MODELING OF DIELECTRIC EROSION AND COPPER DISHING IN COPPER CHEMICAL-MECHANICAL POLISHING

by

Kyungyoon Noh

B.S., Mechanical Engineering
Seoul National University, 1995

M.S., Mechanical Engineering
Seoul National University, 1997

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

Signature of Author: ______________________________

Kyungyoon Noh
Department of Mechanical Engineering
May 25, 2005

Certified by: ______________________________

Jung-Hoon Chun
Professor of Mechanical Engineering
Thesis Supervisor

Certified by: ______________________________

Nannaji Saka
Research Affiliate, Department of Mechanical Engineering
Thesis Supervisor

Certified by: ______________________________

Lallit Anand
Professor of Mechanical Engineering
Chairman, Department Committee on Graduate Students

MODELING OF DIELECTRIC EROSION AND COPPER DISHING IN COPPER CHEMICAL-MECHANICAL POLISHING

by

Kyungyoon Noh

Submitted to the Department of Mechanical Engineering
on May 25, 2005 in partial fulfillment of the requirement for the Degree of
Doctor of Philosophy in Mechanical Engineering

ABSTRACT

The phenomenal success in the manufacture of multi-layer, Ultra-Large-Scale-Integrated (ULSI) semiconductor devices is in part due to the local and global planarization capabilities of the chemical-mechanical polishing (CMP) process. At present, copper is widely used as the interconnect material in the ULSI technology. The greatest challenge in Cu CMP now is the control of wafer surface non-uniformity—primarily due to dielectric erosion and copper dishing at various scales—to within the ever stringent industry specifications.

In this thesis, an integrated non-uniformity model is developed by combining wafer-, die- and feature-scale non-uniformities. A feature-scale pressure calculation scheme based on surface step-height is adopted, and the evolution of the surface in each polishing stage is modeled in terms of geometric, material and process parameters. Various pad/wafer contact mechanics regimes have been considered to model oxide erosion and Cu dishing, from submicron device level to the global wiring level. The plausible causes of erosion and dishing at wafer-, die- and feature-scales were identified and integrated into the feature-scale step-height models. Such parameters include: initial pattern geometry, wafer-scale uniformity, and Cu-to-oxide slurry selectivity, material properties, and surface topography of the pad.

Based on the developed erosion and dishing models, the effects of model parameters on the wafer-surface non-uniformity in Cu CMP are discussed, and parameter sets to satisfy both dishing and erosion specifications are obtained. In single-step polishing, for example, the Cu deposition factor should be less than 0.1 and the wafer-scale uniformity factor needs to be greater than 0.95 to maintain both erosion and dishing within 5% of interconnect thickness across the wafer if the polishing slurry has a selectivity of 15.

Results of polishing experiments on 100 mm patterned Cu wafers validated both the step-height models and the integrated non-uniformity model. Based on the present models, erosion and dishing across the wafer was bounded by predefined parameters. Additionally, as predicted by the models, it was observed that the step-heights of the slowest and the fastest dies evolve in the ratio of the wafer-scale uniformity factor.

Thesis Supervisor: Dr. Jung-Hoon Chun
Title: Professor of Mechanical Engineering

Thesis Supervisor: Dr. Nannaji Saka
Title: Research Affiliate, Department of Mechanical Engineering

Acknowledgments

As I look back, the past five years my life at MIT have been both demanding and exciting. I wish to thank the people who have guided and supported me.

First of all, I wish to acknowledge the financial support of the Singapore MIT Alliance program.

I thank Professor Jung-Hoon Chun, my thesis advisor, for his guidance, encouragement, and advice on professional development. Whenever I faced problems, in research or personal life, his straightforward advice cleared my hesitation, and gave me confidence and resolution to move forward. His continuous emphasis on the improvement of my English, writing everyday, and better management of time has greatly affected me, even though I could not keep up all the time. His approach to solving problems has inspired me, and will be one of the important lessons that I learned as a researcher.

I thank Dr. Nannaji Saka for his effort to transform me into a better researcher. I appreciate the enormous time and energy he has spent on me to discuss my research in detail, to correct my documents, and to instruct me as to what kind of research attitude I should have: honesty, passion, and unwillingness to compromise quality over quantity. He has been tough on me from the start to finish; yet I deeply appreciate it.

I also thank Professor Duane Boning, my thesis committee member, whose continual feedback on CMP research elsewhere and valuable comments on my own research. His gentle comments have significantly improved this thesis. I will always remember him for his kindness.

I would also like to thank Professor Nam P. Suh for his interest in my work and for his generous advice and guidance to my future career.

I thank Krzysztof Kopanski for the time and effort we spent together in the lab everyday for the past two years and wish him the best. I am also grateful to all my previous and current officemates for their support: Hady Joumma, Munhee Sohn, Sam Korb, Dr. Jeanie Cherng and Dr. Wayne Hsiao. Special thanks are due the LMP machine shop people who have helped me build machines and fix apparatus: Gerald Wentworth, Mark Belanger, David Dow and Patrick Mcatamney. I also thank Lisa Falco, Catherine Nichols and Leslie Regan for their cheerful administrative support.

I thank my best friend Soohyung Kim and his family. Thanks are also due my friends at MIT, especially the KGSA Mechanical Engineering and KGSA soccer team members.

Finally, I thank my parents who have always supported and believed in me, no matter what, and my parents-in-law for their support in my life. My parents and in-laws always cheered me up and taught me to be a better person. I also thank my brother and his family, and my brother-in-law. Above all, I thank my loving wife, Soo Hyun Kim. For the past ten years, since we first met, she not only shared the joys of my life but also stood by me every single moment I struggled.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Background

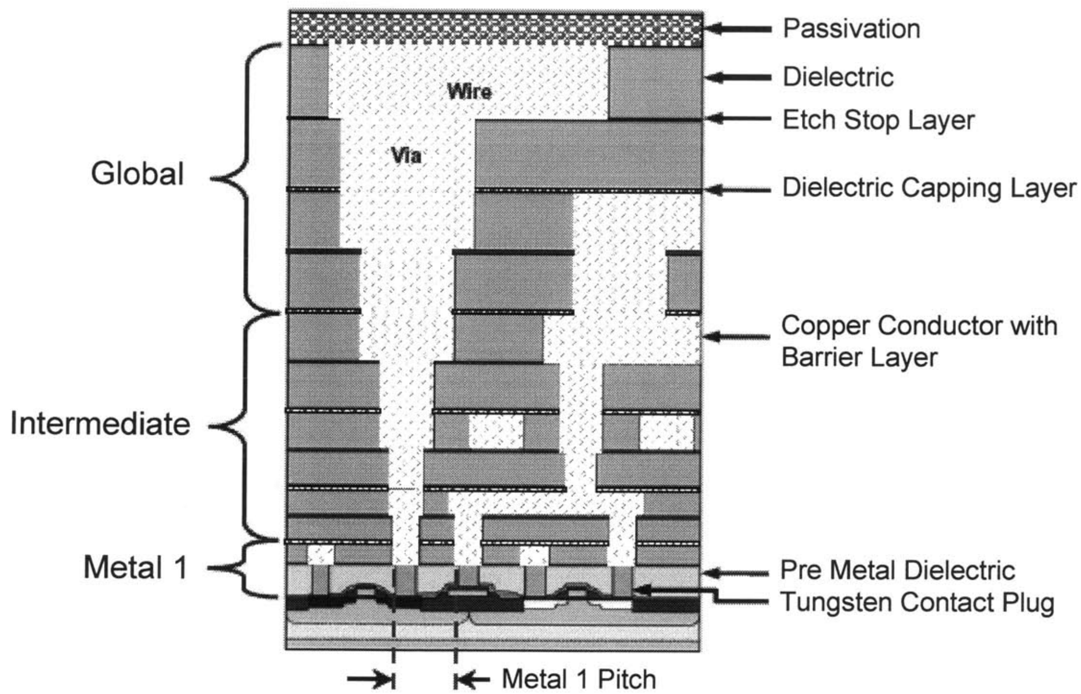
Over the past decade and a half, the semiconductor industry has grown rapidly to meet the ever-increasing demand for high-performance, ultra-large-scale integrated (ULSI) electronics. The number of components on a chip has doubled approximately every two years with more than 10^8 devices currently, as Moore had predicted decades ago [Moore, 1965; Chang and Sze, 1996]. As a result, the semiconductor industry is motivated to design and fabricate submicron features of ever-finer resolution, denser packing and multi-layer structures. Table 1.1 lists the wiring needs of interconnects between 2005 and 2009 for high-performance microprocessors [International Technology Roadmap for Semiconductors (ITRS) – Interconnect, 2003]. For example, in 2005, the requirement for minimum pitch at the submicron device level of microprocessors is 95 nm, and 11 metal layers on a chip.

By virtue of its low electrical resistivity and resistance to electromigration, in the past decade, copper has rapidly emerged as the preferred interconnect material in lieu of aluminum. Figure 1.1(a) is a schematic cross section of a typical chip with Cu interconnects. The metal 1 layer is called the submicron device level, the interconnect levels up to five layers are designated as intermediate levels, and the interconnect layers higher than the sixth as the global wiring levels. The wiring specifications for the submicron, intermediate, and the global wiring levels are listed in Table 1.1. Figure 1.1(b) is a scanning electron micrograph of a sample multi-layer Cu interconnect chip.

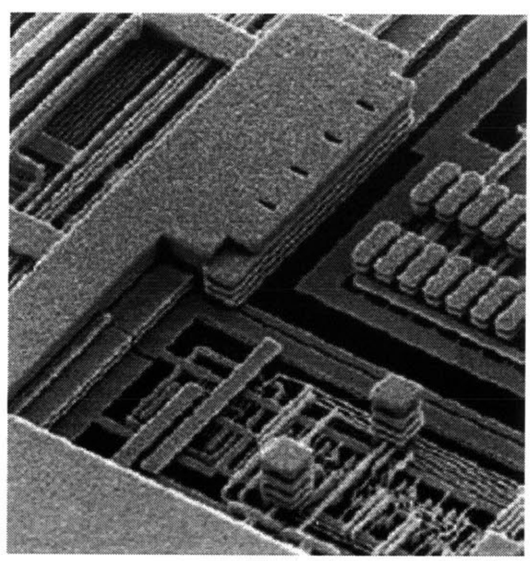
In the relentless endeavor to meet the ever-stringent specifications, the chemical-mechanical planarization or polishing (CMP) process has played a key role due to its local and global planarization capabilities in both Al and Cu techniques as shown in Fig. 1.2. In Al interconnect technology, shown in Fig 1.2(a) – (c), the interconnect lines are generated by metal deposition, patterning and etching processes. Then, an oxide layer is deposited over the interconnect lines and planarized by the CMP process, which is defined as inter-level dielectric (ILD) CMP. In the

Table 1.1 Microprocessor interconnect technology requirements [ITRS Interconnect, 2003].

Year of Production	**2005**	**2007**	**2009**
Technology Node	hp90	hp65	hp65
Dram ½ Pitch (nm)	80	65	50
MPU/ASIC ½ Pitch (nm)	95	76	60
MPU Printed Gate Length (nm)	45	35	28
MPU Physical Gate Length (nm)	32	25	20
Number of metal levels	11	11	12
Total interconnect length (m/cm^2) – active wiring only, excluding global levels	907	1117	1559
Metal 1 wiring pitch (nm)	190	152	120
Metal 1 A/R (for Cu)	1.7	1.7	1.8
Interconnect RC delay (ps) for 1 mm Metal 1 line	284	384	595
Cu thinning at minimum pitch due to erosion (nm), 10% x height, 50% areal density, 500 μm square array	16	13	11
Intermediate wiring pitch (nm)	240	195	156
Intermediate wiring dual Damascene A/R (Cu wire/via)	1.7/1.5	1.8/1.6	1.8/1.6
Interconnect RC delay (ps) for 1 mm intermediate line	182	229	358
Cu thinning at intermediate pitch due to erosion (nm), 10% x height, 50% areal density, 500 μm square array	20	18	10
Minimum global wiring pitch (nm)	360	290	234
Ratio range (global wiring pitches/intermediate wiring pitches)	1.5–6.7	1.5–8.0	1.5–8.0
Global wiring dual Damascene A/R (Cu wire/via)	2.2/2.0	2.2/2.0	2.3/2.0
Interconnect RC delay (ps) for 1 mm global line at minimum pitch	69	92	139
Cu thinning at maximum width global wiring due to dishing and erosion (nm), 10% x height, 80% areal density	176	172	144
Cu thinning global wiring due to dishing (nm), 100 μm wide feature	24	19	15
Barrier/cladding thickness (for Cu intermediate wiring) (nm)	9	7	6
Intermediate metal insulator (minimum expected) – bulk dielectric constant (κ)	<2.7	<2.4	<2.4



(a)



(b)

Figure 1.1 (a) Schematics of typical chip cross section [ITRS Interconnect, 2003] and (b) Scanning electron micrograph of IBM's six-level copper interconnect technology in an integrated circuit chip [IBM Corporation, 1997].

