 8 of a 0.1 μm technology.

It should be kept in mind that the above analysis assumes that the interconnects are always active. This is a valid assumption for global lines, because although the activity period of individual transistors is reducing with scaling and increasing transistor count, block size is also increasing and hence the activity period for the whole block is relatively unchanged. Hence the global wires which communicate with these blocks have a high activity rate. Therefore, duty

cycle of 0.1 used in the study is valid. Also note that the circuit analysis of this section is for signal lines, which carry bi-directional currents. These lines are known to have much higher EM immunity, hence the self-consistent values of j_{rms} and j_{peak}, as shown in Figure 10.1-10.6 for the unipolar case, are lower bounds.

10.5 Thermal Effects in Real 3-D Interconnect Arrays

In this section we will briefly address the penultimate issue pointed out in section 10.3.1. Our self-consistent analysis of self-heating and EM effects presented earlier in section 10.3.2 were based on single isolated interconnect lines. In a real IC there are densely packed layers of interconnect lines which form a 3-D array as illustrated in Figure 10.14. The self-heating of interconnect lines within such array could be significantly more severe due to thermal coupling between neighboring lines. The heat flow analysis for such structures is complicated and must involve numerical simulation techniques. We have recently done a very detailed analysis of self-heating effects in such structures using finite element simulations [159].

The RMS current density can be empirically shown to obey the following relationship [see equation (10.9)]

$$j^2_{rms} \propto \frac{T_m - T_{ref}}{\rho_m(T_m)} \tag{10.18}$$

and the proportionality constant (which is independent of the interconnect material) can be obtained empirically from the finite element analysis presented in [159], and shown in Figure 10.15 for different interconnect configurations, which includes 2-D and 3-D effects.

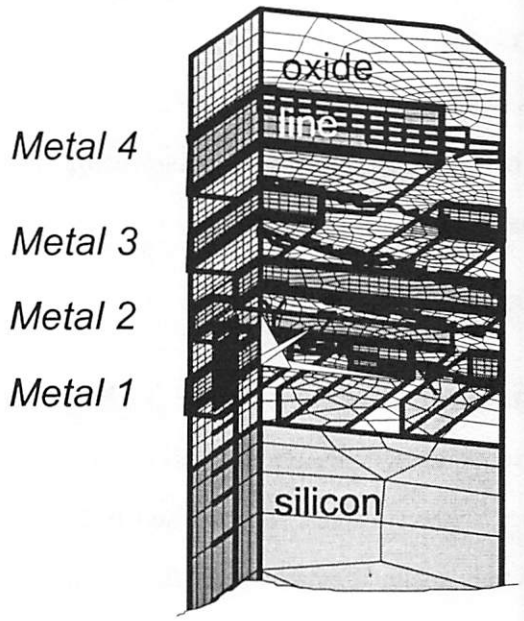


Figure 10.14 Densely packed 3-D interconnect array in a 0.5 μm quadruple level metallization process with SiO_2 as the dielectric material. Self-heating of any line within such an array is strongly affected by thermal coupling from the neighboring lines.

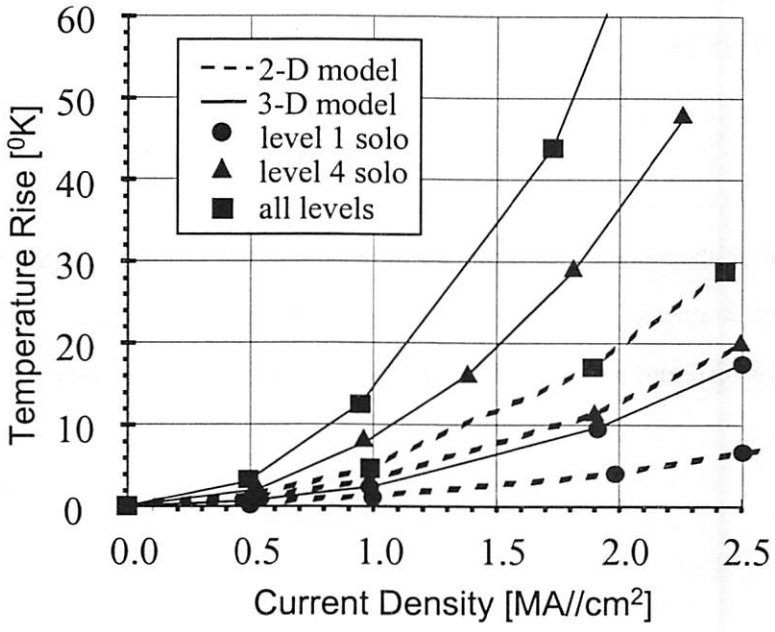


Figure 10.15 Finite element simulation results showing temperature rise under a DC current, test structures: 2-D model (broken lines), real structures: 3-D model (solid lines) [159].

Substituting this equation and the EM reliability equation (10.11) in equation (10.12), we can obtain another self-consistent equation similar to equation (10.13). From this we can calculate the maximum allowed j_{peak} for metal 4 in a 3-D configuration using Cu technology. We also compare this with the j_{peak} obtained from Figure 10.15 for an isolated metal 4. We find that the maximum allowed j_{peak} reduces by nearly 50% for the 3-D case, where all the metal lines are heated with equal current load in all the leads. Hence we can conclude that the maximum allowed j_{peak} for real interconnects would also reduce significantly from the values calculated in Section 10.3.2.

Figure 10.16 shows the j_{peak} values for 3D interconnect structures using this 50% reduction in j_{peak} to illustrate a best case effect of thermal coupling in a 0.1 μm Cu technology using various dielectrics. It can be observed that $j_{peak\text{-}3D}$ comes even closer *to* $j_{peak\text{-}electrical}$. However it should be noted that for deep sub-micron technologies that employ low-k dielectrics, this 50% reduction in j_{peak} is optimistic, since increasing interconnect packing density and low-k materials will increase thermal effects more than that of the 0.5 μm technology considered here.

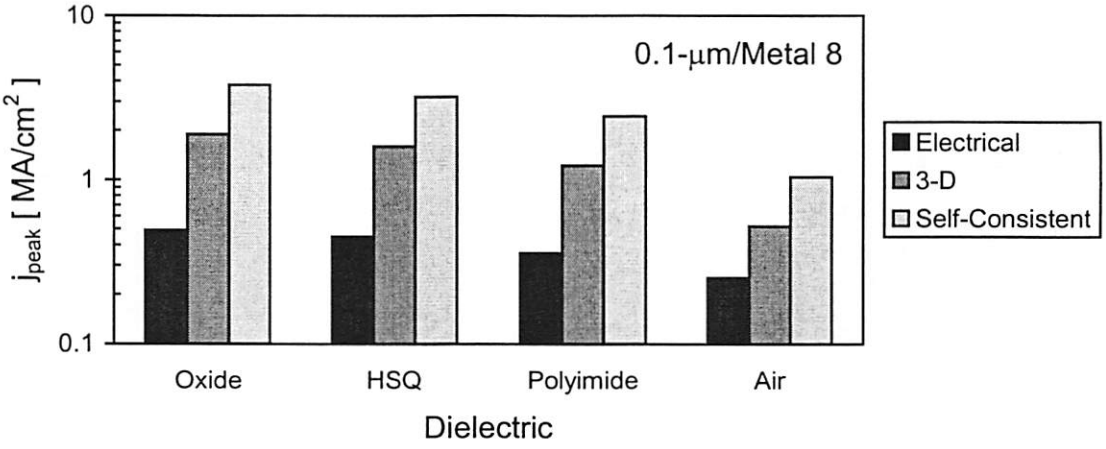


Figure 10.16 The effect of thermal coupling in interconnect arrays on the j_{peak} values is shown here as a best case scenario for various dielectric materials.

