MULTI-SCALE MODELING OF THE SLURRY FLOW AND THE MATERIAL REMOVAL IN CHEMICAL MECHANICAL POLISHING

By

KUIDE QIN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2003

Copyright 2003

by

Kuide Qin

This study is dedicated to my parents, QIN Shengchang and SUN Rongmei.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude and appreciation to Dr. Chang-Won Park for his support, inspiration and guidance in directing this study. His hard work and attention to detail were an excellent example during my studies. He always helped me to learn more and taught me the importance of thinking creatively. Many of the advances achieved during my graduate study would not have been possible without his support.

I would also like to gratefully acknowledge Dr. Brij M. Moudgil, who has set a great example of lifetime success in research and engineering, for his valuable support and guidance throughout this study. His sincere dedication to science and discipline in conducting research made significant contributions to this dissertation.

Moreover, I wish to thank Dr. Richard B. Dickinson, Dr. Jason Weaver, and Dr. Hassen El-Shall for their time, useful discussions and guidance as members of the supervisory committee.

I would like to acknowledge the financial support of the National Science Foundation's Engineering Research Center for Particle Science and Technology. ERC gave me an excellent opportunity to develop and enhance my knowledge of particle science by providing short courses and valuable seminars.

My colleagues, Lav Agarwal, Scott Brown, Won-Seop Choi, Madhavan Esayanur, Ron Sabo, Pankaj Singh, Ivan Vakarelski and Suresh Yeruva, who have contributed to this research by their valuable discussions and friendship, are also gratefully acknowledged.

Finally, I would like to thank my wife, Chenchen Qiu, for her unselfish support and understanding throughout my doctorate work, which would not have been possible without her. I would also like to thank my parents and my brother for their encouragement and inspiration. Even though they were half a planet away from me, they have always been with me in my heart.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MULTI-SCALE MODELING OF THE SLURRY FLOW AND THE MATERIAL REMOVAL IN CHEMICAL MECHANICAL POLISHING

By

Kuide Qin

May 2003

Chair: Chang-Won Park
Cochair: Brij M. Moudgil
Major Department: Chemical Engineering

Rapid advances in the microelectronics industry demand continuously a decrease in the device sizes to produce faster and more functional processors, which results in a stringent requirement of global planarization across the die. Chemical mechanical polishing (CMP) is widely adopted in achieving excellent local and global planarization for microelectronic device manufacturing. It has been demonstrated experimentally that the polishing performance is a result of the synergetic effect of both the chemicals and the particles involved in CMP. However, the fundamental mechanisms of material removal and the interactions of the chemical and mechanical effects are not well understood, which limits a better control and improvement of the CMP process.

In this study, the CMP process is modeled in different length scales based on a systematic study of the pad-particle-wafer interactions during polishing. One of the main problems in CMP is polishing non-uniformity caused by wafer-scale slurry transport. In

this investigation, therefore, the delivery of the slurries by the polishing pad onto the to-be-polished wafer surface is first studied by solving the Navier-Stokes equation between two eccentrically rotating disks. It is shown that a similar rotational speed of both the wafer and the pad is necessary for a uniform delivery of the polishing slurry. In the feature scale, the shearing effect of the slurry flow on different shapes of features is simulated using the finite element method. Although stress concentration is found always in the corners of the features, no direct evidence about feature breakage off the wafer surface by the slurry flow is seen. To investigate the mechanism of material removal, a micro contact model is developed by combining the chemically formed surface layer with mechanical abrasion. Most of the variables involved in the CMP process, such as solid loading, particle size and distribution, pad modulus, asperity size and distribution, down pressure and rotational velocity, are all included in the proposed model. It provides a reliable tool in predicting the effects of these variables, particularly the properties of the surface layer. Furthermore, the model shows that a balance between the chemical effect and the mechanical effect has to be achieved for optimal slurry performance. Finally, the correlation between the material removal and the coefficient of friction is quantified. The results of this development show that a relatively higher friction coefficient between the abrasive particle and the wafer surface is needed for higher polishing rate. The reported findings in this study provide basic understandings of the CMP process in multiple length scales. A slurry design criterion can be developed to achieve optimal polishing performance.

CHAPTER 1
INTRODUCTION

The continuous demand in the semiconductor industry for developing faster and more functional processors has motivated the fabrication of ultra large scale integrated (ULSI) circuits with smaller feature size, higher resolution, denser packing and multi-layer interconnects. The ULSI technology in turn results in stringent demands on global planarity on the inter level dielectric (ILD) layers and shallow trench isolation (STI). Chemical mechanical polishing (CMP) is superior to other planarization technologies in producing excellent local and global planarization at low cost, and thus is widely adopted in many back-end processes for planarizing multi-layer interconnects. The main objectives of CMP process are to smooth surface topography of dielectric deposits to enable multi-level metalization (MLM) and to remove excess coating material to produce inlaid metal damascene structures and shallow isolation trenches. In a typical CMP process, a rotating wafer is pressed against a rotating polishing pad in the presence of slurry in between. The wafer and the pad, both of a circular shape, rotate typically in the same direction but eccentrically. The slurry is an aqueous suspension of colloidal abrasive particles. The slurry also contains chemical reagents, which are used to create an easily removable surface layer on the material to be polished. The surface layer is then removed by mechanical action of the abrasive particles. The planarization is achieved in CMP by maintaining high removal rates at high surface areas and low material removal rates at low surface areas. Consequently, the polishing pad used in regular CMP is made

from relatively rigid polymeric materials (e.g., polyurethane), and it does not conform to the wafer surface. Thus, less force is exerted on the low areas.

So far, the CMP process has been used to polish several types of materials such as SiO_2, Si_3N_4, tungsten, and aluminum. The advents of copper interconnect technology and the introduction of low dielectric constant materials have brought more challenges in the studies of CMP. It is well recognized empirically that many variables such as applied normal force (or down pressure), relative velocity of wafer to the pad, pad properties (elastic modulus, hardness, etc.), and slurry characteristics (solids loading, chemical concentration, particle size and size distribution, etc.), have profound influences on the general effectiveness of the CMP process [Ste97]. However, due to the complexity of CMP by concurrent polishing of multiple materials and lumped parameter conditions, the fundamental polishing mechanisms underlying the process are not yet well understood. Many process design options (e.g., slurry optimization and pad properties) remain in a trial-and-error procedure due to the lack of reliable models. To date, the modeling of the material removal process during CMP appears to have taken two different types of approaches based on the two extremes in dealing with the interactions between the pad and the wafer. One is a purely *fluid mechanical approach*, in which the wafer and the pad are assumed to be separated by a continuous fluid layer of slurry and the material removal is viewed as a consequence of erosion and/or corrosion by the fluid action. The studies of Runnels *et al.* [Run94a, Run94b, Run96] belong to this category in which the shear and normal stresses induced by the lubrication-type of slurry flow were analyzed. These are wafer-scale macroscopic analyses that may provide useful information about the influence of shear and normal stresses on the removal rate. However, they are

incapable of delineating the effects of slurry particles and pad properties including the surface topology despite their significant influence observed experimentally [Mah00]. The other is the approach based on *contact mechanics*, in which the material removal is attributed to the abrasive wear of the wafer surface by the action of the slurry particles directly in contact with the wafer surface mediated by the pad. Since this approach appears to be physically more plausible in describing experimental observations, it is widely accepted by many researchers. However, chemical effects were not considered in most of these attempts despite experimental evidences that chemical effects are significant in determining the material removal rate and the quality of the surface finish. Chemical effects specific to a certain system are evident in various experimental studies. In glass polishing, the polishing rate is observed to be nearly zero if liquid hydrocarbon or liquid without hydroxyl group is used [Coo90]. Surface quality is also observed to be poor when such liquids are used. Bielmann [Bie98] reported that the polishing rate of tungsten obtained with a slurry not containing an oxidizer was negligibly small compared to the polishing rate with a slurry containing an oxidizer. In polishing a copper surface using a slurry containing $Fe(NO_3)_3$, Luo [Luo99] observed a dramatic change in the polishing rate as the $Fe(NO_3)_3$ concentration varied slightly. Such observations cannot be explained by the changes in the surface charge because the surface charge effect is very small. In fact, significant changes in the surface properties of the wafer exposed to the chemicals in the slurry are expected to play a significant role in the CMP process. Thus, the actual material removal mechanism in CMP has to be a result of the combinations of both chemical and mechanical effects.

In this study, fundamental understandings of the mechanics of material removal during CMP are established based on the modeling of wafer-pad-particle interactions. Optimization scheme of high performance slurries and operation parameters to reduce wafer scale variation can be developed on the basis of this study. In addition, the value of the experimental and modeling efforts of current study can provide guidelines for the design of novel polishing processes and for the identification of unexplored process parameters. A synopsis of the efforts constituting this study is organized as follows.

Chapter 2 presents a general review of the CMP process. First, the background of applying CMP in integrated circuits (IC) manufacturing is discussed. This includes a brief description of multilevel metalization (MLM) and other planarization techniques used in the microelectronics industry prior to CMP. The benefits of CMP are concluded from the comparison of these techniques. Then, the CMP process and its three major components, namely surface to be polished, the polishing pad and the polishing slurry, are introduced in detail. The main properties of these components and their impacts on the CMP performance are summarized. In addition, the motivation of introducing new materials to microelectronic devices and the consequent challenges in CMP are also addressed. Furthermore, the complexity of the CMP process is revealed through a brief coverage of the input and output parameters that control the CMP performance. Previous findings related to these parameters are discussed accordingly. Finally, the modeling efforts in predicting polishing performance are summarized. The models are classified into two categories and discussed in terms of their effectiveness and weakness in evaluating particular output parameters of the CMP process. New approaches for developing more effective predictive methodologies are introduced.

Chapter 3 focuses on wafer-scale slurry flow simulation. The flow pattern and shear rate generated by the slurry flow are obtained from hydrodynamic calculations by assuming that the wafer and the pad are flat surfaces and are parallel to each other. According to the simulation results, the fluid stagnant region under the wafer is discussed. The shear rate non-uniformity is defined and compared with the experimental WIWNU (within wafer non-uniformity) results.

Chapter 4 investigates the impact of slurry flow on the stress distribution of different features. The results from feature-scale modeling of both hydrodynamics and solid mechanics, in which the governing equations are solved using finite element method, are presented.

Chapter 5 presents a micro-contact model for the material removal rate. The model includes both the chemical effect due to the reaction between the wafer and slurry additives, and mechanical effect due to particle abrasion. The interactions of the polishing pad, slurry particles and the wafer are investigated based on reasonable simplifications, and then verified with experimental results. At the end of this chapter, the predictions of the proposed model are compared with experimental observations. The guidelines for slurry design and optimization of the CMP process control based on the model predictions can thus be derived.

Chapter 6 is a continuation of the modeling efforts of Chapter 5. However, instead of the indentation of abrasive particles into the wafer in normal direction, which is considered in previous chapters, the shear strength of the wafer surface, which is measured by friction in the lateral direction, is measured. The particle abrasion is incorporated into the material removal rate by the measurement of the coefficient of

friction. Since the coefficient of friction is a function of particle indentation depth, which in turn is related to the surface layer property, the variables needed for the model prediction (e.g., particle indentation depth) are derived through the friction coefficient measurement. The model predictions are compared to the experimental polishing data, and a good agreement is obtained.

Chapter 7 summarizes the findings of this study with the suggestions for future work.

CHAPTER 2
BACKGROUND AND LITERATURE REVIEW

Background

Impact of Interconnect Delay on the Performance of Integrated Circuits

The ever-increasing demand in the semiconductor industry for high performance microelectronics has resulted in fabrication of increasingly complex, dense and miniaturized devices and circuits [Ste97]. As the device dimensions are scaled down, the integrated circuits (IC) face the interconnect delay problem associated with the capacitance (C) and resistance (R) of the metal lines. The interconnect delay is usually measured as RC time delay, which is the time it takes for the voltage at one end of a metal line to reach 63% of its final value when a step input is presented at the other end of the line [Ste97]. The RC time delay is given by the product of the total capacitance and the total resistance of the metal lines. Consequently, higher RC value leads to slower device operating speed. The expression for RC time delay is given as

$$RC = \rho\varepsilon\frac{l^2}{td} \tag{2.1}$$

where ρ, l, and d are the resistivity, length and thickness of the metal line, and ε and t are the relative permitivity and thickness of the insulator respectively. The relationship shows that ρ, ε and l must be decreased, and d and t must be increased in order to reduce the RC time delay. Practically, although the metal line thickness (d) and the dielectric thickness (t) can be reduced when the feature sizes are scaled down, the line length (l) remains

constant due to the chip complexity. Therefore, the RC delay increases as the dimensions (i.e., *t* and *d*) are reduced.

Table 2.1 Options for low-κ dielectric materials

Dielectric Material	**Dielectric Constant (κ)**	**Important Notes**
Silicon Dioxide (SiO_2)	3.9	IC Processing standard
Fluorinated Oxide (FSG)	3.5-3.7	Fluorine interaction issues
Spin-on Glasses (SOG) and Sisesquioxane (HSQ, MSQ)	2.9-3.6	Liquid spin-on deposition, κ varies with H, C content
Carbon-doped Glasses (OSG)	2.7-3.0	Deposition by CVD, κ varies with H, C content
Polymers	2.6-3.2	Mechanically and thermally less stable than oxides
Fluorinated Polymers	2.1-2.3	Adhesion issues and fluorine interaction issues
Porous Media	1.5-3.0	Mehanically weaker than non-porous, can be any porous low-κ (SiO_2, MSQ, polymer)

Source: [Mur00]

From Equation 2.1, the RC delay is directly proportional to the resistivity of the metal line and the dielectric constant of the inter-layer dielectric (ILD). By switching from oxide based ILD to a low ε ILD material, significant gains may be obtained in decreasing the interconnect delay. These materials are commonly referred to as low-κ materials since $\varepsilon = \kappa\varepsilon_o$. Several inorganic and organic materials have been proposed for use as low-κ ILD (Table 2.1). Currently the material used mostly as ILD is vapor deposited SiO_2 with κ between 3.9 and 6.0 depending upon water concentration in SiO_2. By doping the SiO_2 with fluorine or boron and lowering the water content, κ may be reduced to 3.0-3.9. Organic materials such as polyimides, parylenes, fluoro-polymers and teflon exhibit even lower κ and are excellent candidates for low-κ ILD material.

Besides the selection of low-κ dielectric materials, the RC delay can also be reduced by the introduction of metals with low resistivity. In fact, the conventional Al/SiO_2 materials set in current chip manufacturing is being replaced by the Cu/low-κ materials set. These material changes can result in dramatic reduction of the RC time delay. Figure 2.1 shows the comparison between the Cu/low-κ materials set and the Al/SiO_2 materials set on the RC time delay. Using the Cu/low-κ materials set makes it possible to reach smaller device generation at which the interconnect becomes limiting to chip speed.

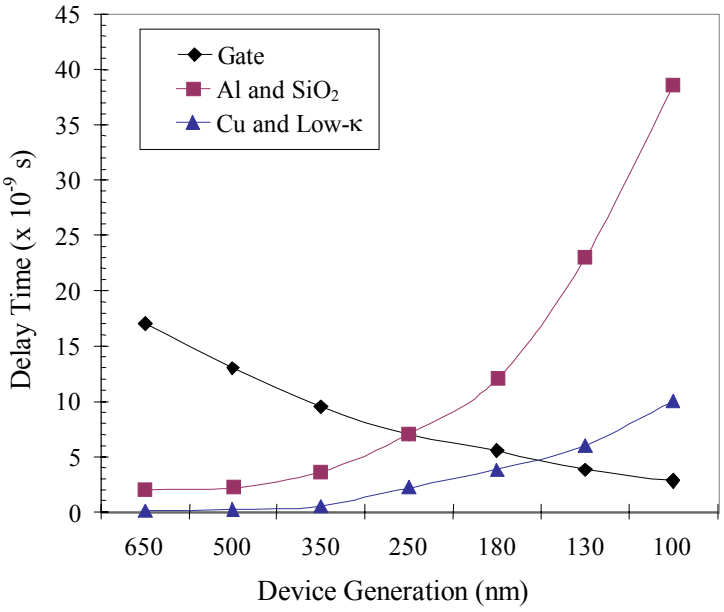


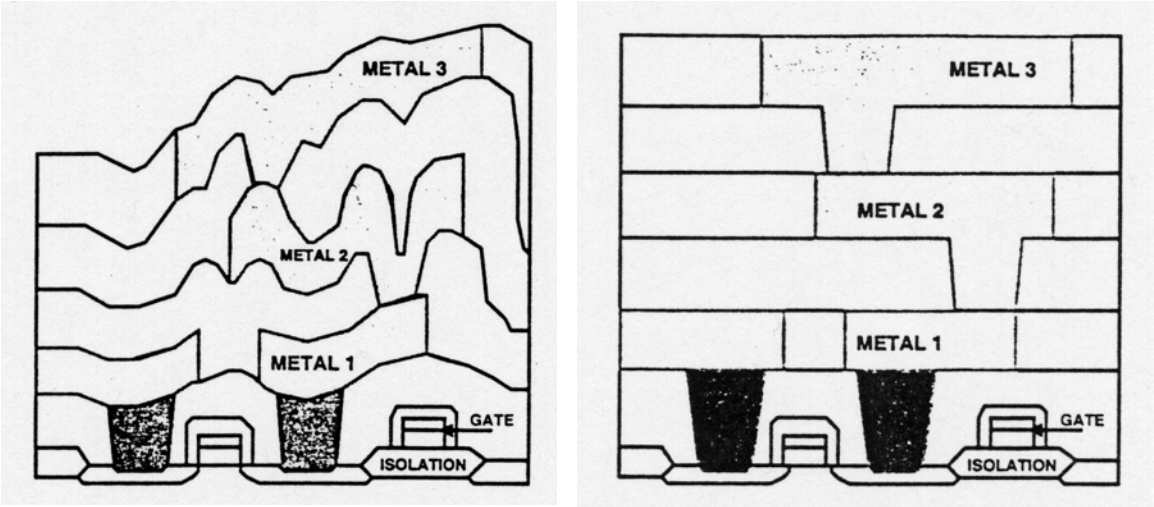
Figure 2.1 Interconnect delay as a function of device generation [Mur00]

Decreasing the line length, *l*, has a dramatic impact on RC delay since it increases with the square of *l*. So to decrease RC delay, long metal lines should be avoided while the width of the line-to-line interconnections should be increased to reduce line-to-line capacitance. Multilevel metallization helps in achieving both.

Multilevel Metallization (MLM)

In a multilevel metallization (MLM) scheme, the metal interconnections are not confined to one plane; they span several planes, isolated by dielectric layers, and are interconnected by wiring in the third dimension through holes in the dielectric planes called "vias." A cross-section of the MLM scheme is shown in Figure 2b. MLM reduces interconnect delay by several means. First, by increasing the number of levels of metal, the packing density of metal lines does not need to keep pace with the packing density of the gate level. MLM avoids long interconnections and frees the space for devices. Second, multiple levels of metallization reduce the length of interconnections by allowing for more direct routing. Finally, in instances where long metal lines cannot be avoided, MLM allows for the possibility of routing these lines at the upper levels which maintain even wider dimensions to prevent large delay times on the long lines [Ste97].

The processes associated with making MLM structures are deposition, exact lithography, etching, and planarization. Particularly, following processes are required to fabricate MLM structures: (1) planarization to flatten the surface, (2) etching the via holes without affecting the underlying metal or increasing the via diameter, and in some cases (e.g., dual damascene process) etching the interconnect wiring channels in the dielectric so that the metal can be deposited into such channels to form interconnections, (3) filling the holes (vertical interconnections) with metals having negligible contact resistance between metals, (4) deposition and patterning of metal into interconnection lines with good line width control, (5) deposition of dielectrics on the metal and planarization if needed, and (6) creation of test structures to evaluate the MLM process effectiveness.



(a) (b)

Figure 2.2 Comparison between (a) a MLM structure without planarization and (b) a planar structure with CMP [Wol99]

To achieve MLM, it is critical that each level is flat so that the lithography and patterning can be precise to allow vertical interconnections to be made. The decreasing device dimensions and the increasing number of metal layers further necessitate the improved planarization of the surfaces for MLM, since non-planarity of the surface becomes cumulative as the layers are built one on top of another. A schematic comparison between a non-planar structure and a planar MLM structure is shown in Figure 2.2. High topographic variation in the non-planar structure leads to poor step coverage and electromigration problems. Moreover, as the circuit dimensions are scaled down to below the sub-0.5 μm regime, the lithography tools require high numerical aperture lenses to print fine line dimensions. The depth of field of these lenses is approximately 270 nm across a 27 by 27 mm stepper field. Due to this finite depth of field and equipment tolerance, the optical steppers require less than 150 nm variation in topography across the stepper field. The uneven surface topography can also cause

photoresist thickness variation, which in turn results in overexposure in the area of thinner resist. Furthermore, the variation in dielectric film thickness across the wafer requires the etching of vias of different depths, leading to more variability and less process yield. On the other hand, a planarized dielectric film will result in vias of the same depth in all metal levels. Therefore, global planarization of the wafer surface must be achieved for MLM in present semiconductor manufacturing technology.

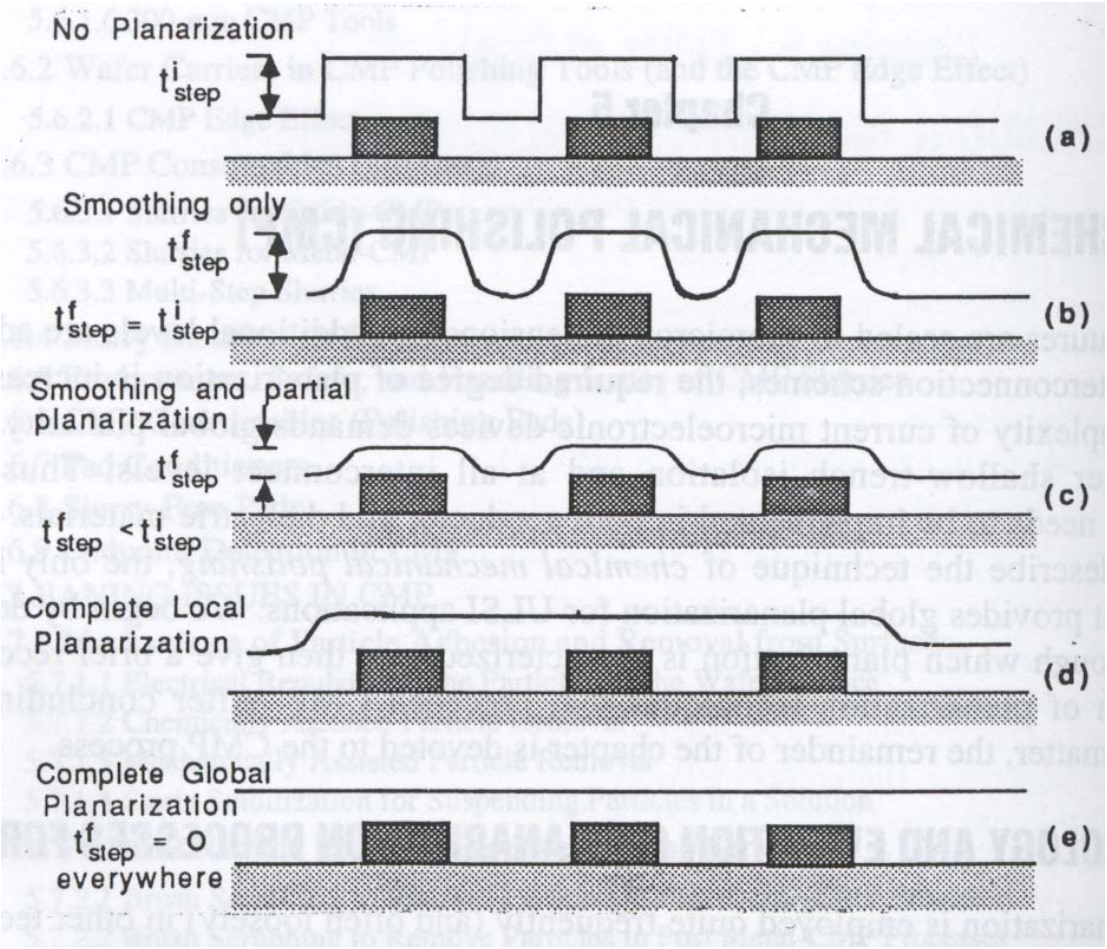


Figure 2.3. Degrees of surface planarity [Wol86].

Planarization Technologies

Several degrees of surface planarity are illustrated in Figure 2.3. The three main scales of planarity are (1) surface smoothing, (2) local planarity, and (3) global planarity.

Surface smoothing refers to the case in which feature corners are smoothed and high aspect ratio trenches or vias are filled. Local planarity refers to the case in which surfaces are flat locally (about 10μm) but the overall surface height may vary across the die. In global planarization, on the other hand, the surface is flat across an entire field (about 20mm). The approaches have been used for achieving planarization include [Oma94] (1) oxide reflow, (2) deposition and etch-back (Dep EB), (3) resist etch back (EB), (4) spin-on glass etch back (SOG EB), (5) electron cycloton resonance (ECR), and (6) chemical mechanical polishing (CMP). Figure 2.4 shows the planarizing ability of these planarization techniques. The techniques will be briefly discussed as below and later more details will be given for CMP.

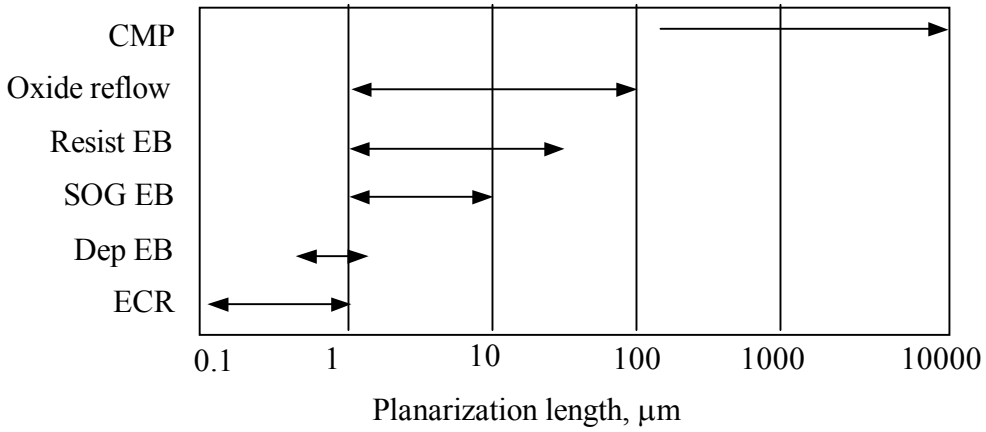


Figure 2.4 Planarizing ability of different planarization techniques [Oma94]

Oxide reflow is one of the simplest processes to achieve leveling of the wafer surface. In this process, phosphorus and/or boron are doped onto a deposited oxide layer such that a glass with lower softening point (about 800^{o}C) than that of silicon dioxide (about 1200 ^{o}C) can be formed. At the softening point, the glass flows and rounds the contour of the top surface. Most IC processes include oxide flow for the first oxide layer.

However, it cannot be used when a metal with a lower melting point has been deposited on the surface. As shown in Figure 2.4, the planarization length of oxide flow is only a few tens of microns. For surfaces over this distance, the surface will be undulate if oxide flow is used as the planarization technique.

Deposition and etch back starts with a deposited thick SiO_2 layer, and then the thickness is reduced by either reactive ion etch (RIE) or sputter etch. In sputter etch, the corners in the SiO_2 film etch faster than the rest of the film because the sputter yield of SiO_2 is greatest at the angle of about 45^o. The planarization length of this method is about 1 micron and it does not planarize a large area.

Resist etch back has two steps. First a photoresist layer is spin-coated onto the oxide, which results in a planar top surface after curing. Then, etching is carried out over the entire wafer surface. The etch process is adjusted such that the etch rates for phtoresist and for oxide are the same. Finally, this process produces an oxide surface that is more planar than it was before. The planarization length of this method is a few tens of microns.

Spin-on glass etch back method starts with an initial CVD SiO_2 deposition, which is spin-coated with the viscous SOG flow. It fills the spaces in the SiO_2 surface, leaving a planar surface. After cured, the SOG/CVD SiO_2 composite is etched back to the desired interlayer dielectric thickness. The planarization length of this method is about a few tens of microns.

Electron cyclotron resonance is essentially a high density plasma oxide deposition method, in which an etching process proceeds simultaneously with deposition. Rates of deposition and etching can be adjusted more or less independently. This method

can be used to fill the gaps of the 2:1 aspect ratio at submicron dimensions, indicating that ECR is an important tool for deep submicron manufacturing. Since its planarization length is small (i.e., less than 1 micron), the planarization is really intra feature.

Chemical mechanical polishing offers more versatility, simplicity and global planarization, and can be applied to all type of materials (e.g. SiO_2, Cu, W, and Al). In the CMP process, the surface to be planarized is pressed against a rotating polishing pad soaked with abrasive slurry. Higher pressure is applied on the raised features, causing a polishing action that is much faster than the lower points on the surface. Eventually, a flat, planar surface results. The process combines aqueous corrosion chemistry and contact mechanics. The planarization length is more than 100 micron.

For all the planarization technologies introduced above, CMP is the only one that can planarize the surface globally to meet the requirements in multilevel metallization. It will be given in greater detail in the next section.

The CMP Process

The CMP process was first introduced by IBM to the development and manufacture of CMOS (complementary metal-oxide-silicon) products in 1985. Its popularity is due to its unique ability in providing global planarization and fewer steps than other planarization methods. It creates an alternative to dry etching for materials like copper, which are not amenable to dry etch [Ste97]. As shown in Figure 2.5, a typical CMP process consists of a polishing pad affixed to a circular polishing table and a carrier to hold the wafer against the pad. Both the carrier and the table are rotating, usually in the same direction and with the same speed, while an aqueous slurry with colloidal particles is dispended onto the pad. Despite its advantages and wide usage, CMP has its own challenges, including (1) an entire new set of metrology tools for its integration into the

IC manufacturing and (2) inadequate understanding of process fundamentals. As a result, CMP process control and stability have not reached the level of maturity of other semiconductor processing technologies such as photolithography and plasma etching. To have a basic understanding of its polishing and planarization process, roles of the three major components of CMP, namely, the surface to be polished, the polishing pad and the slurry, are outlined below.

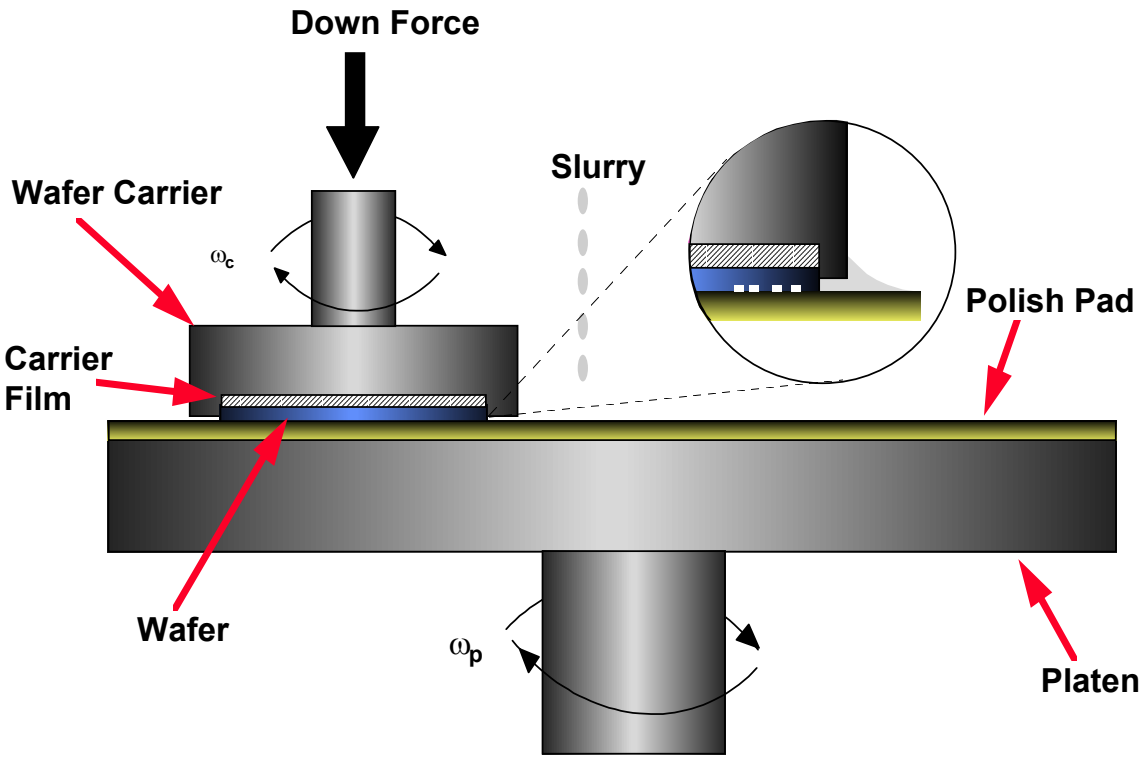


Figure 2.5. Schematic representation of chemical mechanical polishing (CMP) process.

Surface to be Polished

The initial CMP process was applied mainly on the SiO_2 and tungsten layers. However, due to the requirement of reducing the RC time delay, more and more new materials have been introduced to the microelectronic industry. So far, the CMP has been expanded to the planarization of a large variety of metals such as Al, Cu, Ta, Ti, TiN, W

and their alloys and insulators such as SiO_2, Si_3N_4, polysilicon and polymeric low-κ materials.

