Quantitative analysis of the mechanical and electrical properties of Cu-Cu bonds for three-dimensional integrated circuits (3D ICs)

Leong, Hoi Liong

2007

Leong, H. L. (2007). Quantitative analysis of the mechanical and electrical properties of Cu-Cu bonds for three-dimensional integrated circuits (3D ICs). Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/13607>

<https://doi.org/10.32657/10356/13607>

Quantitative Analysis of the Mechanical and Electrical Properties of Cu-Cu Bonds for Three-Dimensional Integrated Circuits (3D ICs)

By

Leong Hoi Liong

B. Engineering (Electrical)
National University of Singapore, 1999

SUBMITTED TO THE SMA OFFICE IN PARTIAL
FULLFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN ADVNACED MATERIALS FOR MICRO- AND NANO- SYSTEMS (AMM&NS)
AT THE
SINGAPORE-MIT ALLIANCE

2007

Signature of author: _____

AMM&NS Programme
September 2007

Certified by: _____

Prof Gan Chee Lip
SMA Fellow, NTU
Thesis Advisor

Certified by: _____

Prof Carl V. Thompson
SMA Fellow, MIT
Thesis Advisor

Certified by: _____

Prof Pey Kin Leong
SMA Fellow, NTU
Thesis Advisor

Certified by: _____

Dr Li Hongyu
Senior Research Engineer, IME
Thesis Advisor

Certified by: _____

Prof Choi Wee Kiong
Programme Co-Chair
AMM&NS Programme

Certified by: _____

Prof Carl V. Thompson
Programme Co-Chair
AMM&NS Programme

Quantitative Analysis of the Mechanical and Electrical Properties of Cu-Cu Bonds for Three-Dimensional Integrated Circuits (3D ICs)

By

Leong Hoi Liong

Submitted to the SMA office
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Advanced Materials for Micro- and
Nano-Systems (AMM&NS)

Abstract

The increasing complexity and the scaling down of feature sizes for devices have led to the increasing dominance of interconnect delays in determining integrated circuit performance. One promising solution is to stack devices vertically, commonly known as 3D ICs. Copper is an attractive candidate for 3D applications as it can be both the bonding and interconnect material. This thesis explores the wafers bonding technique, thermo-compression bonding, to create 3D ICs.

This technique involves the application of pressure and temperature to forge a bond. In this work, copper thin films were used to bond two silicon substrates. Characterization of the bond process focused on the effects of bonding temperature (250°C to 400°C), applied load (400 to 10000 N) and surface roughness (total root-mean-square roughness of 1 nm to 14 nm). The resultant bond was quantified using a four-point bend test technique. High bond strength was obtained and the bond quality was found to improve with increases in the bond temperature and applied load, and with decreases in the surface roughness. However, non-ideality in the load-displacement behavior was observed due to variation in the bond strengths and non-uniformity in the bonding. This is attributed to process issues such as dishing and non-uniform distribution of the true contact area.

The true contact area between two contacting surfaces is found to be a small fraction of the nominal contact area. A contact model was developed to estimate this true contact area, taking into account the applied loads, the nominal contact areas, and the surface roughness of the pre-bond surfaces. The dicing yield of the bonded wafers was found to correlate with the true contact area and is directly proportional up to a maximum yield of 100%. In addition to the applied load, the surface roughness is also critical in obtaining good bonds between the bonded surfaces.

A contact resistance test structure was fabricated to measure the contact resistance of the bonded interface. The contact resistances were shown to increase as surface roughness increases and/or as the applied load decreases. A contact resistance model was also developed to predict the contact resistance of a bonded interface, taking into account of the true contact area. It was shown that the model's prediction is close to the minimum value measured in each case but is about an order lower than the average value. This is likely to be because the model does not account for large-area non-uniformities in the fabrication and bonding processes, such as dishing effects. A contact resistance reduction phenomenon under high current density was also observed and is shown to be due to electro-migration rather than Joule heating effects.

This work has demonstrated the viability of wafer-level thermo-compression bonding of Cu for fabricating 3D ICs. The contact model determines the process controls to achieve

high bonding yields while the contact resistance model allows the monitoring to be done by a non-destructive method. Furthermore, high resistance contacts may be repaired through the application of a high current density before dicing, thus achieving the yields required for manufacturing.

Keywords: 3D ICs, thermo-compression bonding, surface roughness, true contact area, bond strength, four-point bend tests, contact resistance, Joule heating, electro-migration

Thesis Advisors:

1. Prof. Gan Chee Lip, SMA Fellow, NTU
2. Prof Carl Thompson, SMA Fellow, AMMNS Programme Co-Chair, MIT
3. Prof Pey Kin Leong, SMA Fellow, NTU
4. Dr Li Hongyu, Senior Research Fellow, IME

Contents

Abstract	3
List of Figures	7
List of Tables	11
 Chapter 1 – Introduction	 12
1.1 Motivation for 3D ICs	12
1.2 Technology Options for 3D integration	15
1.3 Wafer Bonding Methods	18
1.3.1 Direct Bonding	18
1.3.2 Anodic Bonding	19
1.3.3 Solder/Eutectic Bonding	20
1.3.4 Thermo-compression Bonding	21
1.3.5 Copper Thermo-compression Bonding	22
1.4 Bond Strength Measurement Methods	24
1.4.1 Razor Blade Test	24
1.4.2 Tensile Test	26
1.4.3 Chevron Test	27
1.4.4 Four-Point Bend Test	29
1.5 Contact Theory	32
1.6 Thesis Objectives	35
1.7 Thesis Scope	36
 Chapter 2 – Four-Point Bend Test	 37
2.1 Fabrication Process	37
2.2 Testing Procedure	45
2.3 Factors That Affect Bond Quality	49
2.4 Effects of Bonding Temperatures	50
2.5 Effects of Surface Roughness	51
2.6 Effects of Contact Pressure	53
2.7 Bonding Non-uniformity	56
2.8 Failure Modes	59
2.9 Discussion and Summary	62
 Chapter 3 – Surface Roughness Acquisition Model	 66
3.1 Introduction	66
3.2 The Acquisition Algorithm	68
3.2.1 Calculation of η and σ_s	69
3.2.2 Calculation of asperity curvature radius R	73
3.2.3 Other Results	77
3.3 Summary	79

Chapter 4 – True Contact Area Model	81
4.1 Distribution of the True Contact Area	81
4.2 Effects of Different Bonding Parameters on Dicing Yield	84
4.3 Contact of Surfaces	88
4.3.1 Introduction	88
4.3.2 Contact Modeling	89
4.3.3 Application of Contact Model to Experiments	97
4.4 Effects of Bonding and Surface Parameters on Bonding	104
4.5 Au-Au Bonding in 3D ICs	113
4.6 Summary	116
Chapter 5 – Contact Resistances Measurements and Model	118
5.1 Introduction	118
5.2 Fabrication and Testing of the Contact Resistance Structures	119
5.3 Contact Resistances of Cu-Cu Bonds	125
5.3.1 The Contact Resistance Model	125
5.3.2 Comparison of Theoretical and the Measured Specific Contact Resistances	130
5.4 Summary	134
Chapter 6 – High Contact Resistance Reduction Phenomenon	136
6.1 Introduction	136
6.2 Contact Resistance Reduction	137
6.3 Effects of Joule Heating	139
6.4 Effects of Electro-migration	142
6.5 Summary	145
Chapter 7 – Conclusion and Future Works	146
7.1 Thesis Summary	146
7.2 Recommendation for Future Works	150
References	152
Appendix	159
Appendix A: Algorithm code for the surface roughness acquisition model	
Appendix B: Statistical nature of surfaces	
Appendix C: Algorithm code for identifying the contacting asperities	

List of Figures

1-1: Gate and Interconnect delays plotted against various technology nodes.	13
1-2: Schematic of a 3D chip showing integrated heterogeneous technologies.	15
1-3: Schematic illustration of the razor blade test.	25
1-4: Schematic illustration of tensile test sample attachment.	26
1-5: Schematic of chevron test set-up.	28
1-6: Bimaterial notched four-point bending specimen.	30
1-7: Schematic of an ideal load-displacement plot.	30
2-1: Schematic cross-sections for wafers bonding.	38
2-2: Fabrication process flow for blanket Cu wafers.	39
2-3: Fabrication process flow for patterned Cu wafers.	39
2-4: Setup of the EVG 640 aligner.	42
2-5: Setup of the EVG 520 bonder.	43
2-6: A schematic of the bonding cycle.	43
2-7: Setup of the DISCO DAD651 automated dice saw.	44
2-8: Schematic of crack propagation path.	45
2-9: Schematic of the four-point bend test equipment.	46
2-10: Illustration of non-ideal behavior.	48
2-11: Four-point bend test results of G_C versus bonding temperature.	50
2-12: Four-point bend test results of G_C versus total surface roughness.	52
2-13: Four-point bend test results of G_C versus nominal bonding pressure. (Different applied loads on same nominal contact area)	53
2-14: Four-point bend test results of G_C versus nominal contact pressure. (Different applied load on different nominal contact area)	55
2-15: Illustration of dishing effects.	57
2-16: A yield map of a bonded pair of wafers after the dice-saw process.	57

2-17: SEM micrographs showing dishing effects.	58
2-18: SEM micrographs of contact non-uniformity.	59
2-19: Optical images of fracture interfaces.	60
2-20: EDX analysis on different Cu surfaces.	61
2-21: SEM micrograph of fractured sample showing Cu-Cu and Ta-SiO ₂ interfaces.	62
3-1: Deviations of the measured to the actual surface profile	67
3-2: A 3D surface plot generated from a MATLAB.	69
3-3: A summit exists at z_{ij} if $z_{ij} > z_{i \pm n, j \pm n}$ for n equals 1 and 2 (a total of 16 points).	70
3-4: Variation of average asperity density versus the measured rms surface roughness.	71
3-5: Typical asperity height profiles, following a Gaussian distribution.	72
3-6: Standard deviation of asperity heights versus rms surface roughness.	73
3-7: Asperity profiles along the x, y and two diagonal axes.	74
3-8: An asperity spherical profile, with ‘tapered-off’ characteristics.	75
3-9 Summit radius curvature plotted against their corresponding heights.	75
3-10: Average curvature radius versus rms surface roughness.	76
4-1: Micro-contact spots under increasing loads.	83
4-2: Assumed contact profiles of poor bonded wafers and well bonded wafers	83
4-3: Plot of dicing yield versus applied loads and nominal contact pressures.	85
4-4: Plot of dicing yield versus total surface roughness.	86
4-5: Plots of dicing yield versus (a) applied load and (b) nominal contact pressures.	87
4-6: A rough surface having asperities in contact with a flat plane.	89
4-7: (a) A rigid sphere against a flat plane and (b) a deformed spherical asperity against a flat rigid plane, both with contact radius a and radius curvature R .	92
4-8: Stress-Temperature curves for unpassivated copper films, 0.6 and 1.0 μm thick.	99
4-9: Graph of Dicing yield vs True contact area (for blanket Cu wafers).	100
4-10: Graph of rms surface roughness after each compression versus the applied load.	101

4-11: SEM image of a blanket Cu film that was bonded to a patterned Cu wafer.	102
4-12: Graph of Dicing yield vs True contact area (for all bonding experiments).	104
4-13: Relationship between the true contact area and load for different nominal contact areas.	105
4-14: An applied load on different nominal areas.	107
4-15: Effects of surface roughness and applied load on the actual and normalized values of the true contact area	110
4-16: Effects of surface roughness and applied load on the different nominal areas.	111
4-17: Hypothetical comparisons of (a) Au-Au and (b) Cu-Cu bonds.	115
4-18: Comparison of true contact area versus load for Au-Au and Cu-Cu bonds.	115
5-1: Schematic diagram of a perfect Cu-Cu bonded structure.	118
5-2: Schematic diagram and the measurement of the contact resistance test structure.	119
5-3: Fabrication process flow for the bottom wafer.	120
5-4: Fabrication process flow for the top wafer.	121
5-5: Backgrind and thinning of one side of the bonded wafers.	122
5-6: Planar view of the contact resistance test structure after Si substrate removal	122
5-7: Side view of the contact resistance test structure.	123
5-8: Schematic diagram of (a) our contact resistance test structure and (b) a Kelvin contact resistance test structure.	125
5-9: (a) A rough surface having asperities in contact with a flat plane and (b) Contact of a perfect flat Cu surface with a patterned Cu surface.	127
5-10: Schematics of determining d' .	128
5-11: Plot of L versus the rms surface roughness.	128
5-12: Effects of surface roughness and applied load on the specific contact resistance.	129
5-13: Experimental and model prediction of specific contact resistances of Cu bonded interfaces, for different total surface roughness (2, 4.5 and 7.5 nm).	131
5-14: Experimental and model prediction of specific contact resistances of Cu bonded interfaces, for different applied loads (6, 8 and 10 kN).	133
6-1: Contact resistance test structure and resistance reduction phenomena.	136

6-2: Samples with higher contact resistances exhibit the resistance reduction phenomenon while samples with lower contact resistances did not.	138
6-3: Plot of $\Delta R/R_0$ versus ΔT , with the slope $\alpha = 0.411\%$.	140
6-4: Repeated current stress showed that Joule heating is insufficient to result in permanent structure changes and hence changes in contact resistances.	141
6-5: Current stress and contact resistance measurement of samples with initial contact resistance (a) 3.6 m Ω , (b) 2.8 m Ω and (c) 1.2 m Ω , at different temperatures.	144
B-1: A theoretical height-height correlation function of a Gaussian surface.	166
B-2: Typical surface height profiles, following a Gaussian distribution.	166
B-3: The coordinates for the image mesh used in the model.	167
B-4: A plot of the height-height correlation function of a Cu surface.	168

List of Tables

2-1: Experiments of different applied loads on different nominal areas.	54
3-1: Comparisons of R , σ_s and η for different asperity definitions.	78
4-1: Details of wafer bonding experiments and the corresponding dicing yields.	85
4-2: Calculations of topological parameters and true contact area.	98
4-3: Nominal and true contact pressures for each bonding experiments.	102
6-1: The contact resistance measurements (at 0.2 A) at different temperatures.	140

Chapter 1 – Introduction

1.1 Motivation for 3D ICs

The high growth of the computer and information technology has led to a high demand for very large scale circuits (VLSI) with increasing functionality, increasing speed, lower power dissipation but minimum cost. To meet this demand, VLSI circuits are scaled down aggressively, which in turn, introduced serious problems such as increasing interconnect resistance and delays. Furthermore, an increasing drive to integrate disparate signals (digital, analog, RF) and technologies (SOI, SiGe HBTs and GaAs, etc.) has led to the development of system-on-chip (SoC) concepts, something which conventional wafer fabrication platforms are yet to be able to support.

The current IC design architecture is basically a 2-dimensional (2D) approach, in which interconnect wires are routed to link the millions of transistors on a single planar substrate. Despite the reduction in feature sizes made possible by advances in fabrication technology, such as lithography and etching, the average chip size is increasing continuously [1] due to the ever increasing complexity of chip design and the ever increasing number of transistors per chip [2]. Smaller features have dramatically improved device performances [3-5] but the performance of interconnect wires has been less positive [6-10]. Smaller cross-sections and pitch but longer interconnect wires have increased the interconnect resistances and capacitances, resulting in a significant increase in signal propagation (RC) delays. As scaling continues, RC delays are becoming more dominant in determining the performance of

advanced ICs [1], [6-10]. Figure 1-1 illustrates this problem, where the gate delay and interconnect delay are plotted against various technology nodes.

Figure 1-1: Gate and Interconnect delays plotted against various technology nodes [7].

For the copper (Cu) damascene processes, a barrier material such as Ti or Ta or their nitrides are deposited between the dielectrics and Cu. The barrier is necessary as Cu has poor adhesion to most dielectrics and can drift very quickly through them under an electric bias to cause metal-metal shorts and reach the underlying Si substrate. However, the barrier thickness has certain scaling limitations, as too thin a barrier will not be effective in preventing Cu diffusion into the dielectric. Therefore, a greater fraction of the interconnect area will be consumed by the metal barrier and the interconnect resistance may rise significantly due to higher effective resistivity of Cu and increased surface scattering.

To alleviate the adverse effects of increasing RC delays, Cu with low- k dielectrics was introduced. However, as shown in Figure 1-1, beyond the 130 nm

node, substantial interconnect delays will result in spite of the introduction low- k dielectrics. Further reduction in interconnect delay will be hard to achieve by either scaling or introducing new materials.

There are several challenges to the implementation of SoC concepts. First, large scale integration of functionalities and disparate technologies on a single chip would increase the chip area, which necessitates the use of numerous long global wires. These wires can lead to unacceptable signal transmission delays and increase the power consumption by increasing the total capacitance that needs to be driven by the gates. Second, the integration of disparate technologies would involve very different materials, which conventional fabrication processes might not be able to support. The integration of these different technologies on a single die will require novel design methodologies and tools. Finally, process integration and the noise generated by the different blocks will be a challenging problem.

Three-dimensional (3D) integration is an architecture that could potentially resolve all the issues mentioned above. Since chip size directly affects the interconnect delays, creating multiple active layers will reduce the total chip area, thus shortening critical interconnect lengths and reducing their power consumptions and delays. The 3D architecture can also be used to build SoCs by placing different devices onto different layers, thereby reducing any potential electromagnetic interference between such devices. Furthermore, from an integration perspective, it is less complex and more cost-effective to fabricate such different technologies on separate substrates followed by physical bonding. A schematic diagram of a 3D chip

is shown in Figure 1-2 with logic, memory (DRAM), analog, RF, and optical I/O circuits on different active layers.

Figure 1-2: Schematic of a 3D chip showing integrated heterogeneous technologies [10].

One of the several ways to realize the 3D architecture is wafer-level bonding. A literature survey of several wafer bonding techniques, in particular Cu-Cu thermo-compression bonding, as well as an overview of several bond strength measuring techniques will be presented in the subsequent sections. The work performed in this thesis focused on Cu-Cu thermo-compression bonding and the development of a model to evaluate the bonding quality.

1.2 Technology Options for 3D Integration

Presently, there are several possible fabrication technologies that can be used to realize 3D integration. A brief description of these alternatives is given as follows:

1) Beam Recrystallization: A popular method of fabricating a second active (Si) layer on top of an existing substrate (oxidized Si wafer) is to deposit polysilicon and fabricate thin film transistors (TFTs). To enhance the performance of such

transistors, an intense laser or electron beam is used to induce recrystallization of the polysilicon film to reduce or even eliminate most of the grain boundaries. This technique may not be practical for 3D devices because of the high temperature involved during the melting of the polysilicon and also due to the difficulty in controlling the grain size variations [11-12]. Beam recrystallization polysilicon films can also suffer from lower carrier mobility (compared to single crystal Si) and unintentional doping. However, high-performance TFTs fabricated using low-temperature processing [13], and even low temperature single crystal Si TFTs, have been demonstrated [14] and can be employed to fabricate 3D devices.

2) *Silicon Epitaxial Growth*: Another technique for forming additional Si layers is to etch a hole in a passivated wafer and epitaxially grow a single crystal Si seeded from the open window in the dielectric. The silicon grows vertically and then laterally to cover the dielectric [15]. In principle, the quality of devices fabricated on these epitaxial layers can be as good as those fabricated underneath on the seed wafer surface. However, the high temperature ($\sim 1000^{\circ}\text{C}$) involved in the process cause significant degradation in the quality of the devices on the lower layers. Also, this technique cannot be used over metallization layers.

3) *Solid Phase Crystallization (SPC)*: As an alternative to high temperature epitaxial growth as discussed above, low temperature deposition and crystallization of amorphous silicon (a-Si), on top of the lower active layer devices, can be employed. The amorphous film can be randomly crystallized to form a polysilicon film [16-17]. Local crystallization can be induced using low temperature processes ($< 600^{\circ}\text{C}$) such as patterned seeding of Germanium [18]. In this method, Ge seeds implanted in

narrow patterns made on a-Si can be used to induce lateral crystallization and inhibit additional nucleation. CMOS transistors can then be fabricated within these islands that are almost single-crystal. Another approach is to use Ni seeding to induce simultaneous lateral crystallization and dopant activation after the fabrication of the entire transistor on the a-Si layer. This is known as the metal induced lateral recrystallization (MILC) [19-20], which offers even lower thermal budget ($<500^{\circ}\text{C}$) and can be employed to fabricate high performance devices on upper active layers even with metallization layers below.

4) *Package Level 3D Integration*: This is a technique in which bare dice or packaged chips are stacked vertically and interconnections are formed through solder balls to provide vertical interconnections between stacked multi-chip modules. Using this method, vertical integration of ASICs and memory, signal processing circuits, and microprocessors have been implemented [21]. The drawback is the lower density of vertical interconnects compared to monolithic 3D integration.

5) *Processed Wafer Bonding*: An attractive alternative is to bond two fully processed wafers on which devices are fabricated on the surface, including interconnects such that the wafers completely overlap. This technique is suitable for further processing or the bonding of more pairs in this vertical fashion. Other advantages of this technology include the independence of the processing temperatures since the wafers can be fabricated separately and later bonded. A more in-depth discussion about wafer bonding will be presented in the next section.

1.3 Wafer Bonding Methods

Wafer bonding can generally be categorized into four different techniques with distinct bonding mechanisms: direct, solder/eutectic, anodic and thermo-compression. Direct bonding, also known as fusion bonding, is the bonding of two contacting surfaces without the need of an intermediate bonding layer, in contrast to solder/eutectic bonding, which relies on the low melting temperature of the bonding layer. Anodic bonding, also known as field-assisted bonding, involves the application of an electric field between an insulating substrate such as glass and a conducting substrate. Lastly, thermo-compression bonding involves the application of an external force to the bonding process at elevated temperatures. This section will provide an overview of each type of bonding.

1.3.1 Direct Bonding

The most prominent examples of direct bonding are Si-Si and Si-SiO₂ bonds, which were first demonstrated in 1986 by Shimbo *et al.* [22] and Lasky [23], respectively. Smooth, flat wafers were pressed together and bonded through the formation of hydrogen and Van der Waals bonds and subsequently strengthened by annealing at elevated temperatures. However, this method has stringent requirements on defect control, smoothness and flatness of the bonding surfaces. Furthermore, the annealing temperatures can be as high as 700-1000°C [24], rendering it unacceptable if there are metallization layers beneath the bonding surfaces.

Several groups have studied approaches to lowering the temperature requirement. Gösele *et al.* obtained a bond energy similar to that of a high

temperature bond by contacting the wafers under a few mbar of vacuum and over 60 hours at 150°C [25]. Surface activation with an oxygen or argon plasma (SAB: Surface Activation Bonding) is another possible alternative: Belford and Sood [26] demonstrated that bonded Si wafers that were pre-treated with an oxygen plasma and RCA wet clean have higher bond strengths than wafers which only gone through the RCA clean. However, the plasma is expected to roughen the surface and Gui *et al.* [27] demonstrated the effects of surface roughness on the bondability of Si bonding at room temperature, where the bonding of rougher surfaces was shown to have poorer bond quality.

Direct bonding of Si-Si and Si-SiO₂ is generally followed by an annealing treatment to enhance the bond quality, though as mentioned earlier, the annealing temperature could be unacceptably high. However, a lower annealing temperature is also possible. An hour long post-bonding anneal between 200°C and 400°C resulted in bond energy around 2 J/m² [28], whereas without annealing, the bond energy had a maximum value of 1 J/m² [29]. Finally, the bond quality (of direct bonding of Si/SiO₂ wafers) also depends on other factors such as wafer bow, warpage and patterning of the wafers [30].