10.6 Thermal Effects under ESD Conditions

In this section we will briefly address the last issue pointed out in section 10.3.1. As discussed in Chapter 3 a special case of thermally accelerated interconnect failure can also occur under ESD conditions, which is a high current short-time scale phenomena that can lead to catastrophic open circuit failures, and to latent damage [27], [42]. Integrated circuit failures due to ESD have significantly increased in importance as VLSI technologies continue to shrink towards the deep sub-micron regime [28]. In Chapter 3 we have shown that the critical current density for causing open circuit metal failure in AlCu interconnects is ~ 60 MA/cm^2, and presented a short-timescale high-current failure model for designing robust interconnects to avoid thermal failure under high peak current conditions including those encountered in ESD protection circuits and I/O buffers. As shown in Chapter 3, interconnects can also suffer latent damage if the lines resolidify after melting and this has been shown to degrade the EM lifetime [42]. The model presented in [42], can be used to avoid this reliability hazard and has been shown to be valid for designing Cu interconnects [20]. These interconnect design rules must be obeyed for high current robustness.

10.7 Summary

Thermal effects in advanced high performance interconnect systems arising due to self-heating under various circuit conditions, including ESD has been examined in detail. A self-consistent approach, that simultaneously comprehends self-heating and EM, has been first modified using experimentally measured self-heating data to include quasi 2-D heat conduction in DSM metal lines to allow more aggressive design rules. This modified self-consistent formulation has been applied to analyze the implications of interconnect technology (Cu, low-k etc.) and scaling. Two different Cu based technology nodes (0.25 μm and 0.1 μm) as per NTRS has been used in this study.

Using circuit level simulations we have confirmed that the effective duty cycles for optimized global interconnects remain nearly invariant across metal layers and technologies. A value of 0.12±0.01 was obtained for all cases. This analysis has helped provide a more realistic value of the duty cycle for analyzing thermal and EM effects using the self-consistent approach. Secondly, the implications of these thermal effects on the design (driver sizing) and optimization of the interconnect length between repeaters at the upper-level signal lines have been investigated. We have shown that although the maximum allowed peak current density based on the self-consistent criteria is greater than the peak current density obtained from optimized driver and interconnect configurations (i.e., $j_{peak\text{-}self\text{-}consistent} > j_{peak\text{-}electrical}$), for standard oxide, but they get closer as low-k dielectric materials are introduced. Furthermore, we have shown that in real 3-D interconnect arrays the greater self-heating due to thermal coupling effects could further lower the $j_{peak\text{-}self\text{-}consistent}$ by ~ 50% in a 0.5 μm technology with SiO_2. Hence for DSM technologies that employ low-k dielectrics, thermal effects in interconnects need careful consideration.

Finally, we have pointed out that although the $j_{peak\text{-}self\text{-}consistent}$ values are much lower than the current densities causing open circuit metal failure under ESD type conditions, the interconnects in ESD protection circuits and I/O buffers must be designed separately to meet high current robustness. These results will have important implications for providing robust and aggressive deep sub-micron interconnect design guidelines.

Chapter 11

Conclusions and Future Directions

Conclusions

In this thesis, we have carried out a detailed investigation of high-current and thermal effects in deep sub-micron (DSM) VLSI interconnects and their implications for reliability and performance, and have also investigated their failure mechanisms. We have characterized and modeled high-current behavior of DSM interconnect structures under various circuit conditions including electrostatic discharge (ESD), which according to semiconductor industry surveys, is the largest single cause of failures in ICs.

A very important aspect of our work was the development of a novel thermometry technique, namely the transient resistive thermometry (TRT) together with a variation which employed the Kelvin method, namely the transient Kelvin thermometry (TKT). These techniques allowed for the first time, a systematic study of high-current and thermal effects in various interconnect structures present in a modern VLSI technology, such as metal lines, low-k dielectrics, contacts, vias, and silicide films. These techniques allow precise identification of failure locations under high-current stress conditions and has been employed to study failure mechanisms of the interconnect structures.

Using the TRT technique we have characterized and modeled thermal effects in VLSI metallization arising due to ultra short-duration pulses including ESD events. The TRT technique can provide a spatially averaged temperature rise of interconnect lines with excellent temporal resolution during nanosecond time frames. We have formulated a model that takes into account heat diffusion under non-steady-state conditions, to generate design guidelines for AlCu interconnects used in special protection circuitry and I/O buffers. Several leading IC

manufacturers including Texas Instruments, IBM, and Intel have incorporated this model in their design process. In fact, a recent work from IBM [*Voldmann et al., Proc. IRPS 98*] has applied this model to design damascene-Cu interconnects. In the process, we have also discovered a new latent interconnect failure mode, that significantly degrades electromigration lifetimes of AlCu interconnects, and can have important implications for reliability.

We have also characterized the effects of interconnect scaling and low-k dielectric on the thermal characteristics of metal lines under DC and pulsed current conditions for the first time. The TRT technique has also been shown to be useful for thermal characterization of thin dielectric materials. This technique can be used to evaluate the effective thermal conductivity of various low-k inter-layer dielectric films used in DSM interconnect processes.

In addition, we have also characterized and modeled the high-current behavior of thin Titanium and Cobalt silicide films used in advanced CMOS technologies. We have shown that the sensitivity of these films to high-current conduction is strongly dependent on the processing conditions. In conjunction with the high-current characterization using the TRT technique, we have carried out extensive TEM microstructure characterization to get an insight into the evolution process of silicide film degradation under high-current stress conditions. This study will help the impact of technology design and scaling of silicide films to be defined for the reliability of deep sub-micron CMOS technologies.

The TKT technique has been used to characterize self-heating in deep sub-micron contacts and vias. Using this technique we have identified a new failure mode of silicided barrier contact structures made of W and Al.

We have also characterized and modeled the effects of temperature and current on the $TiSi_2$-Si contact-resistance of small geometry silicided contact structures. We have demonstrated that contact-resistance sensitivity to temperature and current is strongly influenced by the silicide thickness, and is independent of the contact-plug material. We have developed an unified quantitative model which comprehends both field-emission and thermionic field-emission carrier transport mechanisms, and directly relates the contact resistance to temperature, and doping concentration of diffusion.

By close collaborations with the Microscale Thermophysics group at Berkeley we have advanced a technique for measuring the spatial temperature rise in 3-D interconnect links such as vias with sub-100 nm resolution. For the first time, we were able to study the steady state and dynamic thermal characteristics of small geometry W-plug vias and obtained the magnitude and spatial distribution of temperature rise, along with the topographical image using a novel microscale thermometry technique called scanning Joule expansion microscopy.

Furthermore, we have also analyzed the impact of self-heating on interconnect reliability evaluation during highly accelerated EM tests. We have demonstrated a correlation between package-level (moderate stress conditions) EM and wafer-level highly accelerated tests and have shown that accelerated wafer level EM tests can at least be employed for reliability ranking of various interconnect processes provided the self-heating was accurately calibrated. We have proposed a wafer level EM test current density of 10 MA/cm^2.

Finally, we have also done a detailed analysis of the implications of thermal effects and electromigration on Cu/low-k based interconnect performance. We have demonstrated the increasing importance of thermal effects on the design and optimization of the interconnect length between repeaters at the upper level signal lines.

Future Directions

1. Thermal Management of High Performance ICs

Thermal management will be a key issue in the coming generations of CPUs, low-power ASICs, or high-performance ASICs. In high performance microprocessors, different functional blocks have different sleep modes, and significant temperature gradients can develop between these units. These temperature gradients can severely impact circuit performance by increasing the clock skew. It could give rise to a certain percentage increase in skew within a transient cycle or could cause a DC offset of skew under steady-state conditions. Additionally, the design process must comprehend delay variations due to thermal gradients. As VLSI technology progresses towards sub-100 nm feature size these effects are expected to become increasingly dominant. These issues need to be addressed not only at the early design stage but even at the synthesis stage. For example, temperature dependent delay of logic gates would enable timing analysis to take into account the thermal variations.