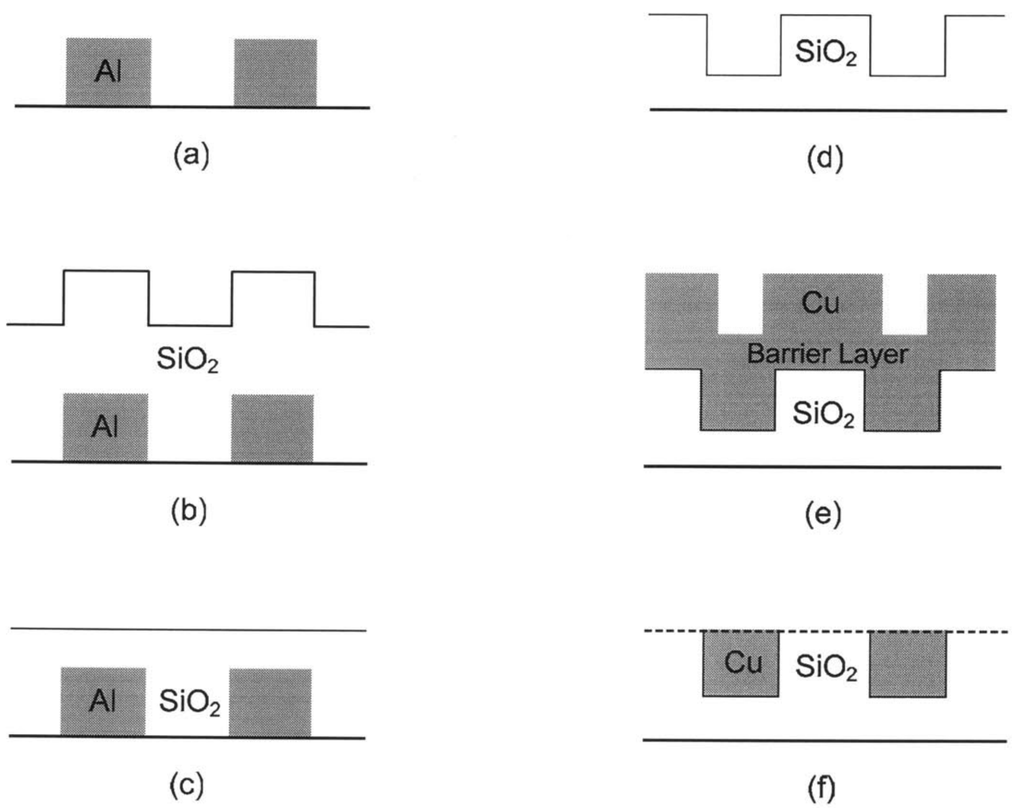


Figure 1.2 Comparison of the Al and Cu interconnect metallization processes. Al metallization: (a) Interconnect deposition, (b) Oxide deposition and (c) Planarization by ILD CMP. Cu metallization: (d) Interconnect pattern etch and barrier layer deposition, (e) Cu deposition and (f) Planarization by Cu CMP.

Cu technology, interconnect lines are produced by a so-called damascene scheme, shown in Fig. 1.2(d)–(f). After a dielectric layer is deposited, interconnect trench patterns are generated by photo-lithography and etching processes. Then, a barrier layer and Cu are deposited over the trenches and the excess Cu and barrier material are removed by the CMP process, which is called Cu CMP. Due to the material and geometric complexities, however, systematic approaches to resolve non-uniformities in Cu CMP are still lacking. In this research, therefore, the primary concern is on Cu CMP. It is hoped that by modeling and optimizing Cu CMP, it is also possible to optimize inter-level dielectric (ILD) and shallow trench isolation (STI) CMP, and the manufacture of novel micro- and nano-scale devices.

1.2 The Copper Chemical-Mechanical Polishing Process

The CMP process is both a material removal and surface planarization process. Based on the relative motion of the pad and the wafer, CMP equipment may be classified as linear, rotary, or orbital. Figure 1.3 shows a schematic of the most common rotary-type CMP process. The wafer to be polished is mounted on a wafer carrier with its polishing surface facing downward, and is pressed against a rotating platen that holds a polishing pad. As the platen and the wafer carrier rotate, abrasive slurry is fed onto the pad at the edge of the wafer carrier and is carried underneath the wafer by the polishing pad.

The purpose of the Cu CMP process is to remove excess Cu and barrier layer in the damascene Cu metallization process shown in Fig. 1.2(d)–(f). Copper is deposited by electroplating or physical vapor deposition (PVD) process over the oxide trenches with various pattern geometries as in Fig. 1.4(a). Ideally, polishing should end when the excess Cu and the barrier layer are completely removed across a wafer while the polishing surface remains perfectly flat as shown in Fig. 1.4(b). In the real CMP process, however, there are always wafer-surface non-uniformities, termed as dielectric erosion and Cu dishing, as shown in Fig. 1.4(c).

In Cu CMP, the material removal and planarization mechanisms are a complex combination of chemical and mechanical interactions dependent on the wafer geometry and materials, slurry and pad, and process parameters as shown in Fig. 1.5. It has been argued that the chemical component of the process is the reaction of the wafer surface with slurry chemicals to form a chemically modified surface for enhancing material removal rate [Cook, 1990; Kaufman *et al.*,

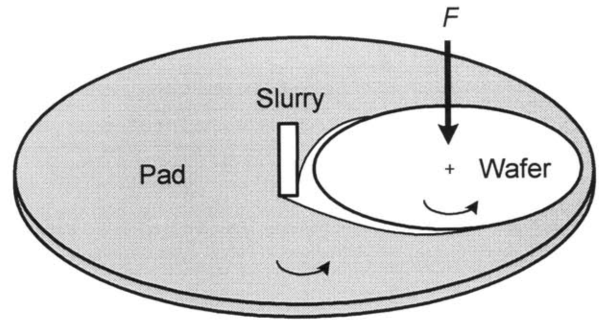


Figure 1.3 Schematic of the conventional face-down, rotary-type CMP process.

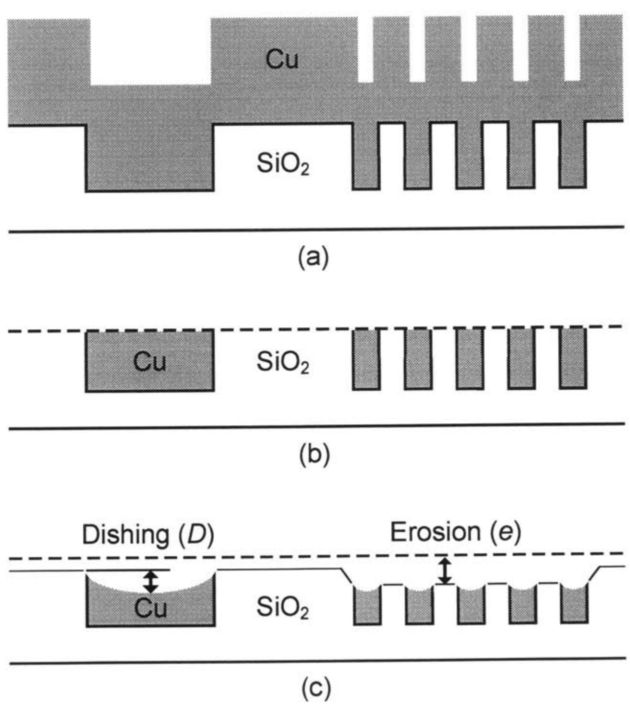


Figure 1.4 Schematics of pattern cross-sections: (a) before CMP, (b) after CMP (ideal case) and (c) after CMP (real case).

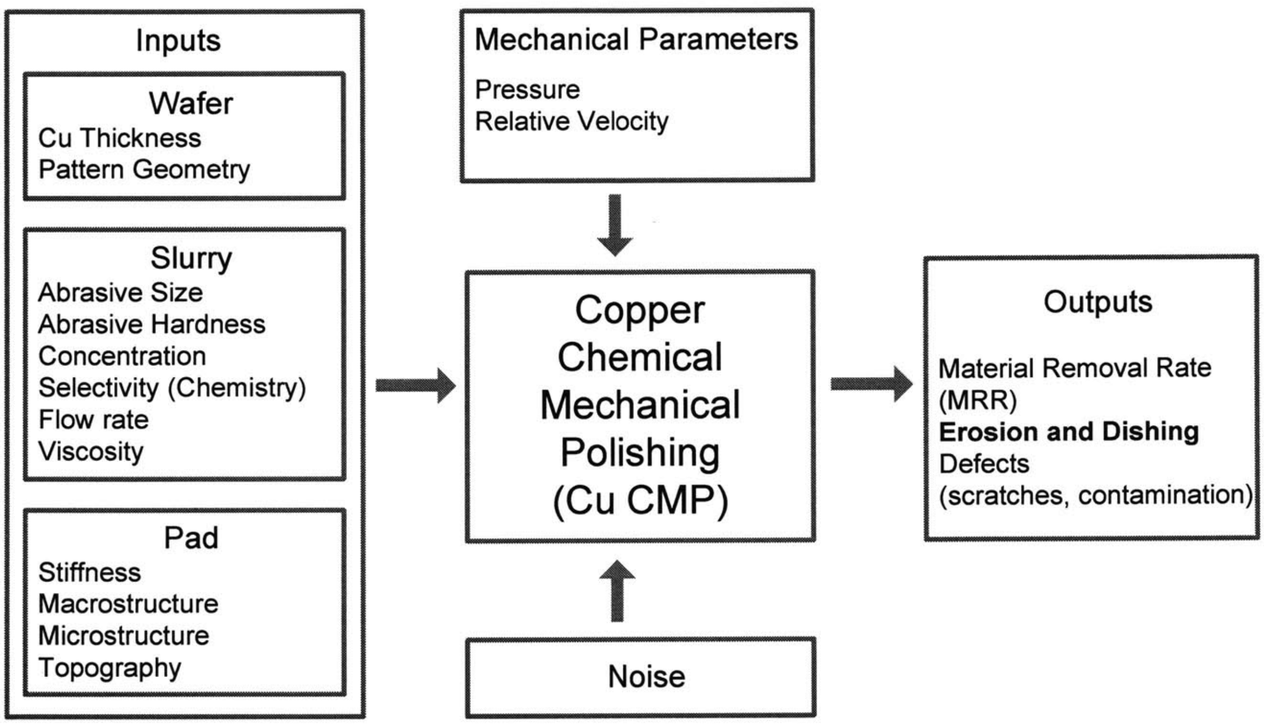


Figure 1.5 Inputs and outputs of the CMP process.

1991; Hariharaputhiran *et al.*, 2000; Singh and Bajaj, 2002; Jindal and Babu, 2004], and the mechanical component of the CMP process is the material removal from the wafer surface by the abrasive particles in the slurry [Liu *et al.*, 1996; Saka *et al.*, 2000; Ahmadi and Xia, 2001; Fu *et al.*, 2001; Luo and Dornfeld, 2003; Seok *et al.*, 2004]. Notwithstanding the enormous research on CMP to meet the ever-stringent demands, fundamental understanding of the process is still inadequate. Although the semiconductor industry has managed to meet the ever-stringent specifications so far, CMP now faces great challenges in the transition to nano-scale IC device fabrication, due to the shrinking device size, and novel materials like low-k dielectrics.

1.3 Scope of Present Thesis

There are three important requirements in the Cu CMP process. First, it is necessary to increase the throughput by increasing the material removal rate of Cu. For this, three approaches are typically attempted: mechanical, chemical and electrochemical enhancements. The mechanical approach is to increase the applied pressure or relative velocity, to decrease the size of abrasives or to use fixed abrasive pads, and to increase the concentration of abrasives [Fu *et al.*, 2001; Luo and Dornfeld, 2003]. In the present industrial application, for example, the size of abrasive varies from 100 nm to a few micrometers. The chemical approach is to use additives that react with Cu and form a softer Cu layer [Singh and Bajaj, 2002]. For example, hydrogen peroxide (H_2O_2) is the most common additive to enhance the material removal rate of Cu in CMP. Recently, an electrochemical approach has been adopted based on electropolishing technique [Padhi *et al.*, 2003; Huo *et al.*, 2005]. As a second key requirement, the wafer surface non-uniformities, dielectric erosion and Cu dishing, should be reduced. As the interconnect linewidth decreases and the number of layers increases, the control of dielectric erosion and Cu dishing has emerged as the greatest challenge in Cu CMP. Generally, dielectric erosion is more prevalent than Cu dishing in the dense sub-micron, copper-line region, whereas dishing is more significant than erosion at the global wiring level, Table 1.1. Third, defects after CMP should be minimized. In this thesis, however, defects during Cu CMP will not be considered.

The main concern in this thesis is to reduce dielectric erosion and Cu dishing across the wafer in Cu CMP. Past efforts to characterize the relationships between dielectric erosion and Cu dishing and process parameters have been primarily empirical. Several semi-theoretical models have been advanced, but such models essentially address the effect of one variable at a time, and are confined to only the feature-scale or at most die-scale. As the size of the wafer increases, however, the significance of non-uniformity, not only at the feature-scale but also at the die- and wafer-scales, on erosion and dishing increases. Therefore, it is timely to develop integrated erosion and dishing models to express wafer surface non-uniformities at wafer-, die- and feature-scales. Throughout this thesis, the following two conditions are considered both in the analysis and in the experiments.

Single-step polishing: Although it is common to use a multi-step polishing scheme in the industry [Dejule, 1998; Moinpour *et al.*, 2002], the single-step polishing scheme is analyzed to understand the physics of surface non-uniformity evolution more clearly and to suggest a possible approach to mitigate both erosion and dishing.

Bimaterial structure: We will assume the patterned wafer as two-material damascene structures, i.e., Cu and oxide. In Cu CMP, the barrier layer is usually very thin and can be treated either as Cu, if the material removal rate of barrier is the same as that of Cu, or as oxide if the material removal rate of barrier is close to that of oxide.

1.4 Thesis Organization

The overall goal of the thesis is to develop dielectric erosion and dishing models that integrate the wafer surface non-uniformities at the wafer-, die- and feature-scales, and to propose practical solutions to mitigate erosion and dishing.

Chapter 1 describes the background and the scope of thesis. In Chapter 2, feature-scale polishing models based on the local pressure calculation scheme with various contact mechanics at the pad/wafer interface and the step-height are introduced. The pad/wafer contact regimes include: smooth, discrete pad; smooth, continuous pad; uniformly rough, continuous pad; randomly rough, continuous pad. In Chapter 3, the non-uniformities at the wafer-, die- and feature-scales are identified, and integrated into the dielectric erosion and Cu dishing model. The effects of physical and geometric parameters on dishing and erosion are discussed based on the parameter sensitivity analysis. The developed Cu dishing and dielectric erosion models are validated by 100 mm patterned wafer polishing experiments in Chapter 4. Finally, conclusions of this thesis and suggestions for future research on Cu CMP are presented in Chapter 5.

References

Chang, C. Y. and Sze, S. M., 1996, **ULSI Technology**, McGraw-Hill, New York.