The most commonly conductor and insulator materials that have been used in IC manufacturing are Al and SiO_2 respectively. Recently, Cu has been used to replace Al as the conductor due to its lower resistivity. It has been well adopted at the 0.18 μm technology and is expected to be the metal of choice for the future ICs [Sin02]. In terms of the selection of low-κ ILD, polymeric materials are the promising candidates. However, several issues have to be resolved before they are incorporated into the IC manufacturing process. First, the selected polymers must be stable at the applied high processing temperatures. Secondly, they have to be mechanically stable such that they can resist to the stress induced deformation and delamination. Their anisotropy in dielectric constant and patterning are other concerns in their integration. SiLK by Dow Chemical and GX-3 by Honywell are some of the commercially available polymeric dielectric products. However, none of these materials are massively used in the IC manufacturing. The bulk of the low-κ implementation is expected at 90 nm device node, where organosilicate and spin-on polymers will dominate at κ=2.7 [Pet02b].

The introduction of new materials brings new challenges in CMP. One of the major problems in Cu CMP is dishing due to greater removal rate of copper features than that of the surrounding oxide (Figure 2.6). Another phenomenon that can reduce metal line thickness and reduce planarity is erosion. Erosion is a thinning of the ILD layer because of non-zero polish rate of dielectric layer. The CMP process, therefore, needs to be designed depending on the materials to be polished on the wafer surface.

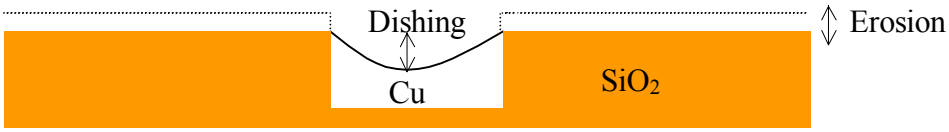


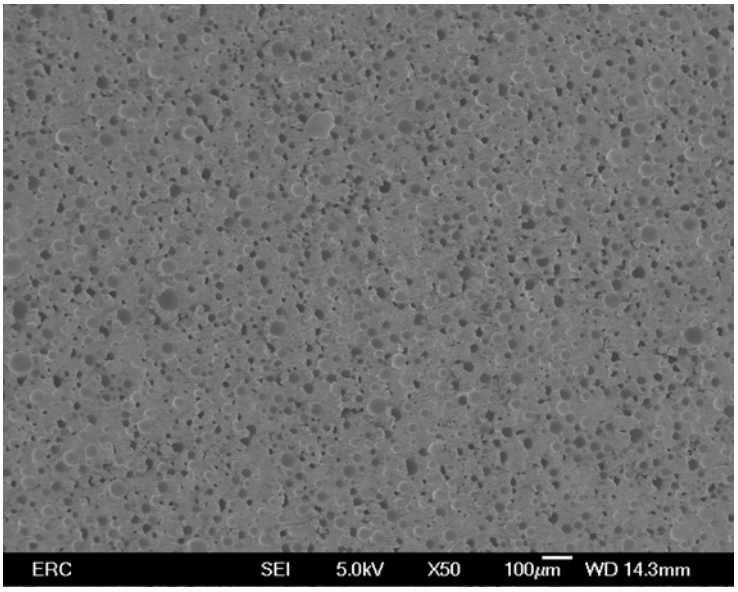
Figure 2.6. Schematic representation of copper dishing and SiO_2 erosion [Aga00].

Polishing Pad

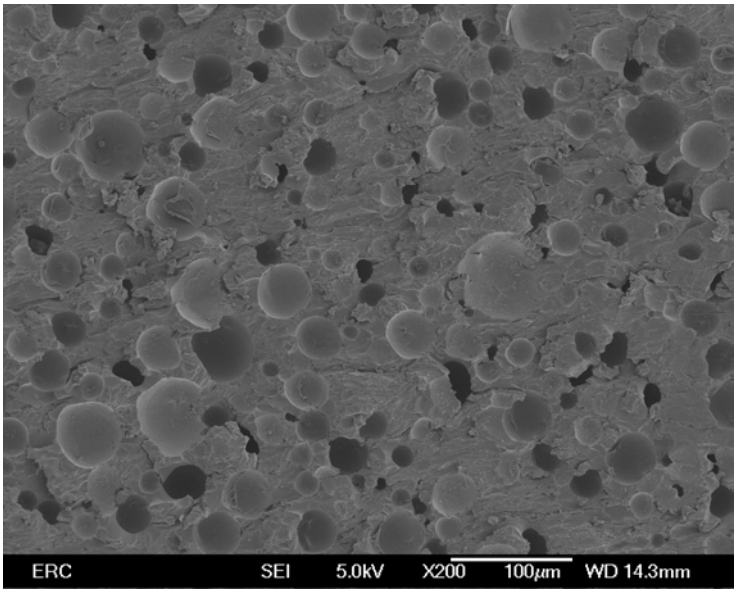
The pad is used to carry abrasives and chemicals (the slurry ingredients) all across the surface being polished and thus enable the transfer of mechanical forces to the wafer surface. An ideal pad should be rigid and chemically inert. In general, the polishing pad is mostly made from polyurethane since the urethane chemistry allows the pad characteristics (e.g., compressibility and porosity) to be tailored to meet specific needs in CMP [Jai94]. The polishing pads are composed of either a matrix of cast polyurethane foam with filler material to control hardness or polyurethane impregnated felts. The urethane foam is a cellular polymer. Its mechanical properties can be derived in part from the cell matrix formed during its manufacture and in part from its intrinsic polymer properties. In addition, filler material may be added to polyurethane to improve its mechanical properties. The surface structure of a commercially available pad (IC1000 from Rodel Inc.) is shown in Figure 2.7, which has a cell size of about tens of microns. The closed cell structure of the pad surface is important because it helps the slurry transport. It has been measured that the pad surface height is not uniform. Instead, it usually has a Gaussian distribution [Yu94]. Because the pad surface is mostly occupied by the closed cells and the surface asperities do not have the same height, only a small portion of the pad surface is expected to be in contact with the wafer surface at a given pressure during polishing. Therefore, the real contact pressure between the wafer and the

pad asperity is expected to be much higher than the apparent down pressure that is applied by the wafer carrier.

Due to the high contact pressure, particle abrasion, and solvent/chemical erosion during polishing, the wearing of the pad is inevitable, which in turn will change the pad surface structure and pad surface roughness. The ability of the pad to deliver slurry decreases as its structure degrades, resulting in a lower material removal rate. Thus, it is necessary to recondition and regenerate the pad surface. Conditioning techniques are performed on the pads to maintain the roughness and porosity of the pads as well as to remove polished products, used slurry and deformed pad material from the pad surface. Diamond grids are usually used as conditioners. Most polishing tools in the market include an abrasive wheel to condition the pad and the conditioning can be performed between polished wafers or in situ during wafer polishing. It is still necessary to understand these changes in pads as a function of the life of the pad.



(a)



(b)

Figure 2.7. Electronic microscopic pictures of IC-1000 polishing pad surface (a) small magnification (50x); (b) large magnification (200x).

Polishing Slurry

A typical CMP slurry is water based and contains both abrasive particles and chemical additives. Slurries provide both the chemical action through the solution chemistry and the mechanical action through the abrasive particles during polishing. High polishing rates, planarity, selectivity, uniformity, post CMP cleaning efficiency (including environmental concerns), shelf-life, and dispersion ability are the factors which need to be considered to optimize the slurry performance [Ste97]. Abrasives in the slurries play a very important role of transferring mechanical energy to the surface to be polished. The most commonly used abrasives in CMP are SiO_2 and Al_2O_3, although CeO_2 is also used for glass polishing.

Even though mechanical grinding by the abrasive particles can provide planarization, this is not desirable because of extensive surface damage after polishing.

Also, in the mechanically dominated polishing regime, the removal rate is low and it is possible to peel off the film from the device directly. The chemistry alone, on the other hand, cannot achieve planarization since most of the chemical reactions are isotropic [Ste97]. In addition, in the chemically etching dominated regime, the process is hard to control, and plug and trench recesses may form. The surface roughness may increase. Therefore, in CMP, the final surface finish has to be a result of the synergy between the chemical and mechanical interactions. Chemical components of the polishing slurries alter the properties of the top surface layer of the substrate, while the abrasive particles interact with this layer under the applied pressure to remove it. Successful polishing is believed to take place when the rates of chemical and mechanical activities are comparable in polishing operation [Ste99]. To achieve global planarity, scratch free surfaces and high removal rate selectivity over the isolation dielectric, both the chemical and mechanical effects have to be controlled.

The primary purpose to have chemicals in the slurry is to dissolve the dislodged materials from the wafer surface and move them away from the surface. Dislodged species are believed to dissolve in the slurry due to their enlarged specific surface area. In addition, the slurry chemicals should be able to ensure local planarity by forming a surface film to prevent the dissolution of material from the low area on the wafer surface where the surface film is not mechanically removed. As the ideal case, the passivation layer or surface film will be removed at a higher rate on the high protruding regions than it is on the recessed areas. A new passivation film will be formed simultaneously on the exposed areas. Again, this layer acts to prevent further passivation and removal of the underlying bulk material. In this way, the planarization is achieved because only the

materials on the high regions are removed, where the pad contacts the surface. In metal CMP, the solubility of the metal surface can be enhanced by the chemical reactions. For instance, the slurry reactions with the surface generate products that have higher solubility than that of the original surface. However, the hydrolysis reaction in oxide CMP is not expected to have a significant effect on surface solubility due to a much lower reaction rate.

The abrasive particles in the slurry are used to transfer mechanical energy to the wafer surface, breaking the bongs that bind atoms to the surface and removing atoms or clusters of atoms. This process is affected by a number of parameters such as the wafer down force, the rotational speeds of the pad and wafer, the solids loading (i.e., abrasive particle concentration), the slurry rheology, the pad surface microstructures and surface roughness, and so on. To address these issues, in the following sections, the impact of characteristics of CMP on the polishing performance is discussed.

Controlling Parameters of CMP

A general list of parameters that are involved in a typical CMP process is shown in Figure 2.8. Since the pad and slurry parameters have been described in previous sections, only major operational variables and CMP outputs will be introduced as following.

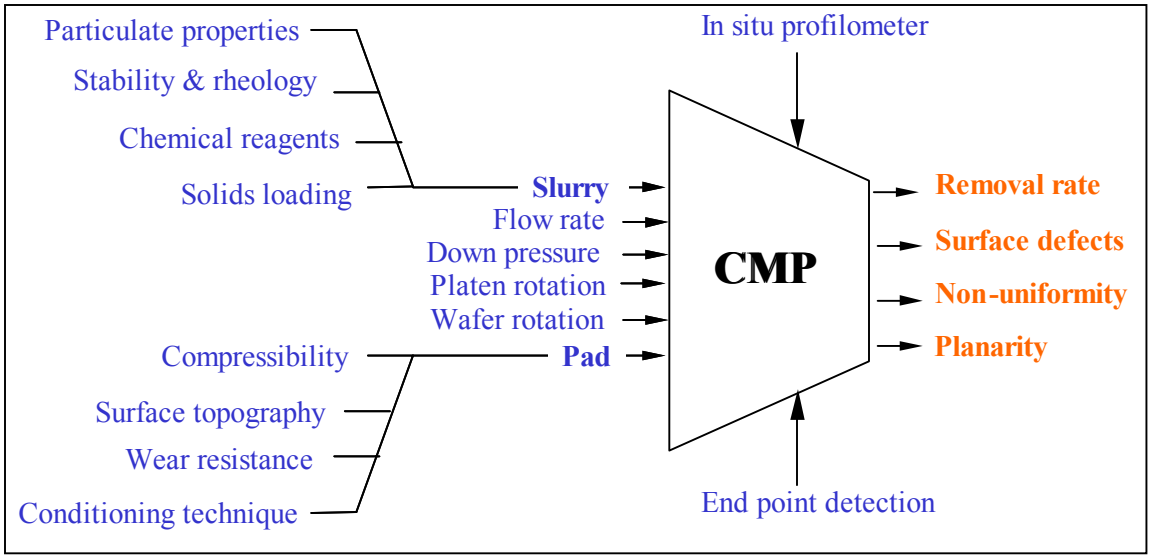


Figure 2.8 General controlling parameters in CMP

Outputs of the CMP process include material removal rate (RR) and planarity. In addition, surface defects and wafer non-uniformity (NU) are also important parameters to judge the CMP performance. The material removal rate is usually expressed in thickness per unit polishing time (e.g., nm/min or Å/min).

The planarity measures the surface topography changes over a given distance. The degree of planarity has been introduced in previous sections. The frequently used metrics of planarity include planarization relaxation distance, planarization angle, and step height ratio (SHR) [Siv92]. A scheme to measure the planarity is shown in Figure 2.9. The planarization relaxation distance is the distance R, at which the difference in height between the lowest and highest feature after planarization becomes equal to the original step height. The planarization angle is given by $\theta = \arctan(T_1/R)$. Step height ratio (SHR) is defined as the ratio of post-planarization step height T_1 to pre-planarization step height T_0 (i.e., $SHR=T_1/T_0$). SHR is zero for ideal long-range planarization and is one when there is no long-range planarization.

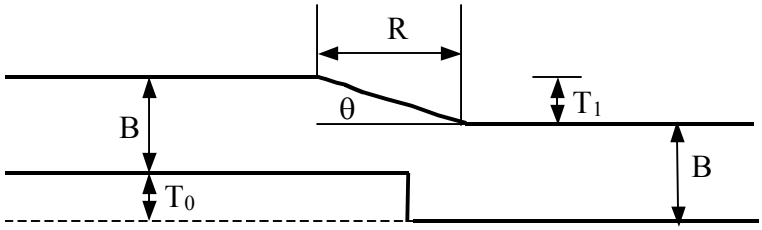


Figure 2.9 Measurement of planarity [Siv92]

Surface defects include slurry particle residuals, scratching, pitting, and delimination. These defects will become nucleation sites for electromigration failure. The micro-scratches in the oxide layer can serve as hiding places for deposited metal. Such metal is hard to remove by plasma etching and can form metal-to-metal shorts with a level [Siv92]. The decreasing device dimensions and the adoption of soft material (e.g., low-κ polymeric material) have made the formation of defects to be more critical. Slurry abrasive particles with sub-micron size can easily form large agglomerates. It has been shown that both the hard agglomerates and the soft agglomerates in the slurry system can cause detrimental scratches on the wafer surface [Bas02]. Practically, dry agglomerates can deposit on the inside wall of the slurry container when the liquid level drops, which is the primary source of over-size particles in the slurry. Another possible source is the particle size distribution, however it can be controlled during particle synthesis and filtration before making slurries.

Non-uniformity (also called within-wafer non-uniformity, WIWNU) is due to the variation of material removal rate across the wafer, which is one of the critical factors that have significant effects on the yield of IC manufacturing. The non-uniformity (NU) problem results from the slurry delivery and stress distribution on the wafer surface. As mentioned earlier, a degraded polishing pad can result in poor slurry distribution across

the wafer and lead to increased non-uniformity. Thus non-uniformity can be reduced by proper conditioning of the pad. More details about the slurry distribution and stress analysis on the wafer surface will be discussed in later chapters.

Down force (or down pressure) is one of the most important parameters that affect the material removal rate. Experimental data have shown that the removal rate increases almost linearly with the down force applied to the wafer [Pre27, Jai94]. The pad will contact with more valley areas of the topography at higher pressure. Thus, the step height ratio is expected to increase with the increasing of down pressure. The down force has no apparent effect on the wafer non-uniformity because the removal rate at the center of the wafer is almost the same as the removal rate on the edges as long as the pressure is evenly distributed across the wafer. Examples will be given in Chapter 3.

Pad and wafer rotational speeds have significant effects on RR, NU and SHR. In order to obtain the same RR at the center and edge, the pad and wafer are usually rotating in the same direction. The RR has been found to be linearly proportional to the linear velocity between the pad and wafer in oxide polishing [Pre27]. The velocity at any point under the wafer is determined by the relative rotating speeds of wafer and pad. When the wafer rotating speed is the same as that of the pad speed, the relative velocity between the pad and wafer is spatially uniform across the wafer [Pat91]. Thus, NU may be reduced by setting the two rotating speeds close to each other. At high pad speed, it contacts mainly the high regions of the topography, resulting in low SHR value because there is no enough time for the pad to conform to the valley regions.

End point detection, as the name suggests, is to determine when the polishing can be stopped. Traditionally, the end of a polishing step can be achieved by setting a

time limit in the process. However, this is only a rough estimation because under- or over- polish may occur due to the apparent reasons such as variations of polishing pad surface property, the fluctuation of slurry feeding rate, pad conditioning issues, and other potential variables. Therefore, in-line monitoring and automatic end-point detection of CMP is needed. However, in-situ end point detection is difficult to implement because of the complicated nature of CMP. To date, the approaches for end point detection include optical, electrical, thermal, and electrochemical method. In oxide polishing, one electrochemical approach that has been used to determine the end point during polishing is to implant an electrode disc into the platen, which serves as one plate of a capacitor with the oxide film on the wafer as the insulator. When the disc passes underneath the wafer, a capacitor will be formed momentarily. The output generated by electronics monitoring the disc is proportional to the oxide thickness. In another approach, the torque that drives the rotating platen and wafer carrier is measured during polishing. It is believed that a detectable change may be measured at the end point when a uniform overlayer is removed to reveal an underlying layer with different properties [Oma94]. The optical approaches include reflectivity and interferometry, both of which involve shining light up through the pad. The reflectivity method measures the intensity of the reflected light and the interferometry method directly measures the layer thickness on the wafer surface. Thermal methods detect changes in temperature of a polishing process due to changes in friction-induced wear mechanisms [Che97]. The heat of chemical reactions between the slurry and the wafer also needs to be measured when it is significant compared to the heat generated by friction. This is a non-contact method and relatively easy to implement. However, surface temperature changes are usually very small, and a

wafer size and actual state of a pad influences the temperature and its variations. Thus, the usage of all the methods mentioned above is limited. More effective end point detection methods are still under development.

Challenges in CMP

Compared to other techniques used in IC manufacturing, CMP is a relatively new process. The problems such as low throughput, inadequate reliability of chemical-mechanical polishers, degradation of removal rate with time, non-uniformity on the wafer surface and high defect density are still with this emerging technology. The interaction between polish pads and slurry and the role of slurry chemistry on the removal rate also need further investigation. In addition, the effect of polish pad characteristics and machine parameters on planarization and the pattern sensitivity to polish rate, is still not completely understood. As shown in previous sections, most of the understandings of the parameters are based on empirical observations. In order for a better control of the process in high volume manufacturing, a detailed understanding of the fundamental principles, which guide material removal and surface planarization, is needed.

Although most of the recipes for CMP, both proprietary and in open literature, have been arrived at empirically, it is unlikely that an optimized process will be achieved empirically because of a large number of process variables in the CMP process. Rather a physical model in conjunction with experimental data and an understanding of CMP fundamentals will be required to obtain optimal performance. Thus, the present challenge of the CMP processing is to develop physical models to explain these processes which then will lead to the development of new and more robust processes.

Modeling Efforts in CMP

Numerous efforts have been put on the modeling of the CMP process. Each model was able to provide at least partial understandings of the CMP process under given conditions. The way to classify the CMP models is based on the characteristic length scale. The modeling of CMP is often classified into two categories, i.e., wafer scale model and feature scale model. The wafer scale models attempt to explain the overall phenomena that happen across the wafer, which is of about hundreds of millimeters in diameter. For example, the slurry distribution and the overall removal rate of the CMP process are the two important variables that have been frequently evaluated. On the other hand, the characteristic length of the feature scale model is the length scale of typical device features on the wafer, which is of few microns or smaller. Thus, it is of orders of magnitude smaller than the wafer scale model in terms of its length scale and the major assumptions made in wafer-scale modeling may not be justifiable in feature scale modeling. Consequently, the focus of feature scale modeling is on local removal rate rather than the overall removal rate.

The other way to classify the existing models is based on the understanding of the interactions among the polishing pad, the wafer to be polished and the slurry. In the presence of the polishing slurry, the pad and the wafer, which are in relative sliding motion, will interact in one of the following three ways: (1) hydroplaning against each other, (2) semi-direct contact, or (3) direct contact. In the first case the load is supported entirely by a continuous fluid layer of slurry between the wafer and pad. In the second case the load is supported partially by pad-wafer contact and partially by hydrodynamic pressure on the slurry fluid. In the final case the load is supported totally through the direct contact between the pad and wafer. Since the semi-contact case is relatively

complicated, the other two extreme cases are usually adopted in the CMP modeling. Because each of these modes is likely to distribute pressure differently, the ability of a CMP process to remove material and to planarize will be affected by which mode a given CMP process operates within.

Preston's Equation and its Modifications

The model developed by Preston, which was initially for glass polishing, is probably the most cited expression for the material removal rate in CMP [Pre27]. According to the material removal rate (*RR*) is the polish rate is directly proportional to the pressure *P* and to the relative velocity *V* between the wafer and the pad, i.e.,

$$(RR) = K_p PV \tag{2.2}$$

This model is purely empirical. All the other variables that may affect the polishing performance such as the abrasive properties, slurry chemistry and the tribological interactions at the wafer-slurry-pad interface are lumped into the so-called Preston coefficient K_p. It should be noted that the pressure *P* is the local contact pressure which is given as $P = L/A$, where *L* is the applied load and *A* is the area of the polish surface in contact with the pad, which is not necessarily the geometric area. If the surface topography of the polishing pad is considered, only the high area of the surface is in direct contact and the contact area, *A*, will be less than the geometric area. This model is an empirical way of relating the operational variables to the material removal rate for a system where the K_p value can be determined by fitting it with experimental measurements. It has been found to be reasonably accurate for SiO_2 and tungsten CMP [Ste97]. However, this model does not explain the effects of other important variables like pad properties, slurry chemicals, and slurry abrasive on the polishing performance.

Furthermore, it only considers overall removal rate in the wafer scale and cannot predict any local removal rate.

Deviations from the Preston's equation have been found in CMP of different wafer materials. By accounting for the contact between the wafer and polishing pad properties Shi *et al*. proposed that the material removal rate is nonlinearly dependent of the applied pressure. The resulting removal rate expression was represented as in Equation 2.3 [Shi98].

$$(RR) = K_p P^{2/3} V \tag{2.3}$$

Tseng and Wang combined the shear stress induced by the slurry flow and the Hertzian particulate penetration approximation to obtain an alternative removal rate equation expressed as [Tse97]

$$(RR) = K_p P^{5/6} V^{1/2} \tag{2.4}$$

In another experimental investigation, Luo *et al*. found that their results of polish rate versus velocity and pressure, respectively, agree well with the Preston's equation. However, upon extrapolating a linear regression of the data to zero pressure or zero velocity, a non-zero polish rate is observed [Luo98]. It is attributed to the chemical effects of the slurry on removal rate. To account for this non-zero polish rate, a modified Preston's equation was proposed:

$$(RR) = (K_p P + B)V + R_c \tag{2.5}$$

where the constants K_p, B and R_C are fitting parameters and can be obtained by least squares analysis of the data. These modifications can provide improved agreement between material removal rate response and the experimental results. However, they were also limited to the simulation of specific polishing systems and cannot be generalized to

become a predictive model due to the incomplete understandings of coefficient K_p. Therefore, it is necessary to develop more comprehensive predictive methodologies by determining an explicit expression of K_p, which should include most of the controlling variables of the CMP process.

Hydrodynamic Analysis

The existence of fluid layer has at least two effects. First the chemical and mechanical forces are transferred through the slurry carried chemicals and particles respectively. Secondly, the fluid layer acts like a lubricating agent and conducts heat from the surface. It is well accepted that there is a fluid layer between the wafer and the pad. However, it is not clear whether the fluid layer is partial or continuous. The model by Runnels and Eyman is probably the first attempt to analyze the slurry flow between the wafer and pad [Run94a]. The model is based on fluid dynamics and primarily focuses on stresses and force balances in a wafer-slurry-pad system by assuming that (1) the wafer and pad are rigid and smooth, (2) the wafer slides at an angle of attack θ upon the slurry film, (3) a continuos slurry layer exists between the pad and the wafer, (4) the wafer has a spherical curvature. The fluid layer thickness and the angle of attack are calculated by an iterative scheme satisfying force and moment balances. The stresses induced on the pad and wafer surface are calculated using finite element method. They also proposed that the material removal rate should be proportional to the normal and shear stresses that developed on the wafer surface instead of the pressure and velocity (Eq. 2.2). The model is thus given as

$$(RR) = C_p \sigma_n \tau \tag{2.6}$$

where C_p is a Preston-like constant, σ is the stress tensor in the normal direction and τ is the shear stress due to slurry flow induced by relative motion between pad and wafer. Besides the material removal rate, the other useful information from this modeling is that a hydrodynamic slurry layer of about 50 μm thick is suggested between the wafer and pad during polishing. The model can be set up on both the wafer scale and the feature scale.

Since the slurry layer is relatively thin compared to the size of the wafer, it is possible to use lubrication approximation to simplify the CMP process. In hydrodynamic lubrication, a fluid wedge is required to provide sufficient fluid pressure to support the wafer. A schematic illustration of this assumption is shown in Figure 2.10. The corresponding governing equation for the hydrodynamic lubrication is the Reynolds equation [Ham94].

$$\frac{d}{dy}\left(h^3 \frac{dp}{dy}\right) = 6\eta V \frac{dh}{dy} \tag{2.7}$$

where p is the pressure profile of the fluid film under the wafer and η is the slurry viscosity.

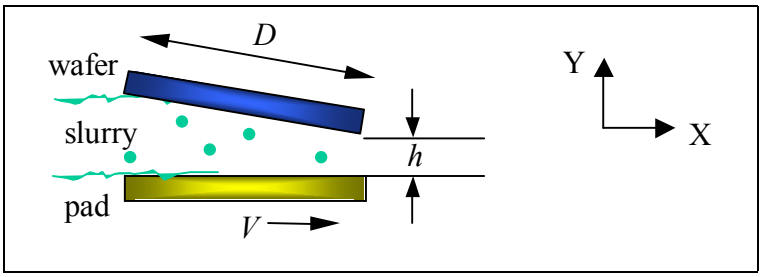


Figure 2.10 Schematic of lubrication CMP

Solutions to the Reynolds equation have been investigated both analytically and numerically in CMP [Sun99, Cho01, Che02]. The observations from these modeling

efforts provide valuable information about the slurry film thickness and pressure distribution under the wafer. However, the material removal mechanism is still not clear. In addition, more experimental data are needed to support the model predictions of these methodologies.

Both the wafer scale and feature scale modeling of the stresses induced by the slurry flow have been started in our group by Sohn [Soh99] and Agarval [Aga00]. The continuous work on these two aspects will be presented in the following chapters. The modeling results will be compared with experimental data, and the possible mechanisms of the effect of slurry flow on the material removal will also be discussed.

Modeling of the Abrasive Wear

When the pad and the wafer are in direct contact and the applied load on the wafe is primarily supported by the deformation of the pad surface, the material removal during CMP is essentially through abrasive wear. The wear due to abrasion can be further subdivided into two-body and three-body abrasive wear. Two-body wear is caused by hard protuberances on the counterface, while in three-body wear hard particles are free to roll and slide between two, perhaps dissimilar, sliding surfaces. As indicated previously, further understanding of the CMP process relies on the delineation of the effects of the process variables involved in CMP. In terms of modeling, an explicit expression of the Preston's coefficient, which should include all the controlling parameters of CMP, is needed. Based on two-body abrasive wear, Brown [Bro81] studied the polishing of metals at a single particle-surface interaction level and modeled the penetration of a single abrasive particle into metal surface using Hertz elastic contact theory. Brown’s model is given as

$$(RR) = \frac{1}{2E} PV \tag{2.8}$$

where E is the Young's modulus of the material being polished. This is essentially a modification of the Preston's equation. The coefficient K_p is replaced with $1/2E$ to account for the material properties of the surface to be polished. It is shown that the model can provide reasonable estimation in calculating the material removal rate for metal polishing like nickel. However, for glass polishing, Cook observed that the model overestimated the polishing rates by over one order of magnitude [Coo90]. To explain this discrepancy he proposed that the material properties of the surface could be different from the bulk material. Formation of such a layer may change the mechanical properties of the surface to be polished significantly.

Due to the over simplification of the Brown's model, its applicability is limited. Moreover, the actual CMP process is a three-body abrasion process, in which the particles are fed continuously onto the polishing pad and carried to the gab between two sliding surfaces, i.e., the polishing pad and the wafer to be polished. The effects of the particulate information and the slurry chemistry have to be included in CMP modeling. Using the theory of contact mechanics, Larsen-Basse and Liang [Lar99] concluded that the particle abrasive wear is the major mechanism in CMP based on experimental data and simple calculations. Xie and Bushan [Xie96], along with other investigators [Luo01, Zha02], took the three-body abrasion approach and modeled the material removal through a micro-cut mechanism. The particles are considered as cutting tools in these models, and the influence of the pad-wafer contact is also included. Although these models are more complicated than the Preston's equation by including more controlling parameters, chemically altered layer formation is not taken into account, which is

important to predict surface quality as well as the material removal rate in CMP. In addition, the slurry flow effect is completely ignored.

Generally, the models reviewed above concentrate only partial of the process variables in CMP. Although these models are adequate for specific applications, they may not be applicable to other systems with different slurry properties. Evidently, there is a need for a systematic study of the effects of the slurry flow, slurry chemistry and mechanical abrasion for developing more fundamental predictive methodologies to control the CMP processes, particularly to keep up with the smaller size device manufacturing.

Scope of the Dissertation

Every model of a physical process must have specific statement of purpose. The overall goal of this study is to provide fundamental understandings of the polishing mechanisms involved in the CMP process. However, CMP is a multi-parameter, multi-scale process. Due to the complexity of the process and the large number of variables involved in CMP, assumptions and other limitations have to be imposed on the models. The two major limitations in modeling the CMP process are the geometric scales and the physics content.

Table 2.2 Different scales involved in CMP

Length Scale		**Time Scale**	
$\sim 10^{-9}$ m	Particle size, Wafer surface roughness	$\sim 10^{7}$ sec	Slurry shelf life, Slurry aging
$\sim 10^{-6}$ m	Pad surface roughness, Feature size, Particle size	$\sim 10^{1}$ sec	Polishing process
$\sim 10^{-1}$ m	Wafer size, Pad size	$\sim 10^{-6}$ sec	Surface layer formation

An example of different scales involved in the CMP process is given in Table 2.2. Both the length scales and the time scales vary in orders of magnitude, making it impossible to include all the input and output variables in one general model. Furthermore, the physical principles involved in the CMP process, such as equation of motion and the contact mechanics, have to be simplified depending on the scales to be investigated. Therefore, the current study concentrates on the modeling of the CMP process in three length scales. First, the wafer scale slurry flow is simulated by assuming the pad and wafer are rigid and parallel to each other. The modeling results are compared with experimental observations and optimal operation conditions of the CMP process can be derived. Then, the stress distribution of different features induced by the slurry flow is calculated through finite element methods. The possible mechanism of slurry flow caused material removal is discussed. Furthermore, a comprehensive model that incorporates both the mechanical and chemical effects in CMP is developed. The optimization of the controlling parameters of the polishing process is obtained through model predictions. It is expected that the reported findings in this study can provide a fundamental understanding of the polishing mechanisms and can provide guidelines for optimization of the CMP applications.

CHAPTER 3
WAFER SCALE SLURRY FLOW SIMULATION

Introduction

Chemical mechanical polishing (CMP) has become a standard process for global planarization of multilevel interconnects in microelectronics industry. Fabrication of increasing complex, dense, and miniaturized integrated circuits without defects necessitates significant improvements in the CMP process. The success of CMP operations relies on the control of a long list of parameters (Fig. 2.8), such as the pressure applied to the wafer, the rotational speeds of the polishing pad and the wafer carrier, the pad condition, and the slurry properties. Although most of the important parameters for each individual CMP process that offers local and global planarity for the semiconductor device fabrications have been established empirically, the fundamental mechanism that elucidates the polishing process is far from a detailed understanding. Optimal performance of the CMP process depends on maximal removal rate and smooth surface finish, which are achieved by the synergistic effects of the chemical and mechanical forces encountered in the CMP process. The delivery of the chemicals to the wafer surface by the slurry flow as it flows in between the pad and the wafer surface under the applied pressure determines the reaction rate of the wafer surface. It in turn affects the material removal rate and polishing uniformity. Therefore, it is crucial to control the slurry flow patterns in wafer scale.

It has been estimated that there is a thin slurry layer of about 20-50 μm thick between the wafer and pad under regular polishing pressure [Sun99]. The wafer scale

slurry flow analysis was first studied by Runnels and Eyman [Run94a]. Finite element method was employed in their analysis to solve the steady-state three-dimensional (3-D) Navier-Stokes equation numerically in the region between the wafer and the pad. However, this analysis can be significantly simplified if one takes into account the fact that the Reynolds number of the slurry flow is rather small during polishing. In addition, the flow film is much thinner than the size of the wafer. Therefore, the lubrication approximations used in the theory of slider bearings are valid [Ham94]. Most of the recent studies have focused on the solution schemes of Reynolds equation using different simplifications. Sundararajan *et al.* [Sun99] solved the Reynolds equation for a 2-D geometry. The pressure distribution developed in the thin slurry film between the pad and wafer and the thickness of the slurry were obtained simultaneously through a semi-analytical procedure. In a later attempt from the same research group, the pad compressibility and porosity were added into the lubrication model. The experimental results of the material removal rate were compared with the shear stress developed by the slurry flow on the wafer surface. A good correlation between the experimental data and the model calculations was obtained. Most recently, Cho *et al.* [Cho01] and Chen and Fang [Che02] expanded the Reynolds equation to 3-D geometry in their individual publications. Similarly, the distributions of thickness and pressure of the slurry fluid on the wafer surface were calculated in both modeling approaches. Cho *et al.* also analyzed the distribution of shear stress of the slurry on the wafer surface. With lubrication theory, the influence of some of the aforementioned process parameters on CMP can be elucidated. However, it assumes the applied load is totally supported by the slurry film, which may not be practical in real polishing process because the process will be hard to

control and it will result in inconsistent polishing rate if it is operated in the lubrication regime. Therefore, the value of the slurry flow analysis does not rely on the calculation of pressure distribution, but on the delivery of the slurries. A uniform distribution of the slurry chemicals and particles between the wafer and pad will lead to uniform surface reaction and material removal, and hence better uniformity.