1.3.2 Anodic Bonding

Anodic bonding, introduced by Wallis and Pomerantz [31], uses electrostatic attraction to pull the bonding surfaces together. It is a method commonly used to bond silicon and metals with glass. A typical glass-silicon bonding consists of applying an electric field of 100-1000 V across the structure, at temperatures between

300-400°C [32, 33]. The silicon substrate is held at a positive potential with respect to the glass, driving away the mobile positive sodium ions from the interface and creates an electric field, pulling the surfaces together by electrostatic forces. Whether a seal (formation of SiO_x between Si and O atoms from opposite surfaces) will form subsequent to contacting will depend on the bonding temperature and length of time the surfaces remain in contact. Thus, anodic bonding is less sensitive to surface roughness or particulates than direct bonding. However, the coefficient of thermal expansion of the materials must be closely matched to minimize residual stresses. The glasses can also outgas at room temperatures. Hence for applications that require hermetic sealing, the glass must be outgassed at elevated temperature prior to bonding, to ensure reliability [34].

1.3.3 Solder/Eutectic Bonding

Solder/Eutectic bonding processes are fairly low temperature processes depending on the use of bonding materials that form eutectic alloys at specific temperatures. Like anodic bonding, they are also less sensitive to surface roughness and particulates than direct bonding [35]. In addition, eutectic alloys can be chosen that are conductive, which provides an electrical connection between the devices and package. Examples of these materials are Au-Sn, Au-Ge and Au-Si, which have eutectic temperatures of 280°C, 361°C and 363°C, respectively [36]. As solder/eutectic bonding involves the formation of eutectic alloys, there must be good wetting between the bonding materials and substrate at the eutectic temperatures. However, oxide formation on any component will hinder proper wetting, which leads to poor bonding [36-37]. As silicon, germanium and tin oxides are readily formed

under ambient conditions, some rubbing or brazing of the surfaces is required to remove the oxides for reliable bond formation. This should be well controlled or else voids or non-uniformity in the bonding layer will be introduced [36].

1.3.4 Thermo-compression Bonding

When surfaces come into contact under the application of an external load, not only will the deformation of the bonding material be enhanced but any existing wafer bow will be reduced, thus further facilitating contact between the surfaces. However, for most bonding materials, the yield stress can be quite high at room temperature and hence any plastic deformation would require a very high load for bonding to occur. When the bonding temperature is increased, the yield stress will decrease and the load requirement for bonding can be relaxed. While surface cleanliness and surface roughness are still important criteria to achieve good bonding, the simultaneous application of load and temperature – thermo-compression bonding – allows bonding to occur under conditions that otherwise could not be achieved using one parameter alone.

Li and Bower [38] combined plasma activation with thermo-compression to achieve low-temperature bonding of silicon. Bonding of glass at lower temperatures using thermo-compression was also demonstrated by Sayah *et al.* [39]. Thermo-compression bonding also allows solder/eutectic bonding to occur below the eutectic temperature. For instance, Au-Sn solder was bonded at 240°C in 20 minutes with 0.4 MPa of pressure [36] and Al-Au bonds were made at 350°C in 30 minutes with 45 MPa [38]. Both were die-level bonds.

More recently, Tsau *et al.* [40] reported the characterizations of wafer-level Au thermo-compression bonding, mainly on effects of bonding temperatures (260°C – 300°C) and pressure (1.25 MPa – 120 MPa). The bond toughness was generally found to increase with bond temperature and pressure but with a large variation that was attributed to effects of plasticity and bonding non-uniformity.

1.3.5 Copper Thermo-compression Bonding

In modern day IC fabrication, Cu has replaced Al as the interconnect material because of its superior electrical properties. Hence the ability to use Cu as both the bonding material as well as interconnect material, presents a strong incentive for development of 3D ICs through Cu wafer bonding. However, Cu is known to oxidize under ambient conditions and its oxide is not self-passivating. Tadeipalli [42] has demonstrated that the presence of Cu oxide resulted in significantly lower Cu-Cu bond strengths compared to bonding of clean Cu surfaces (without oxide). Hence, it is imperative to remove this layer of oxide prior to bonding.

Other factors that could affect the bond quality include thickness of the Cu layer, the geometry and size of the patterns, and the different types of surface pre-bond cleaning. Tadeipalli [42] demonstrated that wafer bows up to 100 μm (tensile and compression) on 100 mm diameter substrates did not cause any reduction of the toughness of thermo-compression Cu-Cu bonds, compared to wafers bonded with nominally zero bow. This is because the applied pressures are sufficiently high to elastically deform the bowed wafers. For bonding of Cu layers with different thicknesses, the Cu-Cu bond toughness has been shown to increase with Cu thickness,

due to increased plastic dissipation in thicker films [40-41]. The Cu-Cu bond toughness is also known to vary with pattern size and geometry of the bonding layer, due to differences in plastic energy dissipation and orientation of the features with respect to the loading direction [42].

On the effects of surface cleaning, Tadepalli [42] found that samples that had gone through an acetic acid pre-clean followed by a reducing gas purge in the bond chamber yielded the highest bond toughness. It was also found that bonds that were created below 250°C exhibited poor adhesion and the wafers routinely separated during post bond die-saw cutting. One of the reasons is that the vacuum during the bonding process is not sufficiently high to prevent surface reoxidation of Cu. The commercial EVG bonder was designed to pump the chamber down to a minimum pressure of about 5×10^{-4} Torr, whereas a higher vacuum of about $10^{-6} - 10^{-8}$ Torr is required to prevent Cu surface oxidation. Hence, the unavailability of high vacuum in commercial wafer bonders is a serious limitation.

Shigetou *et al.* [43] have shown that under a high vacuum of about 10^{-7} Torr and with surface activation using Ar beam bombardment (Surface Activation Bonding: SAB), Cu-Cu bonding at room temperature is possible. However, there is no mention of the effects of the bombardment on the surface roughness and how it could have affected the bonding.

Chen *et al.* [44,45] have shown that the Cu-Cu bond quality could be enhanced with an anneal step of about 30 minutes, after the application of the load, allowing further grain growth, atomic diffusion and recrystallization to occur. Chen *et al.* also observed various morphologies of non-distinct, zigzag, and distinct interfaces

in the Cu bonded layer. It was proposed that these morphologies indicate a strong relationship between the roughness of surfaces and the initial steps of bonding [46].

A main portion of this thesis focuses on the development of a quantitative model, accounting for the effects of surface roughness on the thermo-compression Cu-Cu bond quality.

1.4 Bond Strength Measurement Methods

In order to relate the bonding process parameters to the bond quality, a suitable measurement technique must be developed to quantify the bond strength of the bonded interfaces. The technique must be robust and sensitive enough to capture any effects due to changes in the bonding process. There are several measuring techniques used to measure bond strength and a brief overview of each will be discussed. A comprehensive review of several bond-strength measuring methods is given by Vallin *et al.* [47].

1.4.1 Razor Blade Test

The razor blade test was first described by Maszara *et al.* [48] in which the insertion of a razor blade into the bonded interface, as shown in Figure 1-3, results in a crack length opening that is estimated from IR imaging. The crack length is related to the surface energy γ by:

$$\gamma = \frac{3}{8} \frac{Eh^3 y^2}{L^4} \quad , \quad (1.1)$$

where E is the Young's modulus, h is the thickness of the wafer, y is half the blade thickness and L is the measured crack length. Another measurement parameter is the critical strain energy release rate G_C , which is closely related to γ and simply differ by a factor of two [40], where

$$G_C = 2\gamma \quad (1.2)$$

Figure 1-3: Schematic illustration of the razor blade test [48].

This method is advantageous for its simplicity in equipment and does not require any pre-test preparations of the bonded samples. However, there are several shortcomings to this method. First, the blade cannot be forced in between interfaces with high bond strengths and it is difficult to keep the blade perfectly perpendicular to the interface so that the crack front is straight and parallel to the interface. Both cases would likely result in wafer fracture, at which point the test would be completely qualitative. Second, comparison of results between different research groups can be difficult because measurements such as crack length and how the blade is inserted are

often not reported [49]. Furthermore from equations (1.1) and (1.2), it can be seen that there is a strong dependence of γ and G_C on the crack length, so that slight differences in measurements of the crack length can lead to large errors in values of γ and G_C . Moreover, the observed crack length is also dependent on time and humidity, which can make repeatability difficult [50].

1.4.2 Tensile Test

In the tensile test, also known as the pull test, the bonded samples are glued onto attachments, often called pull studs, on both sides of the sample as shown in Figure 1-4. The sample is then fitted onto a tensile strain test machine and the studs are pulled apart. The maximum force, occurring at the point of separation, is recorded. This maximum force divided by the bonded area is taken as the tensile strength.

Figure 1-4: Schematic illustration of tensile test sample attachment.

Though the tensile test may seem simple to use, there are several factors that if not taken into consideration, will affect the accuracy and the viability of the results

[47]. First, the difficulty in specimen attachment is a major alignment problem. Not only do the studs have to be aligned to each other but the application of the glue has to be uniform or else unwanted bending forces may arise and distort the test results. Second, it is important to minimize defects in the samples since the cracks are initiated at the maximum stress point. Uncontrolled defects may occur at sample edges after operations such as dicing and defects can lead to lower than expected failure stress, with significant degrees of scatter in the data set. Finally, the tensile strength also depends on the geometry of the bonded samples, which makes comparing data between different tests difficult.

1.4.3 Chevron Test

The chevron test method can be seen as a natural development of the tensile test, where a load is applied to a specimen with a chevron-shaped notch, i.e. two edges meet at an angle. The test structure is subjected to a gradually increased load applied normal to the bond, as shown in Figure 1-5. A crack is initially formed at the tip of the chevron structure, exhibiting stable growth since the crack width increases as the crack proceeds. When the crack reaches a critical length, its growth becomes unstable which results in an instantaneous fracture of the sample. The critical load is a measure of the bond strength since the crack length depends only on the specimen geometry and loading condition. Hence, measurement of the fracture load is sufficient for determining the bonded structure's fracture toughness. The challenge is to model the geometry properly.

Figure 1-5: Schematic of chevron test set-up, taken from Bagdahn *et al.* [51].

From the chevron tests, the strain energy release rate G_c can be calculated from the maximum load in the test, F_{max} using the expressions

$$G_c = \frac{1-\nu^2}{E} K_{IC}^2 \quad (1.3)$$

and

$$K_{IC} = \frac{F_{max}}{B\sqrt{W}} Y_{min}, \quad (1.4)$$

where K_{IC} is the critical stress intensity factor, Y_{min} is the minimum value of the geometry function that corresponds to the crack length at which failure occurs. B , W , E and ν are the width, length, Young's modulus and Poisson's ratio of the specimen, respectively.

From equations (1.3) and (1.4), the calculation of the bond strength requires knowledge of the geometry function Y , which is dependent on the thickness of the bonded layers and crack geometry. The geometry function Y can either be experimentally determined [52] or determined using finite-element methods.

There are three exceptional capabilities that can be associated with chevron test method. First, fracture initiation is guided to the bond by the tip of the chevron because of the high stress concentration, hence making the method suitable for measuring high strength bonds. Second, an atomically sharp crack is created at the bond, thus avoiding ambiguity in crack opening angle estimations. Third, the fracture toughness is directly assessed through maximum force measurements. It is not necessary to measure the crack length, thereby reducing the complexity of the measurement equipment. However, it would be laborious to experimentally arrive at and verify the precise specimen geometry as well as determining an expression to take this geometry into account.

1.4.4 Four-Point Bend Test

In 1989, Charalambides *et al.* [53] proposed the four-point bend test for the measurement of interfacial fractural resistance or bond strength. The specimen is a biomaterial beam with a through thickness, central notch in one of the layers, and a pre-crack along the interface, as shown in Figure 1-6. The four loading points provide a constant moment condition in the region between the inner points. At steady state conditions, the crack should propagate at a constant velocity under a constant load, and the strain energy release rate should be constant, as illustrated in Figure 1-7. The crack propagation then decelerates near the inner and outer loading points. Further crack growth would require an increase in load [53].

Figure 1-6: Bimaterial notched four-point bending specimen with symmetrical interfacial cracks, taken from Charalambides *et al.* [53].

Figure 1-7: Schematic of an ideal load-displacement plot.

Based on beam theory, the strain energy release rate of the specimen in steady state can be given by an analytical expression, simplified from the general form in [53] for the case of identical beams:

$$G_c = \frac{3(1-\nu^2)}{2E} \left(\frac{Pl}{b} \right)^2 \left(\frac{1}{h_1^3} - \frac{1}{(h_1 + h_2)^3} \right) \left(\frac{1}{F} \right), \quad (1.5)$$

where P is the load at which crack propagation occurs, l is the distance between the inner and outer loading points, b is the width of specimen, and E , ν , h_1 and h_2 are the Young's modulus, the Poisson's ratio, the thickness of the notched wafer and unnotched wafer, respectively. F is the fraction of the specimen width that is bonded.

An advantage of the four-point bend test is that during debonding, the stress relaxation of the film is constrained and does not contribute to the debonding driving force. There is a small component of G_C due to elastic curvature of the wafer containing the film after debonding. However, it has been shown that this is quite negligible [54]. Another factor to be taken into account is the effect of bonding layer thickness on the dissipation of plastic energy. It has been shown that the effects of plastic deformation are higher in thicker copper stacks which surround the bonded interface [55]. These effects greatly enhance the measured interface toughness beyond the true strength of the bonded interface.

For bond strength measurements, cracks that are initiated and propagate due to a purely tensile mechanism, are commonly referred to as mode I failures. In contrast, when the cracks propagate along an interface due to purely shearing mechanisms, it is referred to as mode II failure. Four-point bend tests involve forces of bending moments and hence the delaminations along the interface are essentially mixed mode fractures. A parameter ψ is used to define this mode mixture, given as

$$\psi = \tan^{-1} \left(\frac{K_{II}}{K_I} \right), \quad (1.6)$$

where K_I and K_{II} are the stress intensity factors in mode I and II, respectively. As ψ varies from 0° to 90° , the failure mode changes from a pure mode I to a pure mode II with respect to the direction of crack propagation. The critical strain energy release rate G_C , is known to vary as a function of the parameter ψ and is typically lower in pure mode I [56]. The mode mixture for four-point bend tests also depends on the ratio of the bond layer thickness and the modulus mismatch between the two layers.

In this work, the four-point bend test is used to measure the interface adhesion in the bonded Cu structures. The Cu thickness is purposely kept low to minimize the effects of plastic energy dissipation. Si wafers with Cu/Ta layers are bonded together using thermo-compression bonding. A pre-machined notch serves as the stress concentration point for the origin of the crack. The crack propagates into the weak interface in the bonded stack and the steady-state value of the strain energy release rate gives the bond strength of the interface. The results are subsequently analyzed to seek a better understanding of the effects of bonding parameters on the bond quality.

1.5 Contact Theory

It is generally accepted that when two nominally smooth surfaces are pressed together, the load will be borne by the contacting surface asperities from both surfaces. The contact area actually covers only certain discrete areas and not the whole nominal contact area, and hence is commonly referred to as the *true contact area* [57]. The ratio of the true-to-nominal contact area can be a very small value and is the basis of explanation propounded for a variety of phenomena, from friction and wear to adhesion and electrical contact resistance [58]. Hence, the study of deformation behavior of the contacting surface asperities and the accurate modeling of the contact between rough surfaces is important for a detailed understanding of these phenomena.

The contact of rough surfaces has been studied by many researchers and the pioneering contribution to this field was made by Greenwood and Williamson [57], who developed a basic elastic contact model (GW model). In their model, a rough

surface was represented by a population of hemispherical-tipped asperities of identical radius of curvature R with their summit heights following a Gaussian distribution. The GW model assumes the surface asperities follow an elastic deformation behavior as proposed by Hertz [59], the contacts are frictionless and there are no interactions between neighboring asperities. Greenwood and Tripp [60] also showed that the contact of two rough surfaces could be modeled as the contact of an equivalent rough surface to a perfectly flat plane. The relationships for the true contact area and applied load, as functions of the nominal contact area and surface roughness (characterized in terms of radius of curvature R , the standard deviation of the asperity heights σ_s and the area density of asperities η , as explained in chapter 3), can be expressed as

$$A_e(w) = \eta A_n \pi R \sigma_s \int_{d/\sigma_s}^{\infty} (z - d) f(z) dz \quad (1.7)$$

and

$$L_e(w) = \frac{4}{3} \eta A_n E R^{1/2} \sigma_s^{3/2} \int_{d/\sigma_s}^{\infty} (z - d)^{3/2} f(z) dz, \quad (1.8)$$

where A_n , E and d are the nominal contact area, Young's modulus, and distance between the mean planes of the contacting surfaces, respectively. z is the asperity height parameter and $f(z)$ is the probability distribution function for the asperity heights.

The GW model has some severe limitations and practicality issues. First, it assumed all surface asperities to have an identical curvature radius R where in reality, one would expect the asperities to have different curvature radius. Furthermore, the GW model assumes elastic deformation. However, most contacting surfaces under

applied load will involve various degrees of plastic deformation of the asperities and without taking into account the plastic deformation component, the true contact area will most likely be underestimated. For instance, Powierza *et al.* [61] performed an experimental verification of the GW model and found that the model was surprisingly accurate at predicting elastic contact phenomena but when the applied load exceeded about half the yield point load, the results predicted by the GW model severely deviated (underestimate) from the experimental results.

In 1933, Abbott and Firestone [62] developed, till today, the most widely used plastic contact model, known as the surface micro-geometry model. This model assumes that the deformation of a rough surface against a rigid smooth flat is fully plastic and is equivalent to the truncation of the undeformed rough surface at its intersection with the flat surface. As a result, the true contact area is simply the geometrical intersection of the flat surface with the original undeformed profile of the rough surface, and the contact pressure is the flow pressure. Pullen and Williamson [63] also established a volume conservation model for the fully plastic contact of rough surfaces. However, these two models are only applicable for fully plastic deformations.

To bridge the two extreme approaches of modeling contacts as fully elastic or fully plastic, Chang [64] put forward an elastic-plastic contact model for rough surfaces on the basis of volume conservation of plastically deformed asperities (CEB model). The model, based on the same principal assumptions as the GW model, is applied to the entire deformation range, from fully elastic to fully plastic. However, this model has several shortcomings. The first is the discontinuity in the contact load

at the critical point of initial yielding. At this point, the average contact pressure is allowed to jump from the yield point Y to an average stress of $1.6Y$ in the plastic regime, where Y is the yield stress of the material. Another shortcoming is that this model only allows two possible states of deformation for a contacting asperity, either fully elastic or fully plastic, the transition from elastic to plastic is not modeled. Zhao *et al.* [65] and Kogut and Etsion [66] have developed models to describe the transition between the fully elastic and fully plastic regimes. Zhao made use of a template function put forth by Francis [67], which satisfies the continuity of the function and its slope at the two transitions, while Kogut and Etsion analyzed the problem with finite-element analyses. Both methods assumed full plasticity beyond a certain critical amount of deformation.

In all these analysis, the average contact pressure, upon the onset of full plasticity, was taken to be equivalent to a fraction of the hardness H , of the material. Tabor [68] showed that this is true for materials without work hardening characteristics or that have already work hardened, but is not valid for work hardening materials, like most metals. A contact model that takes into account of work hardening effects will be proposed in chapter 4.

1.6 Thesis Objectives

The primary goals of this thesis are to characterize the bond quality of wafer-level Cu-Cu thermo-compression bonding and to seek a fundamental understanding of the mechanisms of metal-metal adhesion, in particular on the effects of applied

load and the roughness of the bonding surfaces. Specifically, the objectives of this thesis are

- To understand the dependence of process parameters such as bond temperature, surface roughness and applied loads, on the bond strengths of the Cu-Cu bonds, using the 4-point bend tests.
- To understand the contact and bonding mechanisms of the thermo-compression bonding of metal-metal surfaces
- To measure the electrical resistances of the Cu-Cu bonded interfaces and relate them to the results obtained from the mechanical perspective.

Analytical models are to be developed to characterize the bond quality from both mechanical and electrical perspectives.

1.7 Thesis Scope

In this chapter, an introduction to 3D integration was presented and a literature survey of the various wafer bonding and bond strength characterization techniques was reviewed. Chapter 2 of the thesis describes the fabrication process for the four-point bend test structures and the test results. Chapter 3 discusses the development of a surface roughness acquisition model, which uses the data from atomic force microscopy (AFM) scans. The model determines surface roughness characteristics in terms of radii curvature, standard deviation of the heights and area density of the asperities. Chapter 4 introduces the concept that the dicing yield of bonded wafers can be related to the true contact area between the contacting surfaces and can be treated as an indicator of the bond quality. A contact model is also

developed, and together with the surface roughness model developed in the previous chapter, used to estimate the true contact area between contacting surfaces with different topological characteristics. In Chapter 5, the fabrication of bonded structures for electrical measurements is presented and contact resistances measurements are reported. A contact resistance model is also developed, and using the calculated true contact area, an estimate of the theoretical contact resistance is obtained. A comparison is made between the experimental and theoretical values to test the validity of the calculated true contact area. Chapter 6 reports on the study of contact resistance reduction of high resistance contacts under a high current density. It was determined that the reduction was due to electromigration, rather than Joule heating effects. Finally, conclusions and remarks on the accomplishments of the thesis and future work are made in Chapter 7.

Chapter 2 – Four-Point Bend Test

The mechanical adhesion of Cu bonded wafers were evaluated by the four-point bend test. Test structures were designed for the execution of the testing and a process flow was developed for the fabrication of the structures. The first part of this chapter describes the fabrication and bonding processes, followed by a description of the four-point bend test experimental setup. The effects of different bonding and topological parameters such as applied load, nominal area and surface roughness on the bond quality, as reflected in the four-point bend test results, will then be presented and discussed.

2.1 Fabrication Process

The experiments of thermo-compression bonding of Si wafers with Cu blanket film and Cu patterned lines were performed with different applied loads, bonding temperature, nominal contact area (macroscopic Cu surface area available for bonding) and surface roughness. The wafer bonding configurations can be categorized into two types: blanket-blanket and blanket-patterned, as illustrated in Figure 2-1.

Figure 2-1: Schematic cross-sections for (a) blanket-blanket and (b) blanket-patterned bonding.

200 mm Si wafers of about 700 μm -thick were used as substrates for fabrication of the test structures. The film stack consists of 200 nm-thick PECVD SiO_2 , a 25 nm-thick sputtered Ta barrier/adhesion layer, and an approximately 400 nm-thick electro-copper-plated (ECP) layer. The wafers were bonded either in the as-deposited state or after chemical-mechanical polishing (CMP) and/or plasma etched. The fabrication process flow charts for blanket Cu and patterned Cu lines wafers are shown in Figures 2-2 and 2-3, respectively.

Figure 2-2: Fabrication process flow for blanket Cu wafers.

Figure 2-3: Fabrication process flow for patterned Cu wafers.

For the blanket Cu wafers, the first step was to deposit a layer of 200 nm-thick SiO₂, using plasma enhanced chemical vapor deposition (PECVD) at 300°C. This was followed by depositing the Ta barrier (250 Å) and Cu seed layer (150 nm) using physical vapor deposition (PVD) process. Immediately after the PVD process, a Cu layer of 1 µm-thick was grown using the ECP process and was annealed in N₂ ambient at 200°C for 30 minutes. Cu CMP was subsequently performed to reduce the Cu thickness to about 400 nm. After the CMP process, the Cu is typically very smooth with a root-mean-square (rms) roughness of about 0.6 nm, as measured by AFM and is denoted as blanket (CMP). To bond Cu surfaces with different surface roughness, the Cu surfaces either went through the CMP process followed by a CHF₃ or Ar plasma etching (blanket (CMP and Etched)) or were bonded in the as-deposited state (blanket (ECP)).