Dejule, R., 1998, "CMP Grows in Sophistication," *Semiconductor International*, 11, pp. 56-62.

Fu, G., Chandra, A., Guha, S. and Subhash, G., 2001, "A Plasticity-Based Model of Material Removal in Chemical-Mechanical Polishing (CMP)," *IEEE Trans. on Semiconductor Manufacturing*, vol. 14, pp. 406-417.

Huo, J., Solanki, R. and McAndrew, J., 2005, "A Novel Electroplanarization System for Replacement of CMP," *Electrochemical and Solid-State Letters*, vol. 8, pp. C33-C35.

International Technology Roadmap for Semiconductors (ITRS) – Interconnect, 2003, [Online] http://www.itrs.net.

Luo, J. and Dornfeld, D. A., 2003, "Material Removal Regions in Chemical Mechanical Planarization for Submicron Integrated Circuit Fabrication: Coupling Effects of Slurry Chemicals, Abrasive Size Distribution, and Wafer-Pad Contact Area," *IEEE Trans. on Semiconductor Manufacturing*, vol. 16, pp. 45-56.

Moinpour, M., Tregub, A., Oehler, A. and Cadien, K., 2002, "Advances in Chracterization of CMP Consumables," *MRS Bulletin*, vol. 27, pp. 766-771.

Moore, G. E., 1965, "Cramming More Components onto Integrated Circuits," *Electronics*, vol. 38, pp. 114-117.

Padhi, D., Yahalom, J., Gandikota, S. and Dixit, G., 2003, "Planarization of Copper Thin Film by Electrpolishing in Phosphoric Acid for ULSI Applications," *J. Electrochemical Soc.*, vol. 150, G10-G14.

Singh, R. K. and Bajaj, R., 2002, "Advances in Chemical-Mechanical Planarization," *MRS Bulletin*, vol. 27, pp. 743-751.

CHAPTER 2

FEATURE-SCALE STEP-HEIGHT MODELS

2.1 Introduction

To develop integrated erosion and dishing models, it is necessary to characterize the feature-scale polishing behavior first, since both in CMP are due to differential material removal rates. The material removal rate (MRR) is in general proportional to the pressure and relative velocity. The pressure distribution on the polishing surface is determined by the pad/wafer contact mechanics at a given step-height. Over the past decade, there have been many efforts to describe the evolution of polishing surface based on various contact mechanics models as shown in Fig. 2.1. Previous works may be categorized as:

Smooth pad model: Contact between the wafer and the pad is analyzed based on the assumption that the polishing pad is elastic and smooth [Chekina *et al.*, 1998; Lai *et al.*, 2002]. This approach assumes that the elastic deformation of the smooth, monolithic pad itself as Cu dishing.

Discrete pad model: This approach assumes that the pad deforms as discrete blocks or a series of linear springs. These springs may or may not be connected each other [Runnels *et al.*, 1994, 1999, 2003; Grillaert *et al.*, 1998; Elbel *et al.*, 1998; Chen and Lee, 1999; Yang, 2000; Fu and Chandra, 2003; Guo *et al.*, 2004; Noh *et al*, 2004]. These models account for dishing fairly well, but rely on the physically inadmissible discontinuous deformation of the pad.

Rough pad model: Several contact models have also been proposed based on the contact of random rough surfaces [Yu *et al.*, 1993; Vlassak, 2001, 2004; Borucki, 2002; Nguyen *et al.*, 2003; Seok *et al.*, 2003]. The rough pad/wafer contact is modeled by adopting the classical analysis by Greenwood and Williamson [Greenwood and Williamson, 1966]. In these models, the asperity height distribution is assumed to be either Gaussian or exponential.

In all these models, the general procedure is to calculate the step-height as a function of contact geometry, average pressure, relative velocity, material properties and surface topographies of the pad, and polishing time. Section 2.2 describes step-heihgt evolution. First,

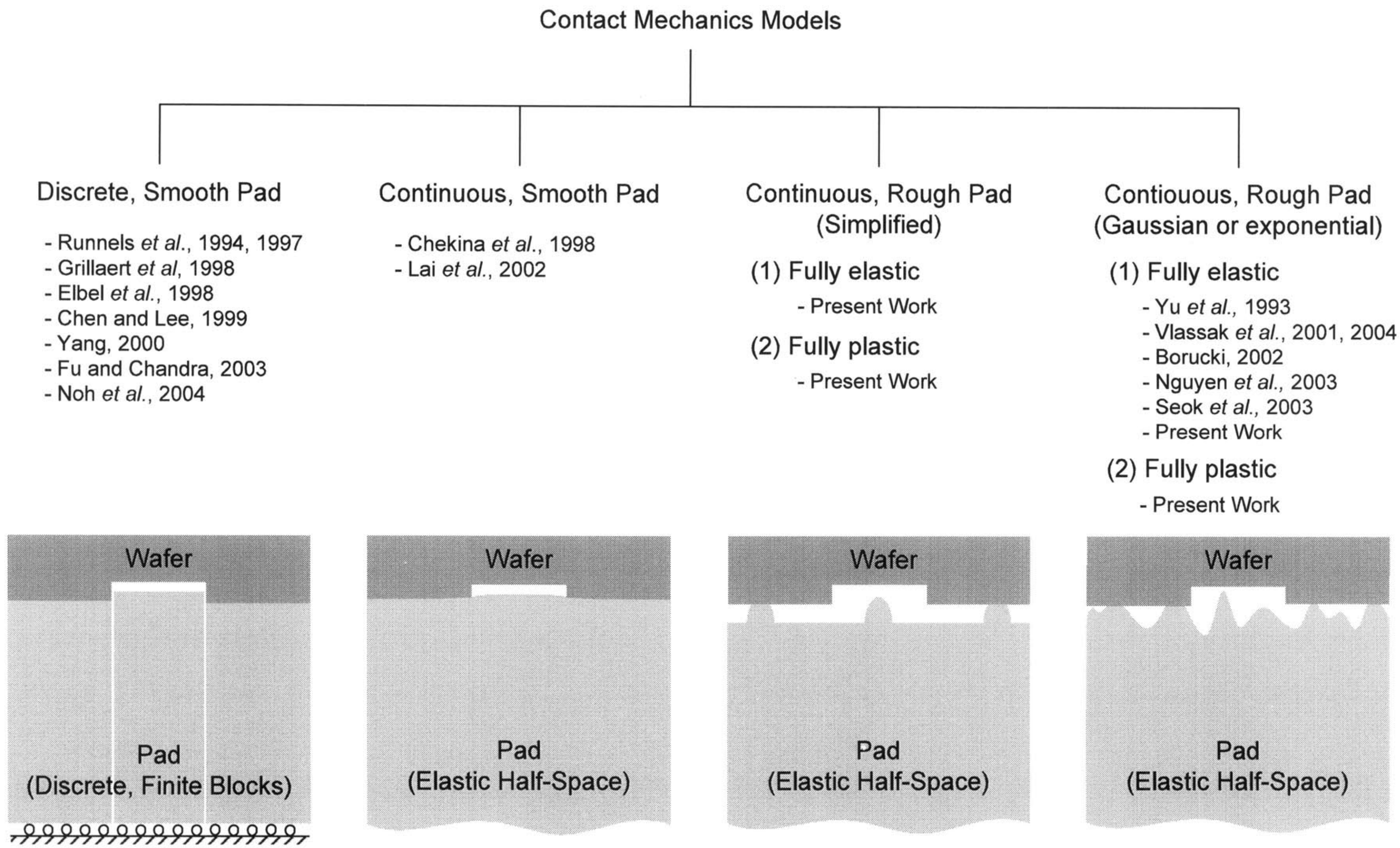


Figure 2.1 Classification of contact mechanics models for patterned wafers in CMP.

the contact pressure distribution at a given step height is related to the elastic deformation of the pad surface in the smooth pad assumption, or to the elastic/plastic deformation of asperities in the rough pad analysis. Then, the evolution of step-height in a patterned wafer is calculated by combining the effects of pad profile and the geometry of Cu interconnects at each polishing stage. Based on the developed step-height model, dielectric erosion and Cu dishing in Cu CMP can be expressed as a function of geometric, material and process parameters. Such parameters include: wafer-scale uniformity factor, pre-CMP wafer surface topography, radius of curvature and spacing of pad asperities, Young's modulus of the pad, yield strength of the pad, the nominal pressure, slurry selectivity, and so on.

Subsequent sections then apply the general step-height model calculation procedure of section 2.2 to various pad/wafer contact conditions. Section 2.3 presents a smooth pad derivation. In this section, the pad deformation based on the contact mechanics model by Lai *et al.* is revisited and extended to the case when interconnect is filled with Cu to a certain step-height. Section 2.4 adapts the previous one-dimensional discrete pad block models and investigates the step-height evolution. Section 2.5 presents a simplified rough pad model. The mean asperity contact radius and the asperity spacing are estimated by the analysis based on the general random rough surface of the pad in section 2.6. Section 2.6 follows the Greenwood-Williamson approach like previous rough pad models. In this section, however, a new approach to extend a rough pad/blanket wafer contact into a patterned wafer analysis is presented. Both elastic and fully plastic deformations of pad asperities are considered and the effect on the evolution of the step-height during polishing is compared. Sample calculations of each model are presented at the end of each section, based on material properties and surface topographies of a commercial polishing pad.

2.2 The General Procedure

Figure 2.2 shows the general procedure for tracking step-height, which is similar to that presented by several previous researchers [Grillaert *et al.*, 1998; Park *et al.*, 2000]. Nevertheless, we formulate the model and define the terminology to enable comparison of the smooth, discrete and rough pad modes in sections 2.3 through 2.6.

To characterize the evolution of the step-height, it is necessary to have an expression for a

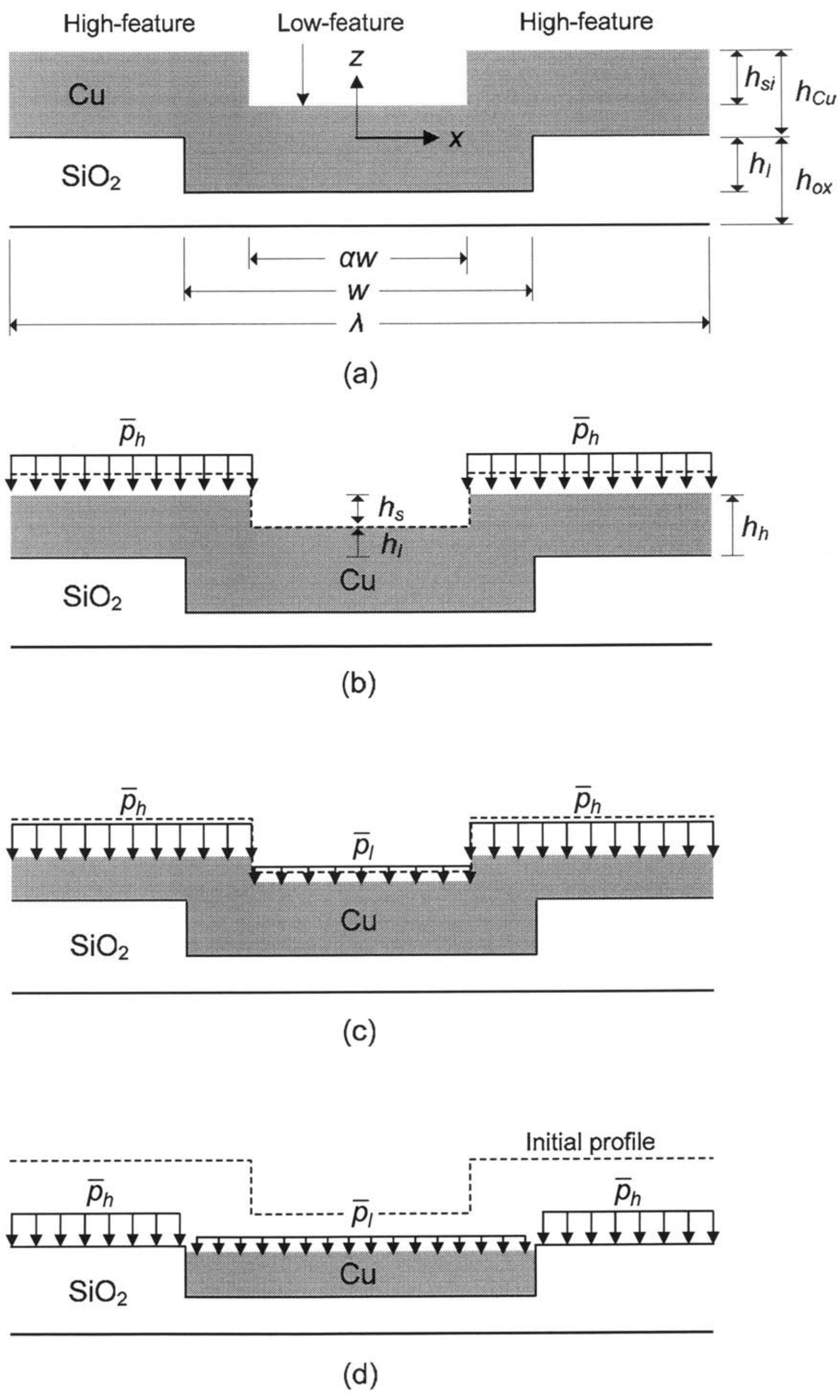


Figure 2.2 Schematic of feature-scale step-height calculation: (a) Definition of high and low features, (b) Stage 1, (c) Stage 2 and (d) Stage 3.

material removal rate at a give feature first. The local material removal rate in CMP is expressed by the Preston equation [Preston, 1926]:

$$\left|\frac{dh}{dt}\right| = k_p \cdot p \cdot v_R \tag{2.1}$$

where h is the thickness of the layer removed, t the polishing time, p the pressure, v_R the relative velocity, and k_p the Preston constant. Although the Preston equation represents the local material removal rate at any point on the wafer, it does not explain the actual material removal mechanism. Nevertheless, several researchers have experimentally demonstrated that the above functional relationship is generally valid in CMP at many scales [Steigerwald *et al.*, 1994; Stavreva *et al.*, 1995 and 1997; Lai, 2000]. The Preston constant, obviously, is not a fundamental constant. It depends on the pad/wafer contact condition, slurry concentration and chemistry, abrasive size and shape, pad stiffness and surface topography, and so on [Liu *et al.*, 1996; Saka *et al.*, 2000; Ahmadi and Xia, 2001; Fu *et al.*, 2001; Luo and Dornfeld, 2003; Seok *et al.*, 2004]. Thus, any variation in these quantities at any scale is expected to result in non-uniformity in material removal rate at that scale. In this thesis, the wafer-scale material removal rate, MRR, is defined by the average applied pressure, p_{av}, and the relative velocity, v_R, in blanket wafer polishing.

$$MRR \equiv k_p \cdot p_{av} \cdot v_R \tag{2.2}$$

where p_{av} is the average pressure defined as a ratio of applied load, F, to the projected area of the wafer, A_w.