The current study starts with the simulation of the hydrodynamic behavior of the slurry fluid trapped between the wafer and the polishing pad. As mentioned in previous chapter, several length scales are involved in the CMP process. When only wafer scale slurry flow in the wafer scale is considered, the other smaller scale effects can be ignored. For example, the polishing pad can be considered as a flat surface because the pad surface roughness, which is in the range of tens of microns [Yu94], is several orders of magnitude smaller than the wafer size, which is in the range of 100-300mm. In this chapter, a wafer scale slurry flow model is built using continuity equation and Navier-Stokes equation. After the equations are solved analytically, the non-uniformity of the shear rate of the slurry flow is defined and compared with experimental results. The effects of several important parameters, including the rotational velocities of the wafer and pad, the distance between the centers of the wafer and pad, on the shear rate non-uniformity are discussed. The optimal operation conditions can be derived from the model calculations.

Formulation of Fluid Motion

The slurry flow behavior is controlled by the relative motion of the pad and wafer. The fluid is considered as incompressible, Newtonian fluid since the non-Newtonian characteristic of the slurry flow is not pronounced at the regular particle concentration level of the slurry (e.g., 1-30 wt%). Assuming that the pad and the wafer are parallel to

each other, we can obtain that the governing equations for the slurry flow in the gap of uniform thickness h are essentially continuity equation and the Navier-Stokes equation, i.e.,

$$\nabla \cdot \mathbf{u} = 0 \tag{3.1}$$

$$\rho(\mathbf{u} \cdot \nabla \mathbf{u}) = -\nabla p + \eta \nabla^2 \mathbf{u} + \rho \mathbf{g} \tag{3.2}$$

where $\mathbf{u}$ is the velocity vector of the fluid flow, ρ and η are the density and viscosity of the slurry respectively. The pressure distribution p of the slurry flow in the gap between the wafer and pad is isotropic and constant. Thus, the gradient of p is zero. To solve these equations, a system is set up as shown in Figure 3.1. The wafer of radius R_w rotates with angular velocity Ω_w about its center, while the pad of radius R_p rotates with its own angular velocity Ω_p. The distance between the center of the pad and the center of the wafer is denoted as L.

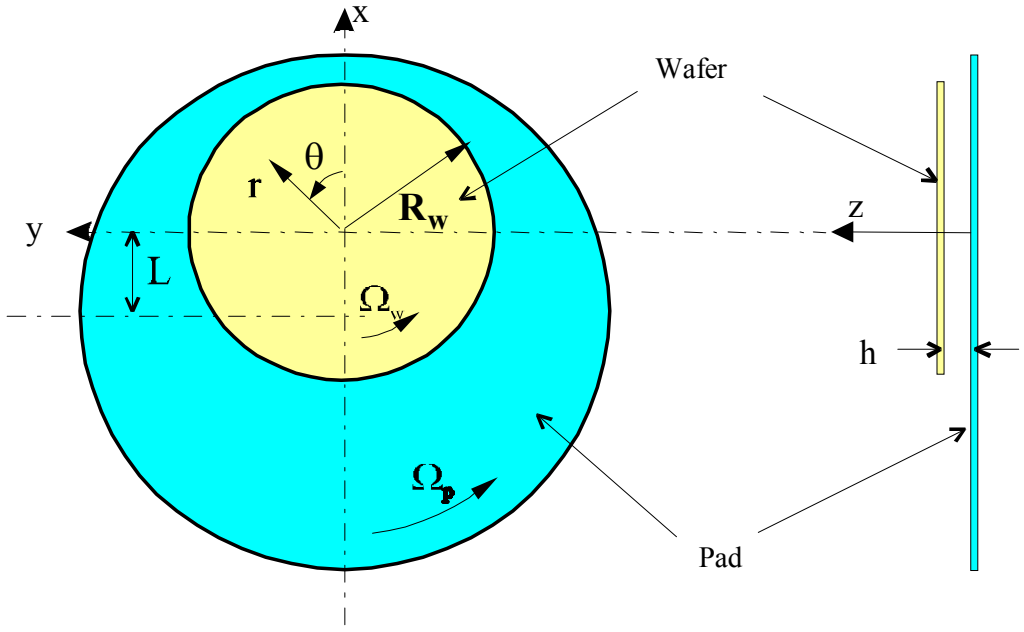


Figure 3.1 Coordinates and parameters for slurry flow simulation [Soh99].

To simplify the equations, the cylindrical coordinate system is adopted. This solving procedure was first done in our research group [Soh99], which is briefly reviewed as below. Since we assume the slurry flows between two parallel plates, the fluid motion is purely shear flow and the velocity component in z-direction is vanished. By non-dimensionalizing the equations 3.1 & 3.2 using the following scaling variables,

$$r^* = \frac{r}{R_w} \quad z^* = \frac{z}{h} \quad u_\theta^* = \frac{u_\theta}{R_w \Omega_w} \quad u_r^* = \frac{u_r}{R_w \Omega_w},$$

we can reduce the governing equations to

$$\frac{1}{r^*}\frac{\partial}{\partial r}\left(r^* u_r^*\right) + \frac{1}{r^*}\frac{\partial u_\theta^*}{\partial \theta} = 0 \tag{3.3}$$

and

$$\begin{aligned} &\mathrm{Re}\,\varepsilon^2\left(u_r^*\frac{\partial u_r^*}{\partial r^*} + \frac{u_\theta^*}{r^*}\frac{\partial u_r^*}{\partial \theta} - \frac{u_r^{*2}}{r^*}\right) \\ &\quad = \left[\varepsilon^2\frac{\partial}{\partial r^*}\left(\frac{1}{r^*}\frac{\partial}{\partial r^*}\left(r^* u_r^*\right)\right) + \frac{\varepsilon^2}{r^{*2}}\frac{\partial^2 u_r^*}{\partial \theta^2} + \frac{\partial^2 u_r^*}{\partial z^2} - \frac{2\varepsilon^2}{r^{*2}}\frac{\partial u_\theta^*}{\partial r^*}\right] \end{aligned} \tag{3.4}$$

$$\begin{aligned} &\mathrm{Re}\,\varepsilon^2\left(u_r^*\frac{\partial u_\theta^*}{\partial r^*} + \frac{u_\theta^*}{r^*}\frac{\partial u_\theta^*}{\partial \theta} - \frac{u_r^* u_\theta^*}{r^*}\right) \\ &\quad = \left[\varepsilon^2\frac{\partial}{\partial r^*}\left(\frac{1}{r^*}\frac{\partial}{\partial r^*}\left(r^* u_\theta^*\right)\right) + \frac{\varepsilon^2}{r^{*2}}\frac{\partial^2 u_\theta^*}{\partial \theta^2} + \frac{\partial^2 u_\theta^*}{\partial z^2} - \frac{2\varepsilon^2}{r^{*2}}\frac{\partial u_r^*}{\partial \theta}\right] \end{aligned} \tag{3.5}$$

where the dimensionless groups used in the above equations are

$$\mathrm{Re} = \frac{\rho R_w^2 \Omega_w}{\eta}, \quad \varepsilon = \frac{h}{R_w}, \quad l = \frac{L}{R_w}, \quad \omega = \frac{\Omega_p}{\Omega_w}$$

Results and Discussion

Velocity Profile of the Slurry Flow

Typical process parameters in CMP are shown in Table 3.1. In addition, it is well known that the standard wafer diameter is either 8 inches or 12 inches in current

semiconductor industry. Accordingly, the order of magnitudes of Re and ε can readily be obtained as 10^4 and 10^{-4} respectively. As a consequence, the terms (Re ε^2) and ε^2 are in the orders of 10^{-4} and 10^{-8} respectively.

Table 3.1 Typical process parameters during CMP

Parameter	**Operational Value**
Platen rotational speed	50 - 100 RPM
Wafer rotational speed	20 - 60 RPM
Wafer pressure	10,000 - 90,000 N/m^2
Slurry flow rate	100 - 300 cm^3/min
Nano-particle concentration in slurry	1 - 30 % by weight
Planarization time	1 - 3 min
Pad life	100 - 1000 polishes
Nano-particle type	silica, alumina, cerium oxide, titania, composite materials
Removal rate	200 - 300 nm/min for SiO_2 300 - 600 nm/min for W or Cu

Source: [Phi01]

Equations 3.4 & 3.5 can therefore be simplified by dropping the terms associated with these small quantities.

$$\frac{\partial^2 u_r^*}{\partial z^{*2}} = 0 \tag{3.6}$$

$$\frac{\partial^2 u_\theta^*}{\partial z^{*2}} = 0 \tag{3.7}$$

Applying non-slip boundary conditions to above equations at the pad surface (i.e., $z^* = 0$) and the wafer surface (i.e., $z^* = 1$) respectively, we have

$$u_r^* = \omega l \sin\theta, \quad u_\theta^* = \omega\left(r^* + l\cos\theta\right), \quad \text{at } z^* = 0 \tag{3.8a,b}$$

$$u_r^* = 0, \quad u_\theta^* = r^*, \quad \text{at } z^* = 1 \tag{3.9a,b}$$

Combining the boundary conditions with equations 3.6 and 3.7, we can obtain the velocity profile for the slurry flow.

$$u_r^* = \omega l \sin\theta\left(1 - z^*\right) \tag{3.10}$$

$$u_\theta^* = \omega\left(r^* + l\cos\theta\right)\left(1 - z^*\right) + r^* z^* \tag{3.11}$$

Streamlines under the Wafer

The local transport of the slurry in the gap between the pad and the wafer is controlled by the slurry flow. With the known velocity profiles, it is possible to find the streamline through the stream functions, i.e.,

$$u_r^* = \frac{1}{r^*}\frac{\partial\psi}{\partial\theta} \tag{3.12}$$

$$u_\theta^* = -\frac{\partial\psi}{\partial r^*} \tag{3.13}$$

where the scale for the dimenstionless stream function ψ is $R_w^2\Omega_w$. Introducing equations 3.10 & 3.11, and integrating equations 3.12 & 3.13, we can determine the stream function ψ as

$$\psi = r^*\omega\left(z^* - 1\right)\left(\frac{1}{2}r^* + l\cos\theta\right) - \frac{1}{2}r^{*2}z^* \tag{3.14}$$

By definition, the contours of any constant ψ form the streamlines of the fluid flow. The streamlines at given height positions under the wafer are shown in Figure 3.2. At each fixed z-plane, the streamlines are concentric circles centered at

$$(x, y) = \left(\frac{-\omega l\left(z^* - 1\right)}{\omega\left(z^* - 1\right) - z^*}, 0\right) \tag{3.15}$$

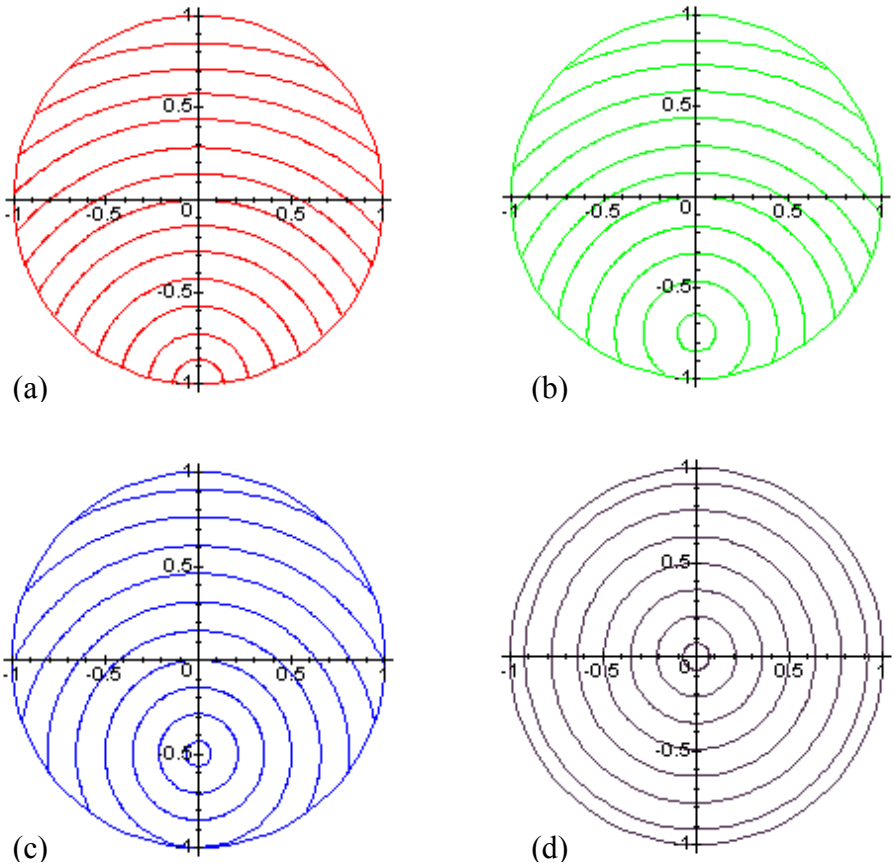


Figure 3.2 Streamlines at constant-z planes for l =1.0 and ω (=Ω_p/Ω_w) = 1.0 (a) z^*=0 (pad surface), (b) z^*=0.25, (c) z^*=0.5, (d) z^*=1 (wafer surface) [Soh99]

The center of the streamlines moves from outside the wafer rim to the center of the wafer when z^* changes from 0 (i.e., pad surface) to 1 (i.e., wafer surface). It is also seen that some of streamlines are cutting across the outer rim of the wafer, while others are closed circles within the wafer rim boundary. Since the pad and wafer are rotating eccentrically, the streamlines across the wafer represent the fluid flow that includes freshly fed slurry, whereas the closed ones provide the location of the stagnant region under the wafer. The slurry in the stagnant region will never be replaced by the fresh slurry during polishing. Thus, it is not desirable and has to be minimized. The stagnant

region can be reduced by increasing the distance l, which will move the center of the streamline away from the center of the wafer as indicated in Eq. 3.15. Likewise, it can also be minimized by increasing ω (e.g., increasing the rotational speed of the pad). The increased motion of the pad can provide efficient sweeping effect of the stagnant slurry. Another practical way is to introduce vertical (or z-direction) motion of the slurry to disturb the regular flow pattern shown in Fig. 3.2. It can be achieved by making grooves or patterns on the pad surface, which is in agreement with most of the commercial pads. The characteristic size of these patterns should be in the order of the gap size for vertical mixing of the slurry.

Shear Rate Non-Uniformity

The delivery and mixing of the slurry are among the most important factors that affect the non-uniformity of the removal rate during polishing, although more factors have to be considered for a complete understanding of the non-uniformity of polishing. The non-uniformity of the shear rate caused by the slurry flow may provide an insight into the non-uniformity of the polishing rate. With the velocities given in Eqs. 3.10 & 3.11, the shear rate can be obtained through $\dot{\gamma} = \sqrt{2\mathbf{D}:\mathbf{D}}$ by definition, and $\mathbf{D}$ is the rate of deformation tensor given as $\mathbf{D} = \frac{1}{2}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{T}}\right]$. Thus the dimensionless shear rate is

$$\dot{\gamma}^* = \sqrt{\left[r^* - \omega\left(r^* + l\cos\theta\right)\right]^2 + \left(\omega l \sin\theta\right)^2} \tag{3.16}$$

The scale of the dimensionless shear rate is $R_w\Omega_w/h$. Following the definition of within wafer non-uniformity (WIWNU) [Bib97], we define the shear rate non-uniformity (SRNU) as

$$SRNU = \frac{1}{\dot{\gamma}^*_{avg}} \left[\frac{1}{n} \sum_{i=1}^{n} \left(\dot{\gamma}^*_i - \dot{\gamma}^*_{avg} \right)^2 \right]^{1/2} \tag{3.17}$$

where $\dot{\gamma}^*_i$ is the mean shear rate over θ ($0 \le \theta \le 2\pi$) the at a given point *i* along the wafer radius, i.e.,

$$\dot{\gamma}^*_i = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{\left[r_i^* - \omega \left(r_i^* + l \cos\theta \right) \right]^2 + \left(\omega l \sin\theta \right)^2} \, d\theta \tag{3.18}$$

and $\dot{\gamma}^*_{avg}$ is the average shear rate in the radial direction which is given as

$$\dot{\gamma}^*_{avg} = \frac{1}{n} \sum_{i=1}^{n} \dot{\gamma}^*_i \tag{3.19}$$

The definition of shear rate non-uniformity in Eq. 3.17 represents the degree of variation of the shear rate along the radius of the wafer. For given variables of ω and *l*, the mean shear rate can be obtained through Eq. 3.18 for any radial position *i*. The ideal case is that all $\dot{\gamma}^*_i$ values are the same along the radius of the wafer. Consequently, the shear rate non-uniformity (SRNU) vanishes and it is expected that the same polish rate can be achieved at any position of the wafer surface.

Recalling that $\omega = \Omega_p/\Omega_w$ and $l = L/R_w$, we can plot the SRNU as a function of wafer and pad rotational speeds (Figure 3.3). The wafer radius R_w is taken as 75mm, and the distance between the wafer and pad centers *L* is set at 150mm. The calculations in Fig. 3.3 along with figures 3.4 & 3.5 are done with numerical integration using Simpson's rule, and the individual variables are discretized accordingly. The SRNU is minimized when a low wafer speed and a high pad speed are applied simultaneously. Moreover, the SRNU is always zero when $\Omega_p=\Omega_w$ which is the aforementioned ideal situation.

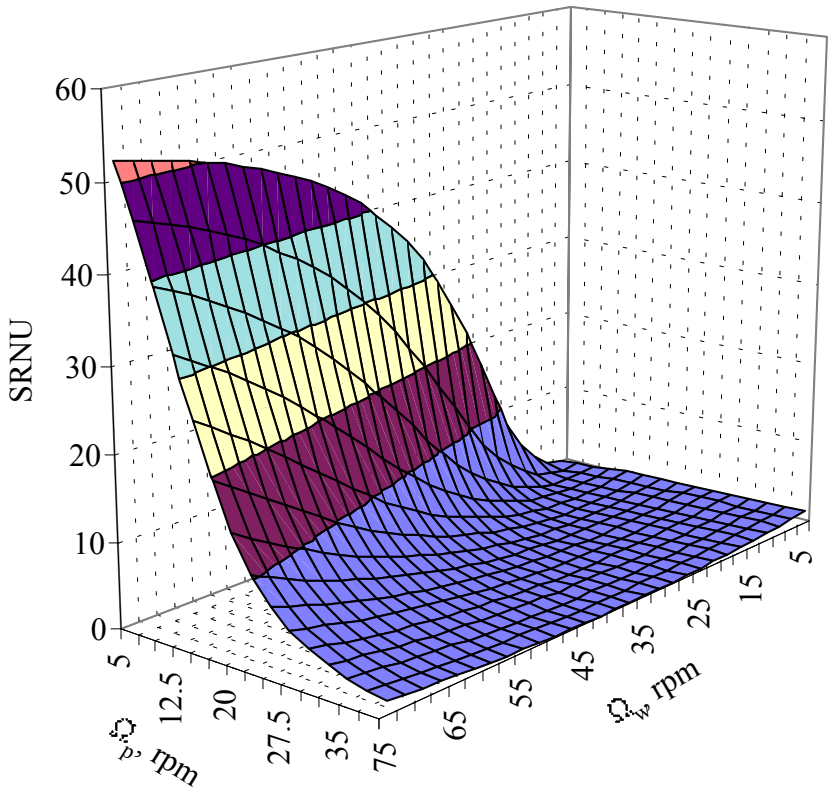


Figure 3.3 Variations of shear rate non-uniformity with the rotational speeds of the pad and wafer for l =2.

Figure 3.4 shows the correlation between the SRNU and the pad rotational speed Ω_p as well as the distance between the wafer and pad centers l. It is seen from the plot that the SRNU can be reduced by placing the wafer away from the center of the pad and, at the same time, increasing the pad rotational speed. However, when the pad rotational speed is close to or higher than the wafer rotational speed, the effect of distance l on the SRNU becomes insignificant as long as it is bigger than 1, which corresponds to a distance between the wafer and pad centers being about the radius of the wafer.

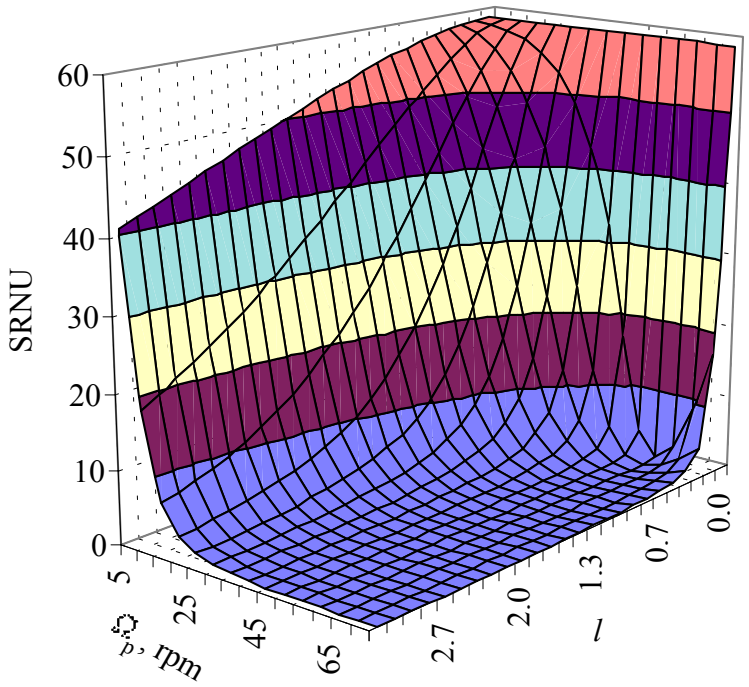


Figure 3.4 Shear rate non-uniformity as a function of the pad rotational speed and the distance between the wafer and pad centers for Ω_w = 42 rpm.

The effects of the wafer rotational speed Ω_w and the distance between the wafer and pad centers l on the SRNU are plotted in Figure 3.5. Again, longer distance between the wafer and pad centers results in less SRNU. Furthermore, the wafer rotational speed does not show pronounced effect on the shear rate non-uniformity, especially when l is bigger than 1.

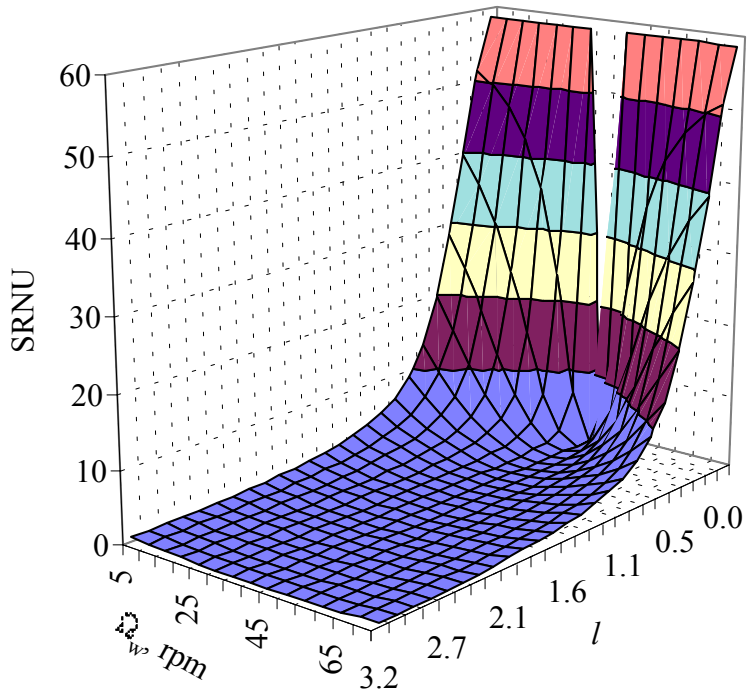


Figure 3.5 Shear rate non-uniformity as a function of the wafer rotational speed and the distance between the wafer and pad centers for Ω_p = 40 rpm.

To compare the calculated SRNU with within-wafer polishing non-uniformity (WIWNU), several sets of experimental data are adopted from the literature. Figure 3.6 shows a comparison between the calculated shear rate non-uniformity and measured within wafer non-uniformity as a function of the pad rotational speed polish. The polishing experiments were performed on a 1 mm thick thermal oxide grown on 150 mm wafers under conditions of Ω_w = 42 rpm, l = 15 cm, and P = 7 psi. The WIWNU was defined similarly as in Eq. 3.17 averaged over those obtained from five wafers [Tse99]. The trend of the model predicted SRNU matches well with the experimental WIWNU values. At low Ω_p, a large difference between the rotational speeds of wafer and pad results in a much higher SRNU and WIWNU. At high Ω_p, on the other hand, a lower

non-uniformity is seen for both measured and predicted values. It agrees with the previous discussion that the rotational speed of the pad has more pronounced effect than the speed of wafer.

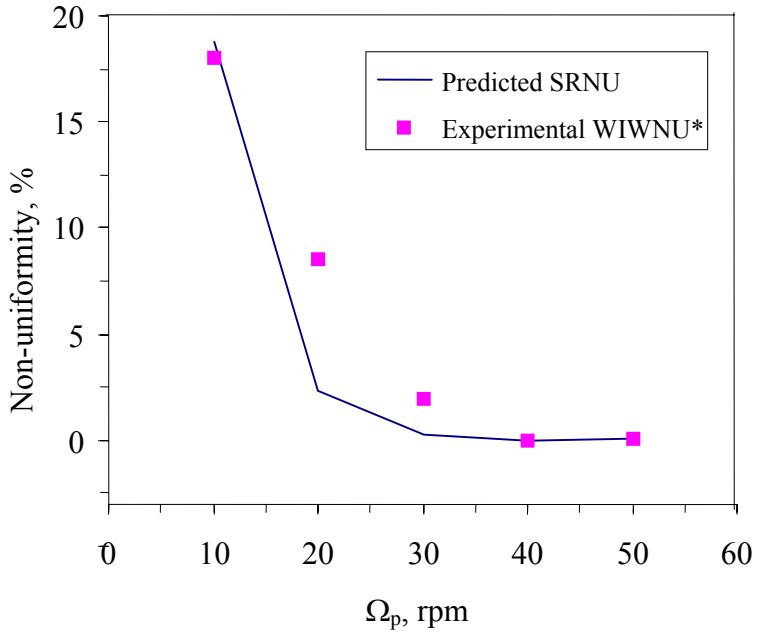


Figure 3.6 Comparison between the calculated shear rate non-uniformity and measured within wafer non-uniformity as a function of the pad rotational speed. *experimental data from Tseng *et al* [Tse99].

Comparison between the model predictions and experimental data are also performed with Ouma's polishing results [Oum99]. It is shown that the down pressure does not have significant effect on the polishing non-uniformity as long as the applied pressure is evenly distributed. The uniformity of the removal rate is mainly controlled by the relative rotational speeds of wafer and pad. Modak *et al*. evaluated removal rate and its nonuniformity as a function of the wafer-to-pad rpm ratio (Ω_w/Ω_p) [Mod97]. They reported that a remarkable increase in WIWNU was observed when $\Omega_w/\Omega_p > 1.14$, which

corresponds to $\omega < 0.88$. Thus a higher pad rotational speed is always desirable, which again confirms the conclusions from Fig. 3.4.

Finally, it is helpful to compare the order of magnitude of the shear stress induced by the slurry flow with the shear strength of the wafer material. If the shear stress of the slurry flow is higher than the shear strength of the material, possible fracture or delamination of the wafer surface may occur and cause material removal. The scale of shear stress in current study is $\eta R_w \Omega_w / h$. Plugging in the typical values of corresponding variables, we can readily obtain that the shear stress is in the order of 10^3-10^4 Pa, which is several orders of magnitude smaller than the shear strength of wafer materials [Sha01]. Therefore, it is unlikely that the wafer material can be removed by slurry flow alone unless there is stress concentration around the cracking point or a much soft layer forms on the wafer surface. Both cases will be investigated in the following chapters.

Summary

The wafer scale flow simulation in current study provides the 3-D flow field of the slurry during CMP, in which the pad and the wafer are assumed to be flat and parallel to each other. The streamline patterns and thus the slurry transport characteristics have been obtained through the simulation. Stagnant regions are found to be a function of the relative rotational velocities and the distance between the pad and the wafer. The modeling results indicate that both the increase of the distance and the increase of the ratio between the pad speed and the wafer speed would minimize the stagnant region. Additionally, the results suggest that the introduction of vertical motion of the slurry be preferable to enhance the CMP process.

The investigation of the shear rate non-uniformity shows that maintaining the same angular velocities of the pad and the wafer is necessary to ensure uniform shear rate distribution across the wafer. In addition, the velocity of pad rotation is found to be more effective in controlling the shear rate non-uniformity. Based on the simulation results, the combination of a high pad rpm, a medium wafer rpm, and a large pad-to-wafer distance should give rise to a minimum non-uniformity. The comparison of the shear rate non-uniformity with the experimental polish non-uniformity shows similar trend, indicating that the slurry delivery is a significant factor in determining the removal rate non-uniformity. Overall, the findings from the above simulation can serve as general guidelines for improving removal rate uniformity under controlled polish conditions.

CHAPTER 4
FEATURE-SCALE MODELING: DEPENDENCE OF STRESS DISTRIBUTION ON THE SHAPE OF SURFACE FEATURES IN CMP

Introduction

In the previous chapter, it was demonstrated that the wafer-scale slurry flow has significant effects on the polishing non-uniformity. Another particular challenge in CMP is that the global planarity requires variation in the removal rate between high and low surfaces, which in turn is relies on the variation in mechanics [Run94a, Run94b]. Thus, the feature scale modeling of the mechanics involved during CMP is needed for the understanding of the polishing mechanism and the design and control of the CMP process.

The most important issue to deal with in CMP modeling is the interaction between the wafer and the polishing pad. As discussed in Chapter 2, there have been two extremes to address this issue, namely hydroplaning and direct contact. If hydroplaning is the dominant process, the wafer is supported by the slurry film. The assumptions of flat wafer and smooth pad are thus justified and the theory of hydrodynamic lubrication can be applied to the process design and control. On the other hand, if intimate contact occurs between the pad and wafer during CMP, the wafer is supported by the pad surface. The evolution of the feature height during polishing has to be considered with contact mechanics. As a consequence of different treatments of the interaction between the pad and wafer, two approaches have been pursued in feature-scale modeling. Shortly after the wafer scale slurry transport model [Run94a], which has been reviewed in Chapter 2,

Runnels [Run94b] published an erosion model to explain the feature height evolution during polishing. It is a semi-physical model, in which the two dimensional Navier-Stokes equation was solved using a finite element method on a mesh that deforms as the feature shape changes. The model related the stress on the features induced by a thin shearing layer of slurry to erosion rate. An empirical relationship between the stress and the erosion rate was hypothesized and used to predict the evolution of the features. However, this relationship has not been justified so far, either theoretically or experimentally. Besides, only simplified geometries (e.g., square, rectangular) were considered.

In another approach, assuming direct contact between the wafer and pad, Warnock [War91] developed a model which allows quantitive analysis of the polish rate for different feature sizes and shapes. It takes into account for all geometrical shapes, and is not limited to two dimensional case. However, this is a purely phenomenological model. By replacing the linear velocity in Preston's equation with the vector-valued shearing stress, Runnels [Run96] later incorporated the stress distribution that is resulted from the deflection of the polishing pad into a model that relates the removal rate to stress distribution around the feature. This is a different approach from his previous models because no slurry flow was considered in this model. The stress distribution was investigated on the edge of the wafer and the curvature of the pad through Hooke's law and the kinematic definitions of strain. This model shows that the removal rate of the wafer material is strongly dependent on the stresses developed. It provides a qualitative analysis of these mechanical effects and the results are comparable with the experimental data. Recently, Chekina [Che98] and others [Noh02] employed contact mechanics

method to the modeling of wafer surface shape evolution during polishing. Those models assume the removal rate is linearly proportional to the contact pressure between the pad and wafer. Thus only the pressure profile, which changes with the feature shape evolution, needs to be investigated.

The models mentioned above rely on, either totally or partially, the empirical linear relations between the material removal and the pressure (or stress). Most of them seem to be able to provide good fit to the experimental data. However, further understanding of polishing mechanisms needs theoretical development on these relations. Moreover, the models were only applied to simple geometrical features, and did not give any information about the interior of a wafer. In this chapter, the slurry flow with the presence of features of different shapes is modeled using finite element method. The stress distribution inside the wafer, which is induced by the slurry flow, is then investigated. By applying a suitable yield criterion for the material being polished, removal of material could be predicted, thereby establishing an understanding for the conditions under which material removal takes place from the wafer surface.

Modeling Procedure

At first, we studied the stress patterns induced by the slurry flow on to the wafer surface using a finite element method. After the stress distribution on the wafer surface with different features was determined, it was then used as a boundary condition to obtain the stress-state inside the wafer. Once the stress distribution inside the wafer was known, yielding of material, hence the removal of the material from the wafer surface could be predicted by applying an appropriate yield criterion.

The asperities are represented by quadrilateral (e.g., square and trapezoid), triangular or semi-circular projections from the wafer surface. In Figure 4.1, 2-D model

systems are shown schematically. Compared to the wafer scale modeling in Chapter 3, the presence of features will generate more complicated local flow patterns, and the stress imposed on the surface of the wafer will vary with the shape of the features. Thus, the main interest in current study is the stress imposed on the lower surface of the wafer as influenced by the fluid motion in the presence of a feature.