In response to the increasing importance of thermal effects at system and device level, we now see the advent of techniques to model thermal effects in designs, and of tools to simulate the thermal profile of a design. However, there haven't been any noticeable efforts at integrating an active knowledge of thermal management into the whole design process, from synthesis all the way down to physical design. Until now the problem has been fixed by appropriate conservative design rule restrictions or by post-design corrections based on a thermal profile simulation. This approach will be very inefficient and cost intensive for future giga-scale ICs. Hence synthesis must incorporate thermal information to improve circuit performance and reliability. Moreover, modeling, simulations, and reliability studies of high performance chips should be done simultaneously and interactively with the design and testing operations. This will allow continuous feedback of thermal information throughout the design process, and hence circuit optimization can be performed around them at every stage.

2. Thermal and Reliability Issues in Novel Device/Circuit Performance

Design and performance of novel circuit architectures, such as 3-D circuits, and several micro-electromechanical systems (MEMS) are also expected to be strongly affected by thermal effects. Thermal effects can also influence the performance of high frequency VLSI and RF-power circuits. Additionally, ESD protection circuit design for high-speed VLSI and RF circuits are going to be very important isssues in the near future. A specific example of thermal problems in 3-D circuits is discussed below.

Material limitations can force a discontinuity in Moore's law as VLSI technology moves towards the nanometer regime. Interconnect is one area where these limitations affect the performance and reliability of deep sub-micron ICs most significantly. In fact interconnection is now one of the most significant factors in determining signal delay, and is expected to become the dominant factor as we move into the nanometer regime. At 0.25 μm technology node, Cu with low-k dielectric was introduced to alleviate the adverse effect of increasing interconnect delay. However, below 130 nm technology node, even these new materials introduce substantial interconnect delays, which in turn will severely limit the chip performance. Further reduction in interconnect delay cannot be achieved by introducing any new materials. This problem is especially acute for global interconnects, which typically comprise about 10% of total wiring, for current architectures. The aggressive scaling of device geometries and increasing die size further aggravates this problem. Also the problem of long-lossy lines cannot be fixed by simply widening the metal lines and using thicker interlayer dielectric since this conventional solution will lead to a sharp increase in the number of metallization layers. Such an approach will increase the complexity, reliability and cost, and will therefore be fundamentally incompatible with the decades-old industry trend of 25% per year improvement in cost per chip function. Therefore innovative solutions beyond mere materials changes are required to meet future IC performance goals. We need to think beyond the current paradigm of system architecture.

3-D integration of circuits to create multiple layers of active devices, is a highly promising technology that can potentially solve the interconnect problem and can simultaneously help reduce chip area and allow higher transistor packing density. This technology can greatly enhance the performance of logic circuits. For instance, logic gates on a critical path can now be

placed very closed to each other using multiple active layers. This technology can be exploited to build systems on a chip, by placing circuits with different voltage and performance requirements in different layers. One such example will be to have logic circuits in the first Si layer and then have memory (SRAM) circuits in the second layer to realize distributed memory systems in a microprocessor. Irrespective of the processing method used to form the upper active layers and the vertical metallization, there are several important issues that need to be quantified to analyze the performance and reliability of these circuits. For example, additional heat generated from the devices in the second active layer will further aggravate thermal problems. Extensive thermal modeling in these structures will be required to analyze their impact on performance and reliability. These models could then be used effectively to develop near-term as well as speculative long-term designs to overcome thermal limitations. For accurate circuit delay predictions an additional capacitive coupling between the top layer metal of the first active layer and the devices on the second active layer will have to be taken into account. The 3-D devices will also open up a whole new world of reliability problems. There will be an increasing need to understand new material interfaces, thin-film material thermal properties, and new failure mechanisms.

APPENDIX A

A.1 Linear Behavior of Interconnect Resistance with Time:

Since it was experimentally observed that the resistance of the metal lines increase linearly with time during all the pulsing events the reason behind such behavior must be quantified. The temperature rise of the interconnect can be expressed as,

$$\Delta T(t)=\frac{E(t)}{C_\theta}=\frac{I^2\int_0^t R(s)ds}{C_\theta}=\frac{I^2}{C_\theta}\int_0^t R_0\left[1+TCR\,\Delta T(s)\right]ds \tag{A.1}$$

Here s is used as another variable for time. Also, since thermal capacity (C_θ) is almost entirely due to the metal for very short pulses it is assumed to be independent of time for simplifying the integration. Differentiating equation (A.1) with time yields,

$$\frac{d\Delta T}{dt}=\frac{I^2 R_0\left[1+TCR\,\Delta T(t)\right]}{C_\theta} \tag{A.2}$$

which can be rearranged to give

$$\int\frac{d\Delta T}{\left[1+TCR\,\Delta T(t)\right]}=\int\frac{I^2R_0}{C_\theta}dt \tag{A.3}$$

Integration yields

$$\frac{1}{TCR}ln\left[1+TCR\,\Delta T(t)\right]=\frac{I^2R_0}{C_\theta}t \tag{A.4}$$

Therefore it follows that

$$\Delta T(t)=\frac{exp\left[\frac{I^2 R_0\,TCR\,t}{C_\theta}\right]-1}{TCR} \tag{A.5}$$

Hence the resistance is given by

$$R = R_0 \, exp\left[\frac{I^2 R_0 TCR\, t}{C_\theta}\right] \tag{A.6}$$

For a 200 ns current pulse of amplitude I = 1.0 A, R_0 = 20 W, and C_θ = 15 nJ/^{0}C (from Figure 3.16) and TCR = 0.00312 $^0C^{-1}$, the exponent equals 0.008, which is a small value.

Therefore equation (A.6) can be expanded to give

$$R = R_0\left[1 + \frac{I^2 R_0 TCR}{C_\theta} t\right] \tag{A.7}$$

Therefore, for very short pulses, the resistance appears to vary linearly with time.

A.2 The Energy Dissipated During a Critical Short-Pulse:

Consider a voltage pulse across a metal line, which is stressed by a critical current pulse of amplitude I_{crit} and duration Δt as shown in Figure A.1. Let R_0 be the initial resistance of the line at time t_0, let R_{Tm} be the resistance of the line just before the melting point at time t_m, and let R_f be the final resistance of the line just before open circuit failure at time t_f. Let the corresponding temperatures be T_0, T_m and T_f. Hence the pulse width, $\Delta t = (t_f - t_0)$. Furthermore, let α_s and α_l be the TCRs during the solid and liquid phases of the line respectively.

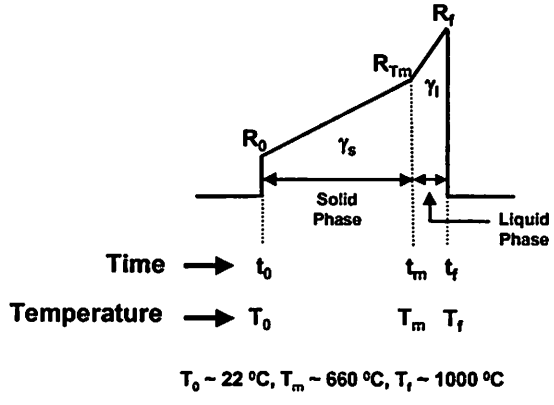


Figure A.1 A voltage pulse across a metal line stressed by a critical current pulse, I_{crit}.