A complex aspect of Cu CMP is that there are at least three different materials – Cu, dielectric and barrier layer – to be polished, sequentially or simultaneously. Therefore, the ratios of material removal rates, or selectivities, are important in characterizing polishing non-uniformity. The selectivities of Cu, oxide and barrier layer are obtained by blanket wafer polishing under the same process or experimental conditions as the patterned wafer polishing. Since the definition of selectivity is based on blanket wafer experiments, with the same nominal pressure and relative velocity, it is the ratio of Preston constants. Thus, the selectivities in Cu CMP are defined as:

$$S_{Cu/ox} \equiv \frac{MRR_{Cu}}{MRR_{ox}} = \frac{k_{p_{Cu}}}{k_{p_{ox}}}$$
$$S_{b/ox} \equiv \frac{MRR_b}{MRR_{ox}} = \frac{k_{p_b}}{k_{p_{ox}}} \tag{2.3}$$

where the subscripts *Cu, b, ox*, respectively, represent copper, barrier and oxide. As described in Chapter 1, throughout this paper, we consider the Cu interconnect structure as two material structure, Cu and oxide. The selectivity depends both on the hardness of the material polished and the chemistry of the slurry [Moinpour *et al.*, 2002; Lai *et al.*, 2002; Jindal and Babu, 2004]. For instance, hydrogen peroxide, a common additive in commercial Cu slurries, reacts with Cu and forms a "soft" layer so that the material removal rate of Cu increases and thus $S_{Cu/ox}$ too increases.

The geometry of interconnects is expressed by Cu linewidth, w, pitch, λ, and interconnect thickness, h_I. Cu is deposited by the amount of the thickness, h_{Cu}, and the pattern geometry of the initial surface profile is represented by the "surface linewidth", αw, and the initial step-height, h_{si}, as shown in Fig. 2.2(a). In the present step-height, erosion and dishing models, the focus is not on the evolution of the complete profile of a feature but on the evolution of the maximum value of the step-height, erosion and dishing. Therefore, the center of two interconnect lines, $x = \pm\lambda/2$, and the center of the individual Cu interconnect line itself, $x = 0$, are designated as the high and low features locations, respectively as shown in Fig. 2.2(a). Therefore,

$$\begin{aligned} h_h(t) &= h(\lambda/2,t) = h(-\lambda/2,t) \\ h_l(t) &= h(0,t) \end{aligned} \tag{2.4}$$

where h_h and h_l, respectively, are the polishing surface heights of the high and low features relative to the top of the oxide.

The step-height, $h_s(t)$, is defined as the height difference between the high and low features at any given time t, as shown in Fig. 2.2(b).

$$h_s(t) \equiv h_h(t) - h_l(t) \tag{2.5}$$

The material removal rates at the high and low features, and thus step-height, can be expressed by the Preston equation as:

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_h}\overline{p}_h(t)v_R \\ \frac{dh_l}{dt} &= -k_{p_l}\overline{p}_l(t)v_R\end{aligned} \tag{2.6}$$

where k_{p_h} and k_{p_h} are the Preston constants, and $\overline{p}_h$ and $\overline{p}_l$ the mean pressures at the high and low features, respectively. The evolution of the step-height can be calculated as:

$$\frac{dh_s}{dt} = \frac{dh_h}{dt} - \frac{dh_l}{dt} \tag{2.7}$$

The material removal rate at any instant is based on the local pressure distribution, which varies as the polishing surface profile changes and can be solved by assuming appropriate pad/wafer contact conditions. Although the real contact pressure may vary along the pitch, λ, and there might be rounding of the edges, it is not considered since the focus of this study is to characterize the maximum dishing and erosion. Therefore, it is assumed that the high and low features remain horizontal during polishing and the mean pressures at the high and low features are defined to calculate the corresponding material removal rates. Both the mean pressure at the high and low features, $\overline{p}_h$ and $\overline{p}_l$, generally vary with time as shown in Fig. 2.2(b) – (d), depending on the pad/ wafer contact conditions and the feature geometries.

To calculate $\overline{p}_h$ and $\overline{p}_l$, two relationships are invoked. One is the force equilibrium equation and the other is compatibility, or the relationship between the pressure and the step-height. The force equilibrium at any stage can be represented as:

$$p_{av} = \overline{p}_h(1 - w/\lambda) + \overline{p}_l(w/\lambda) \tag{2.8}$$

where p_{av} is the average pressure, w the linewidth and λ the pitch.

The relationship between the pressure and the step-height can be solved by expressing the deformation of the pad (asperity) at the high and the low feature, δ_h and δ_l, as a function of the mean pressures, $\overline{p}_h$ and $\overline{p}_l$. Thus,

$$\begin{aligned}\delta_h &= \delta_h(\overline{p}_h, \overline{p}_l) \\ \delta_l &= \delta_l(\overline{p}_h, \overline{p}_l)\end{aligned} \tag{2.9}$$

The functions, δ_h and δ_l, represent the relationship between the deformation of the pad (asperity) and the pressure, and depend on the pad/wafer contact mechanics. Furthermore, δ_h

and δ_l include parameters such as Young's modulus, yield strength, initial thickness, asperity geometry and spacing of the pad. The step-height at any given polishing time can be represented either as the difference between δ_h and δ_l, if both the high and low features are in contact, or as an independent values. From Eqs. (2.8) and (2.9), the mean pressure and thus the material removal rates at the high and the low features can be calculated.

The evolution of pad/wafer contact surface and the step-height are represented in three stages as shown in Fig. 2.2(b) – (d) and Fig. 2.3, respectively.

Stage 1: Initially, pad contacts the high feature only if the relative deformation of the pad between the high and low feature under the given load is smaller than the initial step-height, $\delta < h_{si}$. The material being polished at the high and low featuress is Cu. The end of Stage 1 is designated by t_1: $0 \leq t \leq t_1$.

Stage 2: The pad starts contacting both high and low features. The material being polished at the high and low features is Cu. The end of Stage 2 is designated by t_2: $t_1 \leq t \leq t_2$.

Stage 3: The pad contacts both the high and low features. The materials being polishing at the high and low features, however, are different: oxide at the high feature and Cu at the low feature. The end of Stage 3 is designated as the process endpoint t_{ep}: $t_2 \leq t \leq t_{ep}$.

2.2.1 Stage 1

Initially, i.e., without any load, the pad contacts only the high feature. When load is applied, the pad may or may not contact the low feature. If the load is sufficiently high to deform the pad at the high feature more than the initial step-height, then the pad will touch the low feature. If the load is low, by contrast, the deformation of the pad at the high feature is less than the initial step-height and thus the load will be supported by the high feature only. In Stage 1, the latter case is assumed to apply.

The reference line for the heights of high and low features is the top of the oxide layer. Therefore, initial conditions for $h_h(t)$, $h_l(t)$ and $h_s(t)$ are expressed as:

$$\begin{aligned} h_h(0) &= h_{Cu} \\ h_l(0) &= h_{Cu} - h_{si} \end{aligned} \tag{2.10}$$

$$h_s(0) = h_{si} \tag{2.11}$$

To calculate the mean pressure at the high and the low features, force equilibrium in Stage 1 is

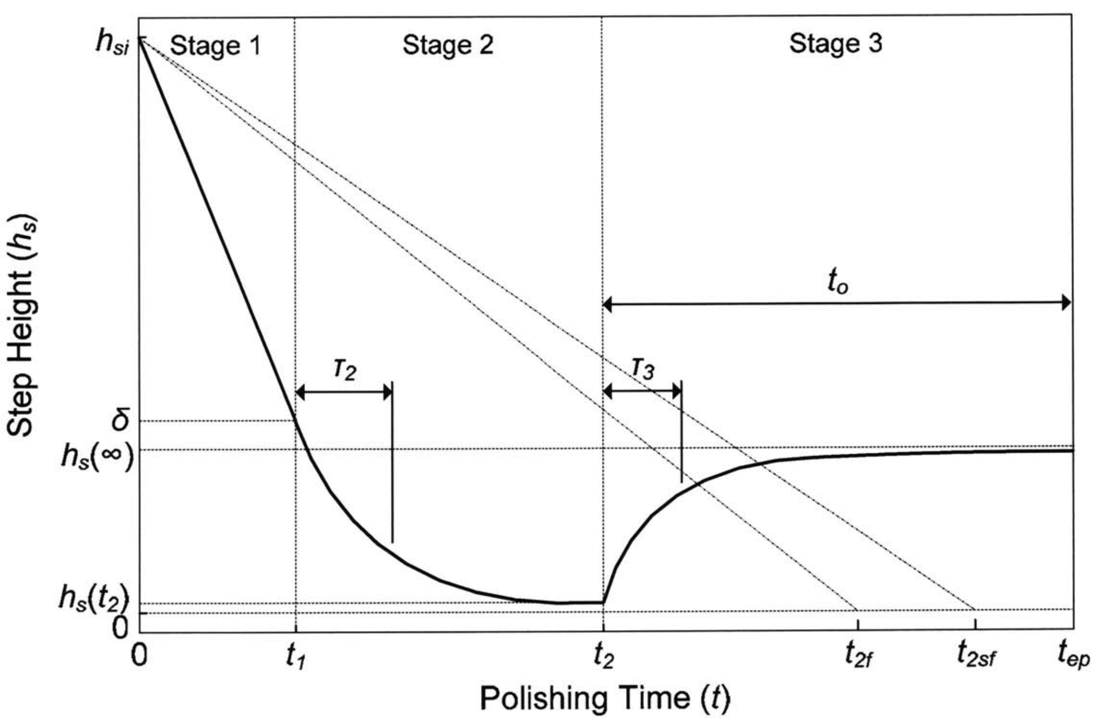


Figure 2.3 Schematic of step-height evolution.

expressed as:

$$p_{av} = \overline{p}_h (1 - \alpha w / \lambda) + \overline{p}_l (\alpha w / \lambda) \tag{2.12}$$

Since only the high feature contacts the pad, the deformation of the pad (asperity) and the step-height do not affect the mean contact pressure. Thus, the mean contact pressure can be easily solved from Eq. (2.12) as:

$$\begin{aligned} \overline{p}_h &= p_{av} \left(\frac{1}{1 - \alpha w / \lambda} \right) \\ \overline{p}_l &= 0 \end{aligned} \tag{2.13}$$

If pressures at the high and low features are constant in Stage 1, material removal rates also remains constant. Additionally, the material being polished at the high feature is Cu: $k_{p_h} = k_{p_{Cu}}$.

Thus, material removal rates at the high and low features, and the step-height are:

$$\begin{aligned} \frac{dh_h}{dt} &= -k_{p_{Cu}} p_{av} v_R \left(\frac{1}{1 - \alpha w / \lambda} \right) \\ \frac{dh_l}{dt} &= 0 \end{aligned} \tag{2.14}$$

$$\frac{dh_s}{dt} = -k_{p_{Cu}} p_{av} v_R \left(\frac{1}{1 - \alpha w / \lambda} \right) \tag{2.15}$$

Combining with the initial conditions, the height at the high and low features and the step-height can be solved as:

$$\begin{aligned} h_h(t) &= h_{Cu} - k_{p_{Cu}} p_{av} v_R \left(\frac{1}{1 - \alpha w / \lambda} \right) t \\ h_l(t) &= h_{Cu} - h_{si} \end{aligned} \tag{2.16}$$

$$h_s(t) = h_{si} - k_{p_{Cu}} p_{av} v_R \left(\frac{1}{1 - \alpha w / \lambda} \right) t \tag{2.17}$$

The end of Stage 1, $t = t_1$, is determined as the moment when the low feature, too, starts supporting load. If the pad were perfectly smooth and rigid, it would contact the low feature only when the height at the high feature decreases by the amount of the initial step-height. A compliant pad, however, contacts the low feature even before the high feature reaches the level of the low feature due to pad (asperity) deformation, δ, under the given load. Thus

$$\delta \equiv \delta_h(\overline{p}_h, 0) - \delta_l(\overline{p}_h, 0) \tag{2.18}$$

The relative pad deformation, δ, depends on the pad/wafer contact assumption and the pattern geometry. If the calculated δ is smaller than the initial step-height h_{si}, the assumption to begin with Stage 1 is valid. But, if $\delta > h_{si}$, the step-height model needs to begin with Stage 2. Thus, the requirement to start with Stage 1 is:

$$\delta < h_{si} \tag{2.19}$$

At the end of Stage 1, $t = t_1$, h_s reaches the pad deformation, δ, and the low feature, too, starts supporting the normal load. Thus the final heights of the high and low features, and the step-height at the end of Stage 1 are:

$$\begin{aligned} h_h(t_1) &= h_{Cu} - h_{si} + \delta \\ h_l(t_1) &= h_{Cu} - h_{si} \end{aligned} \tag{2.20}$$

$$h_s(t_1) = \delta \tag{2.21}$$

Once the pad deformation is calculated from the contact mechanics, t_1 can be expressed as:

$$t_1 = \frac{(h_{si} - \delta)(1 - \alpha w / \lambda)}{k_{p_{Cu}} p_{av} v_R} \tag{2.22}$$

2.2.2 Stage 2

As the pad contacts low features, they too get polished. This stage is designated as Stage 2. At the beginning of Stage 2, $t = t_1$, the initial step-height is δ and the initial heights of the high and the low features are given by Eqs. (2.20) and (2.21). The mean pressures $\overline{p}_h$ and $\overline{p}_l$ can be solved by Eq. (2.12) and the relations between the pressure and the step-height depending on the pad/wafer contact mechanics, δ_h and δ_l. In Stage 2, both the high and the low feature are in contact, and thus the step-height can be rewritten as:

$$h_s = \delta_h - \delta_l \tag{2.23}$$

In this stage, the material being polished at the high and low features is Cu, and thus $k_{p_h} = k_{p_l} = k_{p_{Cu}}$. Therefore, the material removal rates at the high and low features are:

$$\begin{aligned} \frac{dh_h}{dt} &= -k_{p_{Cu}} \overline{p}_h(t) v_R \\ \frac{dh_l}{dt} &= -k_{p_{Cu}} \overline{p}_l(t) v_R \end{aligned} \tag{2.24}$$

The evolution of the step-height can be calculated from the difference between the heights of the high and the low feature as in Eq. (2.7). During Stage 2, material removal rates at high and low feature approach to that of blanket wafer polishing as polishing progresses and the step-height decreases. The rate of step-height reduction depends on the pad/wafer contact mechanics. If the step-height decreases exponentially, for example, the time-constant in Stage 2, τ_2, is the index of how fast the step-height approaches to zero.