For the patterned Cu wafers, 800 nm-thick SiO₂ was first deposited followed by lithography and etching of oxide trenches that were approximately 600 nm in depth. The Ta barrier and Cu seed layer were then deposited by PVD, followed by growing 1 µm-thick of Cu by ECP. After Cu annealing and Cu CMP, an addition 200 nm-thick of SiO₂ was removed to reduce the amount of dishing in the Cu lines. The next step was the deposition of a 500 Å-thick silicon nitride, which acts as a protective layer over the Cu surface, as the backside of the substrate was being patterned and etched to create marks for alignment and dice-saw purposes. The last step of the process was the removal of the silicon nitride using plasma-etch (CHF₃) with over-etch to recess the oxide below the Cu surface, to ensure optimal contact between the Cu lines and the opposite surfaces.

The Cu lines were 50 μm -wide and approximately 400 nm-thick. Different pattern densities were achieved through the use of different line spacing of either 50, 175 or 310 μm , with only one type of spacing on any patterned wafer. The spacing is expressed in the parenthesis, for instance, patterned (175) refers to Cu lines with line spacing of 175 μm . The corresponding nominal contact areas with respect to the 50, 175 and 310 μm line spacings are approximately 0.0126, 0.0057 and 0.004 m^2 , respectively.

The next step in the process flow is wafer bonding. Since Cu forms an oxide film under ambient condition [67], removal of the surface oxide prior to bonding is a key step. Glacial acetic acid clean was used as a pre-bond cleaning step, which was demonstrated to be effective in removing copper oxide and reducing the reoxidation rate of Cu [69]. The wafers were dipped in the acid for 10 minutes to remove the surface oxide and subsequently blown dry with N_2 air gun. It was shown that oxidation of Cu was reduced when surfaces were blown-dried instead of rinsing with water, as the latter provides $(\text{OH})^-$ ions which aid in the oxidation of Cu. Nonetheless, the wafers were exposed to the surrounding ambient for a brief period of time after cleaning and before loading into the bonding chamber. Hence, unless the wafers can be cleaned and bonded without breaking vacuum, Cu reoxidation cannot be totally eliminated.

After the pre-bond cleaning step, the wafers are aligned before bonding. The alignment tool is Electronic Vision GroupTM (EVG) 640 Aligner as shown in Figure 2-4, which makes use of an optical alignment method. For wafer bonding involving blanket Cu wafers, the precision alignment is not crucial and in such cases, the wafer

alignment was performed manually with respect to the wafer notch. However, for bonding of patterned wafers for the contact resistance measurement experiments as described in Chapter 5, the precise alignment of the wafers will be very important.

Figure 2-4: Setup of the EVG 640 aligner.

For patterned wafers, alignment marks were fabricated on both the frontside and backside of the wafers, aligned to one another. The top wafer was placed face-down first and the position of the alignment marks on the frontside was recorded. The bottom wafer was then placed face-up and the alignment marks on the backside were adjusted to the positions recorded for the top wafer. After alignment, the wafers were clamped together in a bond chuck, but separated by three 30 μm -thick metal flaps. The bond chuck was then transferred to a bonding chamber (EVG 520) as shown in Figure 2-5, and the bonding chamber was subjected to three N_2 pump-purge cycles to reduce the ambient oxygen content.

Figure 2-5: Setup of the EVG 520 bonder.

Next, the chamber was pumped down to approximately 5×10^{-4} Torr and a pre-programmed piston load was applied. The temperature was subsequently ramped to 300°C and the wafers were held together for 30 minutes. The piston load was subsequently removed and the wafers were annealed at the same temperature without breaking vacuum for another 30 minutes. The bonding cycle is as summarized in Figure 2-6. The piston load applied for all the experiments ranged from 400 N to 10 kN, which gave a nominal pressure between 0.013 to 2.45 MPa, depending on the nominal contact area.

Figure 2-6: A schematic of the bonding cycle.

The bonded wafers were eventually dice-sawed into rectangular strips of 5.85 mm x 30.4 mm, at a speed of 2 mm/s using a DISCO DAD651 automated dice saw equipment as shown in Figure 2-7.

Figure 2-7: Setup of the DISCO DAD651 automated dice saw.

A notch was also made by cutting into the top wafer, as close as possible to the bonded interface, without cutting the bottom wafer. This is to facilitate crack initiation and propagation during the four-point bend tests. When a load is being applied, the corner of the notch will experience the maximum stress, which resembles a crack tip. The crack will subsequently propagate into and grow along the weakest interface, as shown in Figure 2-8. As will be mentioned in section 2.8, most of the delaminations occurred at the Cu-Cu interface, and when the Cu-Cu bond strength is expected to be high, the cracks would propagate into the Ta/SiO₂ interface.

Figure 2-8: Schematic of test structure showing different layers and the proposed crack propagation path.

During the dicing process, samples which were poorly bonded would be detached. The number of detached samples would depend on the bond quality of the bonded wafers and will be discussed in greater details in the next chapter. The dicing yield, defined as the percentage of diced strips that remained intact after dicing, was recorded for each pair of bonded wafers. If all samples survived the dicing process (100% yield), each bonded wafer pair would yield 154 samples for the four-point bend testing.

2.2 Testing Procedure

Test Apparatus

Testing was carried out on an actuator-controlled test equipment with four-point bend test fixtures as shown in Figure 2-9. The outer and inner loading points can be set to a maximum separation of 35 mm and 27 mm respectively. For the four-point bend tests, the outer and inner loading points were spaced at 25 mm and 17 mm apart respectively. The supports at the loading points are steel pins of approximately 2

mm in diameter, rested on V-groove supports. During the tests, the system measures the amount of displacement and the corresponding load, to a maximum of 180 N. The system can be connected to any computer with installed data acquisition software DTS™ which records the displacement and load data.

Figure 2-9: Schematic of the four-point bend test equipment.

Alignment and Test Procedure

Misalignment in the four-point test fixture occurs when the loadings are not symmetrical about the notch on the sample, resulting in uneven loadings on both ends of the sample [56]. This will probably cause crack initiation and propagation on the side where the loading is stronger and results in erroneous fracture toughness values. On the test fixture, there is an alignment mark (as shown in Figure 2-9) and by aligning the mark to the notch would ensure symmetrical application of load on both ends of the sample. Hence, any errors on the measured fracture toughness due to misalignment would be minimized. The test samples were stressed at a constant displacement rate of 0.3 $\mu\text{m/s}$. The dimensions of the specimen such as thickness and width were measured with a digital micrometer. Physical dimensions of the sample

and the critical load for delamination obtained from the test were substituted into equation (1.5) to calculate values for G_C .

Theoretically, the critical load for delamination P is obtained when the crack propagates at a constant velocity under a constant load as illustrated in the ideal case in Figure 1-7. The ideal load-displacement characteristic was usually not observed in this work. The majority exhibited non-ideal behaviors as shown in Figure 2-10(a). In these non-ideal cases, the load P in equation (1.5) was taken as the average of several load values. Some of the samples showed a load peak followed by a load minimum, upon which elastic loading resumed, as shown in Figure 2-10(b). The load peak probably corresponds to the load required to initiate the crack and the actual load for crack propagation could be lower. Hence, in these cases, the average between the load peak and load minimum was being used to calculate G_C . There are also samples exhibiting a continuous series of load peaks and load minimums and the calculations of G_C for these samples were based on the approximated average load, as shown in Figure 2-10(a).

There are also samples which displayed several constant load regions, as shown in Figure 2-10(c). The first constant load region could be due to the delamination of one side of the sample while the second constant load region could correspond to the delamination of the other side. Hence, to take the average of these two constant load regions as Tsau *et al.* [41] did, would overestimate the bond strength. In our analysis, the first constant load region was taken as the delamination load in the calculation of G_C .

(a)

(b)

(c)

Figure 2-10: Illustration of non-ideal behavior: (a) plateau/multiple load peaks-load minimum, (b) single load peak-load minimum and (c) two constant load regions.

In our experiments, the four-point bend tests were performed on Cu-Cu bonds and hence the modulus mismatch is essentially zero. For the test specimens used in this work, the two layers have identical thickness and hence $\psi \sim 40^\circ$, which implies approximately equal contributions from both mode I and mode II [56].

2.3 Factors That Affect Bond Quality

In thermo-compression bonding, the surfaces of the bonding layers are brought into direct contact. However due to surface roughness, initial contacts will be formed between surface asperities and not the whole surface area (nominal contact area). As the applied load increases, the contacting asperities will deform under the applied pressure and the contact area will increase. At elevated temperatures, metals like Cu will soften due to increasing dislocation mobility, higher diffusivity and lower yield stress. Therefore, the pressure requirement to achieve good bonding can be offset by increasing the bonding temperature. Hence, the bonding quality will also depend on the mechanical properties of the bonding material such as yield stress, Young's modulus, creep and diffusion characteristics. The elastic constants of Cu are known to decrease with increasing temperatures [70] and the creep and diffusion typically have an exponential dependence on temperature as well. It is thus expected that increasing temperature and applied pressure will facilitate the bonding process.

There are many parameters besides material properties of copper that affect the Cu-Cu bonds. For four-point bend tests, Rajappa [42] has investigated the effects of bonding ambient, surface cleaning, Cu thickness and bond temperature. However, the Cu layers were deposited by evaporation whereas in this work, the Cu layers are

electro-plated. In the following sections, the effects of bonding temperature, surface roughness and bond pressure on the measured bond toughness will be presented.

2.4 Effects of Bonding Temperature

The effects of bonding temperature were investigated by bonding wafers with blanket Cu films at temperatures ranging from 250°C to 400°C. The applied load to the bonding process was 2000 N. The bonded wafers were diced and four-point bend tests were performed on the samples. The critical strain energy release rates G_C are plotted versus the bonding temperature as shown in Figure 2-11.

Figure 2-11: Four-point bend test results of G_C versus bonding temperature.

The yield stress and diffusivity of Cu are expected to decrease and increase respectively, as the bonding temperature increases. Hence, there will be greater deformations on the surface asperities and the contact area between the surfaces will increase, enhancing the bond quality.

In general, the four-point bend test results supported the expectation that higher bond temperature would lead to higher bond strength. For wafers bonded at 275°C and below, the majority of the samples delaminated during the dicing process and the four-point bend test results on these samples exhibited bond strengths less than 2 J/m², reflecting the poor bond quality. For samples bonded at 300°C and above, the average bond strengths (of more than 3 J/m²) are evidently higher than those bonded below (300°C). However, within a set of results, the values of the bond strength can vary by as much as $\pm 100\%$ with respect to the average value. Hence, the results between the different sets of experiments have significant amount of overlap and are less distinctive.

2.5 Effects of Surface Roughness

In the fabrication of patterned Cu wafers using the damascene process, one of the last fabrication steps is the recess of the dielectric, such as silicon dioxide, surrounding the Cu structures so that they can be protruded to facilitate contact and bonding with the opposite Cu surface in the subsequent thermo-compression bonding process. This material recess step can be achieved by either plasma or wet etch. For instance, IMEC recently reported the use of plasma etch to recess Si to expose the Cu bonding structures [71] while Tezzaron uses a wet etch to recess oxide to expose the Cu structure [72]. Evidently, these different etch mechanisms will have different impact on the roughness of the Cu surface (the plasma etch will induce a rougher surface with its ion bombarding mechanisms) but the effects of roughness on thermo-compression bonding have never been reported. In this work, the effects of surface

roughness on Cu thermo-compression bonds are one of the main objectives of this thesis.

In this set of experiments, wafers with blanket Cu films were bonded at 300°C with an applied load of 4550 N. The different roughness of the Cu surfaces obtained is process related: CMP, CMP and plasma etched, or in the as-deposited state (ECP). From AFM measurements, the roughness of Cu surfaces after Cu CMP is typically around 0.5 to 0.6 nm while for the ECP wafers, it is an order of magnitude higher of about 5 to 8 nm. For the plasma-etched Cu surfaces, the roughness can range from about 2 nm to 5 nm (depending on the etch process) and the values cannot be replicated exactly, probably due to the non-uniformity of the etching process. Figure 2-12 shows the plot of G_C versus the total surface roughness, which is defined here as the sum of roughnesses of the two surfaces.

Figure 2-12: Four-point bend test results of G_C versus total surface roughness.

The results show that the average bond strength decreases as the surface roughness increases. Like in the previous section, there is also a significant amount of scattering and overlap between the different data sets. However, the general trend

clearly shows that for a given contact load, bonding of rougher surfaces would result in poorer bond quality with low bond strengths.

2.6 Effects of Contact Pressure

The effects of contact pressure were examined from two perspectives: different applied loads on the same nominal contact area or on different nominal contact areas. The first case involves bonding of blanket Cu (CMP) wafers whereas in the second case, it involves bonding of both blanket and patterned Cu (lines) wafers with different nominal contact areas. The bonding temperature in all the experiments was 300°C.

For bonding with different applied loads on the same nominal area, the load ranged from 400 to 10000 N, translating to a nominal contact pressure of 0.013 to 0.32 MPa, respectively. Four-point bend tests were performed and the critical strain energy release rates G_c are plotted versus the nominal bonding pressure as shown in Figure 2-13.

Figure 2-13: Four-point bend test results of G_c versus nominal bonding pressure. (Different applied loads on same nominal contact area)

Blanket Cu (CMP) wafers were used to ensure identical surface roughness. Hence, any difference in the results would be entirely due to the effects of different applied loads. For a given bonding temperature, a higher contact pressure (or higher load) would result in a larger amount of deformation of the surface and hence a larger contact area, which in turn leads to a higher bond strength. Results, as shown in Figure 2-13, support the general expectation that larger contact pressures would lead to larger bond strengths.

From another perspective, the effects of bonding pressure were also examined by considering different nominal contact areas. The applied loads were 4550 and 10000 N and the bonding configurations involved blanket-blanket or blanket-patterned Cu wafers, as explained in section 2.1. The details of each experiment are tabulated in Table 2-1 and the plot of G_C versus the nominal contact pressure is shown in Figure 2-14.

Table 2-1: Details of experiments investigating the effects of different applied loads on different nominal areas.

No.	Bonding Configuration	Applied Load (N)	Nominal Contact Area (m ²)	Nominal Bonding Pressure (MPa)	Surface Roughness by AFM (nm)			Average G_C (J/m ²)
					Blanket	Patterned	Total Nominal Roughness	
1	Blanket-Blanket	4550	0.0314	0.14	0.52	-	1.04	6.53
2	Blanket-Blanket	10000	0.0314	0.32	0.52	-	1.04	7.84
3	Blanket-Patterned(50)	4550	0.0126	0.36	0.52	8.63	9.15	15.54
4	Blanket-Patterned(50)	10000	0.0126	0.80	0.52	8.63	9.15	17.47*
5	Blanket-Patterned(175)	4550	0.0057	0.80	0.52	10.00	10.52	38.63*
6	Blanket-Patterned(175)	10000	0.0057	1.77	0.52	10.00	10.52	12.2
7	Blanket-Patterned(310)	4550	0.0041	1.11	0.52	11.95	12.47	7.55
8	Blanket-Patterned(310)	10000	0.0041	2.45	0.52	11.95	12.47	9.44

* Exp. No 4 and 5 have the same nominal bonding pressure. Hence, the average G_C takes in account all data from these two experiments.

Figure 2-14: Four-point bend test results of G_C versus nominal contact pressure. (Different applied load on different nominal contact area)

In the first analysis, the applied load is the only variable while the nominal contact area and surface roughness did not change. In the second analysis, all the three parameters: applied load, nominal contact area and surface roughness, vary. Results from Figure 2-14, unlike results from Figure 2-13, show that the bond strength did not increase with the nominal contact pressure. In fact, the bond strengths obtained at higher nominal bonding pressure seemed to be lower if not comparable to samples bonded at a lower bonding pressure. The surfaces of these patterned wafers were however, much rougher than that of the blanket Cu CMP wafers, due to the effects of plasma-etch. This shows that neither the nominal contact pressure nor surface roughness, by themselves, could adequately describe the bond quality/bond strengths of the bonded surfaces: all factors have to be taken into account.

2.7 Bonding Non-uniformity

From equation (1-5), the critical strain energy release rate G_C has units of J/m^2 , which implicitly assumes that the bond strength is identical across the bonded area. However, even for blanket Cu wafers, either after ECP or CMP, there is a variation of the Cu film thickness across the wafer to an average standard deviation of about $\pm 5\%$. Hence when two wafers come into contact, certain regions of a wafer would have come into contact earlier than other regions and the contact pressure in these regions could be different. Furthermore, for patterned Cu wafers, Cu CMP processes will result in dishing effects as well as additional degrees of non-uniformity, depending on the geometry and density of the structures. Hence, it is most likely that various degrees of bonding non-uniformity exist in wafer bonding and it is important to take that into account in subsequent analyses. A schematic of the dishing effect and actual profilometric measurements across the Cu lines are shown in Figures 2-15(a) and 2-15(b), respectively. The amount of dishing, as illustrated in Figure 2-15(b), could be as high as 50 nm.

(a)

Figure 2-15: Illustration of dishing effects due to Cu CMP process, across a patterned Cu line: (a) schematics (b) actual profilometric measurements.

As mentioned in section 2.1, the bonded wafers were dice-sawed into rectangular samples for four-point bend tests and those that were poorly bonded would detach. Figure 2-16 shows a typical yield map after the dice-saw process. The colored regions represent samples that remained bonded whereas the uncolored regions represent detached samples. Result shows that the bonding is not uniform across the bonded wafer: some regions will bond better than others.

Figure 2-16: A yield map of a bonded pair of wafers after the dice-saw process.

SEM micrographs of delaminated samples after four-point bend tests are shown in Figures 2-17 and 2-18, showing the effects of dishing and bonding non-uniformity, respectively.

Figure 2-17: SEM micrographs of the surface of a patterned Cu line showing dishing effects.

Figure 2-17 illustrates a delaminated surface of a patterned Cu line. The regions near the edges of the line seemed to have bonded better, having features of ductile failures. In some cases, the bond was so strong that whole chunks of Cu were pulled off. Nearer towards the line center, deformations are evident but not as severe as those seen at the edges. This is most likely due to dishing effects, where Cu material near the edges would have sustained a higher contact pressure than at the center and hence would have bonded better (more deformations).

Dishing effects could account for the bonding non-uniformity on a small scale. However, the Cu CMP process could have polishing non-uniformity and some regions could have been polished more. Figure 2-18 shows the SEM micrographs of a blanket Cu surface delaminated from a patterned Cu lines wafer. The top left portion shows the imprints of the Cu lines from the opposite wafer as evidence of contact

whereas the bottom right portion of the Cu surface shows little evidence of deformations.

Figure 2-18: SEM micrographs of a blanket Cu surface showing effects of contact non-uniformity.

It is shown that for two contacting surfaces, not all the nominal contact area would be in contact and bonded. Within the bonded areas, there is also non-uniformity: areas which have sustained higher contact pressure will probably bond better than those with lower contact pressure. Bonding non-uniformity is a concern, especially for bonding involving patterned structures. It affects the overall bonding quality and is something that should be taken into account of when performing any type of bond strength measurements and data analysis.

2.8 Failure Modes

As mentioned in section 2.2, the observed failure modes are either cohesive, which fails within the bonded Cu layers, or adhesive, which fails along the Ta-SiO₂

interface. Optical images of two opposing delaminated samples after the four-point bend test are shown in Figure 2-19.

Figure 2-19: Optical images of fracture interfaces. Brown regions indicate Cu-Cu interface, light silver (Ta) and greenish silver (SiO_2) indicate Ta- SiO_2 interfaces.

Figure 2-19 suggests that the delamination may initially occur at one interface but could propagate to another interface. Based on color, the brown, light silver and greenish silver regions were determined to be Cu, Ta and SiO_2 respectively, by energy dispersive X-ray spectroscopy (EDX) analysis as shown in Figures 2-20(a) to 2-20(c).

(a)

(b)

(c)

Figure 2-20: EDX analysis on (a) brown region, (b) greenish-silver region and (c) light silver region.

An electron beam is accelerated and directed to the surface to be characterized. The incident beam excites electrons in an inner shell of the various atoms, prompting its ejection and resulting in the formation of electron-hole pairs within the atoms' electronic structure. Electrons from an outer, higher-energy shell then fill the holes, and the excess energy of the electrons are released in the form of an X-ray. The release of X-rays creates spectral lines that are highly specific to individual elements; thus, the X-ray emission data can be analyzed to characterize the sample in question.

The EDX analysis confirmed expectations that the brown region is Cu with small amount of Si counts from the substrate. The greenish silver region shows Si and O peaks, indicating SiO_2 , with the count for Si much higher due to contributions from the substrate. The numerous Ta peaks for the light silver region indicates the presence of Ta. The Si and Cu peaks were obtained from the Cu film and Si substrate beneath the Ta layer.

A SEM micrograph of a fracture interface is shown in Figure 2-21. The Cu-Cu interface is rough, showing ductile failures while the Ta surface is quite smooth, suggesting brittle failure with little plastic deformation [42].

Figure 2-21: SEM micrograph of fractured sample showing Cu-Cu and Ta- SiO_2 interfaces.

2.9 Discussion and Summary

This chapter presented the experimental results and discussed the effects of temperature, contact pressure and roughness of bonding surfaces on the bond strengths of the Cu-Cu bonds. As bond temperature increases, the Cu-Cu bond strength generally increases due to larger contact area through larger Cu deformation

(lower yield stress), higher diffusivity and interfacial diffusion. However below 300°C, the bond quality is evidently low and many samples delaminated during the dice-saw process and most of the measured bond strengths are below 2 J/m². Compared with measurements obtained by Rajappa [42], the bond strengths obtained in this work are low. A significant factor could be due to the absence of a forming gas purge in the bonding chamber, which would have reduced native Cu oxide on the surface and had been demonstrated to enhance the Cu-Cu bond strength. It is also unclear if issues of bonding non-uniformity as mentioned in section 2.7 would be a lesser concern if Cu is deposited by the lift-off method as in [42].

It was also demonstrated that rough surfaces are detrimental to good bonding: the rougher the surfaces, the lower will be the bond strengths. As mentioned in section 2.3 when two surfaces are pressed together, the surface asperities, not the nominal contact area, will make the initial contacts and the contact area will grow under higher temperatures and pressures. Even under high temperature and contract pressure, the number of contacting asperities and hence the contact area is expected to be smaller for the contact of rough surfaces than in the case of smoother surfaces, or vice versa, as verified by the experimental results.

At a given bonding temperature and similar surface roughness, the bond strength of the bonded interface increases as the contact pressure increases (increase in applied load with the same nominal contact area). This is expected as a larger contact force will result in larger deformations, resulting in a larger contact area and hence the bond strength will increase. However, when the contact pressure is varied through different configurations of applied loads, nominal contact area and surface

roughness, this relationship is no longer valid. Due to rougher surfaces, the initial contact area could be much smaller than compared to contacts of smoother surfaces. Even if the nominal contact pressure is much higher, the contact area may not be as high as when a lesser nominal contact pressure is applied to smoother surfaces.

Analysis of fractured surfaces showed that the crack paths can be non-uniform, i.e. both Ta-SiO₂ and Cu-Cu fracture interfaces are observed. This suggests that the crack paths do not necessarily propagate along the Cu-Cu interface but could propagate through the bulk of the Cu stack into the Ta-SiO₂ interface as well.

The G_c results obtained from the four-point bend tests showed that the applied load, nominal contact pressure and surface roughness, by themselves, are inadequate in characterizing the bond quality. The effects of bond quality due to changes in these bonding parameters must be collectively analyzed as a whole. It is proposed that the bond quality can be correlated to the true contact area, which is a function of the applied load, nominal contact area and surface roughness related characteristics such as curvature radii of asperities, standard deviation of asperity heights and asperity density. The amount of contact area between the contacting surfaces can be perceived as an important measure of the bond quality and is expected to co-relate to the bond strengths and dicing yields.

The bond strength results also exhibit a huge amount of variation by as much as $\pm 100\%$ about the average value, creating significant amount of overlaps between the different sets of data, making the effects of the different bonding parameters less distinct. A similar huge scattering of the bond strength measurements have also been observed by Tsau [41] on Au-Au thermo-compression bonded samples and has

attributed this effect to the plasticity that is being introduced in the samples. In this work, it is proposed that the bond strengths can be co-related to the true contact area between the contacting surfaces and in addition to the effects of plasticity, the huge scattering of results could also be due to the non-uniform distribution of the contact area across the contacting surfaces, which will be discussed in greater details in Chapter 4.