The critical energy dissipated can be expressed similar to equation (3.11) as

$$E_{crit} = \int_{t_0}^{t_f} I_{crit}\, V\, dt = I_{crit}^{\,2} \int_{t_0}^{t_m} \left[R_0 + \frac{R_{Tm} - R_0}{(t_m - t_0)} t \right] dt + I_{crit}^{\,2} \int_{t_m}^{t_f} \left[R_{Tm} + \frac{R_f - R_{Tm}}{(t_f - t_m)} t \right] dt \quad \text{(A.8)}$$

Integration of equation (A.8) yields

$$\begin{aligned} E_{crit} = I_{crit}^{\,2} R_0 (t_m - t_0) + \frac{(t_m - t_0)}{2}(R_{Tm} - R_0) \\ + I_{crit}^{\,2} R_{Tm} (t_f - t_m) + \frac{(t_f - t_m)}{2}(R_f - R_{Tm}) \end{aligned} \quad \text{(A.9)}$$

Let

$$\begin{aligned} (t_m - t_0) &= \Delta t_1 \\ and \quad (t_f - t_m) &= \Delta t_2 \end{aligned} \quad \text{(A.10)}$$

Substitution in equation (A.9) yields

$$E_{crit} = \frac{1}{2} I_{crit}^{\,2} \Delta t_1 (R_0 + R_{Tm}) + \frac{1}{2} I_{crit}^{\,2} \Delta t_2 (R_{Tm} + R_f) \quad \text{(A.11)}$$

This can be rearranged to give

$$E_{crit} = \frac{1}{2} I_{crit}^{\,2} \left[(\Delta t_1 + \Delta t_2) R_{Tm} + \Delta t_1 R_0 + \Delta t_2 R_f \right] \quad \text{(A.12)}$$

Now since the resistance rise in the solid and liquid phase have linear temperature dependence, we can write

$$R_{Tm} = R_0 \left[1 + \alpha_s (T_m - T_0) \right] \quad \text{and}$$

$$R_f = R_{Tm} \left[1 + \alpha_l (T_f - T_m) \right] \quad \text{(A.13)}$$

Equation (A.13) can be written in terms of the resistance rise factors γ_s, and γ_l, in the solid and liquid phase respectively as

$$\gamma_s = \frac{R_{Tm} - R_0}{R_0} = \alpha_s \left(T_m - T_0\right) \quad \text{and}$$

$$\gamma_l = \frac{R_f - R_{Tm}}{R_{Tm}} = \alpha_l \left(T_f - T_m\right) \tag{A.14}$$

Substituting R_{Tm} and R_f from equation (A.13) in equation (A.12), and writing in terms of γ_s, and γ_l, it follows that

$$E_{crit} = \frac{1}{2} I_{crit}^{\ 2} R_0 \left[\Delta t_1 \left(2 + \gamma_s\right) + \Delta t_2 \left(1 + \gamma_s\right)\left(2 + \gamma_l\right)\right] \tag{A.15}$$

Comparing equation (A.15) with equation (3.21) it can be observed that the two expressions are very similar if, $[\Delta t\ (2 + \gamma_{crit})]$ in equation (3.21) is replaced by $[\Delta t_1\ (2 + \gamma_s) + \Delta t_2\ (1 + \gamma_s)\ (2 + \gamma_l)]$. Also, the value of $[\Delta t\ (2 + \gamma_{crit})]$ will always be slightly larger than that of $[\Delta t_1\ (2 + \gamma_s) + \Delta t_2\ (1 + \gamma_s)\ (2 + \gamma_l)]$. This can be numerically verified from the self-heating data of Figure 3.10 for a 1 μs pulse, since $\gamma_{crit} \sim 3.5$, $\gamma_s \sim 2.1$, and $\gamma_l \sim 0.2$, and $\Delta t_1 \sim 950$ ns and $\Delta t_2 \sim 50$ ns. Hence, equation (3.21) will always provide conservative estimates of I_{crit}, which is safe and desirable.

APPENDIX B

B.1 NTRS based interconnect technology file for two Cu processes:

Process (μ)		0.25	0.1
V_{DD} (V)		2.0	1.2
L_{eff} (nm)		160	50
T_{ox} (nm)		5.0	2.5
Metal Levels		6	8
Poly	H (μ)	0.2	0.1
	W (μ)	0.25	0.1
	Space (μ)	0.25	0.1
	Sheet ρ (Ω/square)	4	8
M1-2	H (μ)	0.5	0.26
	W (μ)	0.3	0.13
	Space (μ)	0.3	0.13
	Sheet ρ (Ω/square)	0.044	0.085
	t_{ins} (nm)	650	320
M3-4	H (μ)	0.9	0.55
	W (μ)	0.6	0.3
	Space (μ)	0.6	0.30
	Sheet ρ (Ω/square)	0.024	0.04
	t_{ins} (nm)	900	600
M5-6	H (μ)	2.5	1.2
	W (μ)	2.0	1.0
	Space (μ)	2.0	1.0
	Sheet ρ (Ω/square)	0.009	0.018
	t_{ins} (nm)	1400	800
M7-8	H (μ)	-	2.5
	W (μ)	-	2.0
	Space (μ)	-	2.0
	Sheet ρ (Ω/square)	-	0.009
	t_{ins} (nm)	-	1400

Bibliography

[1] M. T. Bohr, "Interconnect scaling-the real limiter to high performance ULSI," *International Electron Device Meeting, Tech. Dig.*, 1995, pp. 241-244.

[2] W. J. Dally, "Interconnect limited VLSI architecture," *International Interconnect Technology Conference Proceedings*, 1999, pp. 15-17.

[3] C. R. Barrett, "Microprocessor evolution and technology impact," *VLSI Technology Symposium, Dig. Tech. Papers*, 1993, pp. 7-10.

[4] C. Hu, "MOSFET scaling in the next decade and beyond," *Semiconductor International*, pp. 105-114, June 1994.

[5] G. A. Sai-Halasz, "Performance trends in high-end processors," *Proceedings of the IEEE*, vol. 83, no. 1, pp. 20-36, Jan. 1995.

[6] B. Davari, R. H. Dennard, and G. G. Shahidi, "CMOS scaling for high performance and low power-the next ten years," *Proceedings of the IEEE*, vol. 83, no. 4, pp. 595-606, April 1995.

[7] A. Masaki, "Deep-submicron CMOS warms up to high-speed logic," *Proc. IEEE Circuits and Devices*, pp. 18-24, 1992.

[8] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous, and A. R. LeBank, "Design of ion-implanted MOSFETs with very small physical dimensions," *IEEE J. Solid-State Circuits*, Vol. SC-9, pp. 256-268, 1974.

[9] J. Ida, M. Yoshimaru, T. Usami, A. Ohtomo, K. Shimokawa, A. Kita, and M. Ino, "Reduction of wiring capacitance with new low dielectric SiOF interlayer film for high speed/low power sub-half micron CMOS," *VLSI Technology Symposium, Dig. Tech. Papers*, pp. 59-60, 1994.

[10] C. Duvvury and A. Amerasekera, "ESD: A pervasive reliability concern for IC technologies," *Proceedings of the IEEE,* Vol. 81, No. 5, pp. 690-702, May 1993.

[11] T. Green, "A review of EOS/ESD field failures in military equipment," *Proc. 10th, EOS/ESD Symposium*, pp. 7-14, 1988.

[12] M. J. Pelgrom and E. C. Dijkmans, "A 3/5 V compatible I/O buffer," *IEEE J. Solid-State Circuits*, vol. 30, No. 7, pp. 823-825, 1995.

[13] The National Technology Roadmap for Semiconductors, 1997.

[14] H. B. Bakoglu, "*Circuits, Interconnections, and Packaging for VLSI*," Addison-Wesley Pub. Co. 1990.

[15] T. Sakurai and K. Tamaru, " Simple formulas for two- and three-dimensional capacitances," *IEEE Trans. Electron Devices*, vol. ED-30, no. 2 pp. 183-185, 1983.

[16] J. Rabaey, "*Digital Integrated Circuits- A Design Perspective*," Prentice Hall Pub. Co. 1996.

[17] J. R. Black, "Electromigration – A brief survey and some recent results," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 338-347, 1969.

[18] B. K. Liew, N. W. Cheung, and C. Hu, "Projecting interconnect electromigration lifetime for arbitrary current waveforms," *IEEE Trans. Electron Devices*, vol. 37, pp. 1343-50, 1990.

[19] S. Ramaswamy, C. Duvvury and S. Kang, "EOS/ESD Reliability of Deep Sub-Micron NMOS Protection Devices," *Proceedings of the IEEE International Reliability Physics Symposium*, pp. 284-291, 1995.

[20] S. H. Voldman, "ESD robustness and scaling implications of aluminum and copper interconnects in advanced semiconductor technology," *Proc. EOS/ESD Symposium*, pp. 316-329, Sept. 1997.

[21] N. W. Ashcroft and N. D. Mermin, "*Solid State Physics*," Saunders College, HRW Pub Co. 1976.