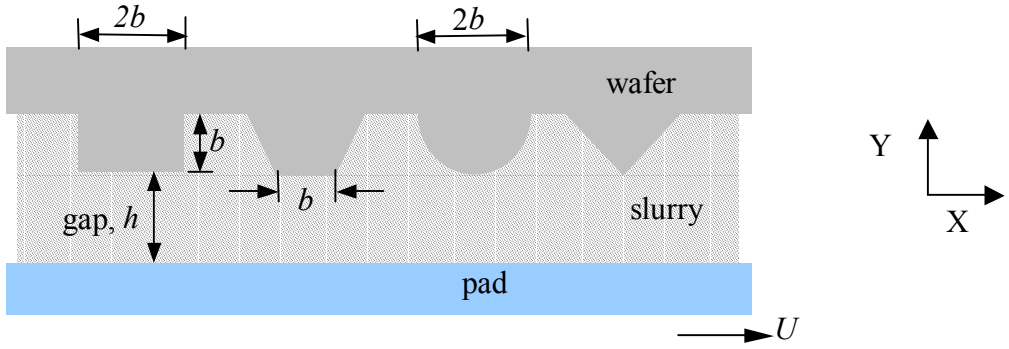


Figure 4.1 Schematic representations of different features

Hydrodynamic Calculations

In this section, a flow field of the slurry in the gap between the wafer and the pad is described. The pad and the wafer are assumed rigid with a square feature of side, b, projecting from the wafer. The channel gap between the feature and pad is, *h* (Fig. 4.1). Without losing the generality, the slurry flow is treated as a Newtonian fluid. The governing equations for the slurry flow in the gap are the same as the governing equations for wafer scale modeling, i.e., Eqs. 3.1 & 3.2. The shape of the feature is varied as shown in Fig. 4.1. For each feature, the aspect ratio and the gap size (*h/b*) also play important roles in determining the surface stress. All these factors as well as the feature shape are taken into consideration. The dimensions and properties of the model system are shown in Table 4.1.

Table 4.1 Modeling conditions

Modeling parameters	**Value**
Feature length *b*, μm	25
Relative velocity of the pad *U*, cm/s	60
Channel length, μm	525
Slurry viscosity η, cp	20
Minimum gap, *h/b*	0.1
Density ρ, g/cc	1

Due to the capability of numerical calculations, no additional assumptions are needed for the governing equations 3.1 & 3.2. The governing equations are first non-dimensionalized as equations 4.1 through 4.4.

$$\frac{\partial u_x^*}{\partial x^*} + \frac{\partial u_y^*}{\partial y^*} + \frac{\partial u_z^*}{\partial z^*} = 0 \tag{4.1}$$

$$\mathrm{Re}\left(u_x^* \frac{\partial u_x^*}{\partial x^*} + u_y^* \frac{\partial u_x^*}{\partial y^*} + u_z^* \frac{\partial u_x^*}{\partial z^*} \right) = -\frac{\partial p^*}{\partial x^*} + \left(\frac{\partial^2 u_x^*}{\partial x^{*2}} + \frac{\partial^2 u_x^*}{\partial y^{*2}} + \frac{\partial u_x^*}{\partial z^{*2}} \right) + \frac{\rho b^2 g_x}{\eta U} \tag{4.2}$$

$$\mathrm{Re}\left(u_x^* \frac{\partial u_y^*}{\partial x^*} + u_y^* \frac{\partial u_y^*}{\partial y^*} + u_z^* \frac{\partial u_y^*}{\partial z^*} \right) = -\frac{\partial p^*}{\partial y^*} + \left(\frac{\partial^2 u_y^*}{\partial x^{*2}} + \frac{\partial^2 u_y^*}{\partial y^{*2}} + \frac{\partial u_y^*}{\partial z^{*2}} \right) + \frac{\rho b^2 g_y}{\eta U} \tag{4.3}$$

$$\mathrm{Re}\left(u_x^* \frac{\partial u_z^*}{\partial x^*} + u_y^* \frac{\partial u_z^*}{\partial y^*} + u_z^* \frac{\partial u_z^{**}}{\partial z^*} \right) = -\frac{\partial p^*}{\partial z^*} + \left(\frac{\partial^2 u_z^*}{\partial x^{*2}} + \frac{\partial^2 u_z^*}{\partial y^{*2}} + \frac{\partial u_z^*}{\partial z^{*2}} \right) + \frac{\rho b^2 g_z}{\eta U} \tag{4.4}$$

The scaling variables are given as below.

$$u_x^* = \frac{u_x}{U} \quad u_y^* = \frac{u_y}{U} \quad u_z^* = \frac{u_z}{U} \tag{4.5a,b,c}$$

$$x^* = \frac{x}{b} \quad y^* = \frac{y}{b} \quad z^* = \frac{z}{b} \tag{4.6a,b,c}$$

Following the conditions of the model system given in Table 4.1, the Reynolds number is an order one quantity whereas the gravity effect is negligible.

$$\mathrm{Re} = \frac{bU\rho}{\eta} \sim 0.75$$

$$\frac{\rho b^2 g}{\eta U} \sim 5.1 \times 10^{-4}$$

It has been demonstrated in Chapter 2 that every point of the slurry flow under the wafer will have the same linear velocity when the wafer and the pad rotate at the same speed. It is the ideal case for reducing the polish non-uniformity. In current study, the pad is assumed to move with a constant velocity equivalent to the relative velocity between the wafer and the pad. On the inlet and the outlet of the flow domain, zero normal force condition is imposed, which is equivalent to imposing a periodic symmetry at the inlet and the outlet. Thus, although only a small part of the flow field spanning the entire wafer surface is considered in the calculation, the current modeling set up, i.e., a symmetry condition at the inlet and the outlet, is plausible if we assume that the flow domain under current consideration occurs repeatedly in a periodic manner. The boundary conditions are the no slip conditions at the pad and wafer surface.

The model system is discretized into a mesh of over 3000 elements. The mesh elements for different features are illustrated in Figure 4.2. The slurry flow profile is then obtained numerically using a finite element method. Polyflow (Version 3.6), a commercial software for solving finite element problems in fluid mechanics is used to solve the governing equations 4.1-4.4. The surface stress induced by the slurry flow can be derived from the definition of stress tensor T_{ij} for a Newtonian fluid:

$$T_{ij} = -p\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{4.7}$$

Here T_{ij} is the i, j^{th} component of the stress tensor, u_i is the velocity of the fluid in x_i direction, p is the pressure and η is the viscosity of the fluid. For Cartesian coordinates, x_i with i =1,2,3 corresponds to x, y, and z respectively. Denoting the outward normal to the surface by n_i, the stress on the surface is given by

$$\sigma_j = T_{ij} n_i \tag{4.8}$$

Solid Mechanics Calculations

The stress distribution on the wafer surface due to the slurry flow obtained using Eq. 4.8 is considered as a boundary condition to the solid mechanics problem. A finite element software was then used for solving stress-strain equations in solids. The distribution of stresses inside the wafer is obtained finally. The feature with dimensions same as that used in the fluid mechanics problem, is modeled along with a portion of wafer around the feature. PATRAN (Version 8.5) software is used for mesh generation while ABAQUS (Version 5.8) is used to solve the equations in solid mechanics. The solid wafer is meshed in such a way that the mesh at the surface matches that used for modeling the fluid flow.

The governing equations for solving stresses inside the solid wafer is

$$T_{ij} = \lambda \varepsilon_{kk} \delta_{ij} + 2\mu \varepsilon_{ij} \tag{4.9}$$

where T_{ij} and ε_{ij} are the stress and strain tensors, respectively. Solving for ε_{ij} we get

$$\varepsilon_{ij} = -\frac{\nu}{E} T_{kk} \delta_{ij} + \frac{1+\nu}{E} T_{ij} \tag{4.10}$$

where λ and μ are Lame's constants and are related to the Young's modulus E and Poisson's ratio, ν, as follows:

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)} \tag{4.11}$$

$$\mu = \frac{E}{2(1+\nu)} \tag{4.12}$$

The Young's modulus, E, for silica wafer is 193 GPa, and the Poisson's ratio, ν, is 0.3 [Siv92]. Equations 4.10, 4.11 and 4.12 are equivalent to the following six equations which constitute the isotropic Hooke's law

$$\varepsilon_x = \frac{1}{E}\left[\sigma_x - \nu\left(\sigma_y + \sigma_z\right)\right] \qquad \gamma_{yz} = \frac{1}{G}\tau_{yz}$$

$$\varepsilon_y = \frac{1}{E}\left[\sigma_y - \nu\left(\sigma_x + \sigma_z\right)\right] \qquad \gamma_{zx} = \frac{1}{G}\tau_{zx}$$

$$\varepsilon_z = \frac{1}{E}\left[\sigma_z - \nu\left(\sigma_x + \sigma_y\right)\right] \qquad \gamma_{xy} = \frac{1}{G}\tau_{xy} \tag{4.13}$$

where the shear modulus, G, is given by

$$G = \frac{E}{2(1+\nu)} \tag{4.14}$$

Boundary conditions for the solid mechanics problem are

(1) Top surface of the wafer is fixed, no translation or rotation at the top surface i.e., $u_x = u_y = u_z = 0$, $\omega_x = \omega_y = \omega_z = 0$

(2) Symmetry along the x-plane at the ends, i.e., $u_x = 0$, $\omega_y = \omega_z = 0$

(3) Bottom surface of the wafer is acted on by stresses imposed from the slurry.

The model was non-dimensionalized using the same scaling factors as in fluid mechanics calculations i.e. length scaled by b and the Young's modulus scaled by the scaling factor for stress $\frac{\eta U}{b}$.

Recall that the purpose of this study is to predict possible yield (or removal) of the wafer material by the slurry flow. With the known stress distribution inside the wafer, an appropriate yield criterion can then be applied.

Yield Conditions

For analysis of continuum stress and strain distribution, a constitutive theory of plasticity must satisfy the yield condition under combined stresses, since the uniaxial condition $|\sigma|$ = yield stress, is inadequate when there is more than one stress component. There are a number of yield conditions described in literature, which correlate yielding in a three-dimensional state of stress with yielding in the uniaxial tensile test. Two of the most commonly used yield conditions, namely von Mises yield condition and Tresca yield condition, will be discussed here.

von Mises yield condition

This condition assumes that yielding can occur in a three-dimensional state of stress when the root mean square of the difference between the principal stresses reaches the same value as it has when yielding occurs in a tensile stress. Letting Y denote the stress at which yielding begins in a simple tensile test, the principal stresses are $\sigma_1 = Y$, $\sigma_2 = 0$, $\sigma_3 = 0$. Thus in a tensile test yielding occurs when the root mean square of the differences between the principal stresses is

$$\sqrt{\tfrac{1}{3}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]}=\sqrt{\tfrac{1}{3}[Y^2+0+Y^2]}=\sqrt{\tfrac{2}{3}}Y \quad (4.15)$$

which can be simplified to

$$\sqrt{\tfrac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]}=Y \quad (4.16)$$

When a stress state is known in terms of stress components, with respect to non-principal axes, the Mises yield criterion equation is transformed to

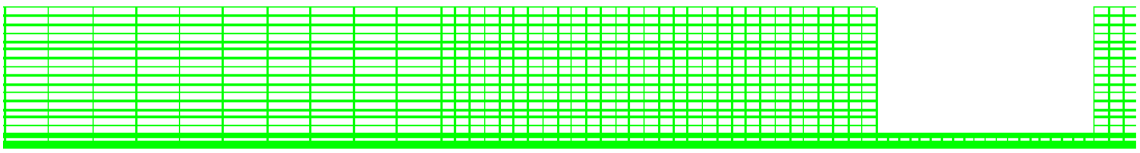
$$\sqrt{\tfrac{1}{2}[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2]+3\tau_{xy}^2+3\tau_{yz}^2+3\tau_{xz}^2}=Y \quad (4.17)$$

Tresca yield condition

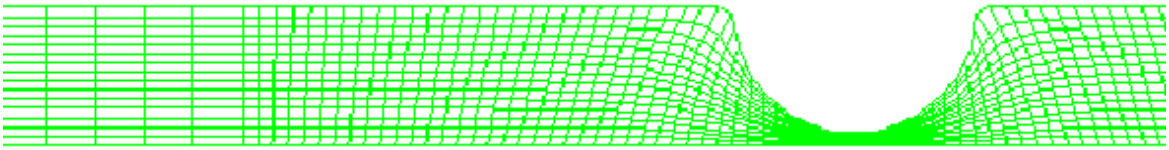
This condition assumes that yield will occur when the maximum shear stress reaches the value it has when yielding occurs in the tensile test. The maximum shear stress is one-half the difference between the maximum and the minimum principal stresses. In a tensile test the maximum shear stress is $Y/2$, so this yield criterion predicts that yielding occurs when

$$\tau_{max} = \frac{\sigma_{max} - \sigma_{min}}{2} = \frac{Y}{2} \tag{4.21}$$

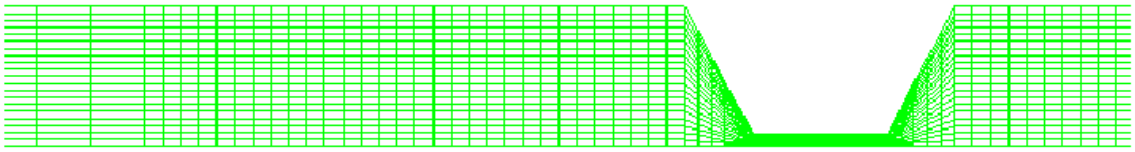
The yield strength of silica glass, which has almost the same composition as oxide layer on the wafer, is estimated to be 5 GPa [Hol86], but the observed strength for silica glass is usually less than a hundredth of this theoretical level due to the microcracks on the surface. Therefore, even though the actual yield strength of silicon oxide under CMP conditions is not available in literature, it is expected to be a lot smaller than silica glass given the harsh conditions the oxide is subjected to during CMP. Thus, the stresses induced by the slurry on the wafer may be sufficient to fracture material off the wafer surface. The objective of current study in solving for the stress distribution inside the wafer surface is to be able to apply a suitable yield criterion, and develop an understanding for the conditions under which the fracture of the wafer material may take place from the wafer surface.



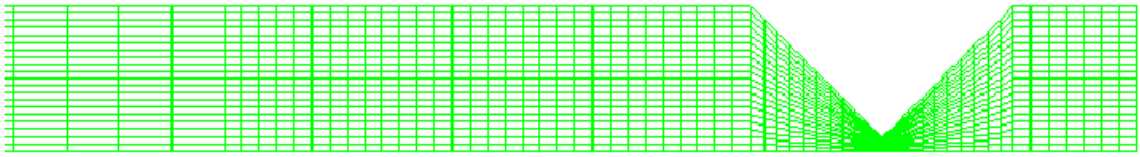
(a)



(b)



(c)



(d)

Figure 4.2. Portion of mesh generated for solving 2-D fluid mechanics problems, (a) rectangular feature, (b) semi-circular feature, (c) trapezoid feature, (d) triangular feature.

Results and Discussion

The modeling results of square shaped features for both slurry flow and solid mechanics calculations have been obtained by Agarwal [Aga00]. It is reproduced in current study for the purpose of comparison. However, the same results will not be presented in this discussion. Only new calculation results will be included. Readers can refer to ref. [Aga00] for the modeling details of square shaped features.

Hydrodynamic Analysis

For a two dimensional feature, the stress that imposed from x-direction and the pressure are the two most important factors induced by the slurry flow. Figure 4.3 shows the pressure fields due to the presence of different features. The profile is generated on each feature with h/b=0.1. The length of the channel is 21 units, and the total height of the channel is 3 units. The relations of the dimensions are shown in Fig. 4.1. It is clear that a pressure build up on the front side whereas a negative pressure is created in the back due to the presence of the feature. A shear flow that acts against a positive pressure gradient in the front region whereas is dragged in the rear region will be generated consequently. Since the gap sizes are kept the same, the maximum pressure build-ups are approximately the same for different features except the triangular feature. Both the peak value and the slope of the triangular shaped feature on the pressure curve are much smaller than the others. Apparently, the slurry is easier to flow through the triangular feature, and less pressure is expected to impose to the feature surface. Thus, it may result in a smaller removal rate.

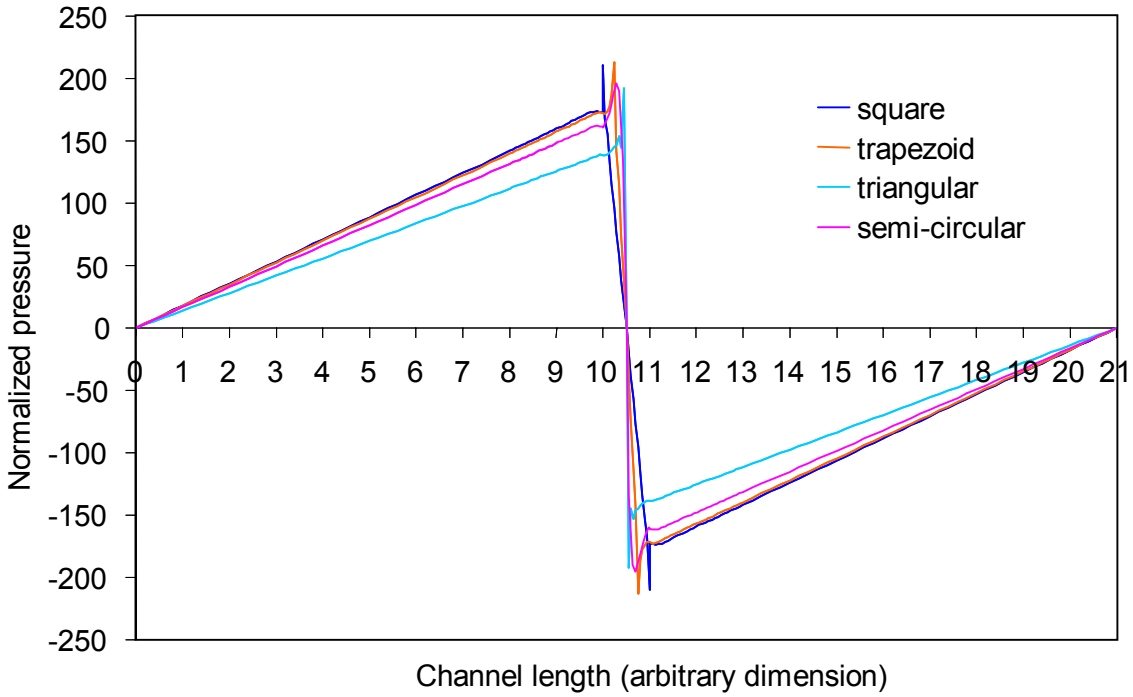


Figure 4.3. Pressure profile developed by slurry flow for different features

Beside the pressure profile, the shear stress in the x-direction is another critical factor to be investigated. In Figure 4.4, the x-directional stresses are plotted for the same features as in Fig. 4.3. A sudden increase in the shear stress is shown for all the features investigated due to the abrupt decrease in the flow channel, and the profiles of the shear stress matches well with the shape of the features. The square shaped feature shows the lowest stress on its surface among all the features, whereas it has maximum stress in the corners due to the sharp change on the edges. It indicates qualitatively that the evolution of the square feature during CMP may start with the rounding of the corner, and then the polishing of the feature surface. For the triangular feature, it has maximum shear on the tip. Consequently, it is possible for it to wear off the tip first during polishing.

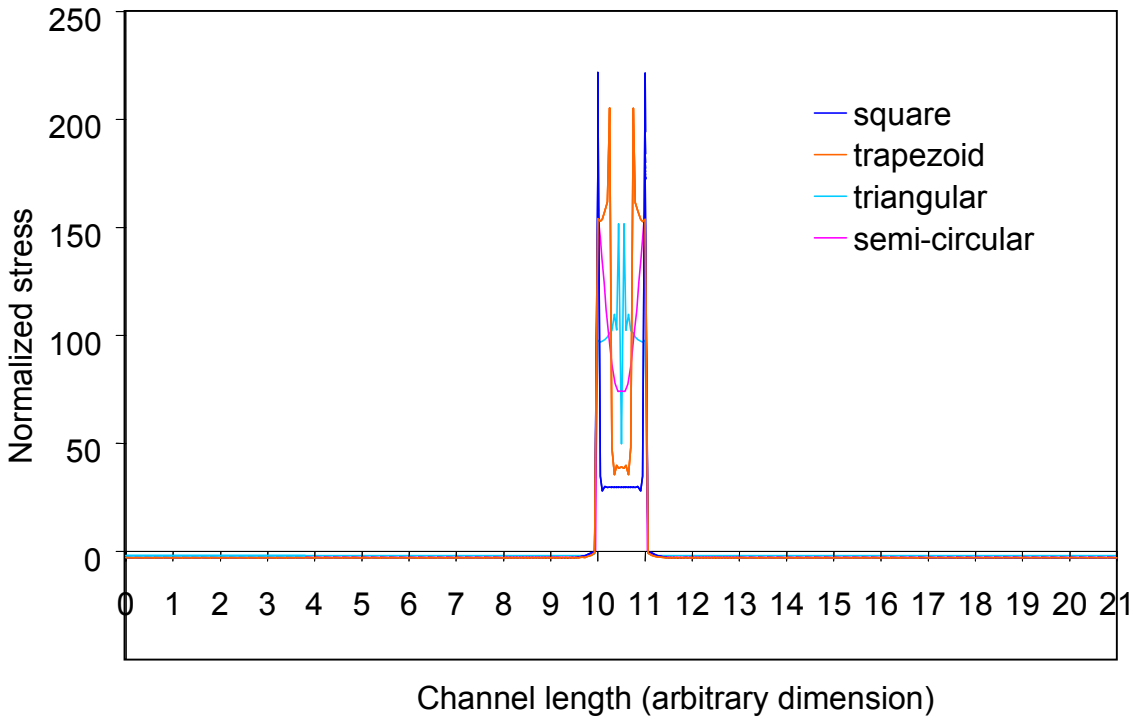


Figure 4.4. Stresses in x-direction imposed by slurry flow for different features

Solid Mechanics Analysis

The results from calculations solving mechanics equations in solid for different features are shown in Figures 4.5-4.8. Interestingly, all the features clearly indicate stress concentration at the corners where the feature protrudes out of the wafer. It also proves, at least indirectly, that the stress concentration is not an artifact of the finite element method. As is evident from the figures, stresses in wafer are negligible away from the feature. There is a high concentration of von Mises stresses at the corners and the propagation of the stresses can be clearly seen through the spanning patterns in the upper region of the features. As is pointed out previously, if the stress corrosion is the mechanism for material removal, the stress concentration may result in “necking” around the feature with particles of small size leaving the feature. This process will continue till the “neck” is no longer thick enough to withstand the stresses and the feature will break

away. On the other hand, if the maximum shear stresses, which span the neck of the feature, are larger than the yield stress of the material, the entire feature can just break off. Thus it is helpful to compare the maximum stress for different features. As shown in Figure 4.9, the rectangular shaped feature has the highest stress at its corners among all the features, while the triangular feature has the least. For the rectangular feature, the maximum von Mises stress value is about 0.27 MPa (Fig. 4.5), which is approximately twice as much as the maximum value of the triangular feature. Three-dimensional modeling results are not presented due to the fact that the stress value inside the wafer is expected to be much smaller than the stress obtained in two-dimensional calculations.

Although the maximum von Mises stress value is orders of magnitude smaller than the aforementioned yield stress of the glass (or silica) wafer, it should be noted that under the harsh CMP environment, the yield strength of SiO_2 on the wafer can be significantly lower than the values reported on the strength of materials, in literature. When the chemicals in the high pH slurry react with the wafer surface, a soft layer may form and cause the decrease of the yield strength of the material. In addition, stress corrosion is still possible due to the development of micro-cracks on the wafer surface around the corners. The increased diffusivity of slurry additives in the wafer under tensile stress makes the process of erosion even faster.

Another insight derived from these modeling results is that the strength of the material in IC manufacturing is critical for the CMP process. The maximum stress of the features imposed by the slurry flow alone during CMP is about one order of magnitude higher than the bulk stress value due to the stress concentration effect. It indicates that, in

the selection of wafer materials, the yield strength of the material has to be compared with the possible concentrated stress value.

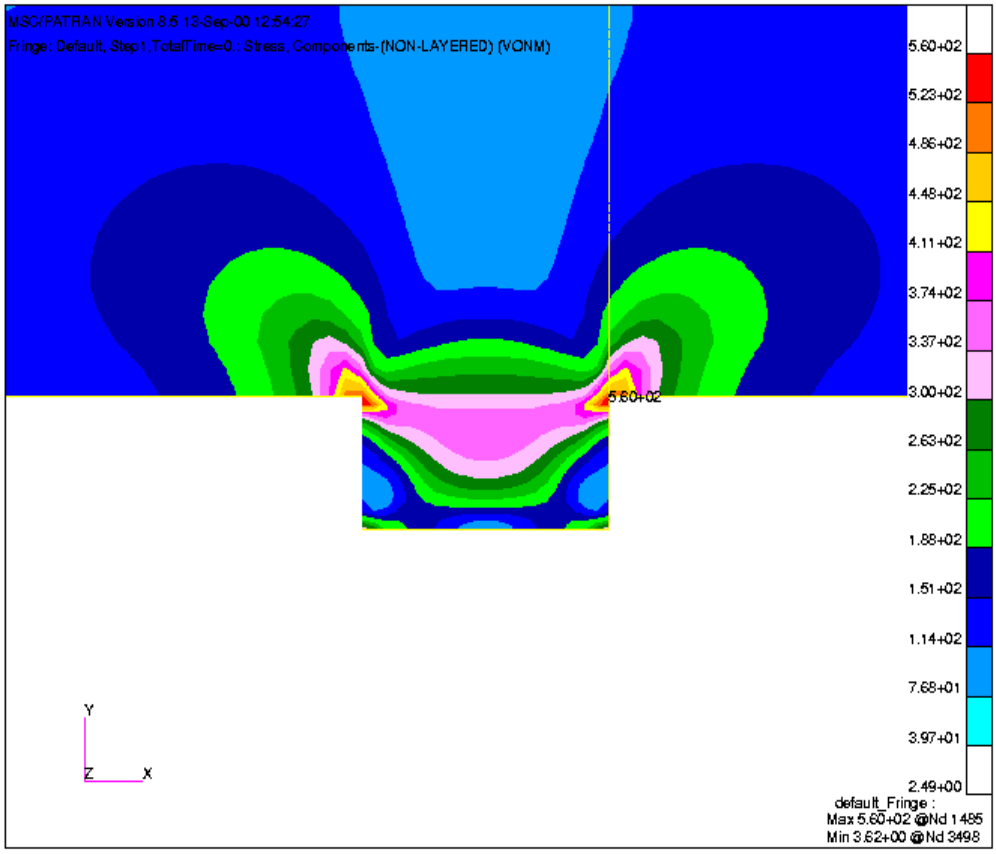


Figure 4.5 von Mises stresses in a rectangular feature induced by slurry in channel with gap width $h/b = 0.1$. The stresses have been non-dimensionalized by a factor of $\eta U/b$ (= 480 Pa). Red color represents maximum stress while light blue represents minimum stress.

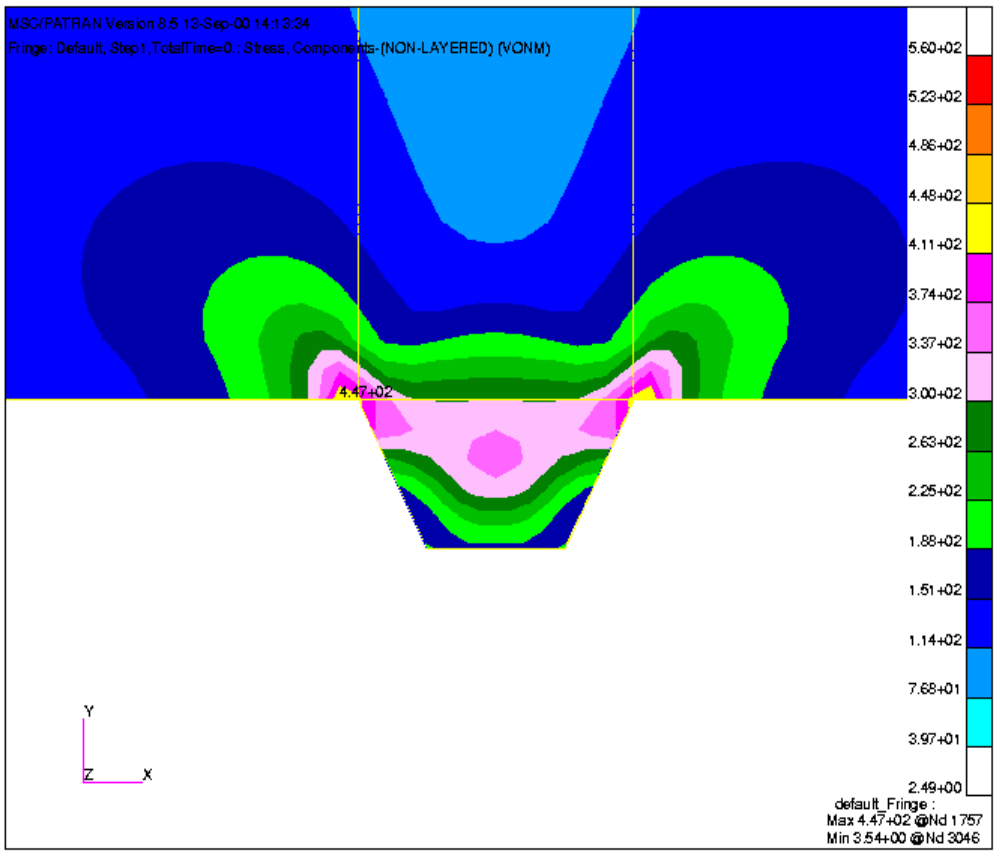


Figure 4.6 von Mises stresses in a trapezoid feature induced by slurry in channel with gap width $h/b = 0.1$. The stresses have been non-dimensionalized by a factor of $\eta U/b$ (= 480 Pa). Red color represents maximum stress while light blue represents minimum stress.

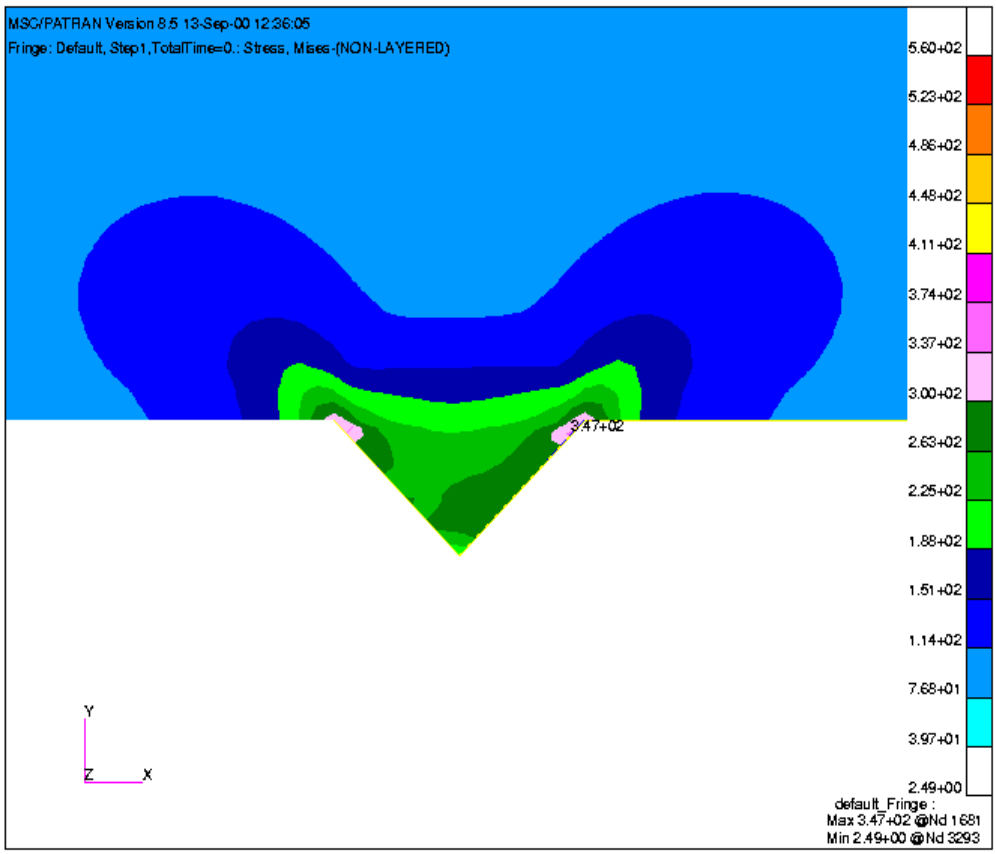


Figure 4.7 von Mises stresses in a triangular feature induced by slurry in channel with gap width $h/b = 0.1$. The stresses have been non-dimensionalized by a factor of $\eta U/b$ (= 480 Pa). Red color represents maximum stress while light blue represents minimum stress.