Chapter 3 – Surface Roughness Acquisition Model

3.1 Introduction

Knowledge of the contact mechanism between contacting surfaces is important in the understanding of both the mechanical and electrical properties of bonded interfaces. Such mechanical and electrical properties include the true contact area, bond strengths, contact resistance (conductivity) and reliability.

Surface roughness plays a crucial role in the contact and friction between surfaces. For describing the roughness of a surface, statistical parameters for the surface height distribution function such as the root-mean-square (rms) height, slope and radius curvature of asperities have been used in several studies. These parameters can be directly related to the density of asperity η , the asperity radii curvature R and the standard deviation of summit heights σ_S [57, 59-66]. However, these parameters, if determined experimentally, can vary with sample size and instrumentation resolution [73].

Examination of rough surfaces shows that they often have multi-scale features. That is, when a section of a rough surface is magnified, smaller scales of roughness appear. This general characteristic of surfaces was recognized long ago by Archard who described an engineering surface as consisting of “protuberances on protuberances” [74]. Archard later proposed a model (WA model) suggesting that the product of R , η and σ_S can be expressed as

$$R\sigma_S\eta = \text{Constant} \quad (3.1)$$

where the constant has a value in the range of 0.03 to 0.05 [75,76]. However, the WA model too, is not an accurate characterization of surface topology for two reasons. First, the minimum sampling interval on their equipment is $0.25\text{ }\mu\text{m}$, which is too large to give an accurate representation of the surface profile. Second, the analysis was based on 2D profile scans instead of 3D, and their method of obtaining R and η was purely mathematical rather than from actual measurements.

Suppose a surface profiler makes a scan across a hypothetical surface profile as shown in Figure 3-1. If the sampling interval is relatively large, as denoted by the red spots, then the generated profile will be greatly distorted from the actual surface profile.

Figure 3-1: Deviations of the measured to the actual surface profiles: Solid line – surface profile. Dots – sampling points. Dotted line – generated surface profile.

In AFM measurements, regardless of the scan size, the number of sampling points is commonly set at 256, though the maximum can be set at 512. Therefore, the larger the scan size, the larger will be the sampling interval and the generated profile will more likely not reflect the actual surface profile. Hence to ensure the generated profile reflects the actual surface profile as close as possible, the scanning equipment should have sufficiently high resolution and the scan size should be kept small.

The motivation here is to create a model to analyze the surface topography and extract the surface roughness parameters based on AFM's 3D surface profile

measurements. The extracted surface roughness parameters (R , σ_S and η) will then be used in the calculation of the true contact area, as will be discussed in Chapter 4.

3.2 The Acquisition Algorithm

AFM surface profile measurements are taken on the pre-bond Cu surfaces with a scan size of $1 \times 1 \mu\text{m}^2$. The sampling interval is approximately 4 nm in both the orthogonally defined x and y directions. Such small scan sizes should provide sufficiently high resolutions to reflect the actual surface profiles and obtain maximum accuracy of the surface roughness parameters in subsequent calculations. To ensure that these small-area characterizations would be representative of the entire surface, a minimum of six scans per surface were taken and the average values were used in subsequent analyses.

Scanned data from the AFM measurements were extracted and analyzed using a MATLAB algorithm. The algorithm generates 3D topographical maps that are representations of the surface in the form of a set of points, each of which is defined in a Cartesian coordinate system by the three ordinates x , y and z as shown in Figure 3-2. With the extracted scan data, surface roughness parameters R , σ_S and η can then be estimated and analyzed.

Figure 3-2: A 3D surface plot generated from a MATLAB algorithm with raw data from AFM scan.

3.2.1 Calculation of η and σ_s

A summit can be defined at a location where the height is a local maximum. In the case of a 2D surface (i.e. a line scan), a sample point is a summit if its height exceeds that of its two neighboring points, while for a 3D surface (i.e. an AFM scan), a sample point is a summit if its height exceeds that of at each of its eight neighboring points. Nayak [77] has shown that a sample point whose height exceeds that of at each of the eight neighboring points might be a ‘false’ summit. To reduce the probability of identifying these ‘false’ summits, the height of the summit point should be larger than the heights of at least two consecutive neighboring points in each of the eight directions, i.e. sixteen neighboring points (Figure 3-3).

Figure 3-3: A summit exists at z_{ij} if $z_{ij} > z_{i \pm n, j \pm n}$ for n equals 1 and 2 (a total of 16 points).

In other words, z_{ij} is a summit height when

$$z_{ij} > z_{i+1,j} > z_{i+2,j} ,$$

$$z_{ij} > z_{i-1,j} > z_{i-2,j} ,$$

$$z_{ij} > z_{i,j+1} > z_{i,j+2} ,$$

$$z_{ij} > z_{i,j-1} > z_{i,j-2} ,$$

$$z_{ij} > z_{i+1,j+1} > z_{i+2,j+2} ,$$

$$z_{ij} > z_{i-1,j-1} > z_{i-2,j-2} ,$$

$$z_{ij} > z_{i-1,j+1} > z_{i-2,j+2} \text{ and}$$

$$z_{ij} > z_{i+1,j-1} > z_{i+2,j-2} .$$

The identification of these asperity summits provides an estimate of the total number of asperities per scan area (asperity density η). It was found that the density of asperities tends to decrease as surface roughness increases, as shown in Figure 3-4.

Figure 3-4: Variation of average asperity density versus the measured rms surface roughness.

As the surface roughness increases from 0.5 nm to 10 nm, the asperity density decreases from about 6×10^{14} to about 2×10^{14} . For surface roughness of between 8 nm and 12 nm, the asperity density η seems to vary about 2×10^{14} . This may be due to rougher surfaces having larger asperities such that the number of identified summits is about the same for this range of surface roughness.

The identification of the asperity summits also provided a distribution of the summit heights. The GW model proposed that the *asperities' heights* would follow an approximately Gaussian distribution (which many subsequent papers quoted). From the AFM scans and the model, we obtained the distribution of asperity heights and results showed that the asperity heights is well fitted into a Gaussian distribution function as shown in Figure 3-5. However, surfaces may have different surface height distributions such as Gaussian, Exponential, Setch, Lorentzian, Power Law, Bessel

and Self-Affine (Fractal) [78]. An analysis to determine if the surface heights (of the Cu surface) follow a Gaussian distribution is presented in Appendix B.

Figure 3-5: Typical asperity height profiles, following a Gaussian distribution. The asperity heights are with respect from a mean height plane.

From the distribution of asperity heights, the standard deviation of summit heights σ_s , of a surface can be determined and was found to scale almost proportionally to the rms surface roughness, as shown in Figure 3-6. This is to be expected as the root-mean-square value and the standard deviation are mathematically related.

Figure 3-6: Variation of the standard deviation of asperity heights with respect to the measured rms surface roughness.

3.2.2 Calculation of asperity curvature radius R

The curvature radius of an asperity summit is calculated along the x , y and the two diagonal axes, and all successive lower z -coordinate data points (with respect to the summit point) are considered. The asperity summits are generally found to be approximately spherical, but the radii of curvature may vary, depending on the direction of consideration. For instance, Figure 3-7 shows the surface profile of a typical asperity characterized along the four different directions. The curvature radius R of the asperity is represented by the best fit curve, calculated using the sum-of-least squares method, taking into account all data points along the four axial directions.

Figure 3-7: Asperity profiles along the x , y and two diagonal axes. The best fit is depicted by the solid curve.

The value of R will also depend on the boundary conditions imposed by the calculating algorithm. For instance, an asperity profile may not be perfectly spherical but may taper off in some directions as shown by the encircled data points in Figure 3-8. Suppose the calculation considers all successive lower z -coordinate points including the encircled data points, the calculated R is approximately 85 nm, compared to 69 nm if the encircled data points are not considered. Out of a random sampling size of ten asperity profiles, only two scans have such ‘tapered-off’ characteristics. The slightly over-estimated values of R are not sufficiently high to affect the overall average value $\bar{R}$ and hence the effects, if any, are minimal. As a rule of thumb, the algorithm imposes the maximum number of successive lower z -coordinates in each direction to 15, which will be used in the calculation of R . The MATLAB code of the algorithm is given in the Appendix A.

Figure 3-8: An asperity spherical profile, with 'tapered-off' characteristics in certain directions.

As mentioned in section 1.6, the summit radius curvature is not expected to be identical over the entire surface and to assume it to be identical will be a major simplification for any contact model. From the AFM scans, a typical spread of the radius curvatures with respect to their corresponding heights is shown in Figure 3-9.

Figure 3-9: Summit radius curvature plotted against their corresponding heights.

The radius curvature ranges from about 20 nm to more than 100 nm, with an average of approximately 50 nm. The radius curvature does not seem to follow any form of distribution and varies randomly with the asperity heights. Some of the asperity heights have negative values because the reference represents the average height which is set to zero.

The radius curvature also seems to vary on a macroscopic level. Figure 3-10 shows the variation of the calculated average radius curvature against the different surface roughness. Results suggest that the average radius curvature tends to decrease as surface roughness increases. This can be qualitatively explained if the surface can be envisaged as one of valleys with various depths and the overall topography can be measured as a degree of roughness. Then, rougher surfaces would have deeper valleys than smoother surfaces and the summit radius curvature of asperities with deeper valleys would be arguably smaller.

Figure 3-10: Variation of the average curvature radius with respect the measured rms surface roughness

3.2.3 Other Results

As mentioned in the introduction, Archard [76] proposed that the product of $\bar{R}$, σ_s and η is a constant ranging from 0.03 to 0.05. However, due to reasons as mentioned in section 3.1, the measured profile may not reflect the actual surface profile and hence the estimated values of $\bar{R}$, σ_s and η are probably incorrect. With this newly developed model, we would be able to test and validate this postulation.

An asperity is identified when the summit's z -coordinate is larger than at least two consecutive neighboring points in each of the eight directions (16 points). This criterion can be relaxed such that an asperity summit exists when its z -coordinate is just larger than its adjacent point (8 points), or be more stringent where its z -coordinate must be at least larger than three consecutive neighboring points (24 points). Values of $\bar{R}$, σ_s and η are calculated for surfaces with different rms roughness ranging from 0.5 to 12 nm (as measured on an area of $1 \times 1 \mu\text{m}^2$). Tables 3-1a, b and c show the results of $\bar{R}$, σ_s and η for these three different criteria for different surface roughness.

Tables 3-1(a), (b) and (c) show that for the same surface (roughness), the number of summits and hence the value of η , may increase or decrease as the summit definition criteria is relaxed or tightened, respectively. The values of $\bar{R}$ and σ_s are found to have no significant changes, which implies that the number of identified summits is sufficient to produce approximately the same statistical average for $\bar{R}$ and σ_s .

Table 3-1: Comparisons of R , σ_s and η when the z-coordinate of a summit is larger than the (a) 8 neighbouring points, (b) 16 neighbouring points and (c) 24 neighbouring points.

Roughness (nm)	$\bar{R}$ (m)	σ_s (m)	η (/m ²)	$\bar{R}\sigma_s\eta$
0.52	1.17E-07	7.20E-10	2.93E+15	0.25
2.85	9.69E-08	2.43E-09	1.03E+15	0.24
4.01	6.54E-08	5.13E-09	7.28E+14	0.24
6.39	3.47E-08	6.59E-09	6.30E+14	0.14
8.67	1.08E-07	1.23E-08	2.51E+14	0.33
10.51	8.16E-08	1.30E-08	4.03E+14	0.43
12.11	6.20E-08	1.38E-08	5.65E+14	0.48

(a)

Roughness (nm)	$\bar{R}$ (m)	σ_s (m)	η (/m ²)	$\bar{R}\sigma_s\eta$
0.52	1.12E-07	6.91E-10	6.91E+14	0.05
2.85	1.09E-07	2.45E-09	2.52E+14	0.07
4.01	6.14E-08	5.32E-09	4.28E+14	0.14
6.39	3.47E-08	6.59E-09	6.30E+14	0.14
8.67	1.16E-07	1.17E-08	8.00E+13	0.11
10.51	5.77E-08	1.12E-08	1.81E+14	0.12
12.11	6.55E-08	1.37E-08	1.34E+14	0.12

(b)

Roughness (nm)	$\bar{R}$ (m)	σ_s (m)	η (/m ²)	$\bar{R}\sigma_s\eta$
0.52	1.22E-07	6.72E-10	4.69E+13	0.004
2.85	1.11E-07	2.24E-09	1.24E+14	0.03
4.01	6.10E-08	5.46E-09	1.20E+14	0.04
6.39	3.50E-08	5.60E-09	2.13E+14	0.04
8.67	1.17E-07	1.16E-08	2.70E+13	0.04
10.51	5.54E-08	1.12E-08	1.29E+14	0.08
12.11	5.67E-08	1.34E-08	6.35E+13	0.05

(c)

The product of $\bar{R}$, σ_s and η were computed and was found that for any given summit definition criteria, the product value is fairly constant with some minor variations. However when comparisons are made against different summit definition criteria, the product of $\bar{R}$, σ_s and η would vary in different ranges. For instance, from Table 3-1(a), (b) and (c), the product varies with an average value of about 0.3, 0.1 and 0.05, respectively. Hence, the product of $\bar{R}$, σ_s and η very much depends on the

definitions that are being applied to the model and also on the resolution of the measuring equipment.

3.3 Summary

A surface roughness acquisition model, based on 3D surface profiles measured by the AFM, was developed to characterize the surface topology in terms of average summit radius curvature $\bar{R}$, standard deviation of asperity heights σ_s and asperity density η .

Our results concur with previous reports that the distribution of the asperity heights is approximately Gaussian in nature and the standard deviation of asperity heights is directly proportional to the measured rms surface roughness. It was also found that the radius curvature does not follow any form of distributions: it varies in a random manner with respect to the asperity heights in any given scan and the average radius curvature tends to decrease as surface roughness increases.

The asperity density also tends to decrease with increasing surface roughness. This does not imply that there are actually less asperities but rather there are many other smaller scale asperities created which may be beyond the detection capabilities of the algorithm and the AFM equipment.

It was also found that for any given surface, the product of $\bar{R}$, σ_s and η is relatively constant for any particular definition of an asperity summit. However, the range of constant is different, depending on which asperity summit definition is adopted in the modeling. The values of $\bar{R}$ and σ_s calculated are quite independent of the definitions of the asperity summit: it is η which will be significantly affected and

it is the changes in η that affected the value of the constant. The accuracy of $\bar{R}$, σ_s and η will depend very much on the methods of calculation and the resolution of the measuring equipment: A high resolution scanning of the surface is needed to reflect the actual surface profile.

Chapter 4 –True Contact Area Model

In this chapter, the bond quality is analyzed from the dicing yield perspective and a contact model will be developed to relate the dicing yield to the true contact area of the contacting surfaces. The first section explains how the dicing yield can be related to the true contact area and hence can be used as a measure of the bond quality. The second section shows the experimental results and the effects of different bonding parameters on the dicing yield. The third section details the development of a contact model to evaluate the true contact area, taking into account bonding parameters such as applied load, nominal contact area and surface topologies of the contacting wafers. The calculated true contact area is shown to correlate to the dicing yield. The last section uses the developed contact model to analyze the effects of different bonding or topological parameters to the bond quality. A hypothetical comparison of Au-Au and Cu-Cu bonding is also presented

4.1 Distribution of the True Contact Area

When two surfaces are pressed together, their initial contact spots are at the surface asperities over several discrete spots, due to surface roughness. As the deformation of surface asperities increases, these contact spots will grow larger and at the same time, new contact spots are created. As interfacial diffusion takes place, the areas in contact will be bonded. The various adhesion measurement techniques, as mentioned in chapter 1, assign a quantitative value to the bond strengths of these bonded interfaces. However, the bond strengths of interfaces should not solely be a measure of the bond quality but should also involve other considerations such as the

contact/bonding uniformity. While the adhesion measurement techniques provide a quantitative measure of the bond strength, they contain no information about the uniformity of the bonding.

The initial contact between contacting surfaces and growth of the contact spots were investigated by Uppal *et al.* [80], using the contact of a polished flat silver-steel surface and a rough aluminum surface. The plastic deformation of the asperity tips was observed by means of Nomarski interference contrast illumination in conjunction with a projection microscope, which revealed the surface contours and differentiated them with different colors. With this method, they were able to observe the growth of both new and existing contact spots, under increasing applied loads, through the change in surface contours. Figures 4-1 (a)-(i) are extracts from [80], illustrating the initiation and growth of contact spots under increasing applied loads from 0.133 to 53.5 N on a circular nominal contact area of $0.5 \mu\text{m}^2$.

Assume that for two pairs of bonded wafers, one poorly bonded and the other well bonded, have the contact area distribution profiles as shown in Figures 4-2(a) and 4-2(b), respectively. If the bonded wafers are diced into rectangular samples along the dotted lines as shown in Figures 4-2, samples with little or no contact spots will be delaminated when diced. For the poorly bonded wafers, the majority of the samples will delaminate during the dice-saw process whereas in the case of the well bonded wafers, the majority of the samples should remain intact. Hence, the dicing yield of the bonded wafers is related to the true contact area between the surfaces and can be used as a semi-quantitative indicator of the bonding quality.

Figure 4-1: Micro-contact spots under increasing loads from (a) to (i). The dotted circle is the periphery of the nominal contact area and the darkened areas are craters which are surrounded by areas of contact [80].

Figure 4-2: Assumed contact profiles of a (a) poor bonded wafers and (b) well bonded wafers [80].

Recall in chapter 2, the four-point bend test results have significant scattering and Tsau [41] attributed it to the effects of plasticity and bonding non-uniformity. Consider the contact area distribution profile of Figure 4-2(b), it can be seen that the amount of true contact area varies from sample to sample. The bond strength of a bonded interface can be assumed to be related to the amount of true contact area between the contacting surfaces. Hence, the larger is the true contact area, the more force will be required to debond the samples and vice-versa. The large scatter of bond strength measurements, as observed from the four-point bend test measurements in Chapter 2, could simply reflect the *macroscopic* non-uniform distribution of the true contact area of the bonded surfaces between individual samples.

4.2 Effects of Different Bonding Parameters on Dicing Yield

Table I summarizes the experiments in which pairs of wafers were bonded with different configurations (blanket-blanket and blanket-patterned) at different applied loads, nominal contact areas, surface roughnesses, and their corresponding dicing yields. All the wafers were bonded at 300°C and this section analyzes the effects of load, nominal contact pressure and surface roughness on the dicing yield.

Figure 4-3 shows the dicing yield of bonded wafers versus different applied loads and nominal contact pressures. Wafers with blanket Cu films were used in these experiments and the wafers were obtained after the Cu CMP process (blanket (CMP)). This ensures that the wafer surfaces have smooth and identical surface roughnesses. Hence, any changes in the dicing yield will be solely due to the effects of the applied load / nominal contact pressure.

Table 4-1: Details of each wafer bonding experiments and the corresponding dicing yield. All wafers were bonded at 300°C.

	Load (N)	Nominal Contact Area (m ²)	RMS Roughness (m ²)	Dicing Yield (%)
Blanket (CMP) - Blanket (CMP)	400	0.03142	1.05, 1.05	11.1
	1000	0.03142	1.05, 1.05	31.9
	2000	0.03142	1.05, 1.05	66.0
	2550	0.03142	1.05, 1.05	76.4
	4550	0.03142	1.05, 1.05	97.2
	8000	0.03142	1.05, 1.05	100.0
	10000	0.03142	1.05, 1.05	100.0
Blanket (CMP) - Blanket (CMP and Etched)	4550	0.03142	1.05, 5.83	81.9
	4550	0.03142	1.05, 2.34	91.0
	10000	0.03142	1.05, 5.83	91.8
Blanket (CMP and Etched) - Blanket (CMP and Etched)	2000	0.03142	4.38, 4.90	1.4
	2000	0.03142	2.50, 1.50	16.7
	4550	0.03142	4.14, 5.53	72.9
Blanket (as deposited) - Blanket (CMP and Etched)	4550	0.03142	5.92, 8.42	66.7
	4550	0.03142	5.92, 2.93	72.2
Blanket (as deposited) - Blanket (as deposited)	6000	0.03142	5.92, 5.92	95.1
Blanket (CMP) - Patterned (50 um)	4550	0.01271	1.05, 8.63	42.4
	10000	0.01271	1.05, 8.63	87.5
Blanket (CMP) - Patterned (175 um)	4550	0.00578	1.05, 10.00	36.0
	10000	0.00578	1.05, 10.00	68.4
Blanket (CMP) - Patterned (310 um)	4550	0.00400	1.05, 11.95	25.6
	10000	0.00400	1.05, 11.95	57.0

Figure 4-3: Plot of dicing yield versus applied loads and nominal contact pressures.

Results suggest that the dicing yield is proportional to the applied load and nominal contact pressure, up to 100% yield. This is expected as a larger applied load would have resulted in larger deformations of the surface asperities, and hence the true contact area and the dicing yield would increase.

The effects of surface roughness on the bond quality of bonded interfaces were also investigated by bonding Cu blanket wafers of different surface roughness with applied loads of 2 kN and 4.55 kN. The dicing yield of the bonded wafers versus the total surface roughness are plotted as shown in Figure 4-4, where the total surface roughness is the sum of surface roughness of the two Cu surfaces before bonding.

Figure 4-4: Plot of dicing yield versus total surface roughness.

For a given applied load, the dicing yield decreases as the total surface roughness increases. Hence, it is evident that bonding of wafers with rough surfaces is detrimental to obtaining good bonding. A more detailed quantitative explanation will be given in sections 4.3 and 4.4 about the effects of surface roughness on the true

contact area. For now, it will be suffice to say that the contact of rough surfaces would result in smaller contact area as compared to that of smooth surfaces, and hence the bonding will be poor, as reflected in the dicing yield.

Results from Figures 2-7 and 4-3 show that, with all other conditions constant such as the surface roughness, an increase in applied load/nominal contact pressure will enhance the bond strength/bond quality. Conversely, Figure 4-4 shows that the bond quality (dicing yield) will decrease as the surface roughness increases. However, the above analyses consider the variation of only a single parameter: applied load or surface roughness. If all the wafer bonding experiments with different applied loads, nominal contact areas and surface roughnesses are to be considered together, the analysis would be more complicated as shown in the plots of dicing yield versus applied load (Figure 4-5(a)) and nominal contact pressure (Figure 4-5(b)).

Figure 4-5: Plots of dicing yield versus (a) applied load and (b) nominal contact pressure for different nominal contact areas and surface roughness.

Results from Figure 4-5 show that the dicing yields have no systematic trends with either the applied load or the nominal contact pressure. Hence, by considering the applied load or nominal contact by itself or together, is inadequate to characterize

the bond quality of the bonded interfaces. Both the applied load and the surface roughness affect the bond quality of the bonded wafers and their effects on the bond quality will have to be analyzed concurrently.

In contact mechanics, the true contact area between two contacting surfaces can be calculated as a function of the applied load, nominal contact area and surface topology characteristics (surface roughness). As mentioned in section 4.1, the true contact area can be related to the dicing yield and used as a measure of the bond quality. Knowing the applied load, nominal contact area and surface roughness parameters, the true contact area can be calculated using a suitable contact model, as developed in the next section.

4.3 Contact of Surfaces

4.3.1 Introduction

In section 1.5, an overview has been presented about contact theories and as mentioned, there is yet a model that considers work hardening effects. In this thesis, a contact model which takes into account of work hardening effects will be developed, as shown in the next section. The developed model will be used to calculate the true contact area, taking into account the different bonding parameters and surface roughness, and is correlated to the experimental results as described in the previous section.