[22] H. S. Carslaw and J. C. Jaeger, "*Conduction of Heat in Solids*," Second Edition, Oxford Univ. Press, 1959.

[23] S. Ramo, J. R. Whinnery, and T. Van Duzer, "*Fields and Waves in Communication Electronics*," Third Edition, John Wiley & Sons Pub. Co., 1994.

[24] SPICE – Simulation Program for Integrated Circuit Emulation, University of California, Berkeley.

[25] M. P. Heisler, *Trans ASME*, 69, pp. 227-236, Apr. 1947.

[26] A. Amerasekera and C. Duvvury, *ESD in Silicon Integrated Circuits*, Wiley Pub. Co., New York, 1995.

[27] C. Duvvury and A. Amerasekera, "State-of-the-art issues for technology and circuit design of ESD protection in CMOS ICs," *Semiconductor Science and Tech.*, pp. 833-850, 1996.

[28] A. Amerasekera and C. Duvvury, "The impact of technology scaling on ESD robustness and protection circuit design," in *EOS/ESD Symp. Proc.*, 1994, pp. 237-245.

[29] A. Amerasekera and J. Verwey, "ESD in integrated circuits," *Quality and Reliability Engineering International*, vol. 8, pp. 259-272, 1992.

[30] M. Ueda, A. Carl, J. Iadanza, H, Oshikawa, M, Yamamoto, M, Nelson, K, Key, Y, Tokuda, T, Saitoh and R, Kilmoyer, "A 3.3 V ASIC mixed voltage applications with shut down mode," in *Proc. IEEE CICC*, 1993, pp. 25.5.1 - 25.5.4.

[31] S. S. Cohen, J. I. Raffel and P. W. Wyatt, "A novel double-metal structure for voltage-programmable links," *IEEE Electron Device Lett.*, Vol. 13, No. 9, pp. 488-490, 1992.

[32] S. Chiang, R. Forouhi, W. Chen, F. Hawley, J. McCollum, E. Hamdy, and C. Hu, "Antifuse structure comparison for field programmable gate arrays," in *Tech. Dig. IEDM*, 1992, pp. 611-614.

[33] S. S. Cohen, E. F. Gleason, P. W. Wyatt and J. I. Raffel, "A flat-aluminum based voltage-programmable link for field programmable devices," *IEEE Trans. Electron Devices*, Vol. 41, No. 5, pp. 721-725, 1994.

[34] S. Ramaswamy, C. Duvvury and S. Kang, "EOS/ESD reliability of deep sub-micron NMOS protection devices," in *Proc. IRPS*, 1995, pp. 284-291.

[35] E. Kinsborn, C. M. Melliar-Smith, and A. T. English, "Failure of small thin-film conductors due to high current density pulses," *IEEE Trans. Electron Devices*, vol ED-26, No. 1, pp. 22-26, 1979.

[36] D. G. Pierce, "Modeling metallization burnout of integrated circuits," *Proc. EOS/ESD Symp.*, 1982, pp. 56-61.

[37] K. Y. Kim and W. Sachse, "Dynamic fracture test of metal thin films deposited on an insulating substrate by a high current pulse method," *Thin Solid Films*, 205, pp. 176-181, 1991.

[38] T. J. Maloney and N. Khurana, "Transmission line pulsing techniques for circuit modeling of ESD phenomena," *EOS/ESD Symp. Proc.*, 1985, pp. 49-54.

[39] T. J. Maloney, "Integrated circuit metal in the charged device model: bootstrap heating, melt damage and scaling laws," *EOS/ESD Symp. Proc.*, 1992, pp. 129-134.

[40] J. E. Murguia and J. B. Bernstein, "Short-time failure of metal interconnect caused by current pulses," *IEEE Electron Device Lett.*, Vol. 14, No. 10, pp. 481-483, 1993.

[41] X. Gui, S. D. Dew and M. J. Brett, "Thermal simulation of thin-film interconnect failure caused by high current pulses," *IEEE Trans. Electron Devices*, Vol. 42, No. 7, pp. 1386-1388, 1995.

[42] K. Banerjee, A. Amerasekera and C. Hu, "Characterization of VLSI circuit interconnect heating and failure under ESD conditions," *Proc. IEEE Int. Reliability Physics Symp.*, 1996, pp. 237-245.

[43] C. Tien, A. Majumdar, and F. M. Gerner, Editors, "*Microscale Energy Transport*," Taylor & Francis Pub. Co. Chapter 7, pp. 229-293, 1998.

[44] A. Miklos, and A. Lorincz, "Transient thermoreflectance of thin metal films in the picosecond regime," *J. Appl. Phys.*, vol. 63, pp. 2391-2395, 1988.

[45] Y. S. Ju, and K. E. Goodson, "Thermal mapping of interconnects subjected to brief electrical stresses", *IEEE Electron Device Letters,* vol. 18, pp. 512-514, 1997.

[46] G. A. Bennett, and S. D. Briles, "Calibration procedure developed for IR surface-temperature measurements," *IEEE Trans. Components, Hybrids Manufacturing Technol.*, vol. 12, pp. 690-695, 1989.

[47] S. Kondo, and K. Hinode, "High-resolution temperature measurement of void dynamics induced by electromigration in aluminum metallization," *Appl. Phys. Lett.*, vol. 67, pp. 1606-1608, 1995.

[48] P. Kolodner, and J. A. Tyson, "Microscopic fluorescent imaging of surface temperature profiles with 0.01C resolution," *Appl. Phys. Lett.*, vol. 40, pp. 782-784, 1982.

[49] P. Caroll, *Cholesteric Liquid Crystals: Their Technology and Applications*, Ovum, London, 1973.

[50] G. Meier, E. Sackmann, and J. G. Grabmeier, *Applications of Liquid Crystals*, Springer Verlag, Berlin, 1975.

[51] J. Hiatt, "A method of detecting hot spots on semi-conductors using liquid crystals", *Proc IEEE Int. Reliab. Phys. Symp.*, 1981, pp. 130-133.

[52] C. C. Williams, and H. K. Wickramasinghe, "Scanning thermal profiler," *Appl. Phys. Lett.*, vol. 49, pp. 1587-1589, 1986.

[53] M. Nonnenmacher, and H. K. Wickramasinghe, "Scanning probe microscopy of thermal conductivity and subsurface properties," *Appl. Phys. Lett.*, vol. 61, pp. 168-170, 1992.

[54] A. Majumdar, J. P. Carrejo, and J. Lai, "Thermal imaging using the atomic force microscope," *Appl. Phys. Lett.* vol. 62, pp. 2501-2503, 1993.

[55] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, "Surface studies by scanning tunneling microscopy," *Physical Review Letters*, vol. 49, pp. 57-61, 1982.

[56] G. Binnig, C. F. Quate, and C. Gerber, "Atomic force microscope," *Physical Review Letters*, vol. 56, pp. 930-933, 1986.

[57] K. Banerjee, A. Amerasekera, N. Cheung and C. Hu, "High-current failure model for VLSI interconnects under short-pulse stress conditions," *IEEE Electron Device Lett.*, vol. 18, No. 9, pp. 405-407, 1997.

[58] E. A. Brandes and G. B. Brook, editors, *Smithells Metals Reference Book*, Butterworth-Heinemann Ltd., 1992, 7th Ed.

[59] E. C. Jordan, and K. G. Balmain, *Electromagnetic Waves and Radiating Systems*, Prentice-Hall, New Jersey, Second Edition, 1990.

[60] D. G. Pierce, W. Shiley, B. D. Mulcahy, K. E. Wagner, and M. Wunder, "Electrical overstress testing of a 256K UVEPROM to rectangular and double exponential pulses," *Proc. EOS/ESD Symp.*, 1988, pp. 137-146.

[61] K. Banerjee, "Thermal Effects in Interconnects - Part 3," Tutorial Notes, *IEEE Int. Reliability Physics Symp.*, 1998.

[62] H. A. Schafft, "Thermal analysis of electromigration test structures," *IEEE Trans. Electron Devices*, vol. ED-34, pp. 664-672, March 1987.