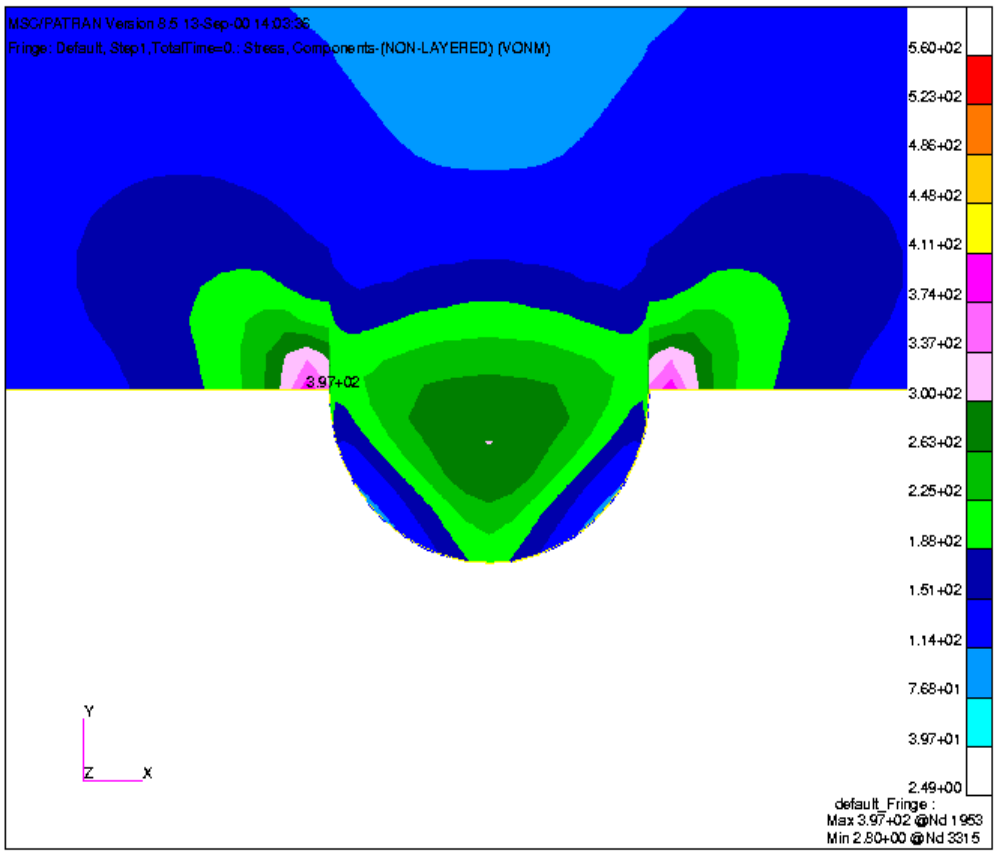


Figure 4.8 von Mises stresses in a hemi-circular feature induced by slurry in channel with gap width $h/b = 0.1$. The stresses have been non-dimensionalized by a factor of $\eta U/b$ (= 480 Pa). Red color represents maximum stress while light blue represents minimum stress.

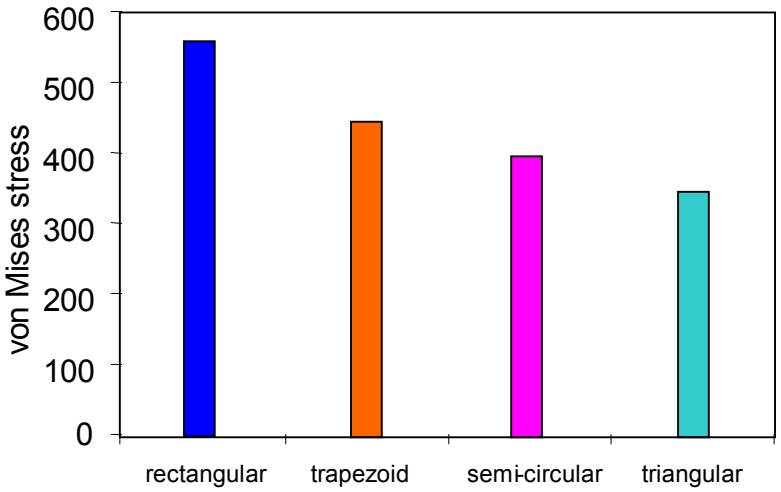


Figure 4.9 Comparison of maximum stress values for different features

Summary

In this chapter, numerical methods are used to solve both the slurry flow in the gap and the stress distribution inside the wafer. A pressure build up is found around the features due to its blockage of the slurry flow, which in turn results in stress concentration in the corners of the feature where it connects to the wafer surface. The square features have the highest von-Mises stress in the concentrated area among the feature shapes studied. When proper yield condition is applied, yield and thus the possible removal of material from the wafer surface could be predicted using the results from this study. However, the calculated von-Mises stress does not satisfy the yield condition if the chemical effects are not considered. Thus, the effects of slurry chemicals and the mechanism of the polishing process need to be further investigated.

CHAPTER 5
A MICRO-CONTACT MODEL OF THE MATERIAL REMOVAL

Introduction

In previous chapters, the wafer scale slurry flow under the wafer and the feature scale stress distribution inside the wafer are investigated respectively. In addition, it has been shown in Chapter 2 that the delamination due to the slurry shear flow alone is highly unlikely. Meanwhile, it is demonstrated in Chapter 3 that stress concentration in the corners of the features can cause possible stress corrosion and then removal of the feature. However, this mechanism is not expected to be the primary material removal mechanism because even the maximum stress is smaller than the yield strength of the wafer material.

As discussed in Chapter 2, the understanding of the interactions between the pad and wafer is crucial to process characterization, process modeling and optimization. The modeling of the CMP mechanism has fallen into two categories, namely hydroplaning-based modeling and contact mechanics based modeling, according to the two extremes in dealing with the pad/wafer interactions. Both approaches have been efficient in addressing some aspects of the CMP process. However, due to the complexity of CMP by multiple involvements of the parameters and concurrent polishing of wide range of materials, there is no explicit methodology to characterize the polishing mechanism with process parameters. Moreover, the chemical effects induced by the slurry reagents have been completely ignored. In this chapter, the chemical effects of different polishing materials (e.g., copper, tungsten and silica) during CMP are discussed. Then the theories

about the polishing mechanisms are reviewed. Based on the understandings of the chemical effects and the mechanisms leading to material removal in CMP, a comprehensive model for calculating the material removal rate of the CMP process is presented in which both chemical and mechanical effects are taken into consideration. The chemical effects come into play through the formation of a chemically modified surface layer on the wafer surface that, in turn, is removed mechanically by the plastic deformation induced by slurry particles.

Slurry Chemistry Effect on the CMP Performance

Metal CMP

Chemical effects specific to a certain system have been evident in various experimental studies. In metal CMP, slurry chemistry is believed to alter the properties of the metal surface to be polished such that high material removal rate, acceptable surface quality and successful planarization can be achieved. The effect of forming the surface layer on the material to be polished is two fold. One is to facilitate the material removal in the higher region on the wafer surface and prevent scratches. The other is to protect the lower region from etching by the slurry chemicals such that the planarization can be obtained. For example, the formation of tungsten oxide on the tungsten metal layer has been shown to be necessary to achieve satisfactory material removal in tungsten polishing. It has been found that the material removal rate is negligible in the absence of the oxidizer (e.g., $K_3Fe(CN)_6$) in tungsten CMP [Bie98]. Furthermore, the passivation effect of the oxide layer formed on the tungsten surface has been demonstrated by Bielmann [Bie98] through wear experiments for tungsten polishing by rubbing the surface using pin-on-disk. The current was measured during the rubbing of the surface. It is found that the current went to near zero when the oxide surface formed, whereas when

the fresh metal surface was exposed to the slurry by rubbing, current increased as the oxidation reaction started on the bare tungsten surface [Bie98]. These findings indicated that the continuous formation and removal of the passivated layer enables the material removal and planarity for metal CMP. Based on the same understandings, Kaufman proposed a polishing mechanism of tungsten, which leads to the planarization of the wafer surface by achieving topographic selectivity [Kau91]. According to this mechanism, the passive oxide layer on the tungsten wafer is removed by the abrasive particles at the higher levels, while the metal at the recessed area is protected from further etching by the formation of passive film. In copper polishing, the same mechanism is expected although more options of appropriate chemistries to form the protective surface layer, which is not necessary to be a native oxide layer, can be applied due to the maturity of copper passivation chemistry. Therefore, the mechanical properties of the chemically modified surface layer have to be studied systematically to achieve a better understanding of the mechanism of material removal, and effective slurry designs can be developed including the chemistries to form a protective layer to result in sufficient material removal with minimal surface defect formation.

Silica CMP

So far, silicon dioxide has been the most commonly used oxide in IC manufacturing as a dielectric material. In silica CMP, formation of a chemically modified layer has been suggested to be necessary to achieve high material removal rate with acceptable surface quality [Izu79]. Particularly, polishing rates were observed to be near zero in hydrocarbon liquids such as kerosene, paraffin or oil as well as in liquids without hydroxyl groups, such as formamide [Coo90]. The surface quality of the polished wafer in such liquids was also detected to be poor. The chemical modification of the surface in

silica CMP is proposed to be provided by the water molecules [Coo90]. It is believed that the interactions between the siloxane bonds (Si-O-Si) and water primarily determine the behavior of silica surface during polishing. A hydrated surface layer is formed as a result of the reaction for the breakage of the network forming siloxane bonds with water [Ile79], i.e.,

$$H_2O + \text{Si-O-Si} \rightleftharpoons 2\text{Si-OH}$$

The reaction is completed in three steps [Bud61]. First, a water molecule from the environment attaches to the Si-O-Si bond by aligning itself due to the formation of a hydrogen bond between its hydrogen and the oxygen atom of the silica surface, and interaction of the lone pair orbitals of the oxygen in water structure with the Si atom. In the second step, a reaction occurs in which proton transfer from water hydrogen to the oxygen of silica is accomplished simultaneously with electron transfer from the oxygen of water to the silica atom. Then, two new bonds are formed by destroying the original bond between the silica and oxygen resulting in Si-O-H bonds on the surface due to these reactions. The diffusion of water into silica structure is believed to be the control factor of the reaction rate [Dor73]. The diffusion coefficient of water in silica at ambient conditions and neutral pH values were measured to be quite low [Nog84, Lan85]. A significant increase in diffusion rate of water into silica structure occurs at high pH values (e.g., pH > 9.0). Iler reported three orders of magnitude increase in silica dissolution rate as the pH value was increased from 2 to 11 [Ile79]. In agreement with these observations, the removal rate of silica polishing was observed to increase significantly with slurries of above pH 9 [Izu84].

Direct evidence was presented by Trogolo and Rajan with transmission electron microscopy (TEM) and Fourier-transform infrared spectroscopy (FTIR) analyses, indicating possible chemical/structural modification of the silica surface after CMP [Tro94]. Two subsurface regions were observed on the wafer surface after polishing, namely, a 2 nm thick surface region possessing lower density than the bulk and underneath, and a 15-20 nm thick region with a higher density according to the grazing angle X-ray density measurements. FTIR analysis confirmed that these were hydrated layers. In a surface forces measurement, Adler has indirectly revealed evidence of the formation of a 'gel' like layer on the silica surfaces in water environment [Adl01]. By evaluating the short-range repulsion on silica surface in a water-silica system under a variety of solution environments he concluded that this behavior could be explained due to the swelling of the silica surface as a result of water diffusion into structure. The surface layer, also referred to as the gel-layer, is suggested to be softer than the silica itself. Besides the surface layer formation, it is the slurry particles that transfer the mechanical forces and achieve the material removal. Thus, a balance between the chemical and mechanical interactions during CMP has to be considered for optimal polishing performance.

Mechanisms of Material Removal

Characterization of the Wafer-Pad Contact Mode

As pointed out in previous chapters, it is critical to understand the contact mode between the polishing pad and the wafer. Three operational regimes have been proposed, namely, hydroplaning, direct contact and mixed mode (or semi-contact), during CMP when the wafer is pressed against the polishing pad and sliding with an intervening fluid layer. Hydroplaning is assumed to be dominant when the sliding speed is sufficiently

high and the applied load is relatively low. On the other hand, direct contact between the pad and the wafer will dominate when the velocity is low and the applied load is relatively high, which causes mechanical deformation of the pad asperities at the contact spots. However, it is reasonable to believe that the practical CMP process involves both hydroplaning of the slurry film and the direct contact between the pad and the wafer. The dominant mode can be derived from the friction force measurement.

Considering the polishing pad shown in Fig. 2.7, the friction between the pad and wafer consists of contributions from two aspects, i.e., the shear between the wafer and the pad asperities and the shear by slurry flow. Mathematically, it can be written as

$$F_f = \sum_{i=1}^{n_c} \iint_{A_{c,i}} \tau_{c,i} dA_{c,i} + \iint_{A_h} \tau_h dA_h \tag{5.1}$$

where F_f is the friction force between the pad and wafer, and τ and A are the shear stress and contact area respectively; the subscripts c and h stand for asperity contact and hydrodynamic fluid film respectively; n_c is the number of asperities in direct contact with the wafer surface during polishing. The summation of the total contact area by the pad asperities A_c and the area occupied by the slurry fluid A_h is essentially the total wafer area A_o, i.e.,

$$A_o = \sum_{i=1}^{n_c} A_{c,i} + A_h = A_c + A_h \tag{5.2}$$

For every asperity i, the friction coefficient is given by the definition of the friction coefficient μ:

$$\mu_{c,i} = \frac{\tau_{c,i}}{p_{c,i}} \tag{5.3}$$

where p is the contact pressure. Without losing the generality, we can assume that the coefficients of friction for all the asperities are the same, i.e., $\mu_{c,i} = \mu_c$. Meanwhile, assuming the friction coefficient of fluid flow is μ_h, the overall coefficient of friction is thus given as

$$\mu = \left(\sum_{i=1}^{n_c} \iint_{A_{c,i}} \tau_{c,i} dA_{c,i} + \iint_{A_h} \tau_h dA_h \right) \Big/ L = \alpha \mu_c + (1-\alpha)\mu_h \tag{5.4}$$

where L is the total load applied on the wafer, and α is the fractional factor defined as the ratio of the load that is supported by the pad asperities to the total load. From Eq. 5.4, it is clear that the operational regime of the CMP process can be derived from the measurement of the overall coefficient of friction. It is well known that the coefficient of friction for solid-solid contact (e.g., pad and wafer) is in the order of magnitude of about 0.1, i.e., $\mu_c \sim 0.1$ [Mah00]. The friction coefficient due to slurry lubrication can be obtained from the derivations of the lubrication equation (i.e., Eq. 2.7). The shear stress induced by slurry flow is given as

$$\tau = \frac{1}{2} h \left(\frac{dp}{dx} \right) + \eta \frac{V}{h} \tag{5.5}$$

Here, all the notations in this equation follow the illustration in Fig. 2.10. The shear stress increases with the increase of fluid viscosity and flow velocity. By scaling the corresponding variables, we can approximate the friction coefficient of slurry flow as

$$\mu_h = \frac{\tau}{p} \sim \frac{h}{D} \approx 10^{-4} \tag{5.6}$$

where, p is the applied normal pressure, D is the wafer diameter. Since the friction coefficient derived from the lubrication approximation is about three orders of magnitude

smaller than it is from pad-wafer direction contact, it is possible to determine the operational mode of the CMP process simply by measuring the overall coefficient of friction during CMP.

Mahajan measured the in-situ friction force by mounting a lateral force measuring instrument to the wafer polisher [Mah00]. To measure the frictional forces, a Sensotec model load cell was employed. Its output was connected to a data acquisition system and data recorded every 250 ms for duration of 60 seconds. The friction coefficients that are calculated from Mahajan's measurements as a function of applied pressure and pad velocity are shown in Figures 5.1 & 5.2 respectively. It is seen from Fig. 5.1 that the friction coefficient is independent of the applied pressure on the wafer when the applied load is higher than 20kPa. The order of magnitude of the friction coefficient is well within the aforementioned solid-solid contact range, indicating that the contact mode is dominant during CMP. Likewise, the same trend can be seen from Fig. 5.2. Although the friction coefficient slightly increased as it is expected when the pad speed is low, the order of magnitude is still the same. No slurry lubrication effect was observed in the range of the experiment. Thus, direct contact between the pad and the wafer will be considered in the following model development.

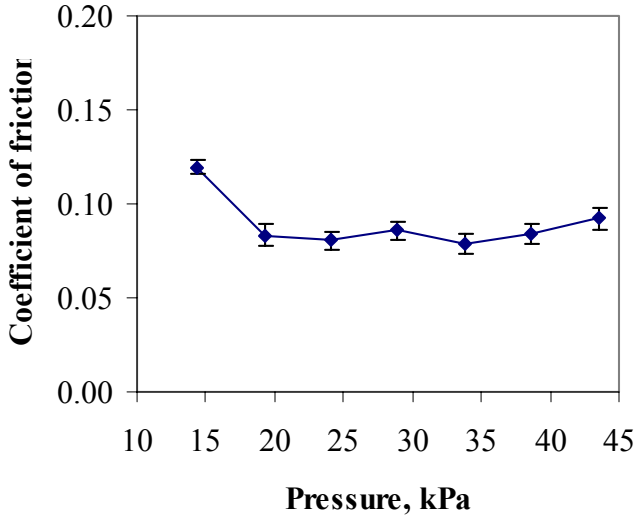


Figure 5.1 Coefficient of friction as a function of applied pressure. The pad velocity is fixed at 2.19 m/s.

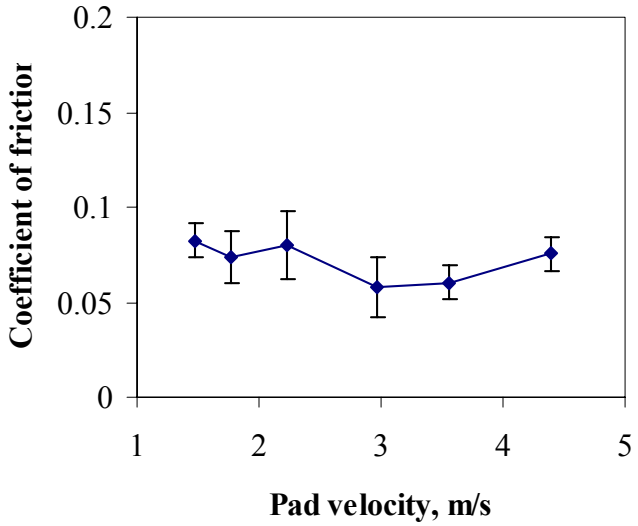


Figure 5.2 Coefficient of friction as a function of pad velocity. The applied pressure is set at 29.0 kPa.

Material Removal Mechanisms

Upon solid-solid contact during polishing, several models have been proposed, e.g., brittle fracture, shear delamination, adhesive wear, and abrasive wear. Each model emphasizes some fundamental mechanism of polishing and attempts to explain the phenomenon of material removal. The material removal of the contact mode in CMP is mainly through the mechanical force transferred by abrasive particles embedded in the pad asperities. Due to the complexity of the CMP process, it is unlikely for all the aspects of polishing to be explained by only one mechanism. The following examination in this section is to elucidate the dominant mechanisms during polishing so that a comprehensive model can be established based on these understandings.

Brittle fracture

Fracture on the surface is possible when the material to be polished is relatively brittle, e.g., glass. The mechanism in metal CMP has been proposed as brittle fracture by relating the Pilling-Bedworth ratio to the cracking of the metal oxide surface [Bas02]. However, it should be noted that plastic deformation is the dominant mode of deformation that induces cracking because the indentation depth by the slurry particles is very small in the order of several nanometers. Cracks initiate when the load on the particle exceeds a transition threshold L_p^* given as equation 5.7 [Eva80].

$$L_p^* = \kappa\left(\frac{K_c^4}{H^3}\right) f\left(E/H\right) \tag{5.7}$$

where κ is a dimensionless constant, K_c the fracture toughness, H the hardness,and E the Young's modulus of the material being polished. For lateral cracks, the function $f(E/H)$ in Eq. 5.7 varies slowly with E/H and the value of $\kappa f(E/H)$ is approximately 2×10^5 for both metal and ceramic materials [Eva80]. Compared to the force applied to each

particle, which is about 10^{-6}N for a typical CMP with 200 nm abrasive particles (see following sections for detail), it is clear that the critical load required for lateral cracking is many orders of magnitude greater. In another calculation, Bifano *et al.* [Bif91] have shown that the ratio between the plastic flow energy and the fracture energy in material grinding is proportional to the indentation depth. With the decrease in the indentation depth, plastic flow becomes an energetically more favorable mechanism for material removal. Studies on ductile grinding based on the bonding property of materials have also indicated that materials are more likely to be machined in a ductile manner if the depth of cut is below 10nm [Sre01]. Therefore, it is highly unlikely that brittle cracking is the dominant mechanism of material removal in CMP. Additionally, small particles employed in CMP can prevent the initiation of cracks and the subsequent fracture. This again suggests that the prevailing mechanism of material removal is due to excessive plastic deformation.

Corrosive wear

Corrosive wear occurs as a result of a chemical reaction on a wearing surface. The most common form of corrosion is due to the oxidation reaction on the metal surface; however, other chemicals in the slurry may also contribute. Corrosion products, usually oxides, have shear strengths different from those of metal surface being polished. The oxides tend to flake away, resulting in the pitting of' wearing surfaces. The wear rate of the corrosion products is believed to be increased significantly when the surface is attacked by the particles in the slurry flow, which is a process commonly referred to as erosion corrosion. However, as it is analyzed in the Chapter 3, the velocity of slurry flow in CMP is relatively low. It suggests that erosion corrosion, even if it occurs, is not the dominant material removal mechanism under CMP conditions.

Adhesive wear

Adhesive wear occurs because of shearing at points of contact or asperities that undergo adhesion or cold welding. Theoretically, this type of wear does not remove material but merely transfers it between wearing surfaces. However, the transferred material is often loosely deposited and eventually flakes away in microscopic particles; these, in turn, cause wear. Surface asperities are "cold welded" together and particles are sheared off as surfaces move. Adhesive wear is related to high shear stresses caused by friction. Shearing occurs through the weakest section, which is not necessarily at the adhesion plane. Adhesive wear is often increased by low hardness, low ductility and cleanness of the surfaces. If cold, solid-phase welding occurs in the micro-contacts, the coefficient of friction is generally very high. In CMP process, however, the measured coefficient of friction (e.g., 0.1) is not sufficiently high to cause welding between surfaces. In addition, slurry additives like surfactants and polymeric dispersants that are used for the stabilization of the slurry particles will absorb on the particle surface and cause lubrication between the wafer and the abrasive particles, which will decrease the probability of adhesive wear. Thus, it is not a viable mechanism under typical CMP conditions.

Abrasive wear

Abrasive wear occurs when a hard surface slides against and cuts grooves from a softer surface. The rough surface acts as a cutting tool to the soft surface. This condition is frequently referred to as two-body abrasion. When particles are introduced between wearing surfaces, the wear of the surfaces is due to the particles embedded in one of the sliding surfaces. This process is referred to as three-body abrasion. Clearly, the particles involved in the polishing process have to be harder than the sliding surface materials for

the abrasion to maintain a high wear rate. In the CMP process, the particles act as single point cutting tools, dig into the wafer and plow a series of grooves in it. Therefore, the CMP process is essentially a three-body abrasive wear. The analysis of the texture of polished surfaces also supports this observation. Figure 5.3 illustrates a typical post CMP surface as examined via an optical surface profiler (Veeco NT 1000 interferometer). The surface texture consists of numerous of randomly oriented nanogrooves. The width and depth of the grooves are determined by the penetration depth of the abrasive particles, which is consistent with traveling Hertzian loaded contact [War91]. Examination of other semiconductor materials after CMP also showed the similar patterns. From the data to date, it appears that abrasive wear is the dominant mechanism for material removal in CMP. Thus, a reliable model can be developed based on this polishing mechanism.

In the following sections, a model that incorporates both the chemical effect, which is due to the slurry chemicals, and the mechanical effect, which is primarily abrasive wear, is presented. Once a wafer is exposed to the chemicals in a slurry, a thin layer is formed whose mechanical properties such as modulus and hardness are quite different from those of the bulk wafer material. Such changes, in turn, affect the removal rate (or wear rate) significantly.

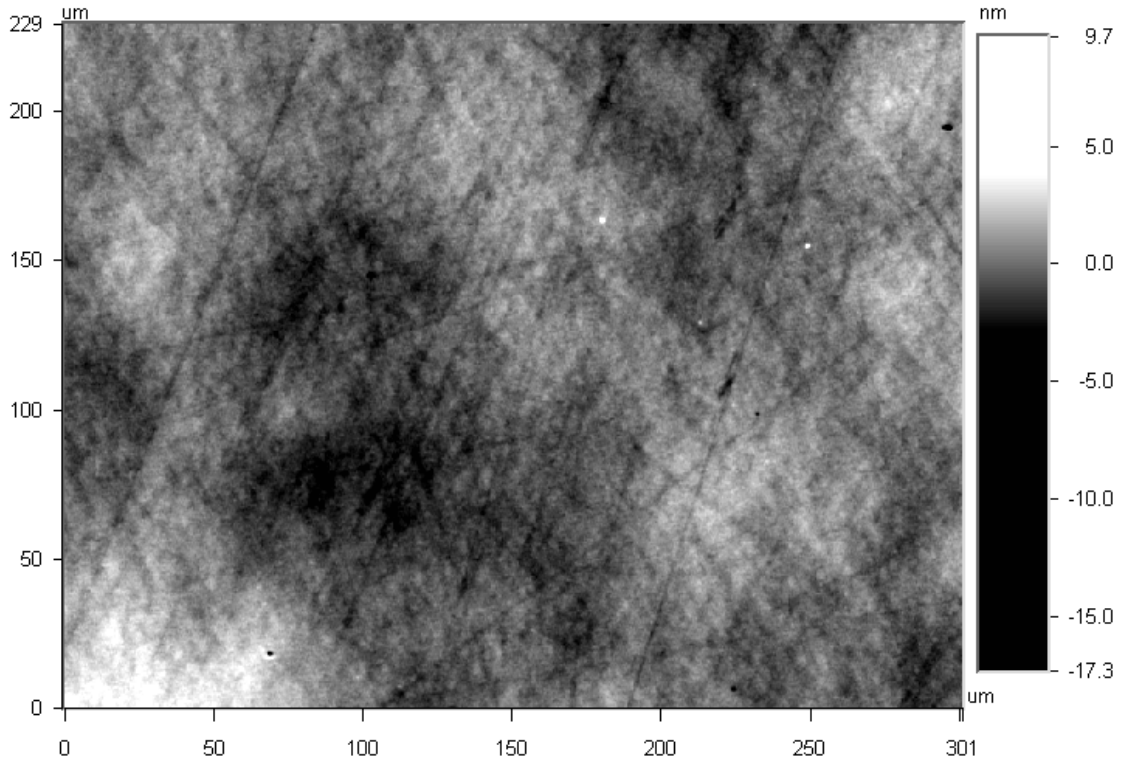


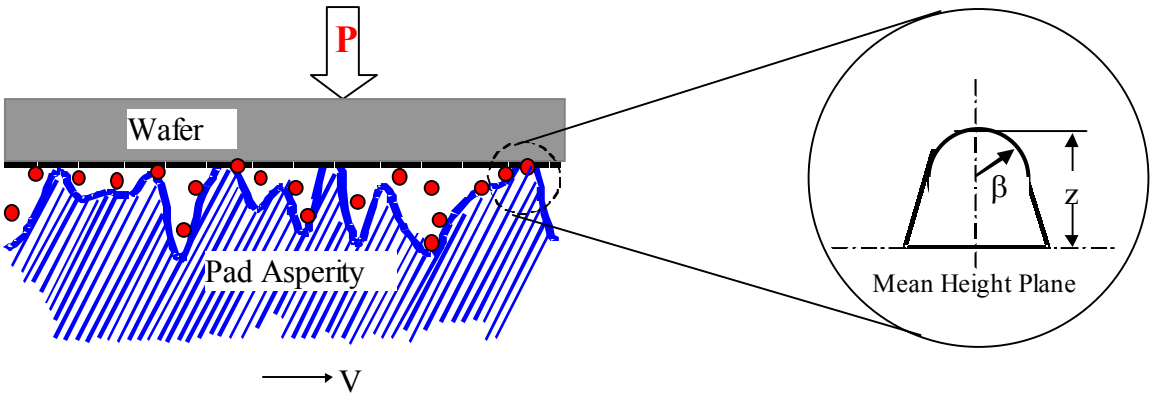
Figure 5.3 An optical image of a silica wafer surface after CMP showing nanoscratches by the slurry particles.

Model Development

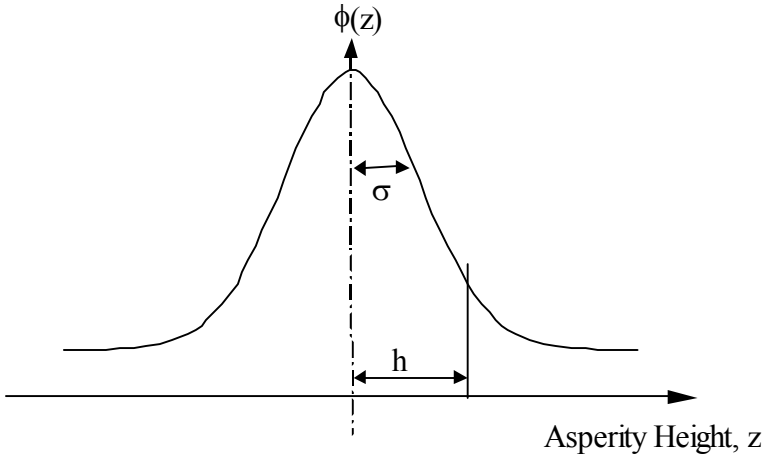
The chemicals in a slurry alter the properties of the wafer surface that comes into contact with them. Since the chemical influence is restricted to a thin region near the wafer surface, it may be plausible to assume the existence of a thin surface layer whose mechanical properties are different from those of the bulk wafer. This surface layer can be formed by chemical reactions at the surface or by the diffusion of ionic species in the slurry into the wafer. The surface layer is then removed by mechanical action of the slurry particles and/or by dissolution into the slurry through a convective diffusion process. Although there may be a few exceptional cases [Luo98], material removal by the

dissolution is small compared to that by the mechanical action of the slurry particles, and it is not considered in this study.

Figure 5.4 describes schematically the physical situation that is considered in the present analysis. The pad asperities in contact with the wafer surface serve as the anchoring sites for the slurry particles. The anchored particles then gouge the chemically altered wafer surface resulting in material removal. In the following, various parameters considered in the present analysis and the underlying assumptions are described followed by the details of the model development.



(a)



(b)

Figure 5.4 Schematic of the pad surface topology (a) that is in contact with the chemically modified layer formed on the wafer surface; (b) Distribution of the pad asperity heights (When the height of an asperity is greater than the average gap h between the pad and the wafer, that asperity is in contact with the wafer surface.)

A Polishing Pad in Contact with a Wafer

A polishing pad is typically made of polyurethane or polyurethane-impregnated felts. The Young's modulus of the pad depends not only on the temperature but also on

the contact time (or soaking time) with the slurry. As the temperature or the soaking time increases, the pad material softens and exhibits smaller modulus [Ste97]. Thus, the pad is much easier to deform than the wafer when they are pressed against each other. Typical values of the modulus of the wafer and pad materials are shown in Table 5.1.

Topology of a pad surface has been characterized and described statistically by Yu *et al.* [Yu94]. Their measurements indicated that pad asperity heights were in the order of tens of microns and appeared to have a normal (or Gaussian) distribution. In Figure 5.4, the topology of the pad surface is shown schematically along with a statistical description characterizing the pad asperities. In this figure, β is the mean radius of curvature of the tip of the pad asperities, ϕ is the distribution function for the height of the asperities, and σ is the standard deviation of the pad asperity heights, respectively.

Since the characteristic length scale for the roughness of the pad surface (e.g., β) is much larger than the feature sizes on the wafer, it is plausible to assume that the wafer surface is smooth and flat compared to the pad surface. This assumption makes it possible to use the Greenwood-Williamson model [Gre66] that describes the contact area, A_r, between a flat surface and a randomly rough surface under a normal force (or load), L:

$$A_r = \pi N A_o \beta \int_h^\infty (z-h)\phi(z)dz \tag{5.8}$$

$$L = \frac{4}{3} N A_o E^* \beta^{1/2} \int_h^\infty (z-h)^{3/2}\phi(z)dz \tag{5.9}$$

Here N is the number density of the pad asperities (i.e., number of asperities per unit planform area), A_o is the total area of the flat wafer surface, and h is the average gap between the smooth wafer and the rough pad under the given conditions. Then, the *down-pressure*, P, applied to the wafer with an area A_o is L/A_o whereas the actual *mean contact*

pressure, P_r, is L/A_r if the normal force applied to the wafer is entirely supported by the pad asperities in contact with the wafer. That is

$$L = PA_o = P_r A_r \tag{5.10}$$

E^* is the composite modulus between the two surfaces in contact that is defined as

$$\frac{1}{E^*} = \frac{1-\nu_p^2}{E_p} + \frac{1-\nu_w^2}{E_w} \tag{5.11}$$

Here E and ν are the Young's modulus and the Poisson's ratio, and the subscripts p and w represent the pad and the wafer, respectively.

Equations 5.8, 5.9 and 5.10 provide an explicit relationship between the contact area (A_r) and the down pressure (P) that is needed in modeling the CMP process. From Eqs. 5.8 and 5.9,

$$\frac{A_r}{L} = \frac{3\pi\beta^{1/2}}{4E^*} \frac{\int_h^\infty (z-h)\phi(z)dz}{\int_h^\infty (z-h)^{3/2}\phi(z)dz} \tag{5.12}$$

Following Yu *et al.* [Yu94], the normal distribution function may be used for the distribution of the asperity height, ϕ. When $\phi(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$ is substituted into Equation 5.12, the integral cannot be obtained in an explicit form. Instead, a numerical integration should be preformed. Johnson showed that the ratio of the two integrals in Eq. 5.12 is nearly a constant for a broad range of h/σ [Joh85]. Thus,

$$A_r = C^{-1}\left(\frac{\beta}{\sigma}\right)^{1/2} \frac{L}{E^*} = C^{-1}\left(\frac{\beta}{\sigma}\right)^{1/2} \frac{PA_o}{E^*} \tag{5.13}$$

Here C is a constant whose value is between 0.3 and 0.4 when h/σ is in the range of 0.5 and 3.0. Recalling that σ is the standard deviation of the asperity height distribution, h/σ=3.0 represents a situation where only 0.13% of the pad asperities are in contact with the wafer whereas about 30.85% is in contact when h/σ=0.5 (Figure 1b). Practical situation in a typical CMP process is well within this range of h/σ value. From Eqs. 5.10 and 5.13,

$$P_r = C\left(\frac{\sigma}{\beta}\right)^{1/2} E^* \tag{5.14}$$

It should be noted that the contact area (A_r) increases linearly with the down-pressure whereas the contact pressure (P_r) remains constant due to the linear increase of A_r with P.