The end of Stage 2 is marked by the polishing time t_2, when the pad surface at high feature reaches the oxide surface. Thus,

$$h_h(t_2) = 0 \tag{2.25}$$

The step-height at the end of Stage 2, $h_s(t_2)$, represents the minimum step-height, if $S_{Cu/ox} > 1$, and thus t_2 is the ideal process endpoint at the feature as shown in Fig. 2.3. However, due to the initial surface geometry variation in a die, the field region in the same die takes more time to reach Stage 3, which is designated as t_{2f}. Additionally, if the die is not the slowest die in a wafer, the time to reach Stage 3 at the field region in the slowest die, t_{2sf}, is greater than t_{2f}.

2.2.3 Stage 3

To calculate the mean pressure $\overline{p}_h$ and $\overline{p}_l$, the force equilibrium is considered first. Since the linewidth is changed to the designed Cu linewidth, w, instead of αw, there are three cases to consider: the high feature with oxide, high feature with Cu and low feature with Cu. In most cases, $h_s(t_2)$ is relatively small compared with the interconnect thickness and thus the high feature with Cu can be assumed to be the low feature with Cu. Therefore, in Stage 3, oxide region is considered as high feature and Cu region as the low feature. Thus

$$p_{av} = \overline{p}_h (1 - w/\lambda) + \overline{p}_l (w/\lambda) \tag{2.26}$$

Second, the relationship between the pressure and the step-height can be obtained from the pad/wafer contact mechanics used in Stage 2.

In Stage 3, the materials being polished at the high and low features are oxide and Cu, respectively: $k_{p_h} = k_{p_{ox}}$ and $k_{p_l} = k_{p_{Cu}}$. Therefore, the material removal rates at the high and low features can be calculated as:

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_{ox}}\overline{p}_h v_R \\ \frac{dh_l}{dt} &= -k_{p_{Cu}}\overline{p}_l v_R\end{aligned} \tag{2.27}$$

Again, the evolution of the step-height is calculated from the difference between the heights at the high and the low features.

Now, the normal load is supported by both high and low features, and the pressure and material removal rates at the high and low features depend on the step-height and slurry selectivity, $S_{Cu/ox}$. If $S_{Cu/ox}= 1$, the step-height decreases, and the pressure and the material removal rate at the high and the low feature approach the same values as those of field region. In most of conventional polishing practices, however, $S_{Cu/ox}> 1$. In the beginning of Stage 3, the material removal rate of Cu is greater than that of oxide, thus the step-height increases. As the step-height increases, the pressure and the material removal rate at the high feature increase and those at the low feature decrease. If the step-height increases in a form of $1-\exp(-t/\tau)$, the time-constant in Stage 3, τ_3, is the index of how fast the step-height increases. If the polishing time is long enough or $(t_{ep}-t_2) \gg \tau_3$, the step-height approaches an asymptotic value, $h_s(\infty)$. As $S_{Cu/ox}$ increases, $h_s(\infty)$ approaches the maximum relative pad deformation, δ_o, which is the pad deformation when the Cu interconnect area is empty, or recessed enough, so that it does not support load.

As the polishing time increases, material removal rates at both high and low features approach the same value. Thus, the asymptotic material removal rate of the high and low features is calculated by equating material removal rates of high and low features.

$$MRR_{\infty} \equiv k_{p_{Cu}} p_{av} v_R \left[\frac{1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right] \tag{2.28}$$

Although it is desirable to set the process endpoint as t_2 at each feature, in the conventional face-down CMP setup, it is impossible to set different endpoints at any two points on a wafer. Thus, there is only one process endpoint across the entire wafer, which is designated as t_{ep}.

2.3 The Smooth Pad Model

In this section, we apply the procedure described in section 2.2, for a pad that is assumed homogeneous, monolithic, elastic and semi-infinite as shown in Fig. 2.4. The smooth, continuous pad models have been developed based on the contact mechanics between the pad and the pattern on a feature. First, the characteristics of the steady state regime are analyzed on the basis of approaches developed in contact mechanics [Chekina *et al.*, 1998]. The polishing pad is assumed to be a perfectly smooth elastic half-space and the displacement of the pad is expressed as a function of local pressure. The effect of pattern geometry is investigated by considering one dimensional periodic feature with a steady-state material removal rate assumption. Another approach was to calculate the maximum pad deformation based on the contact mechanics between the perfectly smooth pad and the pattern trench [Lai *et al.*, 2002]. In this model, the pad deformation was much smaller than the amount of dishing from experiments.

The pad deformation based on the smooth pad assumption does not explain the large pad asperity deformation in the large feature size. Moreover, most commercial polishing pads are rough and there are many pad asperity/wafer contacts. Nonetheless, if the size of the contact between pad and wafer is smaller than the width of Cu interconnects, the smooth pad assumption is valid. As Cu linewidth decreases, it become more important to explain non-uniformities based on the smooth pad condition, for instance, dielectric erosion in the submicron device level.

The key assumptions of the contact model are:

- The pad is an isotropic, elastic, semi-infinite medium.
- The pad surface is perfectly smooth.
- The wafer surface always remains horizontal.
- The deformation is plane-strain.
- The pad surface outside contact region is stress-free: $\sigma_z = \tau_{xz} = 0$.
- The friction coefficient in the contact region is small: $\mu \sim 0.1$.

2.3.1 Theory

The elastic deformation of the homogeneous, monolithic smooth pad, δ, for a periodic structure in the elastic half space with plain strain condition has been solved [Lai *et al.*, 2002]. In this section, that model is employed with the uniform pressure boundary condition as shown

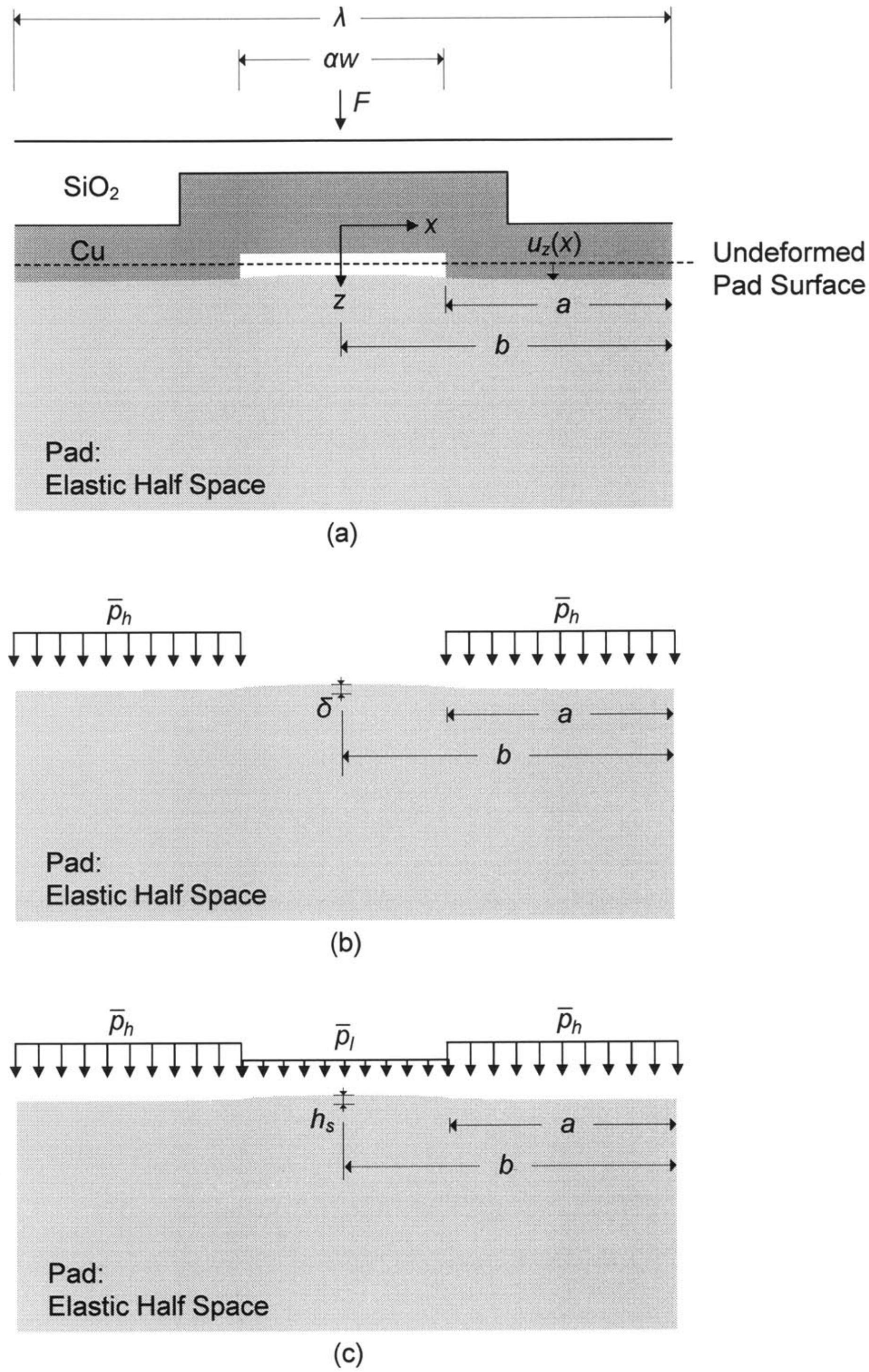


Figure 2.4 Schematics of the pattern/smooth-pad contact interface: (a) initial stage with uniform pressure distribution specified on the high feature and pressures on the high and low features when (b) $h_s > \delta$ and (c) $h_s \leq \delta$. The wafer surface to be polished is facing down.

in Fig. 2.4. The displacement of the pad surface, $\overline{u}_z(x)$, in Stage 1 can be solved as [Johnson, 1985, pp. 11-44]:

$$\begin{aligned}\overline{u}_z(x) = -\frac{\overline{p}_h}{E^*}(2a)\frac{1}{\pi}\Bigg\{&\left[\left(1+\frac{x-b}{a}\right)\ln\left(1+\frac{x-b}{a}\right)+\left(1-\frac{x-b}{a}\right)\ln\left(1-\frac{x-b}{a}\right)\right]\\ &+\left[\left(1+\frac{x+b}{a}\right)\ln\left(1+\frac{x+b}{a}\right)+\left(1-\frac{x+b}{a}\right)\ln\left(1-\frac{x+b}{a}\right)\right]\Bigg\}+C_1\end{aligned} \tag{2.29}$$

where $a=(\lambda-\alpha w)/2$, $b=\lambda/2$, $E^*=E_p/(1-\nu_p^2)$ which represents the elastic modulus in the plane strain problem, and C_1 is a constant determined by any arbitrary reference point.

The model by Lai *et al.* predicts the maximum pad deformation when the Cu line is load-free, and thus can be used in Stage 1. To explain the evolution of step-height under the pattern filled with Cu, the model by Lai *et al.* is extended here with uniform pressure boundary conditions at both high and low features. In this case, the pad displacement can be expressed as:

$$\begin{aligned}\overline{u}_z'(x) = -\frac{\overline{p}_h}{E^*}(2a)\frac{1}{\pi}\Bigg\{&\left[\left(1+\frac{x-b}{a}\right)\ln\left(1+\frac{x-b}{a}\right)+\left(1-\frac{x-b}{a}\right)\ln\left(1-\frac{x-b}{a}\right)\right]\\ &+\left[\left(1+\frac{x+b}{a}\right)\ln\left(1+\frac{x+b}{a}\right)+\left(1-\frac{x+b}{a}\right)\ln\left(1-\frac{x+b}{a}\right)\right]\Bigg\}\\ -\frac{\overline{p}_l}{E^*}2(b-a)\frac{1}{\pi}\Bigg\{&\left[\left(1+\frac{x}{b-a}\right)\ln\left(1+\frac{x}{b-a}\right)+\left(1-\frac{x}{b-a}\right)\ln\left(1-\frac{x}{b-a}\right)\right]+C_2\end{aligned} \tag{2.30}$$

While C_2 varies with the position of datum, since we are interested in relative displacement of the pad, it can be any value. Thus, C_2 is set as zero.