[63] K. Banerjee, S. Rzepka, A. Amerasekera, N. Cheung and C. Hu, "Thermal analysis of the fusion limits of metal interconnect under short duration current pulses," *IEEE Int. Integrated Reliability Workshop, Final Report*, 1996, pp. 98-102.

[64] P. Salome, C. Leroux, P. Crevel, and J. P. Chante, "Investigations on the Thermal Behavior of Interconnects under ESD Transients using a Simplified Thermal RC Network," *EOS/ESD Symp. Proc.*, 1998.

[65] P. Shewmon, *Diffusion in Solids*, TMS, 1989, ch. 1, pp. 14-19.

[66] H. A. Schafft et al, "Thermal conductivity measurements of thin-film silicon dioxide," *Proc. Intl. Conf. on Microelectronic Test structures*, Vol.2, No. 1, pp. 121-125, March 1989.

[67] J. C. Lambropoulos et al, "Thermal conductivity of dielectric thin films," *J. Appl. Phys.*, 66(9), pp. 4230- 4242, 1989.

[68] A. J. Griffin Jr., F. R. Brotzen and P. J. Loos, "Effect of thickness on the transverse thermal conductivity of thin dielectric films," *J. Appl. Phys.*, 75 (8), pp. 3761-3764, 1994.

[69] Y. S. Touloukian et al, *Thermophysical Properties of Matter*, IFI/Plenum, 1970, Vol. 4.

[70] C. Li and J. Yamanis, "Super tough silicon nitride with R-Curve behavior," *Ceram. Eng. Sci. Proc.*, pp. 632-645, 1989.

[71] R. W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials*, Third Edition, Wiley, New York, 1989.

[72] D. A. Porter and K. E. Easterling, *Phase Transformations in Metals and Alloys*, Van Nostrand Reinhold Company Ltd., 1981.

[73] J. M. Poate, K. N. Tu and J. W. Mayer, *Thin Films - Interdiffusion and Reactions*, Wiley-Interscience Publication, 1978, New York. (Ch. 8 by F. M. d'Heurle and P. S. Ho, *Electromigration in Thin Films*), pp. 243-303.

[74] J. Scarpulla, D. C. Eng, S. Brown, and K. P . McWilliams, "Reliability of metal interconnect after a high-current pulse," *IEEE Electron Device Lett.*, vol. 17, No. 7, pp. 322-324, 1996.

[75] R. S. List, A. Singh, A. Ralston, and G. Dixit, "Integration of low-k dielectric materials into sub-0.25-μm interconnects," *MRS Bulletin*, October 1997, pp. 61-69.

[76] W. W. Lee and P. S. Ho, "Low-dielectric-constant materials for ULSI interlayer dielectric applications," *MRS Bulletin*, October 1997, pp. 19-23.

[77] C. Jin, J. D. Luttmer, D. M. Smith, and T. A. Ramos, "Nanoporous silica as an ultra low-k dielectric," *MRS Bulletin*, October 1997, pp. 39-42.

[78] K. Endo, "Fluorinated amorphous carbon as a low-dielectric-constant interlayer dielectric," *MRS Bulletin*, October 1997, pp. 55-58.

[79] J. G. Fleming, E. Roherty-Osmun, and A. J. Farino, "Lowering of intralevel capacitance using air gap structures," *Proc. Advanced Metallization and Interconnect Systems for ULSI Applications*, 1996, Oct., 1-3, Boston, MA.

[80] E. M. Zielinsky, S. W. Russel, R. S. List, A. M. Wilson, C. Jin, K. J. Newton, J. P. Lu, T. Hurd, W. Y. Hsu, V. Cordasco, M. Gopikanth, V. Korthuis, W. Lee, G. Cerny, N. M. Russel, P. B. Smith, S. O'Brien, and R. H. Havemann, "Damascene integration of Cu and ultra-low-k xerogel for high performance interconnects," *Int. Electron Device Meeting Tech. Dig.*, 1997, pp. 936-938.

[81] P. K. Chatterjee, W. R. Hunter, A. Amerasekera, S. Aur, C. Duvvury, P. E. Nicollian, L. M. Ting, and P. Yang, "Trends for deep submicron VLSI and their implications for reliability," *Proc. Int. Reliability Physics Symp.*, 1995, pp. 1-11.

[82] W. R. Hunter, "The implications of self-consistent current density design guidelines comprehending electromigration and Joule heating for interconnect technology evolution," *Int. Electron Device Meeting, Tech. Dig.*, 1995, pp. 483-486.

[83] S-P. Jeng, R. Havemann, and M. Chang, "Process integration and manufacturability issues for high performance multilevel interconnect," *Advanced Metallization for Devices and Circuits – Science, Technology and Manufacturability Symposium*, edited by S.P. Murarka, A. Katz, K.N. Tu, and K. Maex, (Mater. Res. Soc. Symp. Proc., 337, Pittsburgh), 1994, pp. 25-31.

[84] B. Zhao, S.Q. Wang, M. Fiebig, S. Anderson, P.K. Vasudev, and T.D. Seidel, "Reliability and electrical properties of new low dielectric constant interlevel dielectrics for high performance ULSI interconnect," *Proc. Int. Reliability Physics Symp.*, 1996, pp. 156-163.

[85] K. Banerjee, A. Amerasekera, G. Dixit, and C. Hu, "The Effect of Interconnect Scaling and Low-k Dielectric on the Thermal Characteristics of the IC Metal," *Int., Electron Devices Meeting, Tech. Dig.*, 1996, pp. 65-68.

[86] D.K. Schroder, *Semiconductor Material and Device Characterization*, Wiley, New York, 1990.

[87] K. Fu and R. E. Pyle, "On the failure mechanisms of titanium nitride/titanium silicide barrier contacts under high current stress", *IEEE Trans. Electron Devices.*, vol. 35, no. 12, pp. 2151-2159, 1988.

[88] A. Christou, "Electro-thermomigration in Al/Si, Au/Si Interdigitized Test Structures", *J. Appl. Phys.* Vol. 44, No. 7, 1973.

[89] K. Banerjee, A. Amerasekera, G. Dixit, N. Cheung, and C. Hu, "Characterization of contact and via failure under short duration high pulsed current stress," *Proc. International Reliability Physics Symp.*, 1997, pp. 216-220.

[90] T. J. Maloney, "Contact injection: A major cause of ESD failure in integrated circuits," *EOS/ESD Symp. Proc.*, 1986, pp. 166-172.

[91] J. C. Ondrusek, C. F. Dunn and J. W. McPherson, "Kinetics of Contact Wearout for Silicided ($TiSi_2$) and Non-Silicided Contacts," *Proc. International Reliability Physics Symp.*, pp. 154-160, 1987.

[92] J. R. Black, "Mass Transport of Aluminum by Momentum Exchange with Conducting Electrons", *Proc. Ann. Symp. on Reliability Physics*, IEEE Cat. 7-15C58, pp. 148-159, 1967.

[93] H. A. Le, K. Banerjee and J. W. McPherson, "The dependence of W-plug via EM performance on via size," *Semiconductor Science and Technology,* vol. 11, no. 6, pp. 858-864, 1996.

[94] J. Tao, K. K. Young, C.A. Pico, N. W. Cheung and C. Hu, "Electromigration characteristics of Tungsten plug vias under pulse and bidirectional current stressing," *IEEE Electron Device Lett.*, vol. EDL-12, pp. 646-648, 1991.

[95] S. P. Murarka, *Silicides for VLSI Applications*, Academic Press Inc., Orlando, FL, 1983.

[96] K. Maex, "Silicides for integrated circuits: $TiSi_2$ and $CoSi_2$," *Materials Science and Engineering*, R11, nos. 2-3, pp. 53-153, 1993.

[97] M. E. Alperin, T. C. Holloway, R. A. Haken, C. D. Gosemeyer, R. V. Karnaugh, and W. D. Parmantie, "Development of the self-aligned titanium silicide process for VLSI applications," *IEEE Trans. Electron Devices,* vol. 32, no. 2, pp. 141-149, 1985.