4.3.2 Contact Modeling

Contact of a Single Asperity with a Rigid Smooth Surface

Consider the contact between one asperity and a rigid smooth flat, as shown in Figure 4-6. Let z and d represent the asperity height and the mean separation between the two surfaces, respectively. Then the contact interference can be expressed as

$$w = z - d \quad (4.1)$$

Figure 4-6: A rough surface having asperities in contact with a flat plane. The separation is d when the load is L .

The interference w is an important variable that measures the extent of the asperity deformation. The contact pressure and contact area of the asperity are fully determined by this interference. The asperity generally experiences two deformation stages as w increases: elastic and onset (and beyond) of plastic yielding. The following sections develop formulations for the mean contact pressure and the true contact area of the asperity as a function of w and surface roughness parameters as described in the previous section.

(1) *Elastic Contact*. The asperity deforms elastically when w is sufficiently small. According to Hertz theory for the elastic contact of a flat surface with a sphere of radius R , the contact area A_e , contact load L_e , and maximum contact pressure $p_{m,e}$ of the asperity can be expressed in terms of w as

$$A_e(w) = \pi R w \quad (4.2)$$

$$L_e(w) = \frac{4}{3} E R^{1/2} w^{3/2} \quad (4.3)$$

$$p_{m,e} = \frac{3}{2} \frac{L_e(w)}{A_e(w)} = \frac{2E}{\pi} \left(\frac{w}{R} \right)^{1/2} \quad (4.4)$$

where

$$\frac{1}{E} = \frac{1-\nu_1}{E_1} + \frac{1-\nu_2}{E_2} \quad (4.5)$$

and

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}, \quad (4.6)$$

and where E_1 , E_2 , ν_1 , ν_2 , R_1 and R_2 are the Young's moduli, Poisson's ratio, and radius of curvature of the two contacting surfaces, respectively.

Tabor [68] had shown that the initial yielding of an asperity occurs when $p_{m,e} = 1.1Y$, where Y is the yield stress of the deformed material. Substituting $p_{m,e} = 1.1Y$ into equation (4.4), the critical interference w_c can then be expressed as

$$w_c = \left(\frac{\pi p_{m,e}}{2E} \right)^2 R = \left(\frac{1.1\pi Y}{2E} \right)^2 R. \quad (4.7)$$

Hence when $w < w_c$, the contact is elastic and the deformation behavior can be described by Hertz's theory. When $w > w_c$ and plastic deformation starts to set in, the

contact load and contact area dependence on w must be determined by a different model.

(2) *Onset of Plastic Yielding.* Chang *et al.* [64], in their volume conservation model, showed that the contact area for a plastically deformed hemisphere A_p , could be expressed as

$$A_p(w) = \pi a^2 = \pi R w \left(2 - \frac{w_c}{w} \right) = 2\pi R w - \pi R w_c \quad (4.8)$$

where a is the radius of the circular area of contact. For metals like Cu, the yield stress is typically taken at a yield strain of about 0.4% and hence the ratio of Y/E is about 0.004. From Figure 3-7 in the previous chapter, R is typically in the order of 10^{-8} m. Hence from equation (4.7), w_c will be in the order of 10^{-12} m, which would be so small such that w can be assumed to be much larger than w_c . Equation (4.8) can then be approximated as

$$A_p(w) = \pi a^2 \approx 2\pi R w. \quad (4.9)$$

Assuming the contacting asperities to be in fully plastic deformation mode, Tabor [68] showed that the mean contact pressure p_m , would be approximately three times the yield stress and can be expressed as

$$p_m = \frac{L_p(w)}{A_p(w)} = \frac{L_p(w)}{\pi a^2} = 3Y \quad (4.10)$$

The contact load $L_p(w)$ can then be expressed as

$$L_p(w) = 3\pi a^2 Y \quad (4.11)$$

For spherical indentation (rigid sphere-deformable flat surface), Tabor also showed that for work hardening materials beyond the onset of plastic deformation, Y can be replaced by a representative flow stress Y_r , which in turn is related to some representative strain ε_r . This representative strain increases continuously with the indentation size and can be approximated as [68]

$$\varepsilon_r = 0.2 \frac{a}{R}. \quad (4.12)$$

It is assumed that in the case of a deformable sphere against a rigid flat surface, the strain can also be approximated by equation (4.12). This assumption is reasonable as the strain is primarily a function of deformation: in both the case of a rigid sphere on a flat deformable surface and for a deformable sphere on a rigid flat surface, the amount of deformation is approximately the same for a similar spherical radius R and the radius of the circle of contact a ($a \ll R$) (Figures 4-7(a) and 4-7(b)). The same assumption had also been made by Tabor [68] in his analyses.

Figure 4-7: (a) A rigid sphere against a flat plane and (b) a deformed spherical asperity against a flat rigid plane, both with contact radius a and radius curvature R .

For ductile materials, the stress-strain curves are often well described by the power law relationship as proposed by Nadai [81,82]

$$Y_r = Y \left(\frac{\varepsilon_r}{\varepsilon} \right)^x, \quad (4.13)$$

where ε and x are the yield strain and strain hardening index of the material, respectively. This allows calculation of the flow stress Y_r , at some higher strain ε_r , from the values at the onset of yielding, provided the strain hardening index is known.

Replacing Y by Y_r and with equation (4.12), equation (4.11) can be rewritten as

$$L_p(w) = 3\pi a^2 Y \left(\frac{0.2a}{R\varepsilon} \right)^x \quad (4.14)$$

By substituting $a = \sqrt{2Rw}$ from equation (4.9), equation (4.14) can be expanded further as

$$L_p(w) = 3\pi(2)^{1+\frac{x}{2}} R^{1-\frac{x}{2}} w^{1+\frac{x}{2}} Y \left(\frac{0.2}{\varepsilon} \right)^x \quad (4.15)$$

Contact of Two Rough Surfaces

The contact between two rough surfaces can be modeled by an equivalent rough surface in contact with a perfectly flat plane, as mentioned previously, using the parameters

$$\sigma_s = \sqrt{\sigma_{s1}^2 + \sigma_{s2}^2}, \quad (4.16)$$

where σ_{s1} and σ_{s2} are the standard deviations of summit heights for the two contact surfaces and σ_s is the parameter for an equivalent rough surface in contact with a flat plane [83,84]. When this equivalent rough surface is pressed against the flat plane until the mean planes are separated by a distance d , the probability that a randomly selected summit will be in contact is

$$P(\text{Contact}) = P(z > d) = \int_d^{\infty} f(z) dz, \quad (4.17)$$

where $f(z)$ is the Gaussian probability density function for the height distribution of the asperities.

For a given surface with nominal area A_n and density of asperities η , the total number of asperities on the surface will be

$$N = \eta A_n, \quad (4.18)$$

and hence the expected number of asperities in contact N_c , will be

$$N_c = N \int_d^{\infty} f(z) dz = \eta A_n \int_d^{\infty} f(z) dz. \quad (4.19)$$

For a particular asperity, the dependence of the true contact area $A(w)$, and load $L(w)$, on w is determined by the asperity deformation model and once these expressions are known, the expected contact area and expected contact load of each individual asperity can be calculated as [57]

$$A' = \int_d^{\infty} A(z - d) f(z) dz \quad (4.20)$$

and

$$L' = \int_d^\infty L(z-d)f(z)dz, \quad (4.21)$$

respectively.

The total true contact area A_t and total load L_t over the whole surface are obtained by summing the individual asperity contributions as

$$A_t = \eta A_n \int_d^\infty A(z-d)f(z)dz \quad (4.22)$$

and

$$L_t = \eta A_n \int_d^\infty L(z-d)f(z)dz. \quad (4.23)$$

These individual contributions may be calculated by using equations (4.2) and (4.3) when $w < w_c$, and equations (4.9) and (4.15) when $w \geq w_c$. Then, the total true contact area and load can be expressed as

$$A_t = A_e(w) + A_p(w) \quad (4.24)$$

and

$$L_t = L_e(w) + L_p(w), \quad (4.25)$$

respectively, where

$$A_e(w) = \eta A_n \pi R \int_d^{d+w_c} (z-d)f(z)dz, \quad (4.26)$$

$$A_p(w) = 2\eta A_n \pi R \int_{d+w_c}^\infty (z-d)f(z)dz, \quad (4.27)$$

$$L_e(w) = \frac{4}{3} \eta A_n E R^{1/2} \int_d^{d+w_c} (z-d)f(z)dz, \quad (4.28)$$

and

$$L_p(w) = 3(2)^{1+\frac{x}{2}} \eta A_n R^{1-\frac{x}{2}} Y\left(\frac{0.2}{\varepsilon}\right)^x \int_{d+w_c}^{\infty} (z-d)^{1+\frac{x}{2}} f(z) dz. \quad (4.29)$$

It will be subsequently shown that $d \gg w_c$, in which case, equations (4.24) and (4.25) can be rewritten as

$$A_t \approx A_p(w) = 2\eta A_n \pi \bar{R} \int_d^{\infty} (z-d) f(z) dz \quad (4.30)$$

and

$$L_t \approx L_p(w) = 3(2)^{1+\frac{x}{2}} \eta \pi \bar{R}^{1-\frac{x}{2}} Y\left(\frac{0.2}{\varepsilon}\right)^x \int_d^{\infty} (z-d)^{1+\frac{x}{2}} f(z) dz, \quad (4.31)$$

where

$$\frac{1}{\bar{R}} = \frac{1}{\bar{R}_1} + \frac{1}{\bar{R}_2}.$$

The average curvature radius $\bar{R}$ replaces R to take into account that the asperities on a surface do not have identical radii of curvature, as highlighted previously.

It is convenient to introduce standardized variables, and normalized the asperity heights by its standard deviation σ_s [57]. This gives

$$\begin{aligned} A_t &= 2\eta A_n \pi \bar{R} \sigma_s \int_{d/\sigma_s}^{\infty} \frac{z-d}{\sigma_s} \phi(z) dz \text{ or} \\ A_t &= 2\eta A_n \pi \bar{R} \sigma_s F_1\left(\frac{d}{\sigma_s}\right) \end{aligned} \quad (4.32)$$

and

$$\begin{aligned} L_t &= 3(2)^{1+\frac{x}{2}} \eta A_n \pi \bar{R}^{1-\frac{x}{2}} Y\left(\frac{0.2}{\varepsilon}\right)^x \sigma_s^{1+\frac{x}{2}} \int_{d/\sigma_s}^{\infty} \left(\frac{z-d}{\sigma_s}\right)^{1+\frac{x}{2}} \phi(z) dz \text{ or} \\ L_t &= 3(2)^{1+\frac{x}{2}} \eta A_n \pi \bar{R}^{1-\frac{x}{2}} Y\left(\frac{0.2}{\varepsilon}\right)^x \sigma_s^{1+\frac{x}{2}} F_2\left(\frac{d}{\sigma_s}\right), \end{aligned} \quad (4.33)$$

where F_1 and F_2 represents the integral functions in equations (4.32) and (4.33), respectively and $\phi(z)$, the standardized Gaussian distribution, is the height distribution scaled to make its standard deviation unity. The integrals F_1 and F_2 can be estimated using Lagrange's interpolation of parabolic cylindrical functions as explained in [85,86]. With the values of L_t , A_n and the surface roughness parameters known, a value for d/σ_s is obtained such that the equality of equation (4.33) is satisfied. This value is subsequently substituted back into equation (4.32) to estimate the true contact area A_t .

4.3.3 Application of Contact Model to Experiments

The details of the various wafer bonding experiments and their corresponding surface roughness characterizations, calculated true contact area and the dicing yield are summarized in Table 4-2. The values of the applied load L , nominal area A_n , $\bar{R}$, σ_s and η in each experiment were input into equations (4.32) and (4.33) to estimate the corresponding true contact area. Notice from Table 4-2 that the calculated values of d are much larger than w_c , which as discussed earlier is of the order of 10^{-12} m, thus justifying the earlier assumption for equations (4.30) and (4.31).

Table 4-2: Calculations of topological parameters and true contact area for each wafer bonding experiment and the corresponding dicing yield. All wafers were bonded at 300°C.

	Load (N)	Nominal Contact Area (m ²)	$\bar{R}$ (nm)	σ_s (nm)	η (/m ²)	d (nm)	True Contact Area (m ²)	Dicing Yield (%)
Blanket (CMP) - Blanket (CMP)	400	0.03142	47.85	1.13	4.42x10 ¹⁴	3.03	8.98E-07	11.1
	1000	0.03142	47.85	1.13	4.42x10 ¹⁴	2.76	2.21E-06	31.9
	2000	0.03142	47.85	1.13	4.42x10 ¹⁴	2.55	4.16E-06	66.0
	2550	0.03142	47.85	1.13	4.42x10 ¹⁴	2.48	5.58E-06	76.4
	4550	0.03142	47.85	1.13	4.42x10 ¹⁴	2.29	9.77E-06	97.2
	8000	0.03142	47.85	1.13	4.42x10 ¹⁴	2.10	1.70E-05	100.0
	10000	0.03142	47.85	1.13	4.42x10 ¹⁴	2.01	2.13E-05	100.0
Blanket (CMP) - Blanket (CMP and Etched)	4550	0.03142	32.06	5.89	3.59x10 ¹⁴	2.66	6.05E-06	81.9
	4550	0.03142	40.09	2.27	6.30x10 ¹⁴	2.66	7.44E-06	91.0
	10000	0.03142	32.06	5.89	3.59x10 ¹⁴	2.43	1.30E-05	91.7
Blanket (CMP and Etched) - Blanket (CMP and Etched)	2000	0.03142	21.12	5.70	4.54x10 ¹⁴	2.94	2.39E-06	1.4
	2000	0.03142	32.31	2.50	4.50x10 ¹⁴	2.86	2.82E-06	16.7
	4550	0.03142	30.48	3.32	6.57x10 ¹⁴	2.86	5.75E-06	72.9
Blanket (as deposited) - Blanket (CMP and Etched)	4550	0.03142	68.97	8.86	1.28x10 ¹⁴	2.80	6.12E-06	66.7
	4550	0.03142	58.49	6.82	1.28x10 ¹⁴	2.63	5.89E-06	72.2
Blanket (as deposited) - Blanket (as deposited)	6000	0.03142	83.00	6.82	1.28x10 ¹⁴	2.76	8.72E-06	95.1
Blanket (CMP) - Patterned (50 um)	4550	0.01271	30.88	10.2	2.68x10 ¹⁴	2.50	5.08E-06	42.4
	10000	0.01271	30.88	10.2	2.68x10 ¹⁴	2.23	1.15E-05	87.5
Blanket (CMP) - Patterned (175 um)	4550	0.00578	36.11	11.5	2.09x10 ¹⁴	2.23	5.11E-06	36.0
	10000	0.00578	36.11	11.5	2.09x10 ¹⁴	1.95	1.12E-05	68.4
Blanket (CMP) - Patterned (310 um)	4550	0.00400	37.37	15.4	1.35x10 ¹⁴	2.15	4.69E-06	25.6
	10000	0.00400	37.37	15.4	1.35x10 ¹⁴	1.90	1.02E-05	57.0

Calculation of the true contact area using equations (4.32) and (4.33) requires the estimation of yield stress Y at the bonding temperature, yield strain ϵ , and strain hardening factor x . One way to obtain the yield stress is to find the stress as a function of temperature, which can be done with a curvature measurement technique. Curvature measurements during thermal cycling of a blanket 0.6 μm -thick Cu film on oxidized Si wafers was carried out as a function of temperature and stresses were calculated using Stoney's equation [87] as shown in Figure 4.8.

Figure 4-8: Stress-Temperature curves for unpassivated copper films, 0.6 and 1.0 μm thick, from [87].

Result from Figure 4-8 shows that the stress, at 300°C, is estimated to be about 50 MPa. The yield strain is taken to be at 0.4% and the strain hardening index (compression) is estimated to be about 0.5 [88]. Together with the characterized surface roughness parameters $\bar{R}$, σ_s and η as shown in Table 4.2, the true contact area of each experiment can then be calculated.

Bonding of Blanket-Blanket Cu Wafers

Figure 4-9 shows a plot of dicing yield versus the calculated true contact area of bonded blanket Cu wafers with various degrees of surface roughness and applied loads. It can be seen that the dicing yield is now directly proportional to the true contact area, up to a 100% dicing yield, showing that the true contact area captures the best correlation between bond integrity, nominal bonding area, and the applied force.

Figure 4-9: Graph of Dicing yield vs True contact area (for blanket Cu wafers). The dotted line illustrates an approximate best fit.

From equation (4.10), the model assumes the contacting asperities are in fully plastic deformation mode and an attempt is being made here to validate this assumption. A blanket Cu wafer of initial average rms surface roughness of about 3.11 nm was thermo-compressed with increasing applied loads up to 4000 N at 300°C and the surface roughness after each compression was measured. Figure 4-10 shows the plot of surface roughness after each compression versus the corresponding applied load. The change in surface roughness after the compression suggests that while there is significant overlap of data, the average surface roughness after each compression decreases. This implies that the thermo-compression could have resulted in the flattening of the surface asperities and permanent plastic deformations were involved.

Figure 4-10: Graph of rms surface roughness after each compression versus the applied load.

The true contact area, under different bonding conditions as tabulated in Table 4-3, is on the order of 10^{-5} to 10^{-3} times of the nominal contact area, which is an extremely small fraction. To illustrate the validity of these ratios, assume that the true-to-nominal contact area ratio is only 1%. The corresponding contact pressures would range from as low as 1.3 MPa to about 250 MPa, so that in most cases, the contact pressure is much lower than the stress required for plastic deformation (50 MPa). This would be inconsistent with what was observed on the Cu surfaces from the SEM images (Figure 4-11). However, using the calculated true contact areas, the contact pressures would range from about 450 MPa to 900 MPa, so that all asperities would be expected to plastically deform. Hence, a true-to-nominal contact ratio of such low proportions (10^{-5} to 10^{-3}) is reasonable.

Table 4-3: Nominal and true contact pressures for each bonding experiments.

	Load (N)	Nominal Contact Area (m ²)	True Contact Area (m ²)	Nominal Contact Pressure (MPa)	True Contact Pressure (MPa)
Blanket (CMP) - Blanket (CMP)	400	0.03142	8.98E-07	0.01	445.4
	1000	0.03142	2.21E-06	0.03	452.5
	2000	0.03142	4.16E-06	0.06	480.8
	2550	0.03142	5.58E-06	0.08	457.0
	4550	0.03142	9.77E-06	0.14	465.7
	8000	0.03142	1.70E-05	0.25	470.6
	10000	0.03142	2.13E-05	0.32	469.5
Blanket (CMP) - Blanket (CMP and Etched)	4550	0.03142	6.05E-06	0.14	752.1
	4550	0.03142	7.44E-06	0.14	611.6
	10000	0.03142	1.30E-05	0.32	769.2
Blanket (CMP and Etched) - Blanket (CMP and Etched)	2000	0.03142	2.39E-06	0.06	836.8
	2000	0.03142	2.82E-06	0.06	709.2
	4550	0.03142	5.75E-06	0.14	791.3
Blanket (as deposited) - Blanket (CMP and Etched)	4550	0.03142	6.12E-06	0.14	743.5
Blanket (as deposited) - Blanket (as deposited)	4550	0.03142	5.89E-06	0.14	772.5
Blanket (as deposited) - Blanket (as deposited)	6000	0.03142	8.72E-06	0.19	688.1
Blanket (CMP) - Patterned (50 μ m)	4550	0.01271	5.08E-06	0.36	895.7
	10000	0.01271	1.15E-05	0.79	869.6
Blanket (CMP) - Patterned (175 μ m)	4550	0.00578	5.11E-06	0.79	890.4
	10000	0.00578	1.12E-05	1.73	892.9
Blanket (CMP) - Patterned (310 μ m)	4550	0.00400	4.69E-06	1.14	970.1
	10000	0.00400	1.02E-05	2.50	980.4

Figure 4-11: SEM image of a blanket Cu film that was bonded to a patterned Cu wafer.

Bonding of Blanket-Patterned Cu Wafers

Experiments were also conducted on bonding of blanket and patterned Cu films. If the dicing yields from the bonding of patterned wafers (blanket-patterned) are also considered in a plot like Figure 4-9 (as shown in Figure 4-12), significant deviations from the trend-curve are observed. The data points circled in Figure 4-12 represent results from the blanket-patterned bonding experiments, and indicate that the model over-estimates the true contact area. Figure 2-18 shows large-scale bonding non-uniformity that was common for blanket-patterned bonds, but not for blanket-blanket bonds. Surface profilometry on post-CMP patterned films also indicated significant ‘dishing effects’ in the patterned films, with the centers of lines being as much as 50 nm below the surface of line edges. Figure 2-17 shows a de-bond surface of a line, with topography that suggests that plastic deformation occurred at the line edges, but not at the line center, as might be expected in the presence of dishing effects. These results illustrate that there are larger scale variations in bonding and surface flatness that can also affect the true contact area and dicing yield, beyond that effects of micro- and nano-scale roughness captured by the model presented here.

Figure 4-12: Graph of Dicing yield vs True contact area (for both blanket-blanket and blanket-patterned bonded wafers). The circled points are those of blanket-patterned wafers bonding.

4.4 Effects of Bonding and Surface Parameters on Bonding

We have now developed a model, backed with experimental data, which suggests that the dicing yield is proportional to the true contact area and can be used as an indicator of the bond quality between the bonded interfaces. The developed model could be used to investigate the effects of individual bonding, topological parameters and material properties on the bond quality, through changes in the true contact area. This section would use the model and discuss the effects of the applied load, nominal contact area, and surface topological parameters in terms of $\bar{R}$, σ_s and η_s .

Effects of applied load and nominal contact area

From equations (4.32) and (4.33), the relations between applied load L , nominal contact area A_n and true contact area A_t can be determined. For analysis

purposes, the topological parameters of $\bar{R}$, σ_s and η are assigned values of 47.85 nm, 1.13 nm and 4.42×10^{14} respectively. These are measured values of blanket Cu surfaces after the Cu CMP process as shown in Table 4-2. The yield stress Y and strain hardening index x are taken as 50 MPa and 0.5 respectively.

The relationship between the true contact area and applied load, for different nominal contact areas, is shown in Figure 4-13. Results indicate that for a given nominal area, the true contact area is almost directly proportional to the applied load. However, there is only a very small difference between the three curves, for nominal contact areas of 0.03142 m^2 (blanket), 0.02 m^2 (2/3 of blanket area) and 0.01 m^2 (1/3 of blanket area). Hence, the true contact area depends mainly on the applied loads and not on the nominal contact area (or nominal contact pressure).

Figure 4-13: Relationship between the true contact area and load for different nominal contact areas.

As the mean separation d between the surfaces increases / decreases, the integral functions $F_1(d/\sigma_s)$ and $F_2(d/\sigma_s)$ from equations (4.32) and (4.33) will decrease

/ increase, respectively. Consider equation (4.33) at a given load L . As the nominal contact area A_n increases, $F_2(d/\sigma_s)$ will decrease and that implies the value of d and hence the value of $F_1(d/\sigma_s)$ in equation (4.32) will increase and decrease, respectively. However, the decrease in $F_1(d/\sigma_s)$ will be offset by an increase in A_n and hence the overall dependence of the true contact area on the nominal contact area is then calculated to be very small. By using equation (4.33), equation (4.32) can be re-expressed as

$$A_t \approx \frac{L_t \bar{R}^{\frac{x}{2}}}{3(2)^{\frac{x}{2}} Y \left(\frac{0.2}{\varepsilon} \right)^x \sigma_s^{\frac{x}{2}}} \frac{F_1(d/\sigma_s)}{F_2(d/\sigma_s)} \quad (4.34)$$

Equation (4.34) suggests that the true contact area depends mainly on the applied load, surface topological parameters $\bar{R}$ and σ_s , and material properties Y and x . A more qualitative explanation on the independence of the true contact area on nominal contact area can be explained by considering a given load applied on a large nominal area and the resulting true contact area be A_1 . For the same load applied to a smaller nominal contact area, the number of initial contacting asperities will be smaller. However, the contacting pressure and corresponding deformations will be higher (larger contact area per asperity), such that the asperities with lower summit heights will also be in contact. The overall effect would be the true contact area, in this case be A_2 , can be quite close to A_1 , as shown in Figure 4-14.