The underlying assumption for Eq. 5.10 is that the applied normal force on the wafer is supported entirely by the pad asperities. As it is mentioned previously, experimental measurement on the friction coefficient indicates that the hydrodynamic pressure is negligible under typical CMP operational conditions. Furthermore, commercial pads often have grooves or perforations to prevent hydroplaning (i.e., to prevent a large pressure development due to a lubrication flow) and to ensure uniform slurry distribution. Thus, the assumption that the hydrodynamic pressure is negligible is justifiable.

Small Plastic Deformation of the Thin Layer on the Wafer Surface

Considering the fact that the deformation of the surface layer by the slurry particles under the contact pressure is very small in the order of several nanometers, the mode of deformation can be assumed to be plastic deformation. In case of oxide CMP, it

is known that the silica surface is softened when it is in contact with an aqueous solution. A surface layer (or so-called gel layer) is formed due to the diffusion of water molecules into the silica surface altering the surface properties [Tad68].

The situation may be different in case of metal CMP. Since an oxidizer is added in most metal CMP processes, metal oxide is formed on the wafer surface creating a structure mismatch at the interface between the oxide layer and the bulk wafer underneath. Thus, delamination or microcracking of the oxide layer may occur at the interface during the metal CMP process. However, as evidenced in previous sections, it is likely that plastic deformation is still the dominant mode of deformation because the indentation depth by the slurry particles is small (e.g., less than several nanometers). Furthermore, the oxide layer may be soft due to hydration in an aqueous environment. Thus, the assumption that the chemically altered thin surface layer on the wafer exhibits plastic deformation is plausible. This assumption enables us to apply the abrasive wear theory in deriving the expression for the material removal rate. When the indentation depth is beyond the surface layer, the deformation of the unaltered wafer material may be either plastic or elastic depending on the hardness of the wafer material and the slurry particles.

Fully Dispersed Spherical Particles with a Uniform Size Distribution

The slurry particles used for CMP are not of uniform size although the standard deviation of the size distribution is kept to be as small as possible to ensure a consistent polishing rate [Bie98]. While uniform particle size is assumed in the present study for simplicity, inclusion of particle size distribution in the analysis is straightforward although complicated algebraically. In addition, it is assumed that the spherical particles in the CMP slurry are fully dispersed.

Material Removal by Single Particle

In the present analysis, the overall material removal rate is assumed to be the multiple of the removal rate by a single particle and the number of particles in contact with the wafer:

$$(RR) = n\,(RR)_v/A_o \tag{5.15}$$

Here n is the number of slurry particles anchored in the region of the pad that is in contact with the wafer, and $(RR)_v$ is the volumetric removal rate by a single particle. The overall removal rate, (RR), is typically described in terms of the thickness removed per unit time; hence the division by the nominal wafer area A_o.

Schematic illustration of single particle abrasion is given in Figure 5.5. The material removal is viewed as a sliding-indentation process in which the volumetric removal rate by a particle is given as

$$(RR)_v = A_i V \tag{5.16}$$

Here A_i is the cross-sectional area of the particle immersed in the wafer under pressure (hashed area in Figure 5.5a), and V is the relative velocity of the pad to the wafer. Assuming that the hard particle maintains its spherical shape,

$$A_i = -a(R-\omega) + R^2 \arcsin\frac{a}{R} \tag{5.17}$$

Here ω is the indentation depth and a is the radius of the circular indentation on the wafer surface by the particle of radius R. Because ω is very small compared to R, a/R is also very small and Equation 5.17 can be linearized about a/R to give

$$A_i = \frac{4}{3}a\omega\left(1 + O\left(\frac{\omega}{R}\right)\right) \tag{5.18}$$

Thus the volumetric removal rate is

$$(RR)_v = \frac{4}{3} a\omega V \tag{5.19}$$

Because $a^2=(2R\omega-\omega^2)$ and ω is much smaller than R, Equation 5.19 can be rewritten as

$$(RR)_v = \frac{2a^3 V}{3R} \tag{5.20}$$

Once the indentation depth ω (or a) is known, the volumetric removal rate by a single particle can be readily calculated from Eq. 5.20.

Depending on the relative magnitude of the indentation depth (ω) to the surface layer thickness (t), the material removal may be classified into two different regimes as depicted in Figure 5.5; (a) $\omega \le t$ and (b) $\omega > t$.

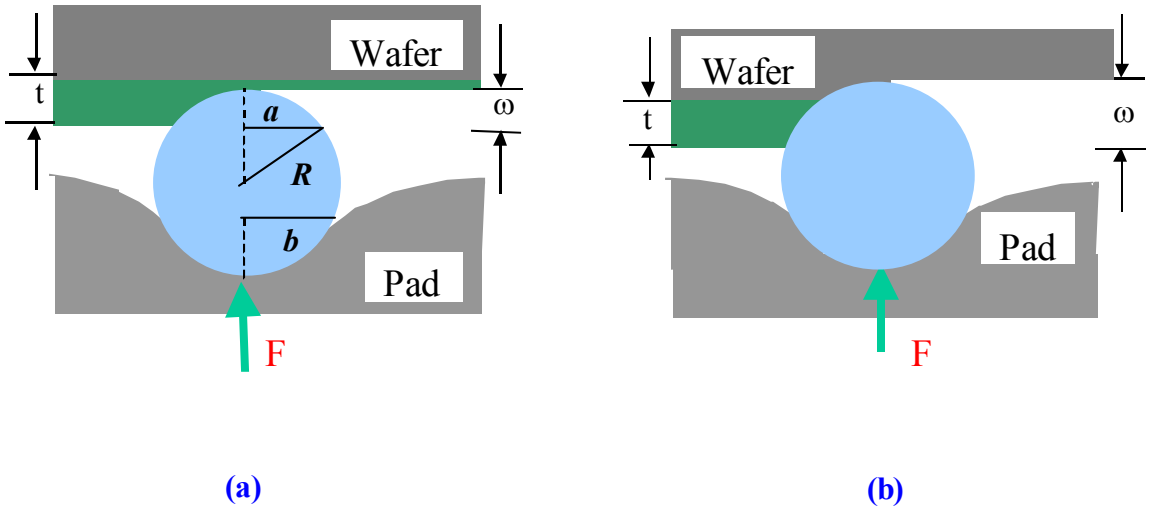


Figure 5.5 Schematic of single particle abrasion showing the removal of chemically modified surface layer, (a). $\omega \le t$; (b). $\omega > t$ with soft wafer materials.

When the indentation depth is smaller than the surface layer thickness

When the chemically modified surface layer is thick enough, the indentation depth ω will be no greater than the surface layer thickness and only plastic deformation

within the surface layer occurs. In this case, ω (or *a*) can be obtained from the force balance imposed on the particle by the pad and the wafer that depends on the indentation depth of the particle into the wafer surface and into the pad. The force imposed on a particle by the pad is given as

$$F_a = \pi b^2 P_r \tag{5.21}$$

Here *b* is the radius of the circular indentation into the pad as described in Figure 5.5a, and P_r is the contact pressure given by Eq. 5.14. The force on the particle by the deforming wafer can be given as

$$F_a = \pi a^2 B_e \tag{5.22}$$

Here B_e is the effective Brinell hardness of the chemically modified surface layer. The term "effective" is used because its value is affected by the thickness of the chemically modified surface layer and the material property of the underlying substrate. The pad is supposed to hold the particles in place thereby transmitting the normal force to the particles that cause indentation of the wafer surface. The pad should be soft enough to provide stable anchoring of the particles on its surface. However, if it is too soft, the particles will be embedded too deeply into the pad diminishing the material removal capability of the particles. Thus, the pad should be sufficiently hard to make the value of *b* in Figure 5.5 to be close to the particle radius *R*. Then, from equations 5.21 and 5.22 with *b*=*R*,

$$a = R\left(\frac{P_r}{B_e}\right)^{1/2} \tag{5.23}$$

This equation along with equations 5.14 and 5.20 gives the following expression for the volumetric removal rate by a single particle of radius *R*.

$$(RR)_v = \frac{2}{3} C^{3/2} R^2 \left(\frac{\sigma}{\beta}\right)^{3/4} \left(\frac{E^*}{B_e}\right)^{3/2} V \tag{5.24}$$

When the indentation depth is greater than the surface layer thickness

When the bulk wafer material beyond the surface layer is relatively soft (e.g., copper) compared to the hardness of the particles, it is possible for the particle to cause deformation into the wafer beyond the surface layer (Fig. 5.5b). As it was described in previous section, the deformation is expected to be plastic for both the surface layer and the bulk wafer, and the force causing the deformation may be given as

$$F_a = \pi\left[2tRB_e + (a^2 - 2tR)B_w\right] \tag{5.25}$$

Here B_w and B_e are the Brinell hardness of the bulk wafer material and the chemically modified surface layer, respectively. Because this force is equivalent to the one imposed on the particle by the down-pressure, Eq. 5.25 is equated with Eq. 5.21 with b=R to give

$$a = R\left[\frac{R^2 P_r + 2(B_w - B_e)Rt}{B_w}\right]^{1/2} \tag{5.26}$$

Then, Eq. 5.20 for the volumetric removal rate by a single particle is reduced to

$$(RR)_v = \frac{2}{3}\left[C\left(\frac{\sigma}{\beta}\right)^{1/2} \frac{E^*}{B_w} R^2 + 2\left(1 - \frac{B_e}{B_w}\right)Rt\right]^{3/2} \frac{V}{R} \tag{5.27}$$

While equation 5.27 describes a situation where the deformation reaches into the bulk wafer beyond the surface layer, deformation of the bulk wafer may not occur if the bulk wafer material is very hard. This limiting case corresponds to the situation of infinitely large B_w in equation 5.27 that is reduced to

$$(RR)_v = \frac{4\sqrt{2}}{3} R^{1/2} t^{3/2} V \tag{5.28}$$

Equations 5.24, 5.27 and 5.28 indicate that the volumetric removal rate by a single slurry particle depends on the geometry of the pad and the particle (i.e., β, σ and R), the material properties (E^*, B_w and B_e), the thickness of the surface layer (t), and the relative velocity of the pad (thus the particle) to the wafer (V). It may be interesting to note that $(RR)_v$ is independent of the down pressure, P. It is due to the fact that the contact pressure, P_r, is independent of P. As equation 5.13 indicates, the contact area of the pad with the wafer (A_r) increases linearly with P. Consequently, the mean contact pressure, P_r, which is PA_o/A_r, is independent of P. In the following section, the number of particles in contact with the wafer surface is estimated to determine the overall removal rate.

Estimation of the Overall Removal Rate

When the wafer is pressed against the pad and is in sliding motion relative to the pad, the particles can be entrapped only in the region where the local distance between the wafer and the pad asperity is smaller than the particle diameter, $2R$. Thus, the number of particles that can be trapped on the pad asperities is

$$n = 2R \cdot A_r \cdot \rho_n \tag{5.29}$$

Here A_r is the real contact area between the wafer surface and the pad asperities as it was described in previous section. ρ_n is the number density of the particles in the slurry (i.e., number of particles per unit slurry volume), that can be calculated from the solid loading C_a (wt%) of the slurry particles as

$$\rho_n = \frac{3}{4}\frac{C_a \rho_s}{\pi R^3 \rho_a} \tag{5.30}$$

where ρ_a and ρ_s are the densities of the particle and the slurry solution, respectively. Substitute of Eqs. 5.30 into 5.29 gives

$$n = \frac{3}{2}\frac{C_a \rho_s}{\pi R^2 \rho_a} A_r \tag{5.31}$$

By combining Eqs. 5.31 and 5.10, the number of particles anchored on the pad asperities becomes

$$n = \frac{3}{2C}\frac{C_a \rho_s}{\pi R^2 \rho_a}\left(\frac{\beta}{\sigma}\right)^{1/2}\frac{PA_o}{E^*} \tag{5.32}$$

Now that $(RR)_v$ and n are given in terms of various processing variables and material properties, the overall material removal rate in a CMP process can be estimated from Eqs. 5.24, 5.27 and 5.32 for the two different cases described above:

(a). *When the indentation depth is smaller than the surface layer thickness (i.e.,* $\omega \le t$*),*

$$(RR) = \left[\frac{\sqrt{C}}{\pi}\frac{C_a \rho_s}{\rho_a}\left(\frac{\sigma}{\beta}\right)^{1/4}(E^*)^{1/2} B_e^{-3/2}\right] PV \tag{5.33}$$

(b) *When the indentation depth is larger than the surface layer thickness (i.e.,* $\omega > t$*),*

$$(RR) = \left[\frac{1}{\pi C}\frac{C_a \rho_s}{\rho_a}\left(\frac{\sigma}{\beta}\right)^{-1/2}\frac{1}{E^*}\left\{C\left(\frac{\sigma}{\beta}\right)^{1/2}\frac{E^*}{B_w} + 2\left(1 - \frac{B_e}{B_w}\right)\frac{t}{R}\right\}^{3/2}\right] PV \tag{5.34}$$

If the bulk wafer material is very hard, equation 5.34 is simplified as

$$(RR) = \left[\frac{2\sqrt{2}}{\pi C}\frac{C_a \rho_s}{\rho_a}\left(\frac{\sigma}{\beta}\right)^{-1/2}\left(\frac{t}{R}\right)^{3/2}\frac{1}{E^*}\right] PV \tag{5.35}$$

As Eq. 5.15 indicates, the overall removal rate, (RR), is the thickness removal rate which is the volumetric removal rate divided by the nominal wafer area A_o. We may recall that C is a constant whose value is about 0.35 when h/σ is between 0.5 and 3.0 [Joh85].

The overall removal rate given above is reminiscent of the Preston's equation in that the material removal rate is proportional to the down-pressure, P, and the relative velocity, V. It should be noted that the Preston's coefficient K_p, whose value relies on empiricism, is now given as a function of various processing variables and material properties. Thus, unlike the Preston's equation, the results of the present analysis may serve as a predictive model for the material removal in a CMP process.

It should be pointed out that the linear dependence of the overall removal rate on the down-pressure P is not due to the increase in the indentation depth but due to the increase in the number of particles in contact with the wafer surface. As the Eq. 5.13 indicates, the real contact area A_r increases linearly with the down-pressure P thus increasing the number of particles in contact. The contact pressure P_r, on the other hand, does not increase with P (Eq. 5.10). Thus, the indentation depth that will increase with P_r does not change with P. Unlike the present analysis, many of the previous models are based on the assumption that the indentation-depth by the particles increases with the down-pressure [Bro81, Tse97, Luo01]. This assumption may be applicable only when the polishing pad is so hard that the contact area of the pad with the wafer is rarely influenced by the down-pressure. The pads used in a typical CMP process, however, are relatively soft with rough surface rather than hard. Furthermore, recent in-situ measurement of the friction force has shown that the friction coefficient during CMP doesn't change with applied load although the friction force linearly increases with the solid loading. This observation is consistent with the current prediction because if the number of particles in contact with the wafer increases linearly with the down-pressure,

the tangential force for the pad rotation will increase linearly whereas the friction coefficient that is the ratio of the tangential force to the down force will remain constant.

The overall removal rate predicted by Eqs. 5.33-5.35 depends on various parameters that represent the geometry of the pad and the particle (i.e., β, σ and R), material properties (E^*, B_w and B_e), thickness of the surface layer (t), and the relative velocity of the pad (thus the particle) to the wafer (V). Among these, the influence of the pad modulus, the size of the slurry particles and the hardness of the surface layer may need further discussion. While Eq. 5.34 may be considered as the most general form for the overall removal rate since it is inclusive of Eqs. 5.33 and 5.35 as a special case of $B_w = B_e$ and $B_w = \infty$ respectively, it may be easier to focus on Eqs. 5.33 and 5.35 in trying to understand the influence of various parameters.

In Eq. 5.11, the modulus of the wafer is typically several orders of magnitude larger than that of the pad (i.e., $E_w \gg E_p$). Thus, the composite modulus (E^*) is essentially represented by the pad modulus. Eq. 5.33 indicates that the overall removal rate increases with the composite modulus (hence with the pad modulus) whereas Eq. 5.35 indicates the opposite. If the pad modulus is larger, the contact area between the pad and the wafer becomes smaller thus resulting in a smaller number of particles in contact with the wafer. However, in case of $\omega \le t$ (i.e., when the indentation depth is smaller than the surface layer thickness), the indentation depth increases due to the increase in the contact pressure P_r and its dependence on E^* is stronger. Thus, the overall removal rate increases with E^*. In case of Eq. 5.35 (i.e., $\omega = t$ due to hard wafer material), on the other hand, the indentation depth is determined not by the contact pressure but by the thickness of the

surface layer. Consequently, larger modulus of the polishing pad results in a smaller removal rate due to the decrease in the number of particles in contact.

Equation 5.33 indicates that the overall removal rate is independent of the particle size R whereas Eq. 5.35 indicates a decrease in *(RR)* with increasing particle size. When the indentation-depth is smaller than the surface layer thickness (i.e., $\omega \leq t$), the volumetric removal rate by a single particle $(RR)_v$ is proportional to R^2 (Eq. 5.24). The number of particles in contact, on the other hand, is inversely proportional to R^2 as the Eq. 5.32 indicates. Consequently, the particle size effect is negated. When the indentation depth is equal to the surface layer thickness t for the hard wafer material (i.e., Eq. 5.35), $(RR)_v$ is not proportional to R^2 anymore and the overall removal rate decreases with increasing particle size.

The effects of the hardness of the wafer and the surface layer (i.e., B_w and B_e) indicated in the Eqs. 5.33 and 5.34 are also readily explicable. If the hardness of the surface layer is large (i.e., large B_e), the indentation depth will be smaller for a given contact pressure P_r. Thus the overall removal rate will be smaller. If the bulk wafer material is soft, the influence of various parameters is rather complicated to understand because removal of not only the surface layer but also the bulk wafer material below the surface layer is involved. With decreasing B_w, removal of the wafer material becomes more significant. The first term in the curly bracket of equation 5.34 represents the removal of the bulk wafer material which increases with decreasing B_w because the particle penetrates deeper into the wafer. The negative part of the second term in the curly bracket, which is also associated with the removal of the bulk wafer, is smaller than the first term because t/R is small.

The present analysis does not provide any guideline regarding which equation among 5.33-5.35 is appropriate to use in comparing the predictions with experimental observations. Nevertheless, it may be pointed out that equation 5.33 should be appropriate when the growth rate of the surface layer is higher than its removal rate, which results in the indentation depth (ω) being smaller than the film thickness (t). In this case, the contribution of the chemical effect to the material removal rate is only through the hardness of the surface layer. On the other hand, when the surface film growth rate is approximately equal to the removal rate, equation 5.35 may be considered. It is appropriate to use this equation when the bulk wafer material is very hard, in which case the particle is unlikely to penetrate into the bulk wafer. Finally, equation 5.34 can be used for wafer materials that are orders of magnitude softer than the abrasives, e.g., the polishing of a copper wafer using alumina particles. It should be noted that transitions between these polishing modes can also be predicted with the proposed model. In fact, the polishing mode transition has been seen in CMP experiments of both oxide [Oum99] and metal wafers [Bie98, Luo97]. As it was pointed out earlier, the chemical action of the slurry is responsible for modifying the wafer surface which is then immediately removed by the mechanical action of the polishing process. Figure 5.6 shows three possible transitions between the predicted polishing modes. The actual polishing regime is essentially dependent on the relative magnitude between surface layer formation rate and the rate of mechanical polishing. The layer formation is affected by the chemicals added in the slurry (e.g., concentration, pH, and ionic strength). The mechanical action is controlled by the particulate properties in the slurry (e.g., solids loading, particle size, and particle shape) and the operational conditions of the CMP (e.g., pressure and velocity).

Based on the comparison between the chemical action and the mechanical action, three scenarios are shown in Fig. 5.6. Curve *a* represents the case when the formation of the surface layer is much faster than the removal of the layer (i.e., $\omega<t$) under the whole range of CMP conditions. As indicated in Eq. 5.33, the material removal rate in this case will be determined by the hardness of the surface layer. The mechanical action dominates the polishing rate. Curve *b* predicts a polishing scenario that a transition may occur when the mechanical action of the CMP process increases to match the chemical action of surface layer formation (i.e., $\omega=t$). The slope shown in curve *b* decreases upon the transition point due to the fact that the material removal rate is controlled by at least two means, namely the increased indentation depth and the increased number of abrasive particle, before the transition; and it is mainly affected by the increased number of abrasive particle after the transition. Curve *c* shows two transition points. Besides the one that is described for curve *b*, it is possible that the mechanical action is bigger than the surface layer formation rate (i.e., $\omega>t$), which results in constant or lower removal rate because the polishing process is essentially controlled by the chemical action only. Curve *c* represents a most complete case that may happen during CMP. However, it is worth to mention that both the segment before the first transition and the segment after the second transition in curve *c* are not beneficial cases in CMP. The former is mechanical abrasion limited and may lead to residual surface layer formation and in turn affect the dielectric constant or the conductivity of the wafer material. Similarly, the later is controlled by the chemical action only. It may result in scratches or defective pits on the wafer surface and damage the IC products. Thus, the polishing process that is shown as the middle segment of curve *c* seems to be the ideal scenario for CMP. It indicates that a careful balance

between the chemical and mechanical components of CMP is needed to provide optimal process performance. In the following section, an attempt is made to compare the predictions of the present model with experimental observations.

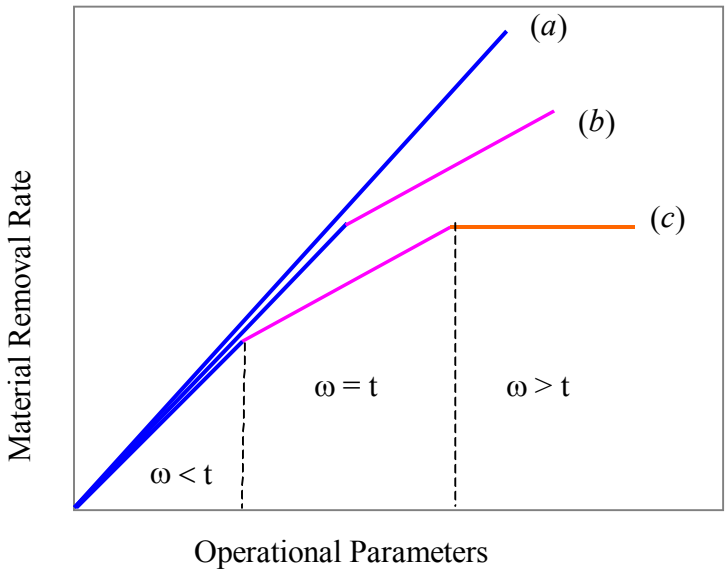


Figure 5.6 Three possible transitions between the polishing modes, (a) no apparent transition when the surface formation rate is always faster than the mechanical polishing rate, (b) transition when the surface formation rate from higher to lower than the mechanical polishing rate, (c) transitions when the surface formation rate from higher to equal to then to lower than the mechanical polishing rate.

Comparison of Model Predictions with Experimental Observations

Pad and Wafer Contact

Before the examination of the model predictions, the validity of Greenwood-Williamson model for the pad and wafer contact has to be verified. A typical profile for the polishing pad surface is provided in Figure 5.6. The profile was obtained using a PDI (Precision Devices, Inc.) surfometer (series 400) with the stylus force of 16mN. It has been shown that the effect of stylus force on the surface topography measured by the

surfometer is negligible [Li95]. The measurement shows that pad asperity heights are in the order of tens of microns and appeared to have a normal (or Gaussian) distribution.

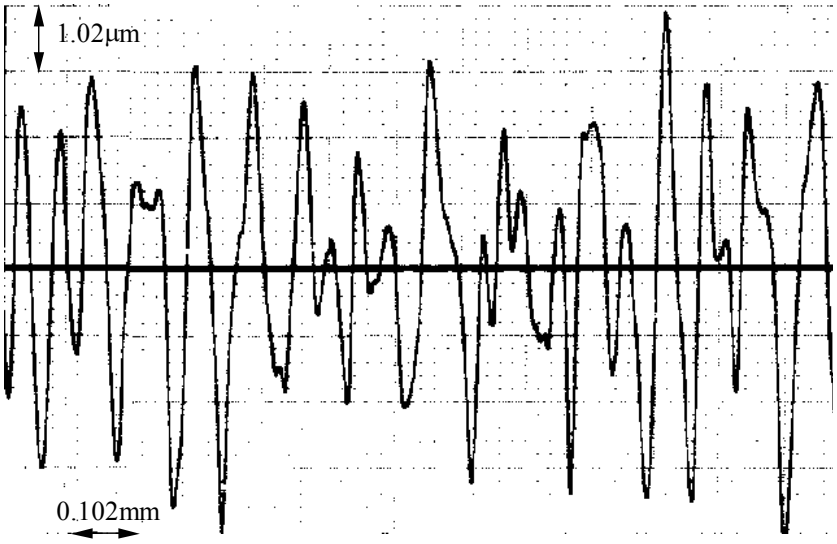


Figure 5.6 The surface profile of a Rodel IC1000 polishing pad

Fig. 5.6 demonstrates that only the higher asperities of the pad surface will be in contact with wafer when it is pressed again the pad, suggesting that the real contact area is small. In current study, Fourier Transform Infrared Spectroscopy/Attenuated Total Internal Reflection Spectroscopy (FT-IR/ATR) technique was selected to measure the contact area between the pad and a flat ATR crystal as a function of applied load. Similar experiment has been done by Basim [Bas02]. The experimental details can be found from [Bas02]. The experimental set up is shown in Figure 5.7 schematically.

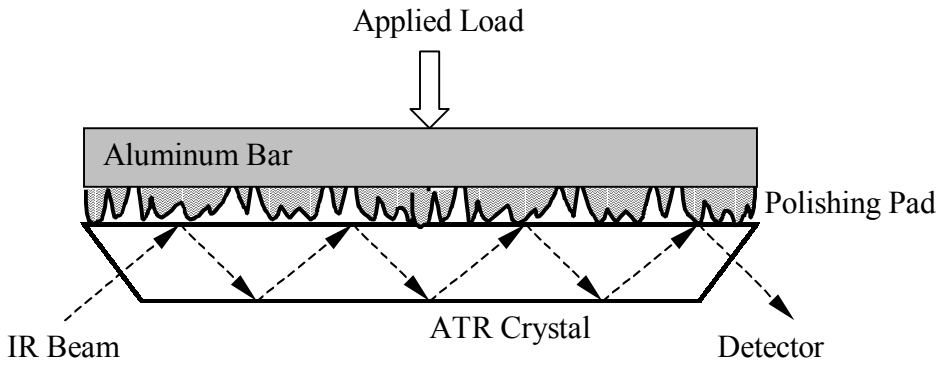


Figure 5.7 Schematic diagram of the pad-ATR crystal contact under applied load

FTIR/ATR data collection was conducted on a Nicolet Magna-IR 760 spectrometer equipped with a N_2 purge gas generator and a MCT detector. ZnSe horizontal crystals (45^o angle of incidence) were used in this study. A number of pieces of the polishing pad were first cut from a 12″ IC1000 pad to match the size of the ZnSe crystal surface. The pad was then attached a thick aluminum bar at the bottom such that uniformly distributed forces can be achieved when it was pressed. During measurement, the pad was brought upside down to the ZnSe crystal surface, and the desired load was directly applied on top of the aluminum bar. Five pieces of pad samples were examined. The crystal surface was cleaned with acetone, methanol, and DI water sequentially between experiments to remove any possible residues from previous sample. The ATR spectra of the samples were collected at a resolution of $4cm^{-1}$ immediately after the load was applied. The weights used in present study were 0g, 100g, 200g, 500g, 700g, and 1000g. The corresponding nominal pressure values were calculated by dividing the load applied (including the weight of the aluminum bar) with the sample surface area. For each sample, the loads were first applied in ascending order. Then the loading order was reversed and the corresponding spectra were also collected. A good match was found on

the spectra for the same load in either loading order, indicating that the pad deforms elastically during the experiments. Both the amide II (C-N stretch and N-H bend at about 1540cm^{-1}) and the amide III (C-N stretch of R-O-CO-NH-R at about 1225cm^{-1}) bands were detected in the pad spectra. The intensities of both bands have been used to characterize the amount of materials involved [Win82, Jac83, Jeo92] and good correlation was obtained between the band intensity and the quantity of the chemicals adsorbed on the ATR crystal. In present study, the peak area of the amide III band was used to calculate the total amount of the pad material that is in contact with the crystal surface. Typical ATR/IR spectra of the polyurethane pad under different applied loads is shown in Figure 5.8. To obtain the percent pad coverage, a film was made from the IC1000 pad by mechanically pressing a piece of pad and putting it in vacuum oven at 120^{o}C for three hours. At that temperature, no degradation of the material was found while the pad was readily compressed and the air in the foam was taken out by vacuum. The film was then cut to desired area, and same analyses were conducted at different applied loads until 100% contact was achieved. Finally, the absorbance of the spectrum at 100% contact was calculated and compared with the absorbance of the pad samples at different loads to obtain the percentage contact value.

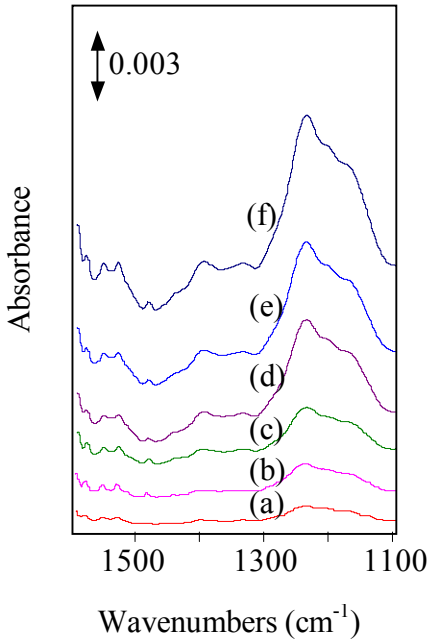


Figure 5.8 FTIR/ATR spectra of the pad sample under different applied pressure: (a) 0.57psi, (b) 1.47psi, (c) 2.38psi, (d) 5.09psi, (e) 6.90psi, (f) 9.61psi.

Compared to the study by Basim, in which the CH_2 peaks were used for the characterization of the contact [Bas02], the current study of measuring CN stretch peaks is more reliable because the CH_2 peaks has the same location as the CO_2 peak, which changes concentration due to the environment difference each time when the samples are introduced. Moreover, instead of using a separately purchased polyurethane film as it was used in other studies (e.g., [Bas02]), which usually has different compositions from the polishing pad materials, a film directly made from the polishing pad was able to be obtained and the percentage contact data calculated from this measurement are expected to be more practical.

It should be noted that the measure IR absorbance value is assumed to be linearly proportional to the contact area between the pad and the ATR crystal. It is justifiable only when the real contact area is small compared to the nominal area of the flat crystal

surface. Otherwise, an exact relationship between the absorbance and the contact area has to be obtained. The calculated percent pad coverage as a function of applied pressure is shown in Figure 5.9.

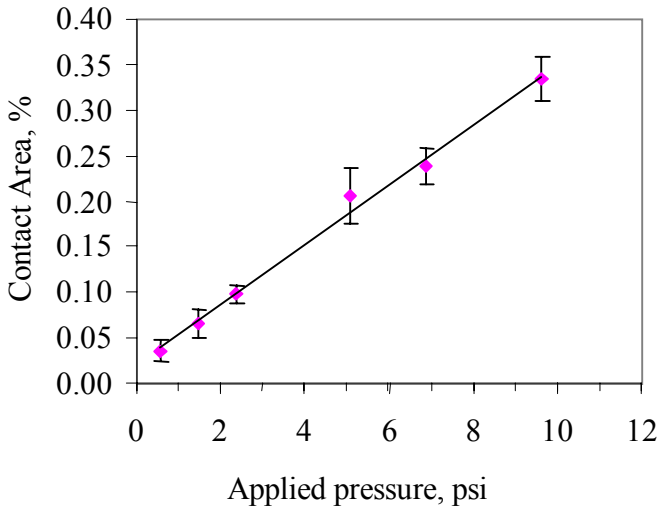


Figure 5.9 The percentage of pad contact area as a function of applied pressure

The ratio of the real contact area to the nominal pad area was calculated to vary from 0 to 0.35% in the selected pressure range, and it is linearly increasing with the applied load. Recalling Eq. 5.13, a good agreement can be concluded between the experimental observation and the application of Greenwood-Williamson model to the pad and wafer contact. Another finding is that the elastic modulus of the pad asperities can be obtained from the experimental contact ratio using Eq. 5.13. For example, at the typical CMP operational pressure of 7 psi, the pad-wafer contact area was determined to be 0.25%, and the elastic modulus value of the IC-1000 polishing pad is predicted to be 40-50 MPa, which is in agreement with the values reported in literature for the same pad material [Tic99]. Moreover, the experimentally determined percent contact area for IC-1000 pad (0.25%) is also in between the values reported for the hard and the soft pads (0.05-0.54

%) [Yu94]. These observations suggested that the adoption of Greenwood-Williamson model for pad-wafer contact is plausible, and only a small portion of the pad was in contact with the wafer during polishing.

Material Removal Rate

For the comparison between the predictions of the present model and experimental material removal data, the values of various parameters in the Eqs. 5.33-5.35 should be specified. Approximate values of most parameters are found in literature except the surface layer properties, B_e and t as summarized in Table 5.1. Although the values of B_e and t are not available, an attempt is made here to estimate the values of B_e and t using the results of a CMP experiment by Ouma [Oum99] and compare the estimated values with an information found in another reference.

Table 5.1 Typical values of the parameters representing the material properties and the surface topology.

Variable	Description	Value	Unit	Reference
E_p	Young's modulus of pad	45.3	MPa	
σ	Standard deviation of the pad asperity height distribution	25	μm	[Yu94]
β	Average radius of curvature of the tip of the pad asperities	30	μm	[Yu94]
ν_p	Poisson's ratio of pad material	0.5		[Ste97]
E_w	Young's modulus of silicon dioxide	66	GPa	[Elk95]
ν_w	Poisson's ratio of silicon dioxide	0.17		[Elk95]
ρ_a	Density of silica particles	2.27	g/cm^3	[Elk95]

In Figure 5.10, experimental results of Ouma for the removal rate of TEOS wafer by a silica slurry are shown as a function of the product of P and V. The removal rate appears to show two different modes in that the slope changes distinctively near $PV \approx 50$.