[98] S. Wolf, *Silicon Processing for the VLSI Era*, vol. 2, Lattice Press, Sunset Beach, CA, 1990.

[99] A. Amerasekera, V. McNeil, and M. Rodder, "Correlating drain junction scaling, salicide thickness, and lateral NPN behavior, with the ESD/EOS performance of a 0.25 μm CMOS process," *Tech. Dig. IEDM*, 1996, pp. 893-896.

[100] A. Amerasekera, M. Chang, J. A. Seitchik, A. Chatterjee, K. Mayaram, and J. Chern, "Self-heating effects in basic semiconductor structures," *IEEE Trans. Electron Devices,* vol. 38, no. 9, pp. 1836-1844, 1991.

[101] J. A. Kittl, Q. Z. Hong, M. Rodder, D. A. Prinslow, and G. R. Misium, "A Ti salicide process for 0.10 μm gate length CMOS technology," *VLSI Technology Symp. Proc.*, 1996, pp. 14-15.

[102] C. T. Huang and T. F. Lei, "Impact of Ge implantation on the electrical characteristics of $TiSi_2$ p+/n shallow junctions with an α-Si (or a poly-Si) buffer layer," *IEEE Trans. Electron Devices*, vol. 44, no. 4, pp. 601-606, 1997.

[103] J. A. Kittl, Q. Z. Hong, C. P. Chao, I. C. Chen, N. Yu, S. O'Brien, and M. Hanratty, "Salicides for 0.10 μm gate lengths: A comparative study of one step RTP Ti with Mo doping, Ti with pre-amorphization and Co processes," *VLSI Technology Symp. Proc.*, 1997, pp. 103-104.

[104] L. A. Clevenger, J. M. E. Harper, C. Cabral, Jr., C. Nobili, G. Ottavini, and R. W. Mann, "Kinetic analysis of C49-$TiSi_2$ and C54-$TiSi_2$ formation and rapid thermal annealing rates," *J. Appl. Phys.*, 72, 4978-4980, 1992.

[105] R. W. Mann, L. A. Clevenger, and Q. Z. Hong, "The C49 to C54-TiSi2 transformation in self-aligned silicide applications," *J. Appl. Phys.*, 73, p. 3566, 1993.

[106] S. Singh, H. Solak, N. Krasnoperov, F. Cerrina, A. Cossy, J. Diaz, J. Stohr, and M. Samant, "An X-ray spectromicroscopic study of the local structure of patterned titanium silicide," *Appl. Phys. Lett.*, 71 (1), pp. 55-57, 1997.

[107] C.-P. Chao, K. E. Violette, S. Unnikrishnan, M. Nandakumar, R. L. Wise, J. A. Kittl, Q.-Z. Hong, and I.C. Chen, "Low resistance Ti or Co salicided raised source/drain transistors for sub-0.13 μm CMOS technologies," *Int. Electron Device Meeting, Tech. Dig.*, 1997, pp. 103-106.

[108] B.-S. Chen and M.-C. Chen, "Formation of Cobalt silicided shallow junction using implant into/through silicide technology and low temperature furnace annealing," *IEEE Trans. Electron Devices*, vol. 43, no. 2, pp. 258-266, 1997.

[109] K. Banerjee, A. Amerasekera, J. A. Kittl, and C. Hu, "High current effects in silicide films for sub-0.25 μm VLSI technologies," *Proc. Int. Reliability Physics Symposium,* 1998, pp. 284-292.

[110] P. L. Hower, and V. G. K. Reddi, "Avalanche injection and second breakdown in transistors," *IEEE Trans. Electron Devices,* vol. ED-17, no. 4, pp. 320-335, 1970.

[111] K. Maex, and M. V. Rossum, eds., *Properties of Metal Silicides*, pp. 31-34, INSPEC, London, U. K., 1995.

[112] K. K. Ng, and W. T. Lynch, "The impact of intrinsic series resistance on MOSFET scaling," *IEEE Trans. Electron Devices*, vol. ED-34, no. 3, pp. 503- 511, 1987.

[113] P. I. Suciu and R. L. Johnston, "Experimental derivation of the source and drain resistance of MOS transistors," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 1846-1848, 1980.

[114] P. K. Chatterjee, W. R. Hunter, T. C. Holloway, and Y. T. Lin, "The impact of scaling laws on the choice of n-channel or p-channel for MOS VLSI," *IEEE Electron Device Lett.*, vol. EDL-1, pp. 220-223, 1980.

[115] D. B. Scott, R. A. Chapman, C. Wei, S. S. Mahant-Shetti, R. A. Haken, and T. C. Holloway, "Titanium disilicide contact resistivity and its impact on 1-μm CMOS circuit performance," *IEEE Trans. Electron Devices*, vol., ED-34, No. 3, pp. 562-574, 1987.

[116] M. Jeng, J. E. Chung, P. Ko, and C. Hu, "The effect of source/drain resistance on deep submicrometer device performance," *IEEE Trans. Electron Devices*, vol. 37, no. 11, pp. 2408-2410, 1990.

[117] C. J. Koeneke, S. M. Sze, R. M. Levin, and E. Kinsbron, "Schottky MOSFET for VLSI," *Int. Electron Device Meeting, Tech. Dig.*, 1981, pp. 367-370.

[118] D. B. Scott, W. R. Hunter, and H. Shichijo, "A transmission line model for silicided diffusions: Impact on the performance of VLSI circuits," *IEEE Trans. Electron Devices*, vol. ED-29, no. 4, pp. 651-661, 1982.

[119] G. K. Reeves and H. B. Harrison, "Obtaining the specific contact resistance from transmission line model measurements," *IEEE Electron Device Lett.*, vol. EDL-3, pp. 111-113, 1982.

[120] C. Y. Ting, and M. Wittmer, "Investigation of the Al/TiSi2/Si contact system," *J. Appl. Phys.*, vol. 54, no. 2, pp. 937-943, 1983.

[121] S. J. Proctor and L. W. Linholm, "A direct measurement of the interfacial contact resistance," *IEEE Electron Device Lett.*, vol. 3, pp. 294-296, 1982.

[122] J. Hui, S. Wong, and J. Moll, "Specific contact resistivity of $TiSi_2$ to p+ and n+ junctions," *IEEE Electron Device Lett.*, vol. 6, pp. 479-481, 1985.

[123] S. J. Proctor, L. W. Linholm, and J. A. Mazer, "Direct measurement of interfacial contact resistance, end contact resistance, and interfacial contact layer uniformity," *IEEE Trans. Electron Devices*, vol. ED-30, no. 11, pp. 1535-1542, 1983.

[124] J. Chern and W. G. Oldham, "Determining specific contact resistivity from contact end resistance measurements," *IEEE Electron Device Lett.*, vol. EDL-5, pp. 178-180, 1984.

[125] H. H. Berger, "Contact resistance and contact resistivity," *J. Electrochem. Soc.*, vol. 119, no. 4, pp. 507-514, 1972.

[126] M. Finetti, A. Scorzoni, and G. Soncini, "Lateral current crowding effects on contact resistance measurements in four terminal resistor test structures," *IEEE Electron Device Lett.*, vol. EDL-5, pp. 524-526, 1984.

[127] W. M. Loh, S. E. Swirhun, T. A. Schreyer, R.M. Swanson, and K. C. Saraswat, "Modeling and measurement of contact resistances," *IEEE Trans. Electron Devices*, vol. ED-34, no. 3, pp. 512-524, 1987.

[128] S. S. Cohen, G. Gildenblat, and D. M. Brown, "Size effect on contact resistance and device scaling," *J. Electrochem. Soc.*, vol., 130, pp. 978-980, 1983.

[129] R. L. Maddox, "On the optimization of VLSI contacts," *IEEE Trans. Electron Devices*, vol. ED-32, no. 3, pp. 682-690, 1985.

[130] J. M. Ford, "Al/Si contact resistance for submicrometer design rules," *IEEE Trans. Electron Devices*, vol. ED-32, pp. 840-842, 1985.