Figure 4-14: An applied load on a (a) larger nominal contact area and on a (b) smaller nominal contact area. The applied pressure in (b) will be larger than in (a) and hence, more deformation. The total contact area (bold lines) A_1 and A_2 could be quite similar.

Effects of surface topological parameters

From equation (4.34), the true contact area is proportional to $\bar{R}^{\frac{x}{2}}$, inversely proportional to $\sigma_s^{\frac{x}{2}}$ and is quite independent of η , like the nominal contact area. With all other parameters kept constant, the true contact area can be maximized for the largest possible $\bar{R}$ and the smallest possible σ_s . From Figures 3-6 and 3-8, this could only be achieved when the surface roughness is very small: any increase in the surface roughness will result in a decrease in $\bar{R}$ and an increase in σ_s , causing a

reduction in the true contact area, as reflected in a decrease of dicing yield as shown in Figure 4-4.

As shown in Figures 3-6 and 3-8, as the surface roughness increases, the corresponding changes in $\bar{R}$ and σ_S are not the same. For instance, an increase of surface roughness from 1 to 10 nm will result in $\bar{R}$ decreasing by about 50% from 100 nm to 50 nm while σ_S (which is mathematically related to rms surface roughness) will increase proportionally with surface roughness by 10 times from about 1 nm to 10 nm. Hence for copper where $x \sim 0.5$, $\bar{R}^{\frac{x}{2}}$ would have decreased by a factor of 0.84 while $1/\sigma_S^{\frac{x}{2}}$ would have decreased by a factor of 0.56. Hence, σ_S can be seen as the dominant surface topological parameter affecting the true contact area. Any analysis on the effects of surface roughness on the true contact area can then be simplified by considering the variations in σ_S alone.

From Figure 4-9, we observe that an approximate true contact area of $3 \times 10^{-6} \text{ m}^2$ is the threshold for zero dicing yield, while to achieve 100% yield, a true contact area of $1 \times 10^{-5} \text{ m}^2$ or above is required. Using the model, a more in-depth analysis of the effects of bonding and topological parameters on the bond quality can be made. A bonding map of the applied load and topological parameters can be established to approximate the conditions that are required to achieve the desired true contact area, and thus the desired dicing yield. As mentioned earlier, the surface roughness can be characterized in terms of $\bar{R}$, σ_S and η . The standard deviation for asperity heights σ_S scales directly with the (rms) surface roughness, which is expected as they are

mathematically related. Hence, the effects of different surface roughness, for a given load, can be simulated by varying the value of σ_S .

Figure 4-15(a) shows a bonding map of applied load and σ_S for true contact areas from $3 \times 10^{-6} \text{ m}^2$ to $1 \times 10^{-5} \text{ m}^2$, which represent the boundary conditions for zero and 100% dicing yields, respectively, for blanket Cu films bonded at 300°C . However, the calculations are based on a blanket 8" wafer size, which makes it difficult to infer to other nominal areas involving different wafer sizes and/or pattern densities. Hence, an alternative representation would be to normalize the true contact area with respect to the wafer area, as shown in Figure 4-15(b) and the normalized values of 3.18×10^{-4} and 9.55×10^{-5} would correspond to a dicing yield of 100% and zero yields, respectively.

It is postulated that these normalized values (proportion of true contact to wafer area) to be proportional to the dicing yield. Hence even if the die sizes were to vary, as long as this ratio is maintained, we could attain a similar level of dicing yield. The underlying assumption is the bonding has to be uniform across the bonding surfaces, which we can expect it to be true for blanket wafers. However, the die size could not be infinitely small and a very small die would have a very minute amount of true contact area such that it is not sufficient to withstand the die-saw process.

Figure 4-15: Effects of surface roughness and applied load on the (a) true contact area and (b) normalized values of true contact to wafer area.

Calculations were also made for wafer areas of 0.01571 m^2 and 0.07068 m^2 , representing 6" and 12" Cu blanket wafers as shown in Figures 4-16(a) and 4-16(b), respectively.

Figure 4-16: Effects of surface roughness and applied load on the normalized values of true contact to wafer area for (a) 6" blanket and (b) 12" blanket wafers.

Results suggest that to achieve the same contact ratio, a smaller load is required for a smaller Cu blanket wafer (i.e. a smaller nominal contact area) than one with a larger size. This is reasonable as for a given load, the contact pressure will be higher on a small wafer than a larger wafer. These results could be further inferred to the bonding of very small nominal areas such as a blanket Cu die to a blanket Cu wafer or die.

The changes in nominal contact areas in Figure 4-16 can also be interpreted as changes in the patterned areas or densities in the die. For instance, Figure 4-16(a) can be seen as a patterned 8" wafer with nominal contact area 0.015 m^2 (compared to Figure 4-15, which the 8" wafer size is the equivalent to the nominal contact area 0.03142 m^2). The analysis can be simplified as in the case of blanket wafers if it is also assumed that the normalized values of true contact to wafer areas can be correlated to the dicing yields. In this case, the dicing yields for different die sizes or pattern densities can be theoretically estimated.

In reality, process issues such as polishing non-uniformity and dishing effects on the Cu patterned structures will result in bonding non-uniformity such that any experimental results will be expected to deviate from theoretical predictions. Furthermore, the layout, such as spacing, of the structures (as revealed by Tezzaron [72]) is critical to the mechanical strength of the bonded interfaces. These process and design considerations are not taken into account in this model and any predictions would be purely hypothetical.

Nonetheless, these bonding maps allow an estimation of the load required to attain a certain level of bond quality, as a function of the surface roughness. The load

required to obtain a 100% dicing yield can also be determined as a function of σ_s . Figure 4-15 shows that for a given applied load, the impact of surface roughness on the true contact area decreases for increasing surface roughness. For instance, at 3 kN applied load, the true contact area decreases from 1×10^{-5} to $9 \times 10^{-6} \text{ m}^2$ as σ_s increases from 0.5 to 0.75 nm. However, σ_s has to increase from 0.75 to 1.35 nm to decrease the true contact area by another $1 \times 10^{-6} \text{ m}^2$. This concurs with the strain hardening characteristics. These results also imply that for a sufficiently high applied load, it is not too critical if the surface is very rough: one can still obtain a sufficiently large true contact area for 100% dicing yield.

From a slightly different perspective, the bonding maps in Figure 4-15 also allow one to estimate, for a given applied load, what is the minimum wafer surface roughness required in order to attain the maximum dicing yield. For instance, consider an applied load of 5 kN and to ensure 100% dicing yield, the value of σ_s must not be more than $6 \times 10^{-9} \text{ nm}$. This is important from a process development perspective as it sets a ‘boundary condition’ of how rough the surface can be. On the other hand from a manufacturing perspective, it would be useful to approximate, for a known applied load and measured wafer surface roughness, the dicing yield to expect.

4.5 Au-Au Bonding in 3D ICs

As mentioned in section 1.2.5, Cu-Cu wafer bonding is an attractive option for the realization of 3D ICs because Cu can act as both the interconnect and the bonding material. Furthermore, the Cu damascene processes as a production platform are well established and it should require minimum effort from a process perspective.

However, the developed contact model shows that the true contact area can be a measure of the bond quality and is dependent on material properties such as the yield stress and work hardening characteristics. Hypothetical calculations of the bond quality can then be made using materials with different values of Y and x , and compare with that of Cu-Cu bonds. Assuming all the processes are feasible, we can then compare which material bonding will give the best quality.

Equation (4.34) suggests that the true contact area is inversely proportional to the yield stress Y . For metals such as Au, Al and Cu, Cu has the highest yield stress among the three metals. However Al, has a higher resistivity (topic on contact resistance will be discussed further in Chapter 5) than Cu and it oxidizes readily in ambient, hence the overall bond quality might not be as good. On the other hand, Au does not oxidize easily and has a lower yield stress than Cu. Hence the bond quality of Au-Au could potentially be better than that of Cu-Cu.

It is assumed that there are no process fabrication issues with the Au structures and a comparison is being made between Cu-Cu and Au-Au bonds (Figure 4-17), in terms of the true contact area. The surface roughness parameters of Au are assumed to be the same as that of Cu as in section 4-4, except that the yield stress Y and x are estimated to be about 30 MPa (at 300°C) [41] and 0.4 [89] for Au, respectively. The nominal contact area is taken to be 0.03142 m^2 and the applied load ranges from 1 kN to 10 kN.

Figure 4-17: Hypothetical comparisons of (a) Au-Au and (b) Cu-Cu bonds.

Figure 4-18 shows the comparison of the true contact area versus applied load for Au-Au and Cu-Cu bonds. Results suggest that for the same applied load and surface roughness, the calculated true contact area for Au-Au bonding can be more than 2 times that of Cu-Cu bonding.

Figure 4-18: Comparison of true contact area versus load for Au-Au and Cu-Cu bonds.

If the true contact area can be taken as a measure of the mechanical strength, then Au-Au bonds could be potentially two times stronger than Cu-Cu bonds. From

an electrical perspective, which will be discussed in Chapter 5, the contact resistance is inversely proportional to the true contact area and directly proportional to the resistivity. Even if the resistivity for Au is about 30% higher than that of Cu ($22 \text{ n}\Omega\text{-m}^2$ for Au versus $17 \text{ n}\Omega\text{-m}^2$ for Cu), the overall bond quality of Au-Au, in terms of contact resistance, could still be better than that of Cu-Cu's by about 40%.

4.6 Summary

It has been suggested that when two surfaces come into contact under an applied load, the distribution of the contact area would not be uniformly distributed: certain regions will have greater amount of contact than others. Hence, when the bonded wafer is subsequently dice-sawed, samples which did not have sufficient amount of contact area will be detached. The dicing yield can then be seen as a measure of the bond quality, in terms of the true contact area of the bonded interfaces. As the amount of true contact area will be different from sample to sample, it could explain why the four-point bend test results exhibit such a large scattering, as shown in chapter 2.

Results also showed that either an increase in applied load or a decrease in surface roughness of the pre-bond surfaces will enhance the bond quality of the bonded interfaces. However, an overall effect on bond quality has to be considered and neither the applied load nor roughness alone is adequate in characterizing the bond quality. It is proposed that the true contact area can be used to characterize the bond quality as it is a function of both the applied load and surface roughness.

A contact model is developed to calculate the true contact area between two contacting surfaces. The true contact area is a function of the applied load, nominal contact area and surface roughness, which is sub-characterized in terms of standard deviation of asperity heights, radius curvature of asperities and density of the asperities. The calculated true contact area in each bonding experiment is subsequently plotted against the dicing yield and was found to be directly proportional up to a maximum yield of 100%. The true contact area is also found to be a small fraction of the nominal contact area, which seems possible as otherwise the contact pressure would not be sufficient to induce plastic deformation for bonding to occur.

Using the developed model, calculations showed that the nominal contact area have little effects on the true contact area and the standard deviation of asperity heights is the dominant topological parameter affecting the true contact area. The surface roughness is also shown to have a diminishing effect on the true contact area: increments in the surface roughness results in smaller decrease in the true contact area. This concurs with the concept of work hardening in which increments in applied load result in lesser deformations, regardless of surface topologies.

Assuming the process is feasible for fabricating Au structures, hypothetical bonding of Au-Au is also considered. Calculations suggested that for the same bonding and surface conditions, Au-Au could achieve significantly higher bond qualities in terms of bond strength and contact resistance, than that of Cu-Cu bonds.

Chapter 5 – The Contact Resistance Measurements and Model

5.1 Introduction

Contact resistance between bonded surfaces is one of the critical parameters for 3D ICs. Data from literature shows that the specific contact resistances for metal-to-metal contacts are in the order of $10^{-8} \Omega\text{-cm}^2$ [90,91]. For a bonded metal-to-metal surfaces of nominal contact area $100 \times 100 \mu\text{m}^2$, the contact resistance would subsequently be in the order of about $10^{-4} \Omega$. Consider the illustration of a Cu-Cu bonded pads as shown in Figure 5-1. For the given dimensions and assuming the bonded interface to be perfect, the resistance of the whole bonded structure would be in the order of $10^{-6} \Omega$, as calculated below.

$$R_c = \frac{\rho L}{A} = \frac{(17 \times 10^{-9})(10^{-6})}{10^{-8}} = 1.7 \times 10^{-6} \Omega$$

Figure 5-1: Schematic diagram of a perfect Cu-Cu bonded structure.

The fact that the contact resistance is approximately in the order of $10^{-4} \Omega$ suggests that the bonded interface is not perfect, and may have a significant contribution to the electrical resistance.

In chapter 4, it was shown that the true contact area between two bonded interfaces can be just a small fraction of the nominal contact area, and hence the higher contact resistance can also be due to the very small true contact area. In the following sections, a 3D bonded Cu-Cu test structure designed for measuring the contact resistance of the bonded interfaces has been fabricated and the contact resistance measured. A contact resistance model has also been developed, based of the calculated true contact area, to predict the contact resistance of the bonded interface. Subsequently, a comparison is made between the theoretical and experimental contact resistance values, to validate both the electrical and the contact models.

5.2 Fabrication and Testing of the Contact Resistance Structures

The schematic diagram of the contact resistance measurement test structure is illustrated in Figure 5-2.

Figure 5-2: Schematic diagram and the measurement of the contact resistance test structure.

The structure was originally designed with two purposes. One of the intentions is to measure the contact resistance of the bonded pads and the other is to study the electromigration behavior of the Cu line in the bottom wafer that is bonded to a top wafer with probing pads. However in this work, only the contact resistance of the bonded interfaces will be measured.

An illustration of the fabrication process flow for the bottom and top wafers are given in Figures 5-3 and 5-4, respectively.

Figure 5-3: Fabrication process flow for the bottom wafer.

Figure 5-4: Fabrication process flow for the top wafer.

Before the bonding process, the oxide surrounding the Cu bond pads was recessed to facilitate contact between the two bonding layers. The oxide recess was performed by either dry or wet etch to achieve different degrees of surface roughness on the Cu bond pads. The wafers were cleaned with acetic acid to remove the Cu oxide, as described in section 2.1. The Cu thickness on the top and bottom wafer is about 400 nm and 1 μm-thick respectively, and the dimensions of the Cu bond pads are either 50 x 50, 75 x 75 or 100 x 100 μm². The wafers were bonded face-to-face at 400°C for 30 min with an applied load ranging from 6 kN to 10 kN, followed by another 30 min of annealing (without the applied load), without breaking vacuum.

After bonding, the Si substrate on the top wafer was back grinded from 700 μm down to approximately 200 μm and the remaining Si material was removed by plasma etch to expose the remaining Cu structure of the top wafer. The schematics of the bonding and backgrind processes, and a planar view of the bonded wafers (after the delayering), are as shown in Figures 5-5 and 5-6, respectively.

Figure 5-5: Backgrind and thinning of one side of the bonded wafers.

Figure 5-6: Planar view of the contact resistance test structure after the removal of the Si substrate of the top wafer. The amount of misalignment was noted for each contact.

The design of the test structure is such that when measuring the contact resistance, the probe pads of the bottom wafer are accessed from the front side while the probe pads of the top wafer are accessed from the back side of the substrate, as illustrated in Figure 5-7.

Figure 5-7: Side view of the contact resistance test structure.

Wafers were also bonded at 300°C, which is the bonding temperature for the experiments described in chapter 4. However, depending on the pre-bond Cu surface roughness, most of the bonded wafers did not survive the back grinding process. Hence, we could not measure the contact resistances of bonded interfaces with different surface roughness and thus unable to make any comparisons.

There are several reasons why some of the bonded wafers could not survive the back grind process. The first reason is as mentioned in section 4.3: Process issues of Cu CMP non-uniformity and dishing on patterned Cu wafers result in a lower than expected true contact area. This effect is further exacerbated as both wafers to be bonded are now patterned Cu wafers. The second reason is a design issue: From discussions with Tezzaron [70], it was learnt that to obtain sufficiently high bond strengths for bonded pattern wafers to survive processes such as back grinding, the design of the structures has to be optimized in terms of arrangement of the bonding

pads. For Tezzaron, the maximum spacing between the bonded dummies or electrical interconnects is no more than 50 μm whereas in our design, as can be seen from Figure 5-7, the maximum spacing can be as large as a few hundred microns. Hence, there are relatively large regions (and regularly spaced due to design) where there is no contact / bonding between the wafers. When one of the wafers is thinned down, the thinned side would sustain a larger stress such that the unbonded areas would become weak points and can be easily cracked and chipped off, acting as precursors for further crack propagations and delaminations.

From chapter 2, it was shown that the bond strength would increase with the bonding temperature. Hence, to produce bonded samples in which electrical measurements could be made, the wafers were bonded at a higher temperature of 400°C. The contact model developed would still be applicable as the only temperature dependent parameter in the equation is the yield stress, which as shown in Figure 4-8, is approximately the same from 300°C to 400°C at a value of 50 MPa.

Our contact resistance test structure (Figure 5-8(a)) is similar to a Kelvin contact resistance test structure [92], as shown in Figure 5-8(b) and the contact resistance can be similarly obtained by

$$R_C = \frac{V_+ - V_-}{I} \quad (5.1)$$

where V_+ and V_- are the voltages on the top and bottom probe pads respectively, that are continuously measured while a current I is ramped from -50 mA to 50 mA from the top to bottom probe pads. The specific contact resistance ρ_C can be obtained from the expression

$$\rho_C = R_C A_n \quad (5.2)$$

where R_C and A_n are the contact resistance and nominal contact area, respectively. (The calculated nominal contact area A_n has taken into account the amount of misalignment in each contact as shown in Figure 5-7)

Figure 5-8: Schematic diagram of (a) our contact resistance test structure and (b) a Kelvin contact resistance test structure.

As an illustration, suppose the measured contact resistance on the test structure in Figure 5-7 is about $5 \times 10^{-4} \Omega$. The dimensions of the interconnect pads are $100 \times 100 \mu\text{m}^2$ with a misalignment of about $25 \mu\text{m}$ in the y -axis and hence the actual nominal contact area will be about $100 \times 75 \mu\text{m}^2$. From equation (5.2), the specific contact resistance for this bonded interconnect will be about $3.75 \times 10^{-8} \Omega\text{-cm}^2$.

5.3 Contact Resistance of Cu-Cu bonds

5.3.1 The Contact Resistance Model

Ohm's law states that the electrical resistance R of a specimen of length L , cross sectional area A and resistivity ρ is expressed as

$$R = \frac{\rho L}{A}. \quad (5.3)$$

Hence, the calculation of the contact resistance in this model would adopt a similar constitutive relationship as that of equation (5.3).

The contact of two rough surfaces, as mentioned in section 4.3.2, can be considered as the contact of an equivalent rough surface to a perfectly flat plane. From a contact mechanics perspective, as illustrated in Figure 5-9(a), the surface asperities are not completely flattened and there will be 'pockets of gaps' distributed along the bonded interface. This can be better visualized as the equivalent of a flat plane in contact with a patterned surface as illustrated in Figure 5-9(b). Assuming that the resistivity of the Cu material beneath the topologies have the same value as the bulk, then it would have negligible contributions to the contact resistance. It is proposed that the length of the 'pockets of gaps' L , together with the amount of true contact area are the main factors affecting the value of the contact resistance of the bond interfaces.

In this contact resistance model, the contact resistance of a bonded interface is defined as

$$R_C = \frac{\rho L}{A_t} \quad (5.4)$$

where ρ is the resistivity of the bonded material, which in this case is that of Cu. A_t is the true contact area and L is a length of interest for a bonded interface as shown in Figures 5-9(a) and 5-9(b).

Figure 5-9: (a) A rough surface having asperities in contact with a flat plane. The distance between the mean planes of the surfaces is d . The length of interest L is a certain distance from the flat plane. (b) Contact of a perfect flat Cu surface with a patterned Cu surface.

For a bonding temperature of 400°C , it is assumed that any defects that may be created into the Cu film during the bonding process will be annealed out and there would be a minimum impact on the value of resistivity, which in our model assumes the bulk value i.e. approximately $16.78 \text{ n}\Omega\text{-m}$. The true contact area A_t is calculated based on our contact model, taking into account of the applied load, nominal contact area and surface roughness, as mentioned in the previous chapter. What is left to be addressed is the definition of L .

From Figure 5-9(a), L is defined as the summation of two components: the distance between the mean planes of the two surfaces d , and the average distance between the mean plane (of the rough surface) to the lowest point of each *contacting* asperities, d' . The former (d) is calculated from the contact model in chapter 4 while the latter (d') is determined with the aid of AFM measurements. For instance, Figure 5-10 illustrates a rough surface in contact with a flat plane. The asperities which are in contact with the flat plane are identified and the distances between the mean plane and the lowest point of each contacting asperity are recorded (the algorithm to identify these contacting asperities is listed in the code shown in Appendix C). The

mean plane is set at zero along the y -axis and the average of these recorded values is represented as d' .

Figure 5-10: Schematic of contact between a rough surface and a flat plane and illustration of determining d' .

Hence for every AFM measured surface profile, the value of L can be determined and plotted against their corresponding rms surface roughness as shown in Figure 5-11, which shows that L has a linear dependence with the rms surface roughness.

Figure 5-11: Plot of L versus the rms surface roughness, suggesting an approximate linear relationship.

A theoretical estimate for the contact resistance R_C , for a Cu-Cu bonded interface can be calculated using equation (5.4) using the calculated values of L and A_t , both of which are dependent on the applied load and surface roughness of the pre-bond Cu layers. As mentioned in section 5.1, the typical specific contact resistance is in the order of $10^{-8} \Omega\text{-cm}^2$. A specific contact resistance map of applied load and surface roughness is shown in Figure 5-12, which is based on each AFM measurement assuming a total nominal contact area of (arbitrarily) 0.02 m^2 . The trend lines of specific contact resistance values range from $10^{-8} \Omega\text{-cm}^2$ to $5 \times 10^{-8} \Omega\text{-cm}^2$, are best fitted curves for each set of calculations and represents arbitrarily defined boundary conditions. This electrical map allows the estimate of the specific contact resistance for a given wafer surface roughness and applied load, or for a given surface roughness, the applied load that is required to attain a certain value of specific contact resistance.

Figure 5-12: Effects of surface roughness and applied load on the specific contact resistance.

Figure 5-12 shows that for a given applied load, the overall effect of surface roughness on the true contact area and length L is such that the specific contact resistance is roughly directly proportional to the surface roughness. For instance, at an applied load of about 2 kN, when the surface roughness increases from 2.5 nm to 5 nm and 7.5 nm, the specific contact resistance increases from $10^{-8} \Omega\text{-cm}^2$ to $2 \times 10^{-8} \Omega\text{-cm}^2$ and $4 \times 10^{-8} \Omega\text{-cm}^2$, respectively. A similar trend is also observed when the surface roughness is fixed: For a given surface roughness, the specific contact resistance is almost inversely proportional to the applied load. For instance, at a surface roughness of about 4 nm, when the applied load increases from about 750 N to 1.5 kN and 3 kN, the specific contact resistance decreases from $4 \times 10^{-8} \Omega\text{-cm}^2$ to $2 \times 10^{-8} \Omega\text{-cm}^2$ and $10^{-8} \Omega\text{-cm}^2$, respectively.

5.3.2 Comparison of Theoretical and Measured Specific Contact Resistance

The contact resistances and specific contact resistances of the test structures were measured and the results plotted against the total surface roughness as shown in Figures 5-13(a) and 5-13(b).

(a)

(b)

Figure 5-13: Experimental and model prediction of specific contact resistances of Cu bonded interfaces, for different total surface roughness (2, 4.5 and 7.5 nm). Wafers were bonded at 400°C and 8 kN.