Accordingly, the deformation of the pad at the high and the low features can be expressed as:

$$\begin{aligned}\delta_h &= \delta_h(\overline{p}_h,\overline{p}_l) = \overline{u}_z'(x=\lambda/2)\\ \delta_l &= \delta_l(\overline{p}_h,\overline{p}_l) = \overline{u}_z'(x=0)\end{aligned} \tag{2.31}$$

Thus,

$$\begin{aligned}\delta_h(\overline{p}_h,\overline{p}_l) &= A_{hh}w\left(\frac{\overline{p}_h}{E^*}\right)+A_{hl}w\left(\frac{\overline{p}_l}{E^*}\right)\\ \delta_l(\overline{p}_h,\overline{p}_l) &= A_{lh}w\left(\frac{\overline{p}_h}{E^*}\right)+A_{ll}w\left(\frac{\overline{p}_l}{E^*}\right)\end{aligned} \tag{2.32}$$

where

$$
\begin{aligned}
A_{hh} &= \frac{1}{\pi}\left(\frac{1+w/\lambda}{w/\lambda}\right)\left\{\left[\left(\frac{3-w/\lambda}{1-w/\lambda}\right)\ln\left(\frac{3-w/\lambda}{1-w/\lambda}\right)-\left(\frac{1+w/\lambda}{1-w/\lambda}\right)\ln\left(\frac{1+w/\lambda}{1-w/\lambda}\right)\right]\right. \\
A_{hl} &= 0 \\
A_{lh} &= \frac{1}{\pi}\left\{\left[\left(\frac{3-w/\lambda}{1-w/\lambda}\right)\ln\left(\frac{3-w/\lambda}{1-w/\lambda}\right)-\left(\frac{1+w/\lambda}{1-w/\lambda}\right)\ln\left(\frac{1+w/\lambda}{1-w/\lambda}\right)\right]\right. \\
A_{ll} &= \frac{1}{\pi}\left\{\left[\left(\frac{1-w/\lambda}{w/\lambda}\right)\ln\left(\frac{1-w/\lambda}{w/\lambda}\right)-\left(\frac{1+w/\lambda}{w/\lambda}\right)\ln\left(\frac{1+w/\lambda}{w/\lambda}\right)\right]\right.
\end{aligned}
\tag{2.33}
$$

Stage 1

The initial relative pad deformation, δ, can be represented as the relative displacement of pad at the high and low features as in Eq. (2.18). By combining Eqs. (2.31) – (2.33) with the condition $\overline{p}_l = 0$ and the surface linewidth αw instead of w, δ can be rewritten as:

$$
\delta = A_1\alpha\left(\frac{\overline{p}_h}{E^*}\right)w \tag{2.34}
$$

where A_1 is a dimensionless coefficient determined by replacing w by αw in Eq. (2.33). Thus,

$$
A_1 = \frac{1}{\pi(\alpha w/\lambda)}\left\{\left[\left(\frac{3-\alpha w/\lambda}{1-\alpha w/\lambda}\right)\ln\left(\frac{3-\alpha w/\lambda}{1-\alpha w/\lambda}\right)-\left(\frac{1+\alpha w/\lambda}{1-\alpha w/\lambda}\right)\ln\left(\frac{1+\alpha w/\lambda}{1-\alpha w/\lambda}\right)\right]\right. \tag{2.35}
$$

By combining with force equilibrium, δ can be rewritten as:

$$
\delta = \frac{A_1\alpha}{(1-\alpha w/\lambda)}\left(\frac{p_{av}}{E^*}\right)w \tag{2.36}
$$

The relative pad deformation, δ, is much smaller than the Cu interconnect thickness, h_I, under most of polishing conditions and pad materials: $p_{av}/E^* \sim 10^{-4}$. In this case, it is reasonable to assume the pad as a flat surface during polishing, especially in the submicron feature level. If the local stiffness value is much lower than the bulk value of the pad, however, the flat pad surface assumption is not valid anymore, and the pad deformation should be into consideration. Nonetheless, in this section, we accept the bulk Young's modulus of the pad and thus, the flat pad assumption is used.

Based on this assumption, the end of Stage 1, t_1, in Eq. (2.22) is expressed as time at which the pad just touches the low feature and starts supporting the normal load. Thus,

$$
t_1 = \frac{h_{si}(1-\alpha w/\lambda)}{k_{p_{Cu}} p_{av} v_R} \tag{2.37}
$$

At the end of Stage 1, $t = t_1$, the polishing surface at high and low feature become the same and thus the step-height becomes zero.

$$h_h(t_1) = h_l(t_1) = h_{Cu} - h_{si} \tag{2.38}$$

$$h_s(t_1) = 0 \tag{2.39}$$

Stage 2

Since the pad is assumed to be flat during polishing in Stage 1, the pressure and material removal rates in this stage are the same as those of the field region, and we do not need to consider the deformation of the pad based on the uniform pressure boundary condition. Thus,

$$\overline{p}_h = \overline{p}_l = p_{av} \tag{2.40}$$

$$\frac{dh_h}{dt} = \frac{dh_l}{dt} = -k_{p_{Cu}} p_{av} v_R \tag{2.41}$$

Therefore, heights at the high and the low features decrease linearly and the step-height remains zero during Stage 2.

$$h_h(t) = h_l(t) = h_{Cu} - h_{si} - k_{p_{Cu}} p_{av} v_R (t - t_1) \tag{2.42}$$

$$h_s(t) = 0 \tag{2.43}$$

Stage 2 ends when the pad reaches the top of the oxide at the high feature. Thus the end of Stage 2, t_2, can be expressed as:

$$t_2 = \frac{h_{Cu} - (\alpha w / \lambda) h_{si}}{k_{p_{Cu}} p_{av} v_R} \tag{2.44}$$

At the end of this stage, $t = t_2$, the polishing surface of both high and low features reaches the top of the oxide, while the step-height remains zero.

$$h_h(t_2) = h_l(t_2) = 0 \tag{2.45}$$

$$h_s(t_2) = 0 \tag{2.46}$$

Stage 3

As the pad surface reaches the top of oxide, the materials being polished at the high and low features are oxide and Cu, respectively. Since the pad is in contact with both high and low features at this moment, and the materials in general have different polishing rates, a step is created by further polishing. Even though the maximum step-height is expected to be very small since the maximum pad deformation is small, this analysis is useful for two reasons. First, by calculating time constant of step-height change, it can verify later that the "steady-state" assumption in Stage 3 is valid in the smooth pad erosion and dishing model. Second, this analysis will be useful to calculate Cu dishing at the submicron features if Young's modulus of a pad is much smaller than that of current polishing pads or if we know the local value of Young's modulus instead of the bulk material property.

The step height can be solved by the pad surface height difference between the high and the low features as:

$$h_s = (A_{hh} - A_{hl})\,w\left(\frac{\overline{p}_h}{E^*}\right) + (A_{lh} - A_{ll})\,w\left(\frac{\overline{p}_l}{E^*}\right) \tag{2.47}$$

where h_s, $\overline{p}_h$ and $\overline{p}_l$ are time-dependent variables. The pressure at the high and low features in Stage 3 can be expressed as the step-height by solving the following linear equations:

$$\begin{bmatrix} 1 - w/\lambda & w/\lambda \\ A_{hh} - A_{hl} & A_{lh} - A_{ll} \end{bmatrix} \begin{pmatrix} \overline{p}_h \\ \overline{p}_l \end{pmatrix} = \begin{pmatrix} p_{av} \\ E^* h_s / w \end{pmatrix} \tag{2.48}$$

We define the inverse matrix of the pressure coefficient matrix as:

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \equiv \begin{bmatrix} 1 - w/\lambda & w/\lambda \\ A_{hh} - A_{hl} & A_{lh} - A_{ll} \end{bmatrix}^{-1} \tag{2.49}$$

where B_{11}, B_{12}, B_{21} and B_{22} are also dimensionless values determined by the area fraction of Cu interconnects, w/λ. Therefore, pressures and material removal rates at the high and low features can be expressed as:

$$\begin{aligned} \overline{p}_h &= p_{av}\left[B_{11} + B_{12}\left(\frac{E^*}{p_{av}}\right)\frac{h_s}{w}\right] \\ \overline{p}_l &= p_{av}\left[B_{21} + B_{22}\left(\frac{E^*}{p_{av}}\right)\frac{h_s}{w}\right] \end{aligned} \tag{2.50}$$

$$\begin{aligned} \frac{dh_h}{dt} &= -k_{p_{ox}} p_{av} v_R \left[B_{11} + B_{12}\left(\frac{E^*}{p_{av}}\right)\frac{h_s}{w}\right] \\ \frac{dh_l}{dt} &= -k_{p_{Cu}} p_{av} v_R \left[B_{21} + B_{22}\left(\frac{E^*}{p_{av}}\right)\frac{h_s}{w}\right] \end{aligned} \tag{2.51}$$

Thus, the step-height can be obtained by solving the following ordinary differential equation:

$$\frac{dh_s}{dt} + \left[k_{p_{Cu}} p_{av} v_R \left(\frac{B_{12}}{S_{Cu\,/\,ox}} - B_{22} \right) \left(\frac{E^*}{p_{av}} \right) \frac{1}{w} \right] h_s = k_{p_{Cu}} p_{av} v_R \left(B_{21} - \frac{B_{11}}{S_{Cu\,/\,ox}} \right) \tag{2.52}$$

At the end of Stage 2, the step-height when the pad touches the top of oxide layer, $t = t_2$, is zero. The general solution for the step-height, therefore, can be expressed as:

$$h_s(t) = \left(\frac{B_{21} S_{Cu\,/\,ox} - B_{11}}{B_{12} - B_{22} S_{Cu\,/\,ox}} \right) \left(\frac{p_{av}}{E^*} \right) w \left[1 - exp\left(-\frac{t - t_2}{\tau_3} \right) \right] \tag{2.53}$$

and

$$\tau_3 = \left(\frac{S_{Cu\,/\,ox}}{B_{12} - B_{22} S_{Cu\,/\,ox}} \right) \left(\frac{w}{k_{p_{Cu}} p_{av} v_R} \right) \left(\frac{p_{av}}{E^*} \right) \tag{2.54}$$

If $(t - t_2) \gg \tau_3$, the step-height approaches $h_s(\infty)$ and the material removal rates at the high and low features approach MRR_∞ very quickly. Thus,

$$h_s(\infty) = \left(\frac{B_{21} S_{Cu\,/\,ox} - B_{11}}{B_{12} - B_{22} S_{Cu\,/\,ox}} \right) \left(\frac{p_{av}}{E^*} \right) w \tag{2.55}$$

$$\begin{aligned} h_h(t) &= -MRR_\infty (t - t_2) \\ h_h(t) &= -MRR_\infty (t - t_2) - h_s(\infty) \end{aligned} \tag{2.56}$$

2.3.2 Sample Calculation

Results of step-height calculation for various pattern geometries based on the smooth pad model are listed in Table 2.1. The model parameters are decided based on conventional blanket and patterned wafer polishing data with commercial pad and slurry, which will be described in Chapter 4. In the conventional CMP practice, the time constant of the smooth pad model, τ_3, is much smaller than the overpolishing time at each feature, $(t_{ep} - t_2) \gg \tau_3$ as listed in Table 2.1. Therefore, the step-height approaches $h_s(\infty)$ soon after it enters Stage 3, and the material removal rates at both the high and low features approach MRR_∞. It may be noted that, however, the commercial polishing pads are rough, and thus the application of smooth pad model is limited to the features for which linewidths are smaller than the contact diameter of pad asperities.

Table 2.2 and Fig. 2.5 show the time evolution of pressure, material removal rate, height of the high and low features and the step-height based on the smooth pad model for comparing with other models described later. The polishing time is normalized by the ideal endpoint of a wafer,

Table 2.1 Step-height based on the smooth pad model: w = 50 μm and $\alpha = 1$.

w/λ	A_1	B_{11}	B_{12}	B_{21}	B_{22}	δ (nm)	t_1 (s)	t_2 (s)	τ_3 (s)	$h_s(\infty)$ (nm)	$h_s(t_{ep})$ (nm)
0.1	1.122	0.957	3.688	1.388	-33.196	4.8	194	338	0.35	3.8	3.8
0.2	1.202	0.871	1.683	1.515	-6.730	5.1	173	317	0.69	4.1	4.1
0.3	1.291	0.939	1.100	1.143	-2.568	5.5	151	295	1.13	4.4	4.4
0.4	1.393	1.073	0.803	0.891	-1.205	6.0	130	273	1.74	4.9	4.9
0.5	1.512	1.283	0.612	0.717	-0.612	6.4	108	252	2.66	5.1	5.1
0.6	1.657	1.621	0.471	0.586	-0.314	7.0	86	230	4.17	5.6	5.6
0.7	1.842	2.215	0.355	0.479	-0.152	7.9	65	209	6.88	6.3	6.2
0.8	2.101	3.476	0.251	0.381	-0.063	8.9	43	187	12.45	7.1	7.0
0.9	2.543	7.587	0.145	0.268	-0.016	10.9	22	166	27.39	8.7	8.5

* Parameters include:

h_{si} = 900 nm
h_{Cu} = 1500 nm
h_I = 1000 nm

$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}
$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}
$S_{Cu/ox}$ = 14.1
p_{av} = 14 kPa
v_R = 0.9 m/s

E_p = 300 MPa
ν_p = 0.3

t_{2f} = 360 s
t_{2sf} = 435 s
t_{ep} = 480 s

Table 2.2 Evolution of step-height based on the smooth pad model: $w = 50$ µm, $\lambda = 50$ µm, $\alpha = 1$.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\overline{p}_h$ (kPa)	$\overline{p}_l$ (kPa)	$\lvert dh_h / dt \rvert$ (nm/min)	$\lvert dh_l / dt \rvert$ (nm/min)
0	900	1500	600	28	0	500	0
30	650	1250	600	28	0	500	0
60	400	1000	600	28	0	500	0
90	150	750	600	28	0	500	0
120	0	546	546	14	14	250	250
150	0	421	421	14	14	250	250
180	0	296	296	14	14	250	250
210	0	171	171	14	14	250	250
240	0	46	46	14	14	250	250
270	1	-9	-10	26	2	33	33
300	1	-26	-27	26	2	33	33
330	1	-43	-44	26	2	33	33
360	1	-59	-60	26	2	33	33
390	1	-76	-77	26	2	33	33
420	1	-92	-93	26	2	33	33
450	1	-109	-110	26	2	33	33
480	1	-125	-126	26	2	33	33