[131] T. A. Schreyer and K. C. saraswat, "A two-dimensional analytical model of the cross-bridge Kelvin resistor," *IEEE Electron Device Lett.*, vol. EDL-7, pp. 661-663, 1984.

[132] S. M. Sze, *Physics of Semiconductor Devices*, Wiley, 1981.

[133] C. R. Crowell and S. M. Sze, "Current transport in metal-semiconductor barriers," *Solid State Electronics*, vol. 9., pp. 1035-1048, 1966.

[134] C. Y. Chang and S. M. Sze, "Carrier transport across metal-semiconductor barriers," *Solid State Electronics*, vol. 13., pp. 727-740, 1970.

[135] F. A. Padovani and R. Stratton, "Field and thermionic-field emission in Schottky barriers," *Solid State Electronics*, vol. 9., pp. 695-707, 1966.

[136] C. R. Crowell and V. L. Rideout, "Normalized thermionic-field (T-F) emission in metal-semiconductor (Schottky) barriers," *Solid State Electronics*, vol. 12., pp. 89-105, 1969.

[137] A. Y. C. Yu, "Electron tunneling and contact resistance of metal-silicon contact barriers," *Solid State Electronics*, vol. 13., pp. 239-247, 1970.

[138] K. Varahramyan and E. J. Verret, "A model for specific contact resistance applicable for titanium silicide-silicon contacts," *Solid State Electronics*, vol. 39., no. 11, pp. 1601-1607, 1996.

[139] C. Y. Chang, Y. K. Fang, and S. M. Sze, "Specific contact resistance of metal-semiconductor barriers," *Solid State Electronics*, vol. 14., pp. 541-550, 1971.

[140] A. M. Cowley and S. M. Sze, "Surface states and barrier height of metal semiconductor systems," *J. Applied Physics*, vol. 36, no. 10, pp. 3212-3220, 1965.

[141] C. R. Crowell, S. M. Sze, and W. G. Spitzer, "Equality of the temperature dependence of the gold-silicon surface barrier and the silicon energy gap in Au n-type Si diodes," *Appl. Phys. Lett.*, vol. 4, no. 5, pp. 91-92, 1964.

[142] J. Bardeen, "Surface states and rectification at a metal semi-conductor contact," *Physical Review*, vol. 71, no. 10, pp. 717-727, 1947.

[143] J. W. Slotboom, "The pn product in Silicon," *Solid State Electronics*, 20, pp. 279-283, 1977.

[144] K. Banerjee, A. Amerasekera, G. Dixit, and C. Hu, "Temperature and current effects on small-geometry-contact resistance," *IEEE Int. Electron Devices Meeting, Tech. Dig.*, 1997, pp.115-118.

[145] K. Banerjee, A. Amerasekera, G. Dixit, and C. Hu, "A new quantitative model for deep submicron contact resistance," *Proceedings of the TECHCON*, 1998.

[146] T. Kwok, T. Nguyen, P. Ho, and S. Yip, "Current density and temperature distributions in multilevel interconnections with studs and vias," *Proc. IEEE Int. Reliab. Phys. Symp,* 1987, pp. 130-135.

[147] K. Weide, and W. Hasse, "3-Dimensional simulations of temperature and current density distribution in a via structure," *Proc. IEEE Int. Reliab. Phys. Symp,* 1992, pp. 361-365.

[148] J. T. Trattles, A. G. O'Neill, and B. C. Mecrow, "Three dimensional finite element investigation of current crowding and peak temperatures in VLSI multilevel interconnections," *IEEE Trans. Electron Devices*, vol. 40, pp. 1344-1347, 1993.

[149] K. E. Goodson, and M. Asheghi, "Near-field optical thermometry," *Microscale Thermophysics Engineering*, Vol. 1, 1997, pp. 225-235.

[150] J. Varesi, and A. Majumdar, "Scanning Joule expansion microscopy at nanometer scales," *Appl. Phys. Lett.*, vol. 72, pp. 37-39, 1998.

[151] K. Banerjee, G. Wu, M. Igeta, A. Amerasekera, A. Majumdar, and C. Hu, "Investigation of self-heating phenomenon in small geometry vias using scanning Joule expansion microscopy," *Proc. IEEE Int. Reliab. Phys. Symp,* 1999, pp. 130-133.

[152] A. Majumdar, and J. Varesi, "Nanoscale temperature distributions measured by scanning Joule expansion microscopy," *ASME J. Heat Transfer*, vol. 120, pp. 297-305, 1998.

[153] A. F. Mills, "*Heat and Mass Transfer*," IRWIN Inc., 1995, Chapter 1, pp. 30-34.

[154] K. Luo, Z. Shi, J. Varesi, and A. Majumdar, "Sensor nanofabrication, performance, and conduction mechanisms in scanning thermal microscopy," *Journal of Vacuum Science and Technology B*, Vol. 15, pp 349-360, 1997.

[155] K. Banerjee, L. Ting, N. Cheung, and C. Hu, "Impact of high current stress conditions on VLSI interconnect electromigration reliability evaluation," *Proc. Int. VLSI Multilevel Interconnection Conference*, 1996, pp. 289- 294.

[156] H. A. Schafft, T. C. Staton, J. Mandel, and J. D. Shott, "Reproducibility of electromigaration measurements," *IEEE Trans. Electron Devices*, vol. ED-34, no. 3, pp. 673-680, 1987.

[157] M. M. Levi and J. B. Matilla, "Thermal gradient effects in electromigration," *Proc. Int. VLSI Multilevel Interconnect Conference*, 1993, pp. 258-264.

[158] W. R. Hunter, "Self-consistent solutions for allowed interconnect current density – Part I: Implications for technology evolution," *IEEE Trans. Electron Devices*, vol. ED-44, pp. 304-309, 1997.

[159] S. Rzepka, K. Banerjee, E. Meusel, and C. Hu, "Characterization of self-heating in advanced VLSI interconnect lines based on thermal finite element simulation," *IEEE Trans. on Components, Packaging and Manufacturing Technology-Part A*, vol. 21, No. 3, pp. 1-6, 1998.

[160] NS Nagaraj, F. Cano, H. Haznedar, and D. Young, "A practical approach to static signal electromigration analysis," *Proc. 35th Design Automation Conf.*, 1998, pp. 572-577.

[161] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, "On thermal effects in deep sub-micron VLSI interconnects," Proc. 36th Design Automation Conferenece, 1999, pp. 885-891.

[162] J. R. Black, "Electromigration failure modes in aluminum metallization for semiconductor devices," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 1587-1594, 1969.

[163] A. A. Bilotti, "Static temperature distribution in IC chips with isothermal heat sources," *IEEE Trans. Electron Devices*, vol. ED-21, pp. 217-226, 1974.

[164] W. R. Hunter, "Self-consistent solutions for allowed Interconnect current density – Part II: Application to design guidelines," *IEEE Trans. Electron Devices*, vol. ED-44, pp. 310-316, 1997.

[165] C. Jin, L. Ting, K. Taylor, T. Seta, and J. D. Luttmer, "Thermal conductivity measurement of low dielectric constant films," in *Proc. Second International Dielectrics for VLSI/ULSI Multilevel Interconnection Conference (DUMIC)*, 1996, pp. 21-28.

[166] *Private Communications*, Professor Kenneth E. Goodson, Thermosciences Division, Mechanical Eng. Department, Stanford University.

[167] R. H. J. M. Otten and R. K. Brayton, "Planning for performance," *Proc. 35th Design Automation Conf.*, 1998, pp. 122-127.

[168] J. Culetu, C. Amir, and J. McDonald, "A practical repeater insertion method in high speed VLSI circuits," *Proc. 35th Design Automation Conf.*, 1998, pp. 392-395.

[169] "Physical design modelling and verification project (SPACE Project)", http://cas.et.tudelft.nl/research/space.html

[170] P. Singer, "Filling contacts and vias: A progress report," *Semiconductor International*, pp. 89-94, February 1996.