The results of the specific contact resistance measurements (Figure 5-13(a)), like the four-point bend test results, show significant scatterings about the average. The reason for the large scattering is the same as in the bond strength measurement results, i.e. the non-uniform distribution of the true contact area results in the deviations of the specific contact resistance data. Nonetheless, the results show that as the total surface roughness increases from about 2 nm to 4.5 nm and 7.5 nm, the average specific contact resistance increases from $2 \times 10^{-8} \Omega\text{-cm}^2$ to 3.25×10^{-8} and $1.04 \times 10^{-7} \Omega\text{-cm}^2$, respectively.

The model's predictions of the specific contact resistances at these different surface roughnesses are about 9×10^{-10} , 2.2×10^{-9} and $3.8 \times 10^{-9} \Omega\text{-cm}^2$, respectively. The prediction is about 20 – 30 times smaller than that of the average experimental values but is in good agreement with the minimum measured value in each case: 1.67×10^{-9} , 3.54×10^{-9} and $6.17 \times 10^{-9} \Omega\text{-cm}^2$, respectively (Figure 5-13(b)). This is reasonable as the model represents the ideal case of bonding of rough surfaces, and has yet to account for the actual process variations such as dishing and bonding non-uniformity.

The effects of the applied load on the contact resistances were also investigated and the plot of specific contact resistances versus applied load is shown in Figures 5-14(a) and 5-14(b).

(a)

(b)

Figure 5-14: Experimental and model prediction of specific contact resistances of Cu bonded interfaces, for different applied loads (6, 8 and 10 kN). Wafers were bonded at 400°C with a total rms surface roughness of 2 nm.

The wafers were bonded with applied loads of 6, 8 and 10 kN. The total surface roughness of the pre-bond surfaces was about 2 nm in each case, to improve the yield from the bonding process. Results show that as the applied load increases, the measured average specific contact resistances would decrease. As in the previous analysis, the model's prediction is about an order lower than the average measured value but is close to the minimum measured value in each case.

Hence the true contact area, which depends on the surface roughness and the applied load, has been shown to affect the contact resistances of bonded interfaces. A high applied load and low surface roughness is needed to ensure the optimal electrical performances in bonded interconnects.

5.4 Summary

A test structure has been designed and fabricated, using thermo-compression bonding, for measuring the contact resistances of the Cu-Cu bonded interfaces. Results show that for a given load, the contact resistances are higher for bonded interfaces with higher pre-bond Cu surface roughness than that with smoother ones. For approximately the same pre-bond Cu surface roughness, a higher applied load would lead to lower contact resistances (due to a larger true contact area) and vice versa. In each experiment, the measured data has a significant spread about the average value. This is attributed to the non-uniform distribution of the true contact area as well as non-uniformity of the bonding process, as described in the previous chapter. Hence to obtain optimal electrical performances (lowest contact resistance),

the surface roughness of the pre-bond surface has to be as low as possible and the maximum possible load has to be applied.

An electrical model is also developed to predict the contact resistance of a bonded interface, taking into account the true contact area (applied loads and surface roughness). Comparison between the modeling and experimental results showed that the theoretical estimate is in the same order as the minimum measured value in each set of experiments, but is about an order lower than the average value. This is because the model represents a situation in which the bonding is ideal and did not account for all the process-related non-uniformities. Nevertheless, the electrical model allows the determination of the bond strengths (through the true contact area) of the bonded interfaces by a non-destructive method through the electrical probing of the bonded pads.

Chapter 6 – High Contact Resistance Reduction Phenomenon

6.1 Introduction

Chen *et al.* [93] first reported the phenomena of contact resistance reduction of a Cu-Cu bonded interface when the current was ramped from 1 to 100 mA on the test structure as shown in Figure 6-1(a).

Figure 6-1: (a) Contact resistance test structure and (b) contact resistance reduction phenomena, extracted from [93].

As the current was increased, the contact resistance decreased and stabilized to a constant value at high stress current. After reaching 100 mA, the stress current was decreased gradually back to 1 mA. During the decrease of the stressing current, the contact resistance remained in the same range and did not change significantly.

The contact resistance reduction phenomenon was observed on samples with high initial value of contact resistances and was postulated to be due to Joule heating effect. It was proposed that the temperature near the bonded interfaces was sufficiently high to enable atomistic rearrangement, thus contributing to the decrease in resistance. When the resistance reached a value that was small enough and the heat

generated around the interface could not support further atomistic rearrangement, the contact resistance would stop decreasing and remained at a stable value. At this point, the interface would have achieved a stable state and therefore the contact resistance did not change after further decrease in stress current [93].

However, a commentary paper by Timsit [94] suggested that the true contact area could be very small such that the current density will be sufficiently high for electro-migration (EM) to occur, even at room temperature. Thus, the contact resistance reduction may not be due to Joule heating but effects of EM. The elevated temperatures around the interface (estimated to be about 300°C) simply serve to enhance this effect.

In this work, the contact resistance reduction phenomenon was also observed and further investigations were conducted to verify if this is due to the effects of Joule heating or EM.

6.2 Contact Resistance Reduction

In this investigation, a current was ramped from 1 mA to 1 A through the contact resistance test structures. The contact resistance reduction in [93] was shown to occur immediately upon current stress. Our results showed that regardless whether the contact resistance would eventually decrease at high current stress, the contact resistance would initially increase as the current increases, as shown in Figure 6-2. This is possible as the electrical resistances in metals are expected to increase with the temperature, which increases with increasing current (density) due to Joule heating.

(a)

(b)

(c)

(d)

(e)

Figure 6-2: Contact resistance behaviour under increasing current stress for samples with different initial contact resistances. Samples with higher initial contact resistances exhibit the resistance reduction phenomenon while samples with lower contact resistances did not.

Our results concur with [93] that the contact resistance reduction phenomenon was observed only if the initial contact resistances are sufficiently high. Figures 6-2(a)-(c) show that when the current was stressed on samples with initial contact resistances above 7 mΩ, the reduction phenomenon was observed whereas for contact resistance of 3 mΩ and below (Figures 6-2 (d)-(e)), the reduction phenomenon seen. Another interesting observation from Figures 6-2 (a)-(c) is that the higher the initial contact resistances, the lower the current value where the reduction phenomenon is observed. With the calculated true contact area of the bonded interface, the current density at which the phenomenon occurs is estimated to be consistently about $2 \times 10^9 \text{ A/cm}^2$. With these results, the next section would strive to examine the origin of this phenomenon.

6.3 Effects of Joule Heating

To verify if the contact resistance reduction is indeed due to Joule heating, we would need to determine the temperature around the bonded interfaces during the reduction to see if it is sufficiently high to result in atomistic rearrangement. The temperature at the contacts can be estimated from the expression

$$\frac{\Delta R}{R_o} = \alpha \Delta T \quad (6.1)$$

where ΔR , R_o , ΔT and α are the changes in contact resistance, the initial contact resistance (at room temperature), the changes in contact temperature from room temperature and the temperature coefficient of resistivity (TCR), respectively.

To determine the value of TCR of the Cu films which were fabricated (which is typically 0.3 - 0.4% for Cu), a test structure with initial contact resistance of 4.65

m Ω was heated up from room temperature (22°C) to 50, 75, 100, 125 and 175°C and the contact resistance was measured at each of these temperatures as tabulated in Table 6-1. The plot of $\Delta R/R_o$ versus ΔT is shown in Figure 6-3, with a slope (α) of 0.41%, which is within expectations. The temperature at the bonded interfaces can now be estimated, given that ΔR and R_o are known.

Table 6-1: Contact resistance measurements (at 0.2 A) at different temperatures.

Temperature /°C	Contact Resistance /m Ω	ΔR /m Ω	ΔT /°C
22	4.65	-	-
50	5.19	0.54	28
75	5.59	0.94	53
100	6.07	1.42	78
125	6.52	1.87	103
150	7.12	2.47	128
175	7.53	2.88	153

Figure 6-3: Plot of $\Delta R/R_o$ versus ΔT , with the slope $\alpha = 0.411\%$.

From Figure 6-2, the values of $\Delta R/R_o$ (up to the point of contact resistance decrease or 1 A) ranges from 0.12 to 0.41, and the temperature at the contacts is estimated to range from 51°C to 122°C, which is much lower than the bonding temperature of 400°C. Several cycles of the current stress were also performed on a

bonded interface with an initial contact resistance of 3 m Ω as shown in Figure 6-4 and results show that there is no change in the initial contact resistances after each cycle if the contact resistance reduction phenomenon was not experienced during the stressing.

Figure 6-4: Repeated current stress showed that Joule heating is insufficient to result in permanent structural changes and hence changes in contact resistances.

Results indicate that the increase in temperature at the contacts due to Joule heating is much lower than the bonding temperature and is thus unlikely to result in any permanent structural rearrangement that could have reduced the contact resistances. This was verified by repeated current stress on the same sample, showing no change in the initial contact resistances and thus the true contact area. Hence, Joule heating is not the mechanism that can explain the contact resistance reduction phenomenon.

6.4 Effects of Electro-migration

It was shown in the previous section that Joule heating is unlikely the mechanism to result in the contact resistance reduction phenomenon. In this section, we will investigate whether electromigration is the mechanism that is responsible for this phenomenon, as suggested by Timsit [94].

Electromigration is the transport of material caused by the gradual movement of the ions in a conductor due to the momentum transfer between conducting electrons and diffusing metal atoms. The effect is important in applications where high direct current densities are used, such as in metallic interconnects. The atomic flux J_a due to electromigration can be expressed as

$$J_a = \frac{z^* e j \rho D_o}{\Omega k T} \exp\left(\frac{-E_a}{k T}\right), \quad (6.2)$$

where D_o is the diffusion coefficient of atom at room temperature, Ω is the atomic volume, k is the Boltzmann's constant, T is the temperature, z^* is the effective valence, e is the elementary charge, j is the current density, and ρ is the resistivity.

From equation (6.2), the atomic diffusion due to EM can be accelerated by increasing the current density or temperature. If EM is indeed responsible for the contact resistance reduction, then it is not that this phenomenon could not be observed in samples where the contact resistances are low (implying a large true contact area), as suggested in [93] and Figures 6-2(d)-(e), but that the current density and/or temperature is yet to be sufficiently high to induce significant atomic flux diffusion in lowering the contact resistance. However, the current density required to trigger a

resistance decrease can be lowered by increasing the stress temperature in which the effects of EM can be accelerated and thus observable.

Three samples with different initial contact resistance were heated to 50, 100, 150, 200 and 250°C. At these temperatures, each sample was stressed by ramping the current from 1 mA to 1 A, and the contact resistances measured are as shown in Figures 6-5 (a)-(c).

(a)

(b)

Figure 6-5: Contact resistance vs current stress measurement of samples with initial contact resistance of (a) 3.6 mΩ, (b) 2.8 mΩ and (c) 1.2 mΩ, at different temperatures.

Results show that the lower the initial contact resistance, the higher the stress temperature needed to observe the contact resistance reduction phenomenon. For the sample with an initial contact resistance of 1.2 mΩ, the phenomenon occurs at 250°C, whereas the sample with initial contact resistance of 3.6 mΩ, the phenomenon occurs at 150°C. Recall from Figure 6-2(c) that the sample with initial contact resistance of 7 mΩ, the phenomenon occurs at room temperature. These show that at low temperatures and low contact resistances (or large contact area, and hence a low current density with the same current value), the EM effects could be small such that the phenomenon is not significant. However at elevated temperatures, the atomic diffusion will be enhanced and contact resistance would eventually be reduced. This could be a potential mechanism whereby the bond quality can be enhanced (lower contact resistances) for 3D ICs.

6.5 Summary

The contact resistance reduction phenomenon has been observed on samples with high initial contact resistances, as reported in the literature. The temperature coefficient of resistivity was determined and showed that for the given change in contact resistance, the temperature at the contacts at the point of resistance reduction is much lower than the bonding temperature and thus unable to cause any permanent microstructural changes to the bonded interface. This is further supported as repeated current stresses (on the same sample) which induced similar temperature rise at the bonded interfaces did not have any impact on the measured initial contact resistance.

The lower the initial contact resistance which implies a larger true contact area, the higher the stressing temperature is needed for the phenomenon to be observed as it requires diffusion of atomic flux to the interface due to the high current density. For samples with even lower initial contact resistances, the contact resistance reduction was not observed at room temperature but at higher temperatures, the phenomenon occurred as the atomic flux diffusion can be accelerated by the temperature, besides the current density. These results suggest that Joule heating could not be the reason for the contact resistance reduction but electromigration is the more likely mechanism.

Chapter 7 – Conclusions

This chapter summarizes the experimental and modeling results and a discussion for the future work concludes the thesis.

7.1 Thesis Summary

The first part of the thesis presented the processes developed for the fabrication of wafer level thermo-compression bonds via copper bonding layers. The processes were designed to generate four-point test specimens so that a quantitative description of the bond strengths could be made. Characterizations of the bond strengths were focused on the effects of bonding temperatures, total surface roughness and contact pressures (in terms of nominal contact area and applied load). The process window examined for the bonding temperatures, total surface roughness and contact pressures ranged from 250°C to 400°C, 1 nm to 14 nm and 0.14 MPa to 2.45 MPa, respectively. Large scatterings in the bond strength measurements were present in all cases. Despite the scatterings, several observations can be made from the data.

- Bonding can be achieved at 250°C but the dicing yields, as well as the bond strengths, are low. The dicing yield and bond strength generally increase with the bonding temperature. The average bond strengths for samples bonded at 250°C is about 1 J/m². As bonding temperature increases, a bond strength of as high as 8 J/m² (at 350°C) was obtained.

- As the surface roughness of the copper bonding layers increases, the bond strength decreases. When the total surface roughness increases from about 1 nm to 14 nm, the average bond strength decreases from 6 J/m² to 1 J/m². Rough bonding surfaces are detrimental to obtaining high bond strengths.
- For the same surface roughness, the bond strength increases as the contact pressure increases. However, this is not necessarily true if surfaces with different roughness are involved.
- Bonding temperatures, surface roughness and contact pressures individually affects the bond strengths of the bonded interfaces. Unless all except one of the parameters are kept constant, they have to be collectively analyzed in any bond strength analysis.

It was also shown that the contact and bonding between opposite copper layers can be non-uniform, possibly due to process (CMP) issues such as dishing and polishing non-uniformity. Hence, copper in some regions could be better bonded than in other regions and that could explain why the delamination path was not confined just to the Cu-Cu interface but can propagate into adjacent interfaces (Ta-SiO₂) in the four point bend tests.

As surface roughness is found to be an important parameter affecting the bond quality, it becomes necessary to understand the surface topology of the pre-bond surfaces in order to predict the subsequent contact and deformation behavior. A surface roughness acquisition model was then developed, using scanned data from AFM. It was found that

- The summit of asperities is generally spherical with radius curvature ranging from 20 nm to over 100 nm. The radius curvature of asperities has no relation with the corresponding asperity heights and varies randomly.
- The average radius curvature and density of the asperities generally decrease as the surface roughness increases.
- The asperity heights have a Gaussian distribution profile.
- The standard deviation of asperity heights scales directly with the measured rms surface roughness from the AFM.

The large scattering in the four-point test measurements is proposed to be due to the non-uniform distribution of the true contact area between individual samples. The dicing yields of the bonded wafers can also be seen as an indicator of the amount of true contact area and bond quality of bonded wafers.

A contact model was developed to estimate the true contact area between a bonded interface, which is a function of the applied load, nominal contact area and surface roughness. The model assumes that the bonding surfaces are plastically deformed upon contact for the range of loads applied, which is reasonable as the critical deformation at which plastic yielding starts to occur is very small. The dicing yield of the bonded wafers was found to scale directly with the true contact area up to a maximum yield of 100%. However, there are deviations involving bonding of the patterned wafers in which the true contact area can be over-estimated. Process issues such as dishing effects on the patterned structures might have resulted in a lower than expected amount of true contact area.

The true contact area was found to be a very small fraction of the nominal contact area, as low as in the order of 0.1%. The resulting contact pressures would then be sufficiently high for the surface asperities to plastically deform and bond. Using the developed contact model, it was shown that the nominal contact area has minimal effects on the true contact area and the standard deviation of asperity heights is the dominant topological factor affecting the true contact area. The surface roughness is also shown to have a diminishing effect on the true contact area: increments in the surface roughness results in corresponding lesser decrease in the true contact area due to work hardening.

The latter part of the thesis presented the processes developed for the fabrication of a contact resistance measurement test structure so that a quantitative description of the bond quality from an electrical perspective could be made. Characterizations of the specific contact resistances were focused on the effects of surface roughness and applied load. Results suggest that the average specific contact resistances increases as surface roughness increases and/or the applied load decreases.

A contact resistance model, taking into account the true contact area (which is a function of the applied load and surface roughness), was developed to predict the specific contact resistances of the bonded interfaces. The model's predictions are about an order below the average experimental results but are otherwise close to the minimum measurements in each case. The discrepancy (compared to the average value) is because the model assumes an ideal bonding scenario and has yet to take into consideration process issues such as dishing and bonding non-uniformity.

The high contact resistance reduction phenomenon was also observed in this work as first reported by Chen *et al.* in [79]. While Chen and Timsit [80] respectively argued (qualitatively) that Joule heating and electro-migration are responsible for the resistance reduction phenomenon, we have experimentally shown that Joule heating effects is insufficient to cause such effects and electromigration is the most likely mechanism responsible for the phenomenon.

In conclusion, a quantitative model of the true contact area of bonded Cu surfaces has been developed from contact theory. This model allows the estimation of the dicing yields and bond strengths of bonded wafers in terms of the bonding parameters (applied load) and wafer conditions (surface roughness and nominal area). A contact resistance model is further developed, based on the contact model. Through this model, a measured electrical contact resistance will give an indication of the true contact area, and thus the bond strength of the bonded wafers. If any of the contact area is low, resulting in a high contact resistance, wafer-level *in-situ* repair can be carried out by applying a high current density, with or without temperature, to enhance the bonding at the interface. Thus, these developments will aid in the adoption of wafer-level bonding of Cu structures for the manufacturing of 3D ICs.

7.2 Recommendations for Future Works

While the work performed in this thesis has led to an improved understanding of the bonding mechanisms of copper wafers for 3D ICs, in particular the quantitative effects of the applied load and surface roughness on the bond quality, it also highlighted several topics that require further work as below:

- The contact model developed in chapter 4 did not account for effects of diffusivity at different bonding temperatures. A more detailed investigation would be required to quantify the effects of the bonding temperature on the mechanical properties of the bonded interfaces.
- The contact resistances measured were of samples bonded at 400°C. The ultimate objective of 3D ICs is to enable bonding at low temperatures and hence to understand the electrical performances of interfaces bonded at lower temperatures is also of great importance.
- EM lifetime of the Cu-Cu bonded interface. The reliability of 3D ICs has to be known before it can be put into practical applications. The general aspects of reliability include electrical, thermal stress cycles and electro-migration. The original purpose of the test structure, as mentioned in section 5.2, was to compare the EM lifetime of the Cu-Cu bonded interface with that of the Cu line. This will provide an idea of how reliable is the Cu-Cu bonded interface compared to the Cu lines.
- Design optimization. As mentioned in section 2.6, it was learnt from Tezzaron that for better mechanical performances, the spacing between the bond pads has to be kept small. For Tezzaron, their maximum spacing is about 50 μm whereas for our design the maximum spacing can be hundreds of microns. This could be one of the reasons why wafers bonded at 300°C could not survive the back grinding process. To bond at a lower temperature, the layout might have to be re-designed to ensure high nominal contact areas with the minimum spacing.

References

- [1] International Technology Roadmap for Semiconductors (ITRS), 1999
- [2] C.R Barrett, "Microprocessor evolution and technology impact", Symp. VLSI Tech. Dig., 7, 1993
- [3] C. Hu, "MOSFET scaling in the next decade and beyond", Semicond. Int., 105, 1994
- [4] B. Davari, R.H Dennard, and G.G Shahidi, "CMOS scaling for high performance and low power – The next ten years", Proc IEEE, 83, 595, Apr 1995.
- [5] G.A Sai-Halasz, "Performance trends in high-end processors, "Proc IEEE, 83, 20, 1995.
- [6] K.C Saraswat and F. Mohammadi, "Effect of interconnect scaling on time delay of VLSI circuits", IEEE Trans. Electron Devices, ED29, 645, 1982.
- [7] M.T Bohr, "Interconnect scaling – The real limiter to high performance ULSI", IEDM Tech Dig., 241, 1995.
- [8] J.D Meindl, "Low power microelectronics: Retrospect and prospect", Proc IEEE, 83, 619, Apr 1995.
- [9] S.Y Oh and K.J Chang, "2001 needs for multi-level interconnect technology", Circuits Dev., 16, 1995.
- [10] K. Banerjee, S.J Souri, P. Kapur and K.C. Saraswat, "3-D ICs: A novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration", Proceedings of the IEEE, 89(5), 602, 2001.
- [11] A.W Wang and S.C Saraswat, "A strategy for modeling of variations due to grain size in polycrystalline thin film transistors", IEEE Transaction of Electron Devices, 47, 1035, 2000
- [12] V. Subramanian, M. Toita, N.R Ibrahim, S.J Souri, K.C Saraswat, "Low-leakage Germanium-seeded laterally crystallized single-grained 100 nm TFTs for vertical integration applications", IEEE Electron Device Lett., 20, 341, 1999
- [13] A. Kohno. T. Samwshima, N. Sano, M. Sekiya and M. Hara, "High performance poly-Si TFTs fabricated using pulsed laser annealing and remote

- plasma CVD with low temperature processing”, IEEE Transaction of Electron Devices, 42(2), 251, 1995
- [14] M.A Crowder, C.Y Chang, T.F Lei, J.Y Cheng, H.C Tseng and L.P Chen, “Low-temperature single-crystal Si TFT’s fabricated on Si-films processed via sequential lateral solidifications”, IEEE Electron Device Lett., 19(8), 306, 1986
 - [15] G.W Neudeck, S. Pae, J.P Denton and T. Su, “Multiple layers of silicon-on-insulator for nanostructure devices”, J. Vac. Sci. Tech. B, 17(3), 994, 1999
 - [16] N. Yamauchi, “Polycrystalline silicon thin films processed with silicon ion implantation and subsequent solid-phase crystallization: Theory, experiments and thin-film transistor applications”, J. Appl. Phys., 75(7), 3235, 1994
 - [17] D.N. Kouvatsos, A.T Voutsas and M.K. Hatalis, “Polycrystalline silicon thin film transistors fabricated in various solid phase crystallized films deposited on glass substrates”, J. Electronic Mater., 28(1), 19, 1999
 - [18] V. Subramanian and K.C Saraswat, “High performance germanium-seeded laterally crystallized TFTs for vertical device integration”, IEEE Transaction of Electron Devices, 45, 1934, 1998
 - [19] S.W Lee and S.K Joo, “Low temperature poly-Si thin film transistor fabrication by metal-induced lateral crystallization”, IEEE Electron Device Lett., 17(4), 160, 1983
 - [20] S.Y Yoon, S.K Kim, J.Y Oh, Y.J Choi, W.S Shon, C.O Kim and J. Jang, “A high performance polycrystalline silicon thin-film transistor using metal-induced crystallization with Ni solution”, Jpn. J. Appl. Phys., pt. 1, 7193, 1998
 - [21] Y.L Low and K.J O’Connor, “Electrical performance of chip-on-chip modules”, IEEE Transaction on Advanced Packaging, 22, 321, 1999.
 - [22] M. Shimbo, K. Furukawa, K. Fukuda, and K. Tanzawa, “Silicon-to-silicon direct bonding method”, J. Appl. Phys., 60(8), 2987, 1986.
 - [23] J. Lasky, “Wafer bonding for silicon-on-insulator technologies”, Appl. Phys. Lett., 48(1), 78, 1986.
 - [24] S. Bengtsson, “Semiconductor wafer bonding: A review of interfacial properties and applications”, J. Electron. Mater., 21(8), 841, 1992.
 - [25] A. Plossl and G. Krauter, “Wafer direct bonding: tailor adhesion between brittle materials”, Mater. Sci. Eng., R25, 1, 1999.