Although the author provided two fitting lines for these data using the Preston's equation, no further discussion was given. As it was pointed out in the previous section, Eq. 5.33 may be appropriate to use for the range of small PV whereas Eq. 5.35 may be applied for larger values of PV considering that TEOS wafer is hard. The linear regression of the data for PV <50 indicates the slope to be 25 (Å/psi·m), and Eq. 5.33 along with the values of other parameters given in Table 5.1 results in $B_e = 3.7\times10^5$ (psi). The data for $PV > 50$, on the other hand, indicates a slope of 10 (Å/psi·m) and Eq. 5.33 gives the film thickness (t) to be 4.7Å. In Ouma [Oum99], the values of C_a, R, and ρ_s were not specified. Thus the most typical values for these variables were used in estimating B_e and t, which were 5wt%, 100nm, and 1.0, respectively. Increasing the solid loading to 10 wt% results in $B_e = 5.8\times10^5$ (psi) and t = 3.0Å, whereas t becomes 9.7Å *if* R = 200nm. Thus, the rather arbitrary choice of these variables does not influence the conclusions described here.

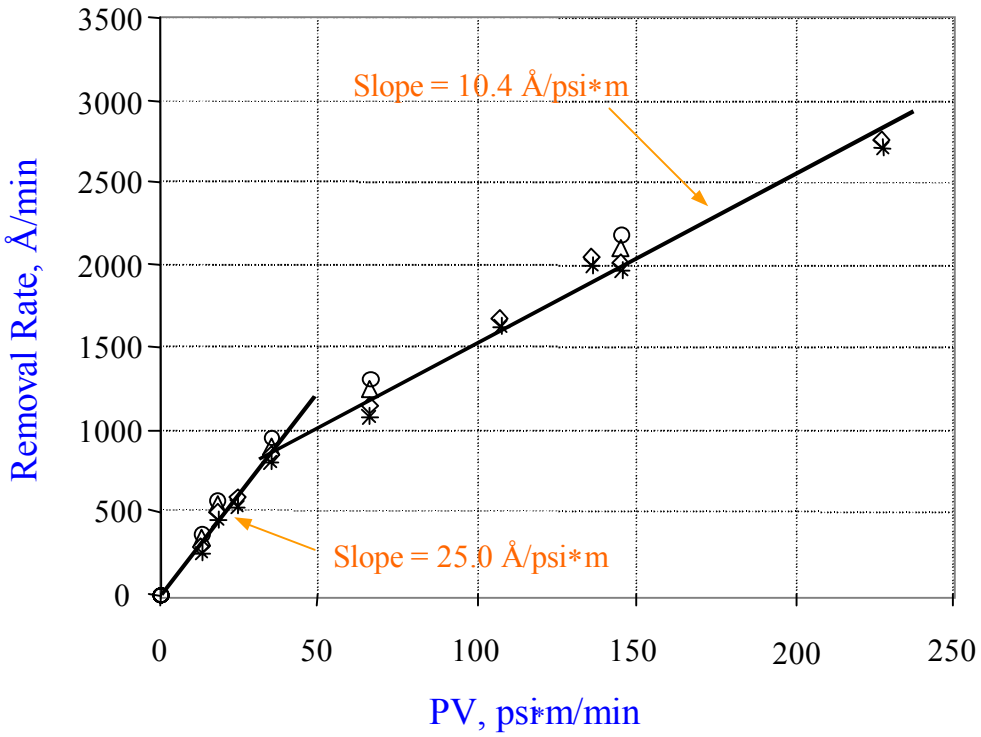


Figure 5.10 Chemical mechanical polishing of TEOS using a silica slurry (Data from Ouma [Oum99]).

Direct measurement of B_e and t under a dynamic condition may be very difficult. However, the order of magnitude of these variables may be estimated. Bartholemew [Bar83] incorporated significant amounts of water into silicate glasses and found that the incorporation of 7wt% water resulted in a reduction of Young's modulus by about 10% from 9.9×10^6 to 8.2×10^6 psi. Since the ratio between the modulus (E) and the hardness (B) for non-metallic materials is known to be approximately 20 [Sre01], the corresponding hardness for $E = 8.2\times10^6$ psi will be approximately 4.1×10^5 psi, which appears to be in the same order of magnitude of the estimated value of 3.7×10^5 (psi) from the data of Ouma. This hardness value may seem to be too large for a gel layer because gel layer typically refers to a soft layer. However, it should be noted that the surface hardness is affected by the substrate as well as the thickness of the surface layer. El-

Sherbiney and Salem [Els79] found that the effective hardness decreased exponentially with increasing layer thickness. Thus the hardness value of a gel layer can be rather large as long as the layer thickness on a hard substrate is very thin. Considering that the predicted value of the layer thickness is as thin as 4.7Å, it may not be surprising to find out that its effective hardness is in the order of 10^5 (psi). The very small value of the predicted layer thickness suggests that the chemically modified surface layer is probably only a couple of atomic layer thick, and this value is well within the range of typical roughness of the polished oxide wafers.

Another experimental results by Sun *et al.* [Sun94] are given in Figure 5.11 for further comparison. Due to the lack of proper information given in the paper, however, only a qualitative comparison is possible in this case. In Figure 5.7, three sets of data are given for the polishing of (1) thermal oxide, (2) PECVD (plasma enhanced chemical vapor deposition) oxide, and (3) BPSG (borophosphosilicate glass), respectively. Considering that the extrapolation of each data set to PV=0 does not approach to zero removal rate and that the oxide wafers are hard, Eq. 5.35 is again used for comparison. It is apparent that the slope for the BPSG glass is greater than that of the PECVD or the thermal oxide films by a factor of about 1.8. Since the same pad and slurry were used for the experiment, the larger slope for BPSG corresponds to a thicker surface layer by a factor of about 1.5 ($=1.8^{2/3}$). The surface reaction caused by the diffusion of water into silica surface is believed to be diffusion-controlled because of very small diffusion coefficient of water. However, in case of BPSG with 5% phosphorus and 4% boron [Sun94], the diffusive exchange of molecules with the surrounding water molecule proceeds at a higher rate leading to the formation of hydroxyls. In addition, the

phosphorus in BPSG film may act as bridging atoms or interposed atoms to loosen the structure of host SiO_2 network [Liu95]. Thus, the diffusion rate of water is expected to be much higher. Consequently, the surface layer formed on the BPSG film is expected to be thicker and softer than those formed on oxide wafer surface. Thus, the experimental results of Sun *et al.* appear to be consistent with the predictions of the present model.

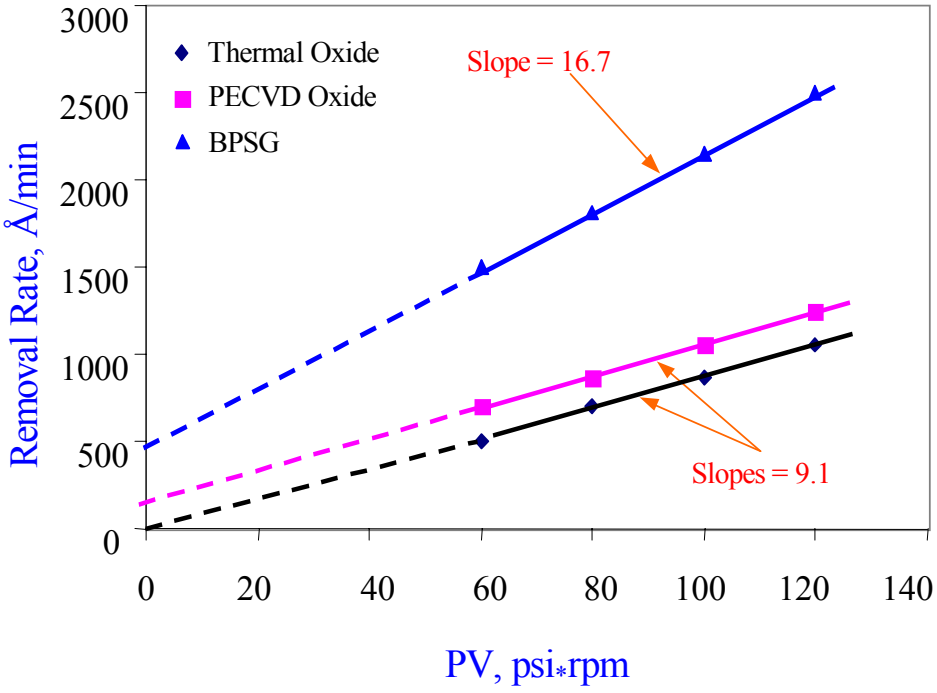


Figure 5.11 Chemical mechanical polishing of three different materials (Data from Sun *et al* [Sun94]).

In oxide CMP, the surface layer is formed due to the hydration of the silica surface. It is reasonable to believe that the thickness of the hydrated layer is a function of solution pH. Thus the material removal rate of oxide polishing will be a result of slurry pH. The next example illustrates the effect of slurry pH on the material removal rate (Figure 5.12). The polishing experiment was performed on thermally grown SiO_2 using an IPEC Avanti 472 polisher [Stei99]. The wafer diameter is 150 mm. The slurry used is

Klebosol 1498-50, which is stabilized with unknown dispersant(s) and has 30 wt% of SiO_2 particles. In Fig. 5.12, the solid dots are experimental results and the continuous lines are model predictions using Eq. 5.35 with the surface layer thickness *t* as the fitting parameter. It is seen that the polish rate is minimized when the slurry pH is close to the natural IEP (isoelectric point) of silica substrate (i.e., pH 3.0). The polish rate increases dramatically with the increase of slurry pH when above the IEP.

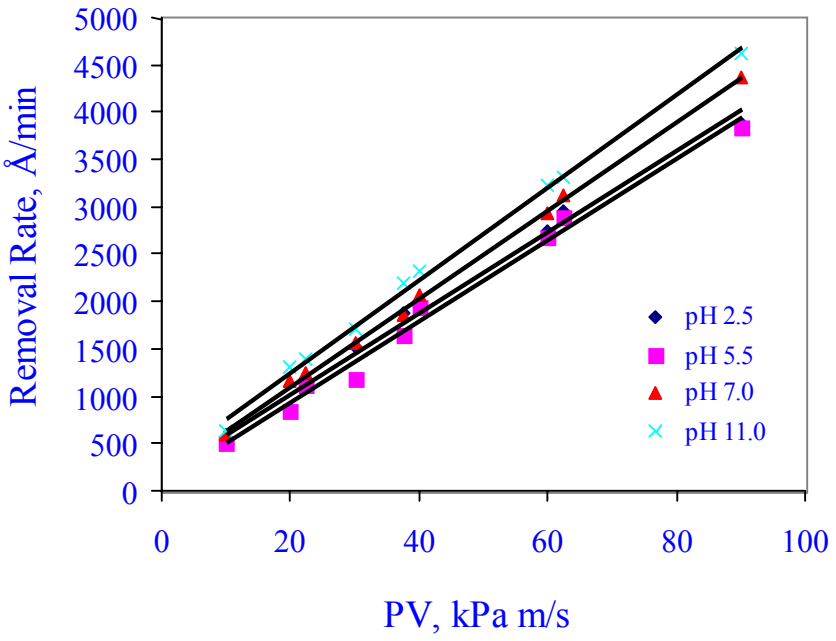


Figure 5.12 Material removal rate of thermally grown SiO_2 CMP as a function of slurry pH. The solid lines are model predictions using surface layer thickness *t* as the fitting parameter. (Experimental data from Stein *et al* [Ste99])

The variation of silica polish rate with the slurry pH can be explained using current model Eq. 5.35. The fitted surface layer thickness data are shown in Table 5.2. The thickness of the hydrated surface layer increases with the increase of slurry pH. This behavior is due to the OH^- ion acting as a catalyst for water attacking on the silica network. The rate of attack is much faster with OH^- ions than it is by water molecules, as

pointed out by Budd, because excess of electrons are created, resulting in higher negative surface potential and consequently more attack by H_3O^+ [Bud61]. The presence of the surface layer not only affects the polishing rate significantly as shown in Fig. 5.12, but also changes the dielectric constant of the silica surface, and hence the surface forces. The water-silica system under different solution pH values has been studied by numerous researchers. A recent measurement of surface forces using AFM by Adler demonstrated that there was a short-range repulsion on silica surface [Ald01]. By evaluating the water-silica system behavior under solution pH of 4, 6, and 9 respectively, he concluded that a gel-layer due to the water diffusion into the silica surface structure was presented. The gel-layer is suggested to be softer than the silica itself. This observation is also in agreement with the polishing results. As mentioned previously, under higher pH (e.g., >9.0), there is a drastic change of the silica surface property and orders of magnitude increase in the dissolution rate of silica [Ile79]. Since the dissolution rate (e.g., 1nm/day) is negligible compared with the polishing rate, the increase in the polishing rate with slurry pH is mainly caused by the surface layer formation and the mechanical abrasion of this surface layer.

Table 5.2 The surface layer thickness as a function of slurry pH. The data are fitted through the polishing data in Figure 5.12.

Slurry pH	2.5	5.5	7.0	11.0
t/R	5.37E-04	5.10E-04	5.46E-04	5.72E-04

In order to examine the validity of the proposed model with metal CMP, a set of experimental results is obtained from Jindal *et al* [Jin01]. While doing CMP on copper with alumina slurries at different pH values, they observed a monotonic increase in the

polishing rate as solids concentration increases from 0.5 wt% to 3.5 wt%. Figure 5.13 shows the correlation between the model predictions and the experimental observations. The experimental data are shown as solid dots and the continuous lines are model predictions. In the prediction, Eq. 5.33 was used because the copper substrate is a relatively soft material compared to the alumina particles used in the slurry. The effective hardness B_e was adopted as the fitting parameter.

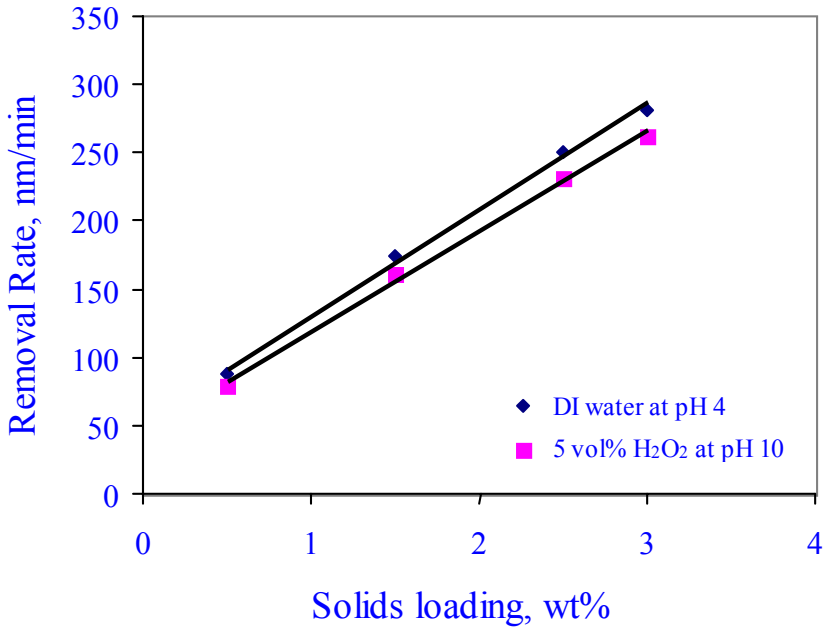


Figure 5.13 Material removal rate of Cu CMP as a function of slurry solids loading. The solid lines are model predictions using surface layer hardness B_e as the fitting parameter. (Experimental data from Jindal *et al* [Jin01])

As shown in Eq. 5.33, the linear increase in Fig. 5.13 as a function of solids loading C_a is due to the fact that the number of particles embedded between the pad and the wafer increases linearly with the solids loading. Additionally, Fig. 5.13 shows that the polishing rate of the slurry at pH 4 is higher than that of the slurry at pH 10 with H_2O_2. In case of copper polishing at pH 4, no surface layer firms on the copper according to

Pourbaix diagram [Pou69]. However, in case of copper polishing at pH 10, a passivation layer, primarily Cu_2O, is expected to form on the copper surface [Pou69]. The polishing rate of Cu_2O is lower than the polishing rate of copper in Fig. 5.13 is attributed to the difference in the surface layer hardness. The hardness increase from copper to Cu_2O is about 25% [Ste95]. However, the corresponding increase in the surface hardness shown in Fig. 5.13 is only approximately 5%. We may consider that the surface layer of Cu_2O is formed in aqueous media, and hydration of the surface film is inevitable. The hydrated Cu_2O is expected to be softer than the Cu_2O that is formed under thermal conditions, and thus a higher polishing rate. One may argue that the etch rate of the copper itself in the slurry of pH 4 has to be taken into account because the dissolution rate of copper is much higher at pH 4 than it is at pH 10 [Luo99]. However, the measured etch rate under acidic conditions during CMP is low [Ste95]. This is because the volume of slurry on the pad is not large enough to support etching of the relatively large surface area. Therefore, the copper polishing is in fact a mechanical abrasion controlled process. This is in agreement with the experimental observations of many researchers [Ste98, Luo99].

Summary

A comprehensive model has been presented for the prediction of the material removal rate in a chemical mechanical polishing (CMP) process. The current model takes into account most variables involved in a typical CMP process. The most important premise of the present model is the existence of a chemically modified thin layer on the wafer surface that is formed by the chemical effects of the CMP slurry and then removed mechanically by the slurry particles. Assuming a sliding-indentation process in which the cross-section of the indentation in the normal direction is a part of a circle caused by non-deforming hard spherical particles, the volumetric removal rate by a single particle is

estimated. The overall removal rate is then determined by multiplying the single particle removal rate to the number of particles in contact with the wafer that is estimated by applying the Greenwood-Williamson model and assuming that the slurry particles are fully dispersed at all times.

The present model developed by the prescribed procedure coincides with the empirical Preston's equation in that the material removal rate is proportional to the down-pressure, P, and the relative velocity, V. However, the Preston's coefficient, which is determined empirically, is now given as a function of various processing variables and material properties. By comparing the model predictions with experimental results, we found that the present model was capable of explaining many experimental observations on both oxide and metal CMP that, otherwise, could not be explained properly. The optimization scheme of the process and the design of a robust polishing process to produce scratch free surfaces can be developed based on the current findings.

One important finding through the comparison between the chemical effect and the mechanical abrasion during CMP showed that the transition between different polishing modes is critical for all types of materials to be polished, indicating the significance in considering the balance the chemical and mechanical effects during robust slurry design. It is noteworthy that the optimal polishing performance (e.g., scratch free surfaces) can be achieved by controlling the balance between both effects. When the growth rate of the chemically modified surface layer is approximately equal to its removal rate (i.e., Eq. 5.35), it is expected that efficient defect free polishing can be achieved. Meanwhile, effective polishing can be achieved by increasing the growth rate of the surface film, and hence higher removal rates.

CHAPTER 6
CORRELATION BETWEEN THE COEFFICIENT OF FRICTION AND THE MATERIAL REMOVAL RATE

Introduction

In previous chapter, a wafer scale material removal rate model was presented. The model started with the basic relation $(RR) = n \times (RR)_a$, where n is the number of particles that actually conduct the abrasion and $(RR)_a$ is the polishing rate of a single particle. The chemical effect during CMP was represented with two variables, the effective hardness and the thickness of the chemically modified layer. It has been demonstrated that the chemically modified layer is very thin (< 1nm) and is abraded off during formation. The common nano-indentation and optical methods (e.g., ellipsometry, FTIR) have been examined on silica wafers to obtained the layer thickness information. However, it is impossible to get reliable measurement data at this stage because the layer to be measured is atomically thin and the resolutions of the instruments are limited. On the other hand, the corresponding effective hardness, as it is discussed in Chapter 5, is close to the bulk hardness value of the material to be polished due to the minimal contribution of the thin surface layer. Thus it can be approximated as the hardness of the bulk wafer, which is a measurable parameter using available techniques like nanoindentation.

Further investigations of the surface layer thickness measurement are beyond the scope of current study. However, it is possible to incorporate the surface layer property to the material removal rate by the measurement of friction coefficients since the chemically modified layer have significant influence on the friction between the abrasive particle and

the wafer surface to be polished. With a soft layer presented between two sliding surfaces, it is well known that lubrication may take over and the friction between the two surfaces is to be reduced. Consequently, the friction due to the interactions among the pad, embedded particles and the wafer surface during CMP can be considered as boundary lubrication [Ham94].

In this chapter, a correlation between the coefficient of friction and the material removal rate is proposed. Instead of the surface layer hardness and thickness, the surface layer properties are characterized with the hardness and shear strength of the layer. The later can be obtained using atomic force microscope (AFM) lateral force measurement. Without losing the advantages of the model developed in the previous chapter, the variables involved in the CMP process such as solids loading, particle size, pad modulus, pad asperity size and distribution, down-pressure, and velocity are integrated in the model as it was done in previous chapter.

Model Development

Recalling the development procedure of the model in Chapter 5, we obtained the number of particles embedded in the pad asperities by considering the pad-wafer contact. It is given as

$$n = \frac{3}{2C}\frac{C_a \rho_s}{\pi R^2 \rho_a}\left(\frac{\beta}{\sigma}\right)^{1/2}\frac{PA_o}{E^*} \tag{5.32}$$

We will still use Eq. 5.32 for the number of particles that actually perform the abrasion of the wafer surface. However, the correlation between the abrasion rate of a single particle, which is related to the indentation depth of the particle into the wafer surface (or the contact radius between the wafer and the particle), and the friction coefficient has to be established. Once the contact radius a (Figure 6.1) is known from the measured

coefficient of friction, the removal rate of a single particle can be calculated from Eq. 5.20.

Particle-Wafer Contact Radius and the Coefficient of Friction

To shear two surfaces by sliding, a friction force [Rab95]

$$F_x = SA = S\pi a^2 \tag{6.1a,b}$$

is required, where S is the shear breaking strength of the contacts, a and A are the contact radius and the contact area between the two surfaces respectively. Assuming Tresca yield surface (Eq. 4.21), the shear strength of the material is related to its yield stress as $S = Y/2$. When a thin film (e.g., oxide layer on metals) is presented between surfaces, it is the surface layer that determines the coefficient of friction [Ash80].

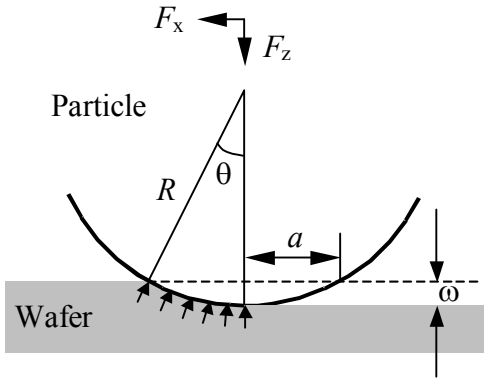


Figure 6.1 Schematic of sliding contact between a spherical particle and a flat wafer

When a spherical particle is indented to a surface and sliding over it, both the normal force F_z and lateral force F_x are supported by the deformation of the substrate. Neglecting the possible pileup around the particle, we can describe the forces in spherical coordinates as

$$F_z = 2\int_0^{\pi/2}\int_0^{\theta} B_e \cos\theta \cdot R^2 \sin\theta d\theta d\varphi \tag{6.2}$$

and

$$F_x = 2\int_0^{\pi/2}\int_0^{\theta}\left(B_e \sin\theta\cos\varphi + S\sin\varphi\right)\cdot R^2 \sin\theta d\theta d\varphi \tag{6.3}$$

respectively. Here ($R^2\sin\theta d\theta d\varphi$) is the infinitesimal area of contact, B_e and S are the hardness and shear strength of the substrate respectively. It is clear that $\sin\theta = \frac{a}{R}$. Thus, equations 6.2 & 6.3 can be readily integrated out, i.e.,

$$F_z = \frac{\pi a^2}{2} B_e \tag{6.4}$$

$$F_x = B_e R^2 \left\{ \sin^{-1}\left(\frac{a}{R}\right) - \left(\frac{a}{R}\right)\left[1-\left(\frac{a}{R}\right)^2\right]^{1/2} \right\} + 2R^2 S\left\{1-\left[1-\left(\frac{a}{R}\right)^2\right]^{1/2}\right\} \tag{6.5}$$

Note that the normal load F_z is half of the indentation force given in Eq. 5.22 because only half of the contact area is supported by the wafer surface during particle cutting or plowing. The coefficient of friction for this sliding process, according to Amonton's law, can be obtained by dividing equation 6.5 with equation 6.4, as

$$\mu = \frac{2}{\pi}\left(\frac{a}{R}\right)^2 \left\{ \sin^{-1}\left(\frac{a}{R}\right) - \left(\frac{a}{R}\right)\left[1-\left(\frac{a}{R}\right)^2\right]^{1/2} + 2\frac{S}{B_e}\left[1-\left(1-\left(\frac{a}{R}\right)^2\right)^{1/2}\right] \right\} \tag{6.6}$$

Although the coefficient of friction is one of the most important parameters in the study of lubrication and wear, and similar development of equation 6.6 can be found from [Kom85], it has not been used previously to model the CMP process. In a CMP process, the coefficient of friction for a spherical particle sliding over the wafer surface is a strong function of the interfacial shear strength S as shown in Eq. 6.6. For example, the shear strength of the surface will be very small in comparison to the bulk shear strength of the wafer in the extreme case when a thick and soft layer is presented. The second term on the left hand side of Eq. 6.6 in the bracket is negligible since $S/B_e \approx 0$. The coefficient of

friction is therefore expected to be minimized (Figure 6.2, curve *a*). On the other hand, if there is no surface layer, the surface shear strength will be the same as the bulk shear strength of the wafer material, and the coefficient of friction is maximized (Figure 6.2, curve *b*). In this case, the ratio of S/B_e is approximately 1/6 because yield stress is roughly 1/3 of the material hardness. However, the practical coefficient of friction as a function of the ratio of contact radius to the radius of the particle (i.e., *a/R*) should lie between the curves (*a*) and (*b*) as shown in Figure 6.2, depending on the ratio of S/B_e. Therefore, it is critical to obtain reliable measurements of the shear strength of the wafer surface, especially when the chemically modified layer is presented.

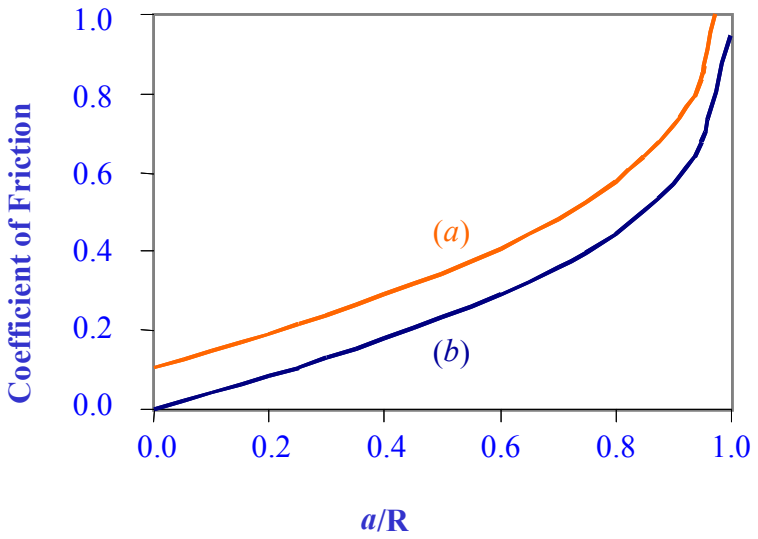


Figure 6.2 Coefficient of friction for a spherical particle sliding over a flat wafer surface as a function of the ratio of contact radius to the radius of the particle, (*a*) without surface layer, (*b*) with a thick and soft surface layer.

With the known ratio of S/B_e, it is shown in Figure 6.2 that the coefficient of friction increases monotonically with the increase of contact radius between the particle and the surface. Thus, there is a unique contact radius value corresponding to a given

coefficient of friction. By fitting the contact radius to the experimentally measured coefficient of friction, the removal rate of each particle can readily be calculated as

$$(RR)_a = \frac{2a^3 V}{3RA_o} \tag{6.7}$$

Here $(RR)_a$ is in the units of thickness per unit time, which is an alternative equation of Eq. 5.20.

Overall Material Removal Rate

As mentioned previously, the overall material removal rate is actually the product of the number of particles embedded in the pad asperities and the removal rate of each particle. It is given by combing Eq. 5.32 with Eq. 6.7, i.e.,

$$(RR) = \frac{1}{C}\frac{C_a \rho_s}{\pi E^* \rho_a}\left(\frac{\beta}{\sigma}\right)^{1/2}\left(\frac{a}{R}\right)^3 PV \tag{6.8}$$

Following the discussions in the previous chapter about the parameters involved in equation 6.8, only the contact radius a needs to be determined. It varies with the formation of surface layer on the wafer to be polished. As shown in Fig 6.2, the variation is possible to be captured through the measuring the coefficient of friction.

Results and Discussion

In this section, the experimental measurements of surface shear strength and wafer hardness shown in Eq. 6.6 are discussed for silica wafers. Then, the predicted material removal rates from the coefficient of friction measurements are compared against existing experimental observations.

Surface Shear Strength and Wafer Hardness

In the expression of equation 6.6, two of the important parameters that need to be determined are the surface shear strength S and the wafer hardness B_e as shown in Eq.

6.6. In current study, the surface shear strength was calculated through the measurement of Cain *et al* [Cai00]. They measured the friction force of a silica particle, which has a size of 10.5 μm, on a silica glass flat using lateral force microscope (LFM). With the measured friction force F_x and the lateral contact stiffness between the particle and the flat surface k_{con}, the shear strength of the contact is given as [Car97]

$$S = \frac{64G^{*2}F_x}{\pi k_{con}} \tag{6.9}$$

where

$$G^* = \left(\frac{2-\nu_1}{G_1} + \frac{2-\nu_2}{G_2} \right)^{-1} \tag{6.10}$$

This is an equation that is similar to Eq. 5.11. Here, G_1 and G_2 are the shear moduli of the particle and flat surface respectively. Following these equations and adopting the measurements of Cain *et al*, we can obtain the shear strength of the contact between the silica particle and the silica wafer is about 0.635 GPa (Table 6.1). Considering that the environment of LFM measurement is not a direct simulation of the CMP condition and that the uncertainties involved in the friction force measurement are far from fully controlled [Cai00], one should note that this calculated shear strength is only an approximation of the surface shear strength that is used in equation 6.6. In the following sections, the application of this surface shear strength value on the prediction of CMP polishing rate shows that the LFM method is possible in giving a reliable measurement of the surface shear strength under CMP conditions although detailed discussions are beyond this study.

The next parameter measured in current study is the surface hardness B_e of the polished wafer. It has been shown in previous chapter that the surface hardness B_e is

close to the bulk hardness of the material since the surface layer is thin. To measure the wafer hardness, a nanoindenter and Berkevich tip provided by Hysitron were used. The wafer sample was supplied by Silicon Quest International, which has 2 μm p-type PECVD silica coating of approximately 0.6nm RMS surface roughness on a 1.8 mm thick silicon wafer. The wafer sample was first dipped in CMP solution for 12 hours in order to simulate the surface layer formation, then was dried using nitrogen gas before measurement. This measurement was also compared with the direct measurement of an as received dry sample. No apparent difference on the measured hardness was detected. By systematically increasing the indentation depth, the hardness was measured as a function of indentation depth. Table 6.1 shows the average value of the hardness in the depth range of 10nm-100nm.

Table 6.1 Surface hardness and shear strength of silica wafer

Surface shear strength, GPa	0.635±0.031
Surface hardness, GPa	6.60±0.82

Once the surface shear strength and wafer hardness are obtained, the calculated coefficient of friction from Eq. 6.6 is equalized to its corresponding measured value by adjusting the contact radius *a*. The fitted contact radius *a* is then introduced into Eq. 6.8 for the calculation of material removal rate, which is in turn verified by its comparison with the experimentally obtained polishing rate.

Material Removal Rate of Silica Polishing

Experimental data for removal rate of silica CMP as a function of solids loading are shown in Figure 6.3 along with the theoretical predictions from Eq. 6.8. The experimental data are adopted from silica wafer polishing of Mahajan [Mah99]. The

coefficient of friction used in Eq. 6.6 to calculate the contact radius of a single particle indentation was measured by mounting a lateral force measuring instrument to the Struers Rotopol 31 polisher [Mah99]. Friction forces were first measured at a 3.5 psi down pressure and 150 rpm platen speed, which are the same conditions as used in CMP. The friction force was then divided by the down force to obtain the coefficient of friction. The coefficient of friction measured in this experiment is 0.15 with experimental error of less than 5%. It is used in the calculations of Eq. 6.8, and only the concentration of particles C_a in the equation is changed corresponding to the experimental conditions shown in Fig. 6.3. It may be observed that the predictions match the experimental data very well. This is a further proof of the fact that the number of particles trapped between the pad and the wafer surface increases linearly with the increase of particle concentration of the slurry, and hence the linear increase of the polishing rate. The coefficient of friction by a single particle is not expected to change with the concentration of particles when the solids loading of the slurry is relatively low because the force acting on each particle is constant (Eq. 5.21). At higher solids loading, however, the increase of the removal rate as a function of solids loading may become sub-linear due to possible aggregates of the slurry particles and the spatial saturation of the trapped particles between the pad and the wafer surface. This phenomenon has been seen in both silica CMP [Bas02] and copper CMP [Luo98]. Similarly, the predicted result in Fig. 6.3 is higher than the experimental removal rate at solids loading of 10wt%. Beyond this point, the occupation of the trapped particles in the space between the wafer and the pad asperities may soon become saturated.