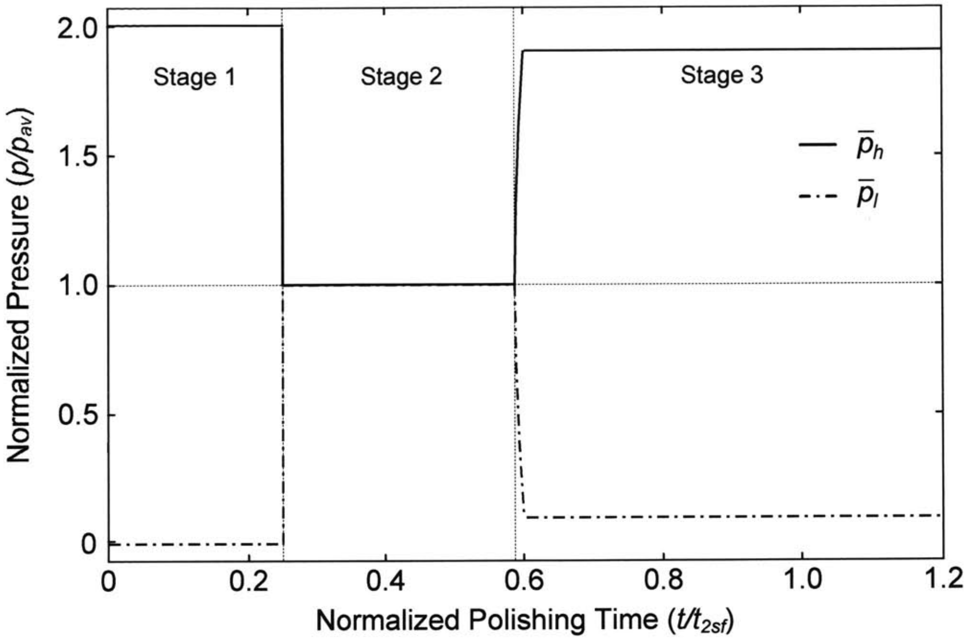
* Parameters include:

$h_{si} = 900$ nm
$h_{Cu} = 1500$ nm
$h_I = 1000$ nm

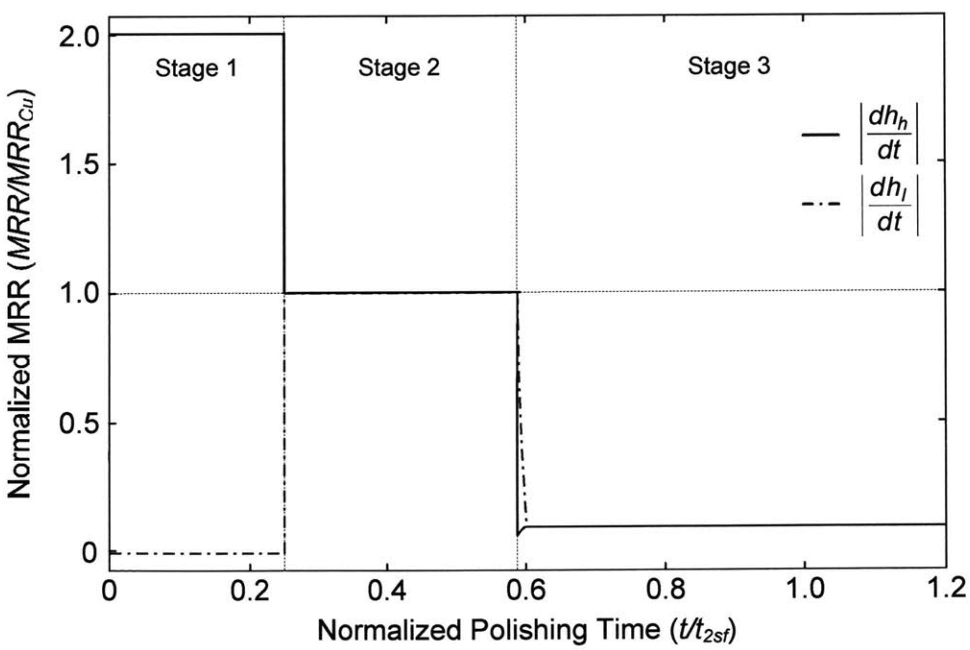
$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}
$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}
$S_{Cu/ox} = 14.1$
$p_{av} = 14$ kPa
$v_R = 0.9$ m/s

$E_p = 300$ MPa
$\nu_p = 0.3$

$t_{2f} = 360$ s
$t_{2sf} = 435$ s
$t_{ep} = 480$ s

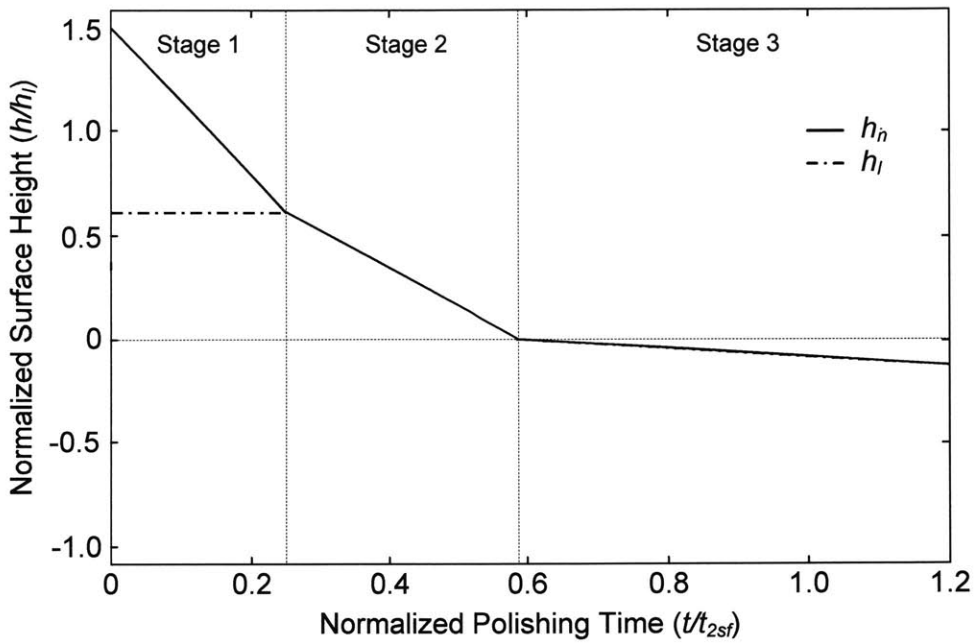


(a)

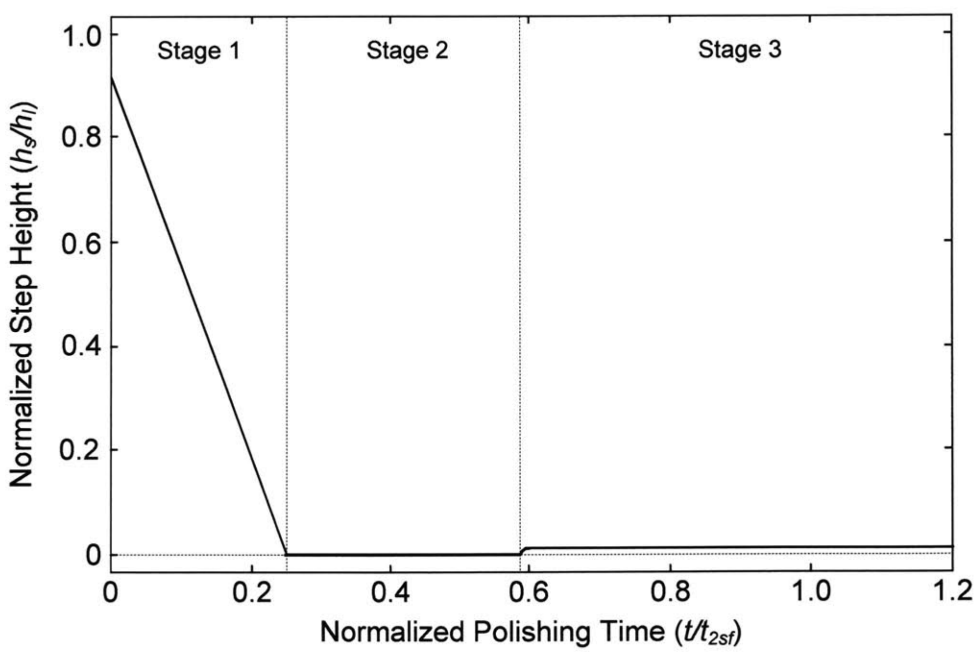


(b)

Figure 2.5 Time evolution of various parameters in the smooth pad model: (a) pressure, (b) material removal rate, (c) polished surface height at the high- and low-feature and (d) step-height (continued).



(c)



(d)

Figure 2.5 Time evolution of various parameters in the smooth pad model: (a) pressure, (b) material removal rate, (c) polished surface height at the high- and low-feature and (d) step-height.

t_{2sf}, at the wafer reference point, i.e., field region of the slowest die in a wafer.

In Stage 1, the pressure at the high feature is twice the average pressure, for $w / \lambda = 0.5$, and the pressure at the low feature is zero. Because the pad is not in contact, the material removal rate at the low feature is zero and the height at the low feature remains unchanged. The pad can be approximated as a perfectly smooth and flat surface, and thus as soon as the low feature contacts the pad, the pressure and the material removal rates at both high and low features become that of field region. When the pad at the high feature contacts the top of the oxide, material removal rate at the high feature is changed to that of the oxide. In this sample calculation, $S'_{Cu/ox}$ is about 14. Since the step-height reaches $h_s(\infty)$ very quickly, the material removal rate reaches MRR_{∞} almost instantly, and the pressure at the high feature is greater than that at the low feature by the ratio of selectivity.

2.4 The Discrete Pad Model

The previous section describes a smooth pad model derivation. The second approach is to relate the pressure on the Cu interconnect to pad deformation by assuming that the pad deforms as discrete, uniaxially loaded elastic blocks.

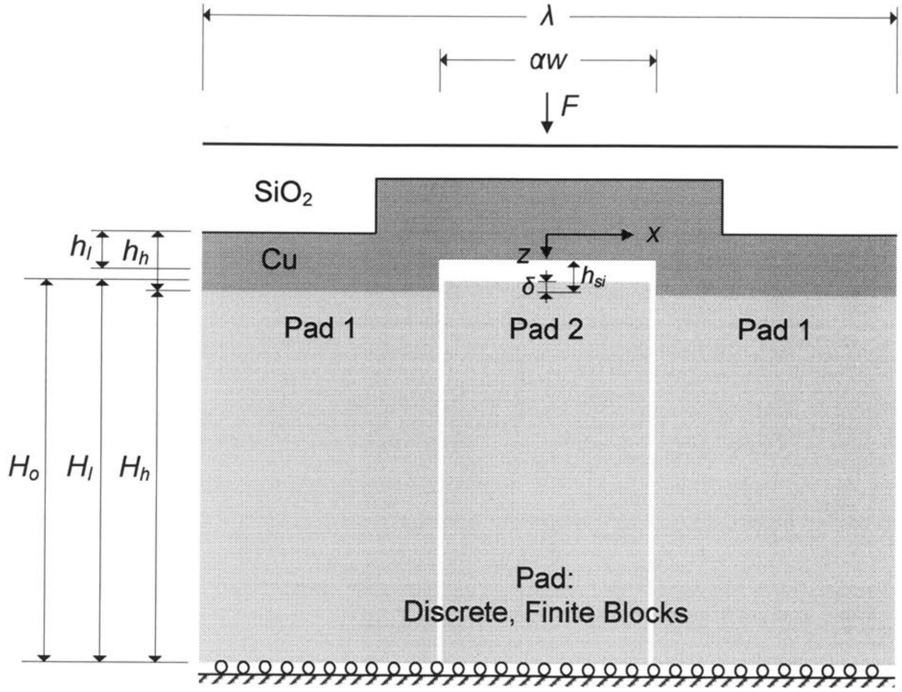
Due to its simplicity, the discrete pad assumption has been adopted by many researchers. Runnels *et al.* [1994, 1999 and 2003] presented a model by idealizing the polishing pad as a series of vertical springs, and horizontal springs connected to the vertical springs. Combined with Preston equation, the evolution of feature- and die-scale surface profile was calculated. Grillaert *et al.* [1998] studied the step-height planarization behavior in ILD CMP based on the assumption that pad deforms as discrete blocks and separated the step-height reduction into two stages: linear and exponential reduction. Elbel *et al.* [1998] proposed a mathematical approach to describe erosion and dishing for tungsten CMP. In this model, the polishing pad is represented as a network of both vertical and horizontal springs to simulate the profile of dished tungsten line. They assumed a linear relationship between the pressure and the step-height at Cu and oxide regions. The stiffness of vertical and horizontal spring was determined by experimental results of maximum dishing.

Chen and Lee [1999] presented pattern planarization model based on the assumption that the pad completely conforms wafer pattern and the pressure difference at the upper and the lower surface is proportional to the step-height. The proportional constant is defined as the loading density coefficient and used to explain the stiffness of the pad. However, since the pattern geometry effect is integrated into the loading density parameter, it requires large amount of experiments to apply for various pattern geometries. Yang [2000] also proposed a copper planarization model by assuming a perfectly conformal pad. In this model, the pressure at the high and low features was determined by the vertical deformation of the pad, and the step-height was calculated by combining with Preston equation. However, he related the pad deformation to the Cu interconnect linewidth only by assuming that the relative change of pad compression distance is proportional to the relative change of the feature size. Fu and Chandra [2003] assumed that the pad behaves like an elastic foundation or linear springs with certain bending ability, while the influences of pad viscosity and pad asperities are ignored. Based on this assumption, step-height reduction and dishing model was developed and effects of pattern geometries were investigated. It was assumed that the pad always touches both the high and the low features. In this model, however, the stiffness of the pad is related to the initial thickness of a polishing pad, and how to determine the bending factor value is not fully explained. Later, Guo *et al.* [2004] extended the previous model to apply viscoelastic pad material properties, i.e., Young's modulus and bending factor.