- [26] R.E Belford and S. Sood, "Surface Activation Using Remote Plasma for Hydrophilic Bonding at Elevated Temperature", *Electrochem. Solid-State Lett.*, 10(5), H145, 2007
- [27] C. Gui, M. Elwenspoek, N. Tas, and J. Gardeniers, "The effect of surface roughness on direct wafer bonding", *J. Appl. Phys.*, 85(10), 7448, 1999.
- [28] G. Kissinger and W. Kissinger, "Void-free silicon-bond strengthening in the 200-400 degrees C range", *Sensor Actuator*, 36(2), 149, 1993.
- [29] S. Bengtsson and P. Amirfeiz, "Room temperature wafer bonding of silicon, oxideized silicon and crystalline quartz", *J. Electron. Mater.*, 29(7), 909, 2000.
- [30] K.T Turner and S.M Spearing, "Modeling of direct wafer bonding: Effect of wafer bow and etch patterns", *J. Appl. Phys.*, 92, 7658, 2002.
- [31] G. Wallis and D. Pomerantz, "Field assisted glass-metal sealing", *J. Appl. Phys.*, 40(10), 3946, 1969.
- [32] S. mack, H. Baumann, and U. Gosele, "Gas tightness of cavities sealed by silicon wafer bonding", *MEMS 1997 (Nagoya, Japan)*, 488, IEEE, 1997.
- [33] M. Schmidt, "Wafer-to-wafer bonding for microstructure formation", *Proc IEEE*, 86(8), 1575, 1998.
- [34] S. Cho and F. Erdmann, "An on-chip hermetic packaging technology for micromechanical devices", *Solid-State Sensor and Actuator Workshop*, 229, 1998.
- [35] H. Tilmans, D. Van de Peer, and E. Beyne, "The indent reflow sealing (IRS) technique – A method for the fabrication of sealed cavities for MEMS devices", *J. Microelectromech. Syst.*, 9(2), 206, 2000.
- [36] C. Wang and C. Lee, "An eutectic bonding technology at a temperature below the eutectic point", *42nd Electronic Components and Tech. Conference*, 502, 1992.
- [37] C. Hoge and S. Thomas, "Some considerations of the gold-silicon die bond based on surface chemical analysis", *18th Annual Proc of Reliability Physics*, 301, IEEE, 1980.
- [38] Y. Li and R. Bower, "Systematic low temperature silicon bonding using pressure and temperature", *Jpn. J. Appl. Phys.*, 37(3A), 737, 1998.

- [39] A. Sayah, D. Solignac, T. Cueni, and M. Gijs, "Development of novel low temperature bonding technologies for microchip chemical analysis applications", *Sensor Actuat. A-Phys.*, 84(1-2), 103, 2000.
- [40] M. Waelti, N. Schneeberger, O. Paul, and H. Baltes, "Low temperature packaging of cmos infrared microsystems by Si-Al-Au bonding", 4th Int. Symp. On Semicond. Wafer Bonding: Science, Technology and Applications., 147, 1998.
- [41] C.H Tsau, "Fabrication and characterization of wafer-level gold thermocompression bonding", PhD Thesis, MIT, 2003.
- [42] R. Tadepalli, "Characterization and requirements for Cu-Cu bonds for three-dimensional integrated circuits", PhD Thesis, MIT, 2007.
- [43] A. Shigetou, N. Hosoda, T. Itoh, and T. Suga, "Room temperature direct bonding of CMP-Cu films for bumpless interconnection", *Electron. Components and Tech. Conference*, 2001.
- [44] K.N Chen, C.S Tan, A. Fan, and R. Reif, "Morphology and bond strength of copper wafer bonding", *Electrochem. Solid-State Lett.*, 7(1), G14, 2004.
- [45] K.N Chen, C.S Tan, A. Fan, and R. Reif, "Copper bonded layers analysis and effects of copper surface conditions on the bonding quality for three-dimensional integration", *J. Electronic Mater.*, 34(12), 1464, 2005.
- [46] K.N Chen, A. Fan, and R. Reif, "Interfacial morphologies and possible mechanisms of copper wafer bonding", *J. Mater. Sci.*, 37, 3441, 2002.
- [47] Ö. Vallin, K. Jonsson, and U. Lindberg, "Adhesion quantification methods for wafer bonding", *Mater. Sci. Engrg Review*, 50, 109, 2005.
- [48] W. Maszara, G. Goetz, A. Caviglia, and J. McKitterick, "Bonding of silicon wafers for silicon-on-insulator", *J. Appl. Phys.*, 64(10), 4943, 1988.
- [49] D. Pasquariello and K. Hjort, "Mesa-spacers: Enabling non-destructive measurement of surface energy in room temperature wafer bonding", *J. Electrochem. Soc.*, 147(6), 2343, 2000.
- [50] T. Martini, J. Steinkirchner, and U. Gösele, "The crack opening method in silicon wafer bonding – how useful is it?", *J. Electrochem. Soc.*, 144, 354, 1997.
- [51] J. Bagdahn and M. Petzold, "Fatigue of directly wafer-bonded silicon under static and cyclic loading", *Microsys. Tech.*, 7, 175, 2001.

- [52] D.S Hurd, R. Caretta and W.W Gerberich, "An experimental fracture mechanics study of a strong interface: The silicon/glass anodic bond", J. Mater. Res., 10, 387, 1995.
- [53] P.G Charalambides, J. Lund, A.G Evans and R.M McMeeking, "A test specimen for determining the fracture resistance of bimaterial interfaces", J. Appl. Mechanics, 56, 77, 1989.
- [54] R.H Dauskradt, M. Lane, Q. Ma and N. Krishna, "The adhesion and debonding of multi-layer thin film structures", Engineering Fracture Mechanics, 61, 141, 1998.
- [55] M. lane, A. Vainchtein, H. Gao and R.H Dauskradt, "Plasticity contributions to interface adhesion in thin-film interconnect structures", J. Mater. Res., 15(12), 2758, 2000.
- [56] S.J Howard, A.J Phillips and T.W Clyne, "The interpretation of data from the four-point bend delamination test to measure interfacial fracture toughness", Composites, 24(2), 103, 1993.
- [57] J.A Greenwood and J.B.P Williamson, "Contact of nominally flat surfaces", Proc. Of Roy. Soc. London Series A, 295, 300, 1966.
- [58] F.P Bowden and D. Tabor, "Friction and Lubrication of Solids Part I", Oxford Univ. Press, London, 1954.
- [59] H. Hertz, J. Regine Angew. Math., 92, 156, 1881; Translated and reprinted in English in "Hertz's Miscellaneous Paper" (Macmillan & Co., London, 1986) Ch. 5.
- [60] J.A Greenwood and J.H Tripp, "The contact of two nominally flat rough surfaces", Proc. Instn. Mech. Engrs., 185, 625, 1971.
- [61] Z.H Powierza, T. Klimczak and A. Polijanniuk, "On the experimental verification of the Greenwood-Williamson model for the contact of rough surfaces", Wear, 154, 115, 1992.
- [62] E.J Abbott and F.A Firestone, "Specifying surface quality", Mech. Eng., 55, 569, 1933.
- [63] J. Pullen and J.B.P Williamson, "On the plastic contact of rough surfaces", Proc. Of Roy. Soc. London Series A, 327, 159, 1972.
- [64] W.R Chang, I. Etsion and D.B Bogoy, "An elastic-plastic model for the contact of rough surfaces", J. Tribology, 109, 257, 1987.

- [65] Y. Zhao, D.M Maietta, L. Chang, “An asperity microcontact model incorporating the transition from elastic deformation to fully plastic flow”, J. Tribology, 122, 86, 2000.
- [66] L. Kogut and I. Etsion, “Elastic-Plastic contact analysis of a sphere and a rigid flat”, Trans. ASME, J. Appl. Mech., 69(95), 657, 2002.
- [67] H.A Francis, “Phenomenological analysis of plastic spherical indentation”, ASME J. Eng. Mater. Tech. 98, 272, 1976.
- [68] D. Tabor, “The Hardness of Metals”, Clarendon Press, Oxford, 1951.
- [69] K.L Chavez and D.W Hess, “A novel method of etching copper oxide using acetic acid”, J. Electrochem. Soc., 148(11), G640, 2001.
- [70] Y.P Varshni, “Temperature dependence of the elastic constants”, Phys. Rev. B, 2(10), 3952, 1970.
- [71] B. Swinnen, W. Ruythooren, P. De Moor, L. Bogarets, L. Carbonell, K. De Munck, B. Eyckens, S. Stoukatch, D. Sabuncuolu Tezcan, Z. Tokei, J. Van Aelst and E. Beyne, IMEC, IEDM 2006.
- [72] Visit to Tezzaron (Singapore), 2006.
- [73] R.S Sayles and T.R Thomas, “Surface topography as a non-stationary random process”, Nature, 271, 431, 1978.
- [74] J.F Archard, “Elastic deformation and the laws of friction”, Proc. Of Roy. Soc. London Series A, 243, 190, 1957.
- [75] D.J Whitehouse and J.F Archard, “The properties of random surfaces of significance in their contact”, Proc. Of Roy. Soc. London Series A, 316, 97, 1970.
- [76] J.F Archard, “Surface topology and tribology”, Tribology Int., 7, 213, 1974.
- [77] P.R. Nayak, “Random process model of rough surfaces”, ASME J. Lubr. Tech., 93(3), 398, 1971.
- [78] Y.P Zhao, G.C Wang and T.M Lu, “Volume 37 – Characterization of Amorphous and Crystalline Rough Surface: Principles and Applications”, Academic Press, 2001.
- [79] Discussion with Prof C.C Wong (NTU, MSE)

- [80] A.H Uppal, S.D Probert, and T.R Thomas, "The real area of contact between a rough and flat surface", *Wear*, 22, 163, 1972.
- [81] J.S Field and M.V Swain, "A simple predictive model for spherical indentation", *J. Mater. Res.*, 8(2), 297, 1993.
- [82] A.I Nadai, "Theory of Flow and Fracture of Solids", McCraw-Hill, Vol. 2, p.221, 1963.
- [83] K. Willner, "Elasto-plastic contact of rough surfaces", *Trans. Engineering Sci.*, 14, 13, 1997.
- [84] K.L Johnson, "Contact Mechanics", New York: Cambridge University Press, 1985.
- [85] John I. McCool, "Comparison of models for the contact of rough surfaces", *Wear*, 107, 37, 1986.
- [86] M. Abramowitz and I.A Stegun, "Handbook of Mathematical Functions", Vol. 55, 1964.
- [87] R-M Keller, P.S Baker and E. Arzt, *J. of Mater. Sci.*, 13(5), 1307, 1998.
- [88] D. Jia, K.T Ramesh, E. Ma, L. Lu and K. Lu, *Scripta Materialia*, 45, 613, 2001.
- [89] B. Wu, A. Heidelberg and J.J Boland, "Mechanical properties of ultrahigh-strength gold nanowires", *Nature Materials*, 4, 525, 2005.
- [90] G.C. Schwartz, "Handbook of Semiconductor Interconnection Technology", New York: Marcel Dekker, p. 187, 1998.
- [91] J. Rabaey, "Digital Integrated Circuits", Englewood Cliffs, NJ: Prentice-Hall, p. 465, 1996.
- [92] D.K Schroder, "Semiconductor Material and Device Characterization", New York: Wiley, p.121, 1990.
- [93] K.N. Chen, C.S. Tan, A. Fan and R. Reif, "Abnormal contact resistance reduction of bonded copper interconnects in three-dimensional integration during current stressing", *Appl. Phys. Lett.*, 86, 011903, 2005.
- [94] R.S. Timsit, "Comment on: Abnormal contact resistance reduction of bonded copper interconnects in three-dimensional integration during current stressing", *Appl. Phys. Lett.*, 88, 066101, 2006.

Appendix A: Algorithm code for the surface roughness acquisition model

```
% start from x=3 and y=3 because criteria set for peak is it must be larger than two
% consecutive points in all four directions

x=3;
y=3;

% Initialize parameters used in the calculation to zero

r=0;
k=0;
sum_sq_err=0;
sum_sq_err_a=0;
counter1=0;
xm_count=0;           %count of x in negative direction
xp_count=0;           %count of x in positive direction
ym_count=0;           %count of y in negative direction
yp_count=0;           %count of y in positive direction
d1_count=0;           %count in one diagonal direction
d2_count=0;           %count in another diagonal direction
d3_count=0;           %count in another diagonal direction
d4_count=0;           %count in another diagonal direction
peak_counts=0;        % calculate no. of peaks that satisfy the criteria

delete results.txt;

%Z is matrix to read in data from the AFM scan

Z=textread('t1.txt','delimiter',' ','emptyvalue',NaN);

while x<=254

%This section begins identifying the asperities, depending on the definition of the asperities

for y=3:254

    while ((x-r-1)>=1) & (Z(x-r,y)>Z(x-r-1,y))
        xm_count = xm_count+1;
        r=r+1;
    end
    r=0;

    while ((x-r-1)>=1) & ((y-r-1)>=1) & (Z(x-r,y-r)>=Z(x-r-1,y-r-1))
        d1_count = d1_count+1;
        r=r+1;
    end
    r=0;
```



```

while ((x+r+1)<=256) & (Z(x+r,y)>=Z(x+r+1,y))
    xp_count=xp_count+1;
    r=r+1;
end
r=0;

while ((x+r+1)<=256) & ((y+r+1)<=256) & (Z(x+r,y+r)>=Z(x+r+1,y+r+1))
    d2_count=d2_count+1;
    r=r+1;
end
r=0;

while ((y-r-1)>=1) & (Z(x,y-r)>Z(x,y-r-1))
    ym_count=ym_count+1;
    r=r+1;
end
r=0;

while ((y-r-1)>=1) & ((x+r+1)<=256) & (Z(x+r,y-r)>=Z(x+r+1,y-r-1))
    d3_count=d3_count+1;
    r=r+1;
end
r=0;

while ((y+r+1)<=256) & (Z(x,y+r)>=Z(x,y+r+1))
    yp_count=yp_count+1;
    r=r+1;
end
r=0;

while ((y+r+1)<=256) & ((x-r-1)>=1) & (Z(x-r,y+r)>=Z(x-r-1,y+r+1))
    d4_count=d4_count+1;
    r=r+1;
end
r=0;

```

%The criteria is ≥ 2 , as an asperity is defined as a point where it is higher the two
 %consecutive points in any single direction

```

if xm_count>=2 & xp_count>=2 & ym_count>=2 & yp_count>=2 & d1_count>=2 &
d2_count>=2 & d3_count>=2 & d4_count>=2
    peak_counts=peak_counts+1;

```

%begin of radius calculation

```

delete dummy1.txt
delete dummy2.txt
dum1=1;      %dum1 is an temp counter

while dum1<=xm_count & dum1<=15
    dummy1=[dum1*3.92 Z(x-dum1,y)];

```

% dummy1 is matrix to store points involve in the computation. 3.92nm is interval
% between each scan point.

```
fid=fopen('dummy1.txt','a');
```

%dummy1.txt is file to store the points in the computation

```
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=xp_count & dum1<=15
    dummy1=[dum1*3.92 Z(x+dum1,y)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=ym_count & dum1<=15
    dummy1=[dum1*3.92 Z(x,y-dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=yp_count & dum1<=15
    dummy1=[dum1*3.92 Z(x,y+dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=d1_count & dum1<=15
    dummy1=[dum1*sqrt(2*(3.92^2)) Z(x-dum1,y-dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=d2_count & dum1<=15
```

```

    dummy1=[dum1*sqrt(2*(3.92^2)) Z(x+dum1,y+dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=d3_count & dum1<=15
    dummy1=[dum1*sqrt(2*(3.92^2)) Z(x+dum1,y-dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end
dum1=1;

while dum1<=d4_count & dum1<=15
    dummy1=[dum1*sqrt(2*(3.92^2)) Z(x-dum1,y+dum1)];
    fid=fopen('dummy1.txt','a');
    fprintf(fid,'%8.4f %8.4f\n',dummy1);
    fclose(fid);
    dum1=dum1+1;
end

```

%Y is matrix to read in data from dummy1.txt
 %The maximum number of points in consideration is set to 15

```

Y=textread('dummy1.txt','','delimiter',' ','emptyvalue',NaN);

if xm_count>=15
    dum_xm=xm_count-15;
    xm_count=xm_count-dum_xm;
end

if xp_count>=15
    dum_xp=xp_count-15;
    xp_count=xp_count-dum_xp;
end

if ym_count>=15
    dum_ym=ym_count-15;
    ym_count=ym_count-dum_ym;
end

if yp_count>=15
    dum_yp=yp_count-15;
    yp_count=yp_count-dum_yp;
end

if d1_count>=15

```

```

    dum_d1=d1_count-15;
    d1_count=d1_count-dum_d1;
end

if d2_count>=15
    dum_d2=d2_count-15;
    d2_count=d2_count-dum_d2;
end

if d3_count>=15
    dum_d3=d3_count-15;
    d3_count=d3_count-dum_d3;
end

if d4_count>=15
    dum_d4=d4_count-15;
    d4_count=d4_count-dum_d4;
end

```

%dum2 represents total number of points per case

```

dum2=xm_count+xp_count+ym_count+yp_count+d1_count+d2_count+d3_count+d4_count;
rad=350;      %rad as counter for radius

```

```

while rad>=1
    theo_x=0;
    sum_sq_err=0;
    dum4=rad-Z(x,y); %dum4 set as offset for Z scale
    X=Y(:,2)+ dum4; %X holds the 'total' Z value
    for p=1:dum2
        theo_X=sqrt((rad^2)-(X(p,1)^2));
        sum_sq_err=sum_sq_err+((Y(p,1)-theo_X)^2);
        p=p+1;
    end
    dummy2=[rad sum_sq_err];
    fid=fopen('dummy2.txt','a');
    fprintf(fid,'%12.4f %12.4f\n',dummy2);
    fclose(fid);
    W=textread('dummy2.txt','','delimiter',' ','emptyvalue',NaN);
    rad=rad-1;
end

for h=2:350
    if W(h,2)>W(h-1,2)
        radius=W(h-1,1);
        break
    end
end
end

```

%end of radius calculation

```
%display1 set as matrix to display data after each run
    display1=[x y Z(x,y) xm_count xp_count ym_count yp_count d1_count d2_count
    d3_count d4_count radius];
    fid=fopen('rt1.txt','a');

    fprintf(fid,'%6.2f %6.2f %8.4f %3.1f %3.1f %3.1f %3.1f %3.1f %3.1f %3.1f
    %6.2f\n',display1);
    fclose(fid);
end
k=0;
xm_count=0;
xp_count=0;
ym_count=0;
yp_count=0;
d1_count=0;
d2_count=0;
d3_count=0;
d4_count=0;
end

x=x+1;
end
display2=textread('rt1.txt','delimiter',' ','emptyvalue',NaN); %display2 set as matrix to
display summary
disp(peak_counts)
```

Appendix B: Statistical Nature of Rough Surfaces

It was shown in section 3.2.1 that the asperity heights of a rough surface follow a Gaussian distribution profile. However, different rough surfaces may have different height distributions such as Gaussian, Exponential, Setch, Lorentzian, Power Law, Bessel and Self-Affine (Fractal) [78]. The most commonly used height distribution is the Gaussian height distribution,

$$f(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right), \quad (\text{B.1})$$

due to its mathematical simplicity and its place in the central limit theorem of the sum of large amount of random variables. This section would examine if surface heights, like the asperity heights, would also follow a Gaussian distribution profile.

There are several surface characteristics to be ascertained if the surface is to be Gaussian. Some of these characteristics include:

- Probability distribution function of the surface heights must be a bell curve and symmetrical about the mean, as in Figure 3-5, and
- The height-height correlation function $H(x)$, which is to study the properties of random surfaces [78], would have the form and expression as shown in Figure B-1.

Figure B-1: A theoretical height-height correlation function of a Gaussian surface.

The distribution of surface heights is shown in Figure B-2 and results showed that the surface heights can be fitted well into a Gaussian distribution function and is symmetrical about the mean surface height which is referenced to zero.

Figure B-2: Typical surface height profiles, following a Gaussian distribution.

As implied in sections 3.1 and 3.2, the surface image lattice can be represented by a 256 x 256 mesh as shown in Figure B-3, and the height-height correlation function can be expressed as equation (B.2)

Figure B-3: The coordinates for the image mesh used in the model.

$$H(r) \approx H(p) = \frac{1}{256(256-p)} \sum_{l=1}^{256} \sum_{n=1}^{256-p} [h(p+n, l) - h(n, l)]^2 \quad (\text{B.2})$$

The height-height correlation function $H(r)$ was applied to the Cu surfaces and the plot of $H(r)$ versus r is shown in Figure B-4.

Figure B-4: A plot of height-height correlation function of a Cu surface. The dashed line represents an extrapolation as r increases.

The plot in Figure B-4 has the same form as the theoretical curve in Figure B-1. However, there are insufficient data points to establish the trend further and hence we could only propose what the curve might eventually turn out as represented by the dashed line.

Curve fitting of the experimental data (to the expression in Figure B-4) shows that parameters a , b and d would vary randomly and did not conform to expectations to that of a Gaussian surface. This does not indicate that the surface is non-Gaussian but rather it is not conclusive from a height-height correlation perspective. If there were more data points, the trend (and expression) could then be better established [79].

Appendix C: Algorithm code for identifying the contacting asperities

```

Z=textread('s23m2.txt','','delimiter',' ','emptyvalue',NaN); %Z is matrix to read in data from
AFM
Y=textread('rs23m2.txt','','delimiter',' ','emptyvalue',NaN); %Z is matrix to read in data from
AFM
X=textread('F127.txt','','delimiter',' ','emptyvalue',NaN); %Z is matrix to read in data from
AFM

delete elec.txt;

load=6000;
dprime=0;
tdprime=0;
count=0;
Noma=0.006; % nominal contact area
R=50e-9; % curvature radius
eta=4.4e14; % density of asperities
flag=1;
n=1;
sigmas=4.063e-9; % standard deviation of asperity heights

F=load/22.72/(R^0.73)/(sigmas^1.27)/50e6/eta/Noma/(50^0.54);

while flag==1
    if F/X(n,3)<1.02
        n=n+1;
    else
        flag=0;
    end
end

tca=2*3.142*R*sigmas*eta*Noma*X(n,2); % calculate true contact area

d=X(n,1)*sigmas;
f=length(Y);

for m=1:f

    x=Y(m,1);
    y=Y(m,2);
    n1=x-Y(m,4);
    n2=x+Y(m,5);
    n3=y-Y(m,6);
    n4=y+Y(m,7);
    d1=Y(m,8);
    d2=Y(m,9);

```

```

d3=Y(m,10);
d4=Y(m,11);

if Y(m,3)*1e-9>d
a=[Z(n1,y),Z(n2,y),Z(x,n3),Z(x,n4),Z(x-d1,y-d1),Z(x+d2,y+d2),Z(x+d3,y-d3),Z(x-
d4,y+d4)];
g=min(a);

if g>0
    tdprime=(tdprime+(d-g));
end
if g<0
    tdprime=(tdprime+(d+abs(g)));
end
count=count+1;

end
end

if count>0
    dprime=tdprime/count*1e-9;
    l=d+dprime;
else
    l=d;
end

scr=16.78e-9*l/tca*Noma*10000; % calculate specific contact resistance

display1=[load scr count d dprime l]; %display1 set as matrix to display data after each run
fid=fopen('elec.txt','a');
fprintf(fid,'%6.2f %e %3f %e %e %e \n',display1);
fclose(fid);

```