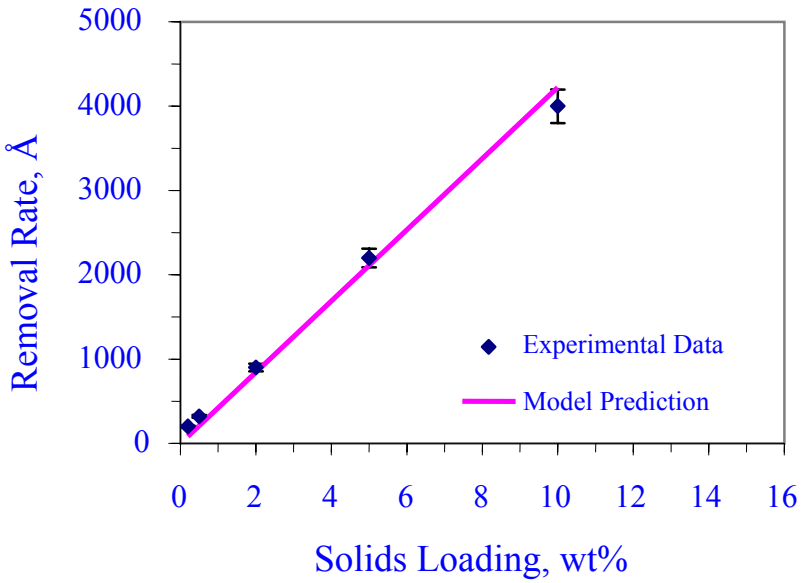


Figure 6.3 Material removal rate of silica wafer polishing as a function of solids loading with particle size of 0.2 μm.

The next set of experimental data was obtained by Mahajan [Mah99]. The CMP conditions were the same as previous example. However, the solids loading of the slurry was fixed at 2wt% and the material removal rate were studied by varying the particle size of the slurry. Figure 6.4 shows the correlation between the model prediction and the experimental observations of Mahajan when the particle size is varied from 0.2 μm to 1.5 μm. The estimated friction coefficients used in the model prediction are shown in Table 6.2. The contact radius of slurry particles with different sizes were obtained using Eq. 6.6, and the ratio of S/B_e was taken as a constant value of 0.096 (=0.635/6.6). The variations of estimated coefficients of friction with particle size, along with the experimental friction coefficients are compared in Table 6.2.

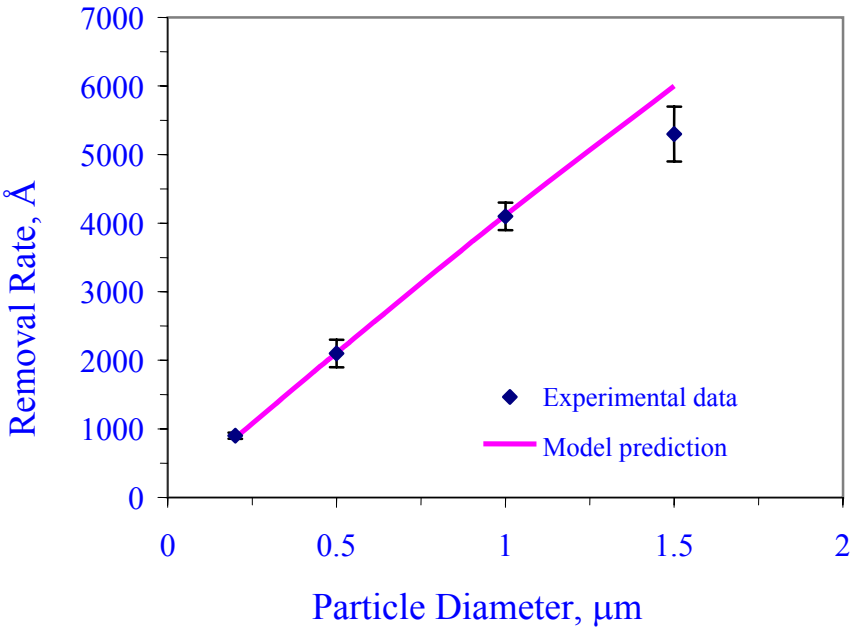


Figure 6.4 Material removal rate of silica wafer polishing as a function of slurry particle size with solids loading of 5wt%.

Mahajan observed the polishing rate to increase considerably as the particle size is increased (Fig. 6.4) when the solids loading of the slurry is relatively low (e.g., < 5wt%). It was explained assuming that the CMP was dominated by the particle indentation at low particle concentrations of the slurry. This is consistent with the model predictions in current study, and a good agreement between the model prediction and the experimental observations can be found from Fig. 6.4. Since the friction between a rigid particle and a flat substrate is determined by the shear strength of the substrate and the contact area between the particle and the substrate (Eq. 6.1a), the coefficient of friction should not change with the particle size if the surface layer is not presented or it is thick enough to eliminate its effect on friction because the normal force and the friction force are both proportional to the square of particle diameter (Eq. 5.23). However, comparing Eq. 5.23 with Eq. 5.26, we know that the contact area between the particle and the substrate is not

directly proportional to the square of particle diameter, indicating a higher friction for bigger particles, when a soft surface layer is presented. The comparison of Table 6.2 shows that this trend can be reliably depicted by the proposed model. The slightly mismatch between the fitted value and the experimental observation for particle size of 1.5 μm in Table 6.2 may be due to the large fluctuation of the experimental data.

Table 6.2 Comparison of the fitted coefficients of friction with the experimental data as a function of particle size. The error of experimental data is within 5%.

Particle size, μm	0.2	0.5	1.0	1.5
Fitted coefficient of friction	0.155	0.173	0.191	0.199
Experimental coefficient of friction	0.150	0.170	0.199	0.226

The essential observation from above analysis is that more friction force is needed for a larger material removal rate because the friction force is proportional to the particle contact area and hence the material removal. When dispersing agents are used in slurry production, a non-native surface layer due to the adsorption of these dispersants is formed on the particle surface as well as the wafer surface. This adsorption layer will act as lubricant between the particle and the wafer to be polished, which not only reduces the friction between them but also the material removal rate. Meanwhile, the competitive adsorption between the dispersants and the salt ions onto the wafer surface may result in reduced rate of chemical layer formation as discussed in previous chapters. Table 6.3 demonstrates these observations with a set of experimental data obtained by Basim [Bas02]. Two kinds of slurries were made with electrostatically dispersed (baseline) and poly(ethylene oxide) (PEO of MW 8,000,000) dispersed silica particles before the polishing experiment. The same operational conditions were applied for both slurries during CMP. The data in Table 6.3 shows that the removal rate of the silica wafer

polishing is reduced by more than 20% with the addition of PEO. Apparently, the contact radius between the abrasive particle and the wafer surface, as *a* in Eq. 6.8, is minimized due to the lubrication effect of the PEO adsorption layer. One would expect reduced friction force between the particle and the wafer substrate.

Table 6.3 Silica wafer polishing rate with electrostatically and PEO dispersed slurries of 12wt% and 0.2μm colloidal silica particles.

Slurry	Removal Rate, Å/min
Baseline (pH 10.5, DI water)	3800±410
0.5mg/g PEO dispersed	3090±240

Figure 6.5 is a direct friction force measurement using the same technique as Cain *et al* [Cai00]. The experiment was conducted with a Digital Instrument Nanoscope III atomic force microscope (AFM) by attaching a 7.3 μm silica particle on its cantilever tip. As the applied load was systematically increased, friction force response was measured as a function of the normal force. These measurements were performed with both solutions. The coefficient of friction can be readily calculated from the slope of the curve in Fig. 6.5 using Amonton's law [Bow50]. Both friction measurements with slurry in the presence of PEO and with the baseline slurry in the absence of any dispersants are shown in Figure 6.5 at a single particle-substrate interaction level. It is observed that the coefficient of friction reduces from 0.25 to 0.12 due to the presence of PEO, indicating the formation of the adsorption layer on the wafer surface. Following the proposed model of Eq. 6.8, we believe that the contact radius between the abrasive particle and the wafer substrate is reduced in the presence of PEO. Basim also investigated the effect of surfactant dispersants in the CMP slurry [Bas02]. When cetyl trimethyl ammonium bromide (CTAB), a cationic surfactant, was used to provide stability to silica slurry at

high ionic strength and extreme pH, it is found that the material removal rate of silica polishing was almost negligible due to the strong adsorption of the surfactant onto the wafer surface. A similar trend has also been reported for tungsten polishing in the presence of surfactants [Bie98], in which study the alumina particles in the polishing slurry were stabilized using a mixed surfactant system (anionic and nonionic). The material removal rate was reduced by at least 30% with the presence of surfactants.

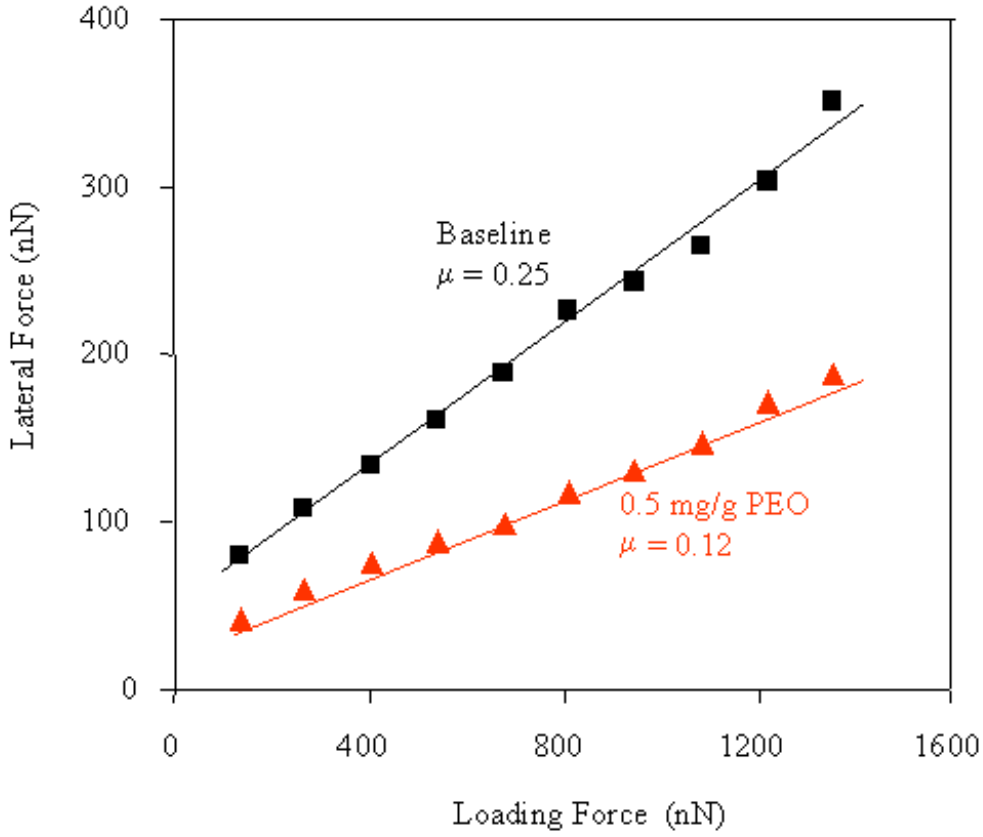


Figure 6.5 AFM friction coefficient measurements for the baseline and PEO containing solutions [Bas02].

The adoption of polymeric or surfactant dispersants is necessary for the stability of the CMP slurry and it may provide a better performance in terms of surface quality. However, the reduction of coefficient of friction and hence low material removal rate is

not desirable because the lubricating layer by the adsorption of the dispersants on the wafer surface and/or the particle surface minimizes the indentation of the abrasive particle into the wafer surface. On the other hand, the concept of reducing polishing rate by the adsorption of dispersants can be utilized to design material selective slurries and thus to achieve selectivity of the CMP process. By enabling the selective adsorption of surfactant molecules to a selected material while the other materials on the substrate remains uncoated, particle-substrate engagement can be achieved for the uncoated material leading to its polishing, whereas the polishing rate of the coated material is minimized due to the adsorption of the lubricating layer. Thus, the correlation between the coefficient of friction and the material removal rate during CMP has to be considered.

Summary

The material removal rate during CMP was quantitively correlated to the coefficient of friction, which can be measured readily with less effort than the polishing itself. A successive model of Chapter 5 was developed in this chapter in order to elucidate the effect of the chemically modified layer on the material removal rate of CMP, which can be estimated using available techniques like friction force measurement. In summary, the proposed correlation has been shown to be a reliable tool to predict the material removal rate through the measured coefficient of friction. Meanwhile, the effect of stabilizing agents that have been commonly used in the CMP slurry dispersions on the friction force between the abrasive particle and wafer surface was discussed, and it was observed that a relatively higher coefficient of friction is needed for better particle-wafer contact and consequently higher material removal rate. In addition, the successful application of the surface shear strength obtained using AFM lateral force measurement shows that it may be one of the feasible methods to probe the properties of the chemically

modified layer formed during CMP. Further investigation can be done for more reliable results of the surface shear strength measurements.

CHAPTER 7
SUMMARY AND SUGGESTIONS FOR FUTURE WORK

Summary

As the sizes of the microelectronic devices keep on shrinking, significant improvements are needed in the chemical mechanical polishing (CMP) process to achieve planarization requirements of atomic level material removal and minimal number of defects. To meet the desired performance criteria due to the rapid advances of microelectronic industry, it is necessary to understand the fundamentals of polishing mechanisms through the investigation of the pad-particle-substrate interactions in CMP. Rather than relaying on experimental results that are coming from trial and error on the tool related operational variables alone, in current study, the CMP process was modeled in multiple geometric scales by investigating the pad-particle-substrate interactions during polishing. To provide fundamental understanding in full physical scale, the modeling efforts in this study were broken down to three geometric scales, namely wafer-scale, feature-scale and micro-scale.

One of the major challenges in CMP is the uniformity of slurry solution distribution in the gap between the wafer and the polishing pad. This issue has to be investigated in the wafer scale. In the beginning of the current study, the wafer scale slurry flow is simulated, in which the pad and wafer are assumed to be rigid and parallel to each other. By solving the Navier-Stokes equation and calculating the shear rate distribution across the wafer, we found that it was necessary to maintain the same angular velocities of the pad and the wafer in order to achieve a uniform transport of the slurry

solutions. In addition, the modeling results indicated that both the increase of the distance and the increase of the ratio between the pad speed and the wafer speed would minimize the stagnant region under the wafer. Finally, the modeling results were compared with experimental observations and it was found that the combination of a high pad rpm, a medium wafer rpm, and a large pad-to-wafer distance should lead to a minimum non-uniformity across the wafer.

During the manufacturing process of microelectronic devices, the presence of features has a significant effect on the slurry flow of chemical mechanical polishing. To evaluate the stress distribution along and inside the wafer in the presence of features, finite element analysis was employed in the feature scale to solve both the slurry flow in the gap and the stress distribution inside the wafer. The modeling results showed an apparent pressure build up around the features due to its blockage of the slurry flow. It in turn resulted in stress concentration around the corner areas of the feature regardless of the feature shapes. It was found that the square feature had the highest von-Mises stress in the concentrated area among the feature shapes studied. However, the stress was not big enough to break the feature off its connection with the bulk wafer, indicating that the slurry flow itself was not likely to be the main factor to cause material removal during CMP.

A complete model for material removal has to take into account the interactions among the pad, the wafer and the abrasive particles in the slurry. Furthermore, due to the complexity of the process parameters on the material removal rate, the modeling should be able to predict the effect of these parameters for a better understanding of the CMP process. In the particle scale (or micro-scale) of current study, a comprehensive model

has been presented for the prediction of the material removal rate in a chemical mechanical polishing (CMP) process. The major contributions of the current model is that it took into account the chemical effect of the polishing slurry, considering the fact that a chemically modified layer forms through the chemical reaction (or diffusion of chemical ions in the slurry) and is being removed immediately by the abrasive particles during CMP. Most variables involved in CMP were also included in the proposed model. The model was based on a sliding-indentation process. The cross-section of the particle indentation and hence the volumetric removal rate by a single particle was first estimated. It was then multiplied with the removal rate of a single particle to obtain the formulation of the overall removal. Three polishing modes were considered according to the comparison between the indentation depth and the thickness of the chemically formed surface layer. One of the important findings of this particle-scale modeling is a balance between the surface layer formation rate and the single particle polishing rate has to be achieved for a defect free polishing performance. Meanwhile, the present model development was described as a function of various processing variables and material properties. Following an elaborate experimental verification, the present model appears to be capable of explaining many experimental observations on both oxide and metal CMP that, otherwise, could not be explained properly.

In the final attempt of this investigation, the correlation between the coefficient of friction and the material removal rate was established in the particle scale. Since the material removal of the CMP process is resulted from the mechanical contact between the particle and the wafer surface, the CMP is essentially a friction and wear process. A model for identifying the significance of the friction coefficient has been formulated.

Without losing its generality in describing the effects of various parameters involved in the CMP process, the role of chemistry in CMP was imposed on the model through the surface strength between the contact of abrasive particles and the wafer to be polished. Due to the availability of the experimental silica polishing results, the silica CMP was adopted as the simulation system to verify the model. An estimated value of the surface shear strength was presented and it has been seen to be possible to measure the surface shear strength using the lateral force microscope under CMP conditions. The modeling results showed that a relatively higher coefficient of friction is needed for better particle-wafer contact and thus higher material removal rate. The friction coefficient can be reduced by the addition of stabling agents due to the adsorption of the lubricating layer. However, it may be used for the design of material selective slurries and to achieve selectivity of the CMP process.

In general, the optimization of the controlling parameters involved in the polishing process can be obtained through model predictions. The reported findings in this study can also provide a fundamental understanding of the polishing mechanisms and can provide guidelines for optimization of the CMP applications.

Suggestions for Future Work

The current study can be expanded from several aspects. For example, the shape of the abrasive particles and the distribution of particle sizes can be added to the micro-contact model formulation of the material removal rate to make more predictions. Throughout this investigation, spherical particles were used in order to simplify the modeling. However, some of the commercial polishing slurries may not be spherical, e.g., fused silica, in which the primary particles are composed of chains of much smaller particle beads. It also creates a broad size distribution instead of uniform size. These

types of slurries are known to polish effectively. The size distribution of the abrasive particles may eventually alter the tendency of trapping of the particles between the pad and wafer contact, and thus the number of particles. The shape of abrasives can result in varied normal force on each particle and then change the indentation of the particle, hence the material removal during CMP. Therefore, the impact of slurry particle shape and the particle size distribution are valuable parameters that may be included in the model.

One of the major challenges in this modeling effort is the prediction of the chemically modified surface layer. It has been shown that successful CMP performance has to be achieved when the surface layer formation is balanced with the abrasions of the single particle. The control of the chemically formed surface layer depends on the formulation of the slurry, i.e., the chemical species added in the slurry solution. Thus the prediction of surface layer formation under various CMP conditions may be achieved by considering the principle of the chemical reactions. Due to the versatility of the materials to be polished and the chemical options that may be used, more challenges are expected besides polishing solution itself. However, silica system can always be picked as a simulation system because of the massive experimental data available in the literature.

Another challenge following this study is the experimental measurement of the surface layer. Although it has been tried out during current study that the conventional techniques like nanoindentation and ellipsometry fail to make in-situ measurement of the formation of the chemically modified layer due to its extreme thinness and limitations of the equipment, the lateral force measurement using AFM seems a feasible tool to capture the thickness information of the surface layer. However, extra instrument like amplifier

has to be installed to a conventional AFM and extremely careful control over the equipment and measuring environment have to be maintained during measurement. The repeatability of these measurements is another challenge that needs to be resolved before taking on the reliable experiment. Additionally, the friction force measurement using AFM can provide valuable information on the prediction using the model developed in the previous chapter for material removal rate estimation. It can also be used for the evaluation of the slurry design based on the measured coefficient of friction.

The correlation between the material removal and the coefficient of friction may be further expanded. It has been shown in this study that the slurry chemistry and the dispersants adsorption can lead to significant variation of the friction, essentially the engagement, between the particle and the wafer surface. The material removal rate is proven to be determined by the friction between the abrasives and the wafer surface. Therefore, a robust slurry design relies on the understanding of the mechanism of tribological issues involved in CMP. For example, the effect of surface chemical reactions and/or surface adsorptions on the coefficient of friction may be experimentally examined for specific systems. Moreover, modeling effort on these tribological issues should also be encouraged.

Finally, the suggestion is on the optimal slurry design. A fully dispersed slurry suspension has evidenced to be beneficial for the CMP performance. Thus properly selected dispersants are desirable for a robust slurry. However, the CMP process is usually conducted under high shear rate, which may result in different adsorption from that under low shear rate of regular slurry making process (e.g., agitation). Therefore, the

adsorption property under high shear rate may be another topic that deserve to be investigated.

LIST OF REFERENCES

Adl01 J. J. Adler, Interaction of Non-Ideal Surfaces in particulate Systems, Ph.D. Dissertation, University of Florida, Gainesville, FL (2001).

Aga00 L. Agarwal, Hydrodynamic and Solid Mechanics of the CMP Process, Master's Thesis, University of Florida, Gainesville, FL (2000).

Ahm01 G. Ahmadi and X. Xia, J. Electrochem. Soc., 148, G99 (2001).

Ash80 M. F. Ashby and D. R. H. Jones, Engineering Materials 1, Pergamon Press, Oxford (1980).

Bar83 R. Bartholemew, J. Non-Cryst. Solids, 56, 331 (1983).

Bas02 G. B. Basim, Formulation of Engineered Particulate Systems for Chemical Mechanical Polishing Applications, Ph.D. Dissertation, University of Florida, Gainesville, FL (2002).

Bib97 T. F. A. Bibby, R. Harwood, D. Schey and K. McKinley, Thin Solid Films, 308-309, 512 (1997).

Bie98 M. Bielmann, Chemical Mechanical Polishing of Tungsten, Master's Thesis, University of Florida, Gainesville, FL (1998).

Bie99 M. Bielmann, U. Mahajan, and R. Singh, Electrochem. Solid State Lett., 2, 148 (1999).

Bif91 T. Bifano, T. A. Dow and R. O. Scattergood, ASME Trans. J. Eng. Industry, 113, 184 (1991).

Bow50 F. P. Bowden and D. Tabor, The Friction and Lubrication of Solids, Clarendon Press, Oxford (1950).

Bro81 N.J. Brown, P.C. Baker and R.T. Maney, Proc. SPIE, 306, 42 (1981).

Bud61 S. M. Budd, Phys. Chem. Glasses, 2, 111 (1961).

Cai00 R. G. Cain, S. Biggs and N. W. Page, J. Colloid Interf. Sci., 227, 55 (2000).

Car97 R. W. Carpick, D. F. Ogletree and M. Salmeron, Appl. Phys. Lett., 70, 1548 (1997).

Che97 Chen et al., US Patent No. 5,597,442 (1997).

Che98 O. G. Chekina, L. M. Keer and H. Liang, J. Electrochem. Soc., 145, 2100 (1998).

Che02 J. M. Chen and Y.-C. Fang, IEEE Trans. Semiconduct. Manuf., 15, 39 (2002).

Cho01 C.-H. Cho, S.-S. Park and Y. Ahn, Thin Solid Films, 389, 254 (2001).

Coo90 L.M. Cook, J. Non-Cryst. Solids, 120, 152 (1990).

Cop00 J. Coppeta, C. Rogers, L. Racz, A. Philipossian and F. B. Kaufman, J. Electrochem. Soc., 147, 1903 (2000).

Dor73 R. Doremus, Glass Science, Wiley, New York, NY (1973).

Elk95 B. El-Kareh, Fundamentals of Semiconductor Processing Technologies, Kluwer Academic Publishers, Boston, 1995.

Els79 M. G. El-Sherbiny and F. B. Salem, Wear, 54, 391 (1979).

Eva80 A. G. Evans and D. B. Marshall, Fundamentals of Friction and Wear of Matreials, ASM, pp. 439-452, (1980).

Gre66 J. A. Greenwood and J. B. Williamson, Proc. Royal Soc. A, 295, 300 (1966).

Ham94 B. J. Hamrock, Fundamentals of Fluid Film Lubrication, McGraw-Hill, New York, NY (1994).

Hol86 D. G. Holloway, Glass Technology, 27, 120(1986).

Ile79 R. Iler, The Chemistry of Silica: Solubility, Polymerization, Colloid and Surface Properties, and Biochemistry, Wiley, New York, NY (1979).

Isr92 J.N. Israelachvili, Intermolecular and Surface Forces, Academic Press, New York, NY (1992).

Izu79 T. Izumitani, in Treatise on Materials Science and Technology, M. Tomozawa, R. Doremus, Editors, p. 115, Academic Press, New York, NY (1979).

Izu84 T. Izumitani, Paper TuB-A1, Tech. Digest, Topical Meeting on the Science of Polishing, Optical Society of America (1984).

Jac83 R. J. Jackobsen, L. L. Brown, R. I. Leininger and R. M. Gendreau, J. Biomed. Mat. Res., 16, 199 (1983).

Jai94 R. Jairath, M. Desai, M. Stell, R. Tolles and D. Scherber-Brewer, in Advanced Metallization for Devices and Circuits -- Science, Technology and Manufacturability, S.P Murarka, A. Katz, K.N. Tu and K. Maex, Editors, 337, p. 121, Materials Research Society Proceedings Series, Pittsburgh, PA (1994).

Jeo92 J. S. Jeon, R. P. Sperline and S. Raghavan, Appl. Spectrosc. 46, 1644 (1992).

Jin01 A. Jindal, S. Hegde and S. V. Babu, Proceedings of the Eighteenth International IEEE VLSI Multilevel Interconnection Conference (VMIC), Santa Clara, CA, p.297 (2001).

Joh85 K. L. Johnson, Contact Mechanics, Cambridge University Press, Cambridge (1985).

Kau91 F. B. Kaufman, D.B. Thomson, R.E. Broadie, M.A. Jaso, W.L. Guthrie, D. J. Pearson and M. B. Small, J. Electrochem. Soc., 138, 3460 (1991).

Kom85 K. Komvopoulos, N. Saka and N. P. Suh, ASME J. Tribology, 107, 452 (1985).

Lar99 J. Larsen-Basse and H. Liang, Wear, 233-235, 647 (1999).

Lev98 J. A. Levert, F. M. Mess, R. F. Salant, S. Danyluk and A. R. Baker, Tribol. Trans., 41, 593 (1998).

Li95 W. Li, D. W. Shin, M. Tomozawa and S. P. Murarka, Thin Solid Films, 270, 610 (1995).

Liu95 C. W. Liu, B. T. Dai and C. F. Yeh, Thin Solid Films, 270, 607 (1995).

Luo98 Q. Luo, S. Ramajan and S.V. Babu, Thin Solid Films, 335, 160 (1998).

Luo99 Q. Luo, Chemical Mechanical Polishing of Copper Thin Films, Ph.D. Dissertation, Clarkson University, New York (1999).

Luo01 J. Luo and D. A. Dornfeld, IEEE Trans. Semiconduct. Manuf., 14, 112 (2001).

Mah99 U. Mahajan, M. Bielmann and R.K. Singh, Electrochem. Solid State Lett., 2, 80 (1999).

Mah00 U. Mahajan, Fundamental Studies on Silicon Dioxide Chemical Mechanical Polishing, Ph.D Dissertation, University of Florida, Gainesville, FL (2000).

Mod97 A. Modak, P. Monteith and N. Parekh, in Proceedings of the 2nd International Chemical Mechanical Planarization for ULSI Multilevel Interconnection Conference (CMP-MIC), p. 169, Santa Clara, CA (1997).

Mur93 S. P. Murarka, Metallization Theory and Practice for VLSI and ULSI, Butterworth-Heinemann, Boston, MA (1993).

Mur00 S. P. Murarka, I. V. Verner and R. J. Gutman, Copper-Fundamental Mechanisms for Microelectronic Applications, John Wiley, New York (2000).

Noh02 K. Noh, J.-Y. Lai, N. Saka, and J.-H. Chun, Proceedings of SMA Symposium, Singapore (2002).

Oma94 W. C. O'mara, Planarization by Chemical Mechanical Polishing for Multilevel Metal Integrated Circuits, O'mara & Associates, Palo Alto, CA (1994).

Oum99 D.O. Ouma, Modelling of Chemical Mechanical Polishing for Dielectric Planarization, PhD Dissertation, MIT, Boston, MA (1999).

Pat91 W. J. Patrick, W. L. Guthrie, C. L. Standley and P. M. Schiable, J. Electrochem. Soc., 138, 1778 (1991).

Pet02a L. Peters, Semiconductor International, 25 (1), 48 (2002).

Pet02b L. Peters, Semiconductor International, 25 (5), 55 (2002).

Phi01 A. Philipossian, PARTEC 2001, Nuremberg, Germany (2001).

Pre27 F. Preston, J. Soc. Glass Technol., 11, 247 (1927).

Rab95 E. Rabinowicz, Friction and Wear of Materials, Wiley, New York (1995).

Run94a S. R. Runnels and L.M. Eyman, J. Electrochem. Soc., 141, 1698 (1994).

Run94b S. R. Runnels, J. Electrochem. Soc., 141, 1900 (1994).

Run96 S. R. Runnels, J. Electron. Mater., 25, 1574 (1996).

Sha01 J. F. Shackelford and W. Alexander, Eds., CRC Materials Science and Engineering Handbook, CRC Press, Boca Raton, FL (2001).

Shi98 F.G. Shi, B. Zhao and S.-Q. Wang, Proc. of International Technology Conference, San Francisco, CA, 73 (1998).

Sin02 P. Singer, Semiconductor International, 25 (5), 46 (2002).

Siv92 S. Sivaram, M. H.M. Bath, E. Lee, R. Leggett and R. Tolles, Proc. SRC Topical Research Conference on Chem-Mechanical Polishing for Planarization, SRC, Research Triangle Park, NC, 8 (1992).

Soh99 I.-S. Sohn, S. Sohn, B. Moudgil, R. Singh, C.-W. Park, in Chemical-mechanical polishing, fundamentals and changes, S.V. Babu, S. Danyluk, M. Krishnan, M. Tsujimura, Eds., 566, p. 181, Materials Research Society Symposium Proceedings, San Francisco, CA (2000).

Sol99 M. J. Solomon, T. Saeki, M. Wan, P. J. Scales, D. V. Boger and H. Usui, Langmuir, 15, 20 (1999).

Sre01 P. S. Sreejith and B. K. A. Ngoi, Int. J. Mach. Tools Manu., 41, 1831 (2001).

Ste97 J.M. Steigerwald, S.P. Murarka and R.J. Gutmann, Chemical Mechanical Planarization of Microelectronic Materials, John Wiley and Sons, New York, NY (1997).

Ste99 D. J. Stein, D. L. Hetherington and J. L. Cecchi, J. Electrochem. Soc., 145, 1934 (1999).

Sun94 S. C. Sun, F. L. Yeh and H. Z. Tien, in: S.P. Murarka (Eds.), Advanced Metallization for Devices and Circuits -- Science, Technology, and Manufacturability, 337, pp. 139, Materials Research Society Symposium Proceedings, San Francisco, CA (1994).

Sun99 S. Sundararajan, D. G. Thakurta, D. W. Schwendeman, S. P. Murarka, and W. N. Gill, J. Electrochem. Soc., 146, 761 (1999).

Sze81 S. M. Sze, Physics of Semiconductor Devices, John Wiley Sons Inc., New York, NY (1981).

Tad68 Th. F. Tadros, and J. Lyklema, J. Electroanal. Chem., 17, 267 (1968).

Tic99 J. Ticky, J. A. Levert, L. Shan and S. Danyluk, J. Electrochem. Soc., 146, 1523 (1999).

Tro94 J. A. Trogolo and K. Rajan, J. Mater. Sci., 29, 4554 (1994).

Tse97 W.-T. Tseng and Y.-L. Wang, J. Electrochem. Soc., 144, 14 (1997).

Tse99 W.-T. Tseng, J.-H. Chin, L.C. Kang, J. Electrochem. Soc., 146, 1952 (1999).

War92 J. Warnock, J. Electrochem. Soc., 138, 2398 (1991).

Win82 S. Winters, R. M. Gendreau, R. I. Leininger and R. J. Jackobsen, Appl. Spectrosc. 36, 404 (1982).

Wol86 S. Wolf and R. N. Tauber, Silicon Processing for the VLSI Era, Lattice Press, Sunset Beach, CA (1986).

Wol99 S. Wolf and R. N. Tauber, Silicon Processing for the VLSI Era, 2nd Edition, Lattice Press, Sunset Beach, CA (1999).

Xie96 Y. Xie and B. Bhushan, Wear, 200, 281 (1996).

Yu94 T. K. Yu, C. C. Yu and M. Orlowski, Proceedings of the 1993 International Electronic Devices Meeting, 4.1, 35 (1994).

Zha02 Y. Zhao and L. Chang, Wear, 252, 220 (2002).

BIOGRAPHICAL SKETCH

Kuide Qin was born on March 12, 1971, in Shandong, China. In 1993, he graduated with a Bachelor of Engineering degree in chemical engineering from Tianjin University, Tianjin, China. From September 1993 to March 1996, he was with the Chemical Engineering Research Center at Tianjin University, where he performed research on complex distillation systems and received his Master of Engineering degree. He then joined Huan-Qiu Chemical Engineering Corporation in Beijing, China, where he worked as a process engineer for two and a half years. In pursuit of a high quality education, he came to the University of Florida in August 1998 and joined the Department of Chemical Engineering with Professor Chang-Won Park as his advisor. He first worked on the study of graded-index polymer optical fibers. The contamination in fiber processing triggered his interest for further exploration of knowledge in particle science and technology. He then joined the Engineering Research Center for Particle Science and Technology, where his research is concentrated on multi-scale modeling of chemical mechanical polishing and design of optimal polishing slurries under the co-guidance of Professor Brij M. Moudgil. He graduated from the University of Florida with a Ph.D. degree in chemical engineering with electronic materials and particle science and technology specialties in May 2003.