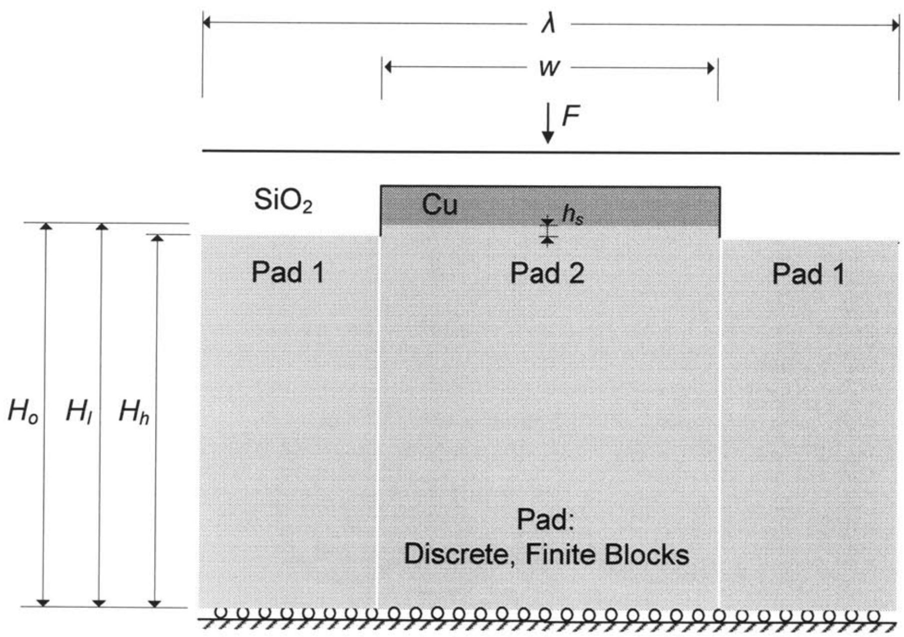
In this section, we specialize the model structure of section 2.2 by adopting the discrete pad model in the following derivation. The high and low features are assumed as shown in Fig. 2.6.

The key assumptions of the model are:

- The pad is an isotropic, elastic material.
- The wafer surface always remains horizontal.
- The pad deforms as separate uniaxially-loaded blocks under the uniform pressure boundary condition.
- The deformation is plane-strain.
- The pad surface is perfectly smooth across each loaded block.
- The pad-surface in x-direction is stress-free and does not expand: $\sigma_x = \tau_{xy} = \tau_{xz} = 0$, $\nu = 0$.
- The bottom of each block remains at the same level and is free to move horizontally.



(a)



(b)

Figure 2.6 Schematics of the pattern/discrete pad contact interface when (a) $h_s > \delta$ and $A_f = \alpha w / \lambda$, and (b) $h_s \leq \delta$ and $A_f = w / \lambda$.

2.4.1 Theory

Stage 1

Again, the initial relative pad deformation, δ, can be represented as the relative displacement of the pad between the high and the low features as in Eq. (2.18).

$$\delta = \left(\frac{\overline{p}_h}{E_p}\right) H_o \tag{2.57}$$

By combining with force equilibrium, (2.57) can be rewritten as:

$$\delta = \left(\frac{1}{1 - \alpha w / \lambda}\right)\left(\frac{p_{av}}{E_p}\right) H_o \tag{2.58}$$

In this section, we start with the condition $\delta < h_{si}$ so that the low feature does not support the load initially. If $\delta \geq h_{si}$, the analysis starts with Stage 2.

The final heights at the high and low features, the step-height, and the time at the end of Stage 1, t_1, can be obtained by substituting δ into Eqs. (2.20) – (2.22), respectively. Thus, the discrete pad model discussed here provides an expression for the height at which low area contact first occurs, as previously mentioned in the general modeling framework of section 2.2.

Stage 2

In Stage 2, the pad is in contact with both the high and low features. Deformations of the pad at the high and the low features, δ_h and δ_l, in Stage 2 can be expressed as:

$$\begin{aligned} \delta_h &= \delta_h(\overline{p}_h, \overline{p}_l) = \left(\frac{\overline{p}_h}{E_p}\right) H_o \\ \delta_l &= \delta_l(\overline{p}_h, \overline{p}_l) = \left(\frac{\overline{p}_l}{E_p}\right) H_o \end{aligned} \tag{2.59}$$

where E_p is the Young's modulus of the pad material and H_o the undeformed pad thickness. Therefore, pad blocks pressing the high and low features act like springs.

Although the discrete pad model relies on finite pad thickness and physically inadmissible discontinuous deformation of the pad, it gives a simple conceptual picture of how a rough pad would behave if the pad asperities act as individual contacts.

The step-height can be expressed by the thickness difference of the high and low features, and the pressure difference by combining Eqs. (2.23) and (2.59) as:

$$h_s = \left(\frac{\overline{p}_h}{E_p}\right)H_o - \left(\frac{\overline{p}_l}{E_p}\right)H_o \tag{2.60}$$

where h_s, $\overline{p}_h$ and $\overline{p}_l$ are time dependent variables. By combining Eq. (2.60) with the force equilibrium, the pressure and the material removal rate at the high and low features at any given time t can be expressed as:

$$\begin{aligned}\overline{p}_h &= p_{av}\left[1 + (\alpha w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s}{H_o}\right]\\ \overline{p}_l &= p_{av}\left[1 - (1 - \alpha w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s}{H_o}\right]\end{aligned} \tag{2.61}$$

and

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_{Cu}} p_{av} v_R\left[1 + (\alpha w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s}{H_o}\right]\\ \frac{dh_l}{dt} &= -k_{p_{Cu}} p_{av} v_R\left[1 - (1 - \alpha w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s}{H_o}\right]\end{aligned} \tag{2.62}$$

Thus, the step-height, h_s, is expressed by the first-order ordinary differential equation

$$\frac{dh_s}{dt} + \left(\frac{k_{p_{Cu}} E_p v_R}{H_o}\right)h_s = 0 \tag{2.63}$$

At the onset of Stage 2, $t = t_1$, the low feature barely contacts the pad, and thus $h_s(t_1) = \delta$. For $t > t_1$, the high feature is polished faster than the low feature, thus the step-height gradually decreases. The general solution for step-height in Stage 2, $t_1 < t < t_2$, is expressed as:

$$h_s(t) = \delta\, exp\left(-\frac{t - t_1}{\tau_2}\right) \tag{2.64}$$

where

$$\tau_2 = \left(\frac{H_o}{k_{p_{Cu}} E_p v_R}\right) \tag{2.65}$$

From (2.61), (2.62) and (2.64), the pressure and material removal rates at the high and low features can be rewritten as:

$$\begin{aligned}\overline{p}_h(t) &= p_{av}\left[1 + \left(\frac{\alpha w/\lambda}{1 - \alpha w/\lambda}\right)exp\left(-\frac{t - t_1}{\tau_2}\right)\right]\\ \overline{p}_l(t) &= p_{av}\left[1 - exp\left(-\frac{t - t_1}{\tau_2}\right)\right]\end{aligned} \tag{2.66}$$

and,

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_{Cu}} p_{av} v_R \left[1 + \left(\frac{\alpha w/\lambda}{1-\alpha w/\lambda}\right) exp\left(-\frac{t-t_1}{\tau_2}\right)\right] \\ \frac{dh_l}{dt} &= -k_{p_{Cu}} p_{av} v_R \left[1 - exp\left(-\frac{t-t_1}{\tau_2}\right)\right]\end{aligned} \tag{2.67}$$

By applying initial values of Stage 2 at the high and low features, $h_h(t)$ and $h_l(t)$ for Stage 2 can be solved as:

$$\begin{aligned}h_h(t) &= h_{Cu} - h_{si} + \delta - k_{p_{Cu}} p_{av} v_R (t - t_1) - (\alpha w/\lambda)\delta\left[1 - exp\left(-\frac{t-t_1}{\tau_2}\right)\right] \\ h_l(t) &= h_{Cu} - h_{si} - k_{p_{Cu}} p_{av} v_R (t - t_1) + (1 - \alpha w/\lambda)\delta\left[1 - exp\left(-\frac{t-t_1}{\tau_2}\right)\right]\end{aligned} \tag{2.68}$$

At the beginning of Stage 2, the high and low features have a step-height δ. Material removal rates at high and low features approach that of blanket wafer polishing as polishing progresses.

Stage 2 ends when the pad reaches the top of the oxide at the high feature. Thus, the dimensionless time interval, t_2^*, defined as $(t_2 - t_1)/\tau_2$, can be obtained by solving the following equation

$$t_2^* = \left(\frac{h_{Cu} - h_{si}}{H_o}\right)\left(\frac{E_p}{p_{av}}\right) + 1 + \left(\frac{\alpha w/\lambda}{1-\alpha w/\lambda}\right) exp\left(-t_2^*\right) \tag{2.69}$$

The step-height when $t = t_2$ is written as:

$$h_s(t_2) = \delta exp\left(-t_2^*\right) \tag{2.70}$$

The step-height at the end of Stage 2, $h_s(t_2)$, is determined by the initial pad deformation, δ, and the ratio of the polishing interval to the time constant $(t_2 - t_1)/\tau_2$, which needs to be solved numerically. However, if $(h_{Cu} - h_{si})E_p / H_o p_{av} > 3$, which is valid for most commercial polishing pads, then $t_2^* > 4$ and thus the exponential term may be neglected for most pattern geometries except when $\alpha w/\lambda$ is close to unity. Additionally, even though t_1 for the general pad is comparably smaller than that for perfectly rigid pad due to its elastic deformation δ, the time at the end of Stage 2, t_2, is almost the same as the flat pad case, as t_2^* increases. Finally, as t_2^* increases, the step-height decreases exponentially. For example, when $t_2^* > 4$, $h_s(t_2)$ is less than 2% of the maximum pad deformation, δ, which may be ignored and considered as a flat surface.

Stage 3

The pressure and the material removal rate in the oxide and Cu regions in Stage 3 can be represented by the step-height in the same way as in (2.61) and (2.62).

$$\begin{aligned} p_h(t) &= p_{av}\left[1+(w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s(t)}{H_o}\right] \\ p_l(t) &= p_{av}\left[1-(1-w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s(t)}{H_o}\right] \end{aligned} \tag{2.71}$$

and

$$\begin{aligned} \frac{dh_h}{dt} &= -k_{p_{ox}} p_{av} v_R\left[1+(w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s(t)}{H_o}\right] \\ \frac{dh_l}{dt} &= -k_{p_{Cu}} p_{av} v_R\left[1-(1-w/\lambda)\left(\frac{E_p}{p_{av}}\right)\frac{h_s(t)}{H_o}\right] \end{aligned} \tag{2.72}$$

The step-height, h_s, in Stage 3 is expressed by the first-order ordinary differential equation as:

$$\frac{dh_s}{dt}+\left\{[(1-w/\lambda)k_{p_{Cu}}+(w/\lambda)k_{p_{ox}}]p_{av}v_R\left(\frac{E_p}{p_{av}}\right)H_o\right\}h_s=[k_{p_{Cu}}-k_{p_{ox}}]p_{av}v_R \tag{2.73}$$

At the onset of Stage 3, $t=t_2$, there is an initial step-height, $h_s(t_2)$. Therefore, the general solution for step-height in Stage 3 is:

$$h_s(t) = h_s(t_2)+[h_s(\infty)-h_s(t_2)]\left\{1-exp\left[-\left(\frac{t-t_2}{\tau_3}\right)\right]\right\} \tag{2.74}$$

$$h_s(\infty) = \left[\frac{S_{Cu/ox}-1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left(\frac{p_{av}}{E_p}\right)H_o \tag{2.75}$$

$$\tau_3 = \left[\frac{S_{Cu/ox}}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left(\frac{H_o}{k_{p_{Cu}}E_p v_R}\right) \tag{2.76}$$

where $h_s(\infty)$ is the asymptotic step-height when $t\rightarrow\infty$, and τ_3 the time constant of Stage 3. The maximum pad deformation in Stage 3, δ_o, defined as the pad deformation when the Cu interconnect area is empty, is:

$$\delta_o \equiv \left(\frac{1}{1-w/\lambda}\right)\left(\frac{p_{av}}{E_p}\right)H_o \tag{2.77}$$

In this stage, the normal load is supported by both high and low features, and the pressure and material removal rates at the high and low features depend on the step-height and the slurry selectivity. If the selectivity is large, the material removal rate of Cu is greater than that of oxide, and the step-height increases. As the step-height increases, the pressure and the material removal rate at the high feature increase and those at the low feature decrease and the material removal rates at both high and low features approach MRR_{∞} in Eq. (2.28).

At the onset of Stage 3, $t = t_2$, the surface heights of the high and low features are:

$$\begin{aligned} h_h(t_2) &= 0 \\ h_l(t_2) &= -h_s(t_2) = -\delta\, exp\left(-t_2^*\right) \end{aligned} \tag{2.78}$$

Therefore, the general solution for $h_h(t)$ and $h_l(t)$ are solved as:

$$\begin{aligned} h_h(t) &= -MRR_{\infty}(t - t_2) + \left[\frac{w/\lambda}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right][h_s(t) - h_s(t_2)] \\ h_l(t) &= -h_s(t_2) - MRR_{\infty}(t - t_2) - \left[\frac{(1 - w/\lambda)S_{Cu/ox}}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right][h_s(t) - h_s(t_2)] \end{aligned} \tag{2.79}$$

In Stage 3, pressure in the oxide region, now the high feature, is greater than that in the Cu region, low feature. Since the material removal rate depends both on the Preston constant and pressure, the Cu-to-oxide selectivity, $S_{Cu/ox}$, between oxide and Cu comes into consideration.

2.4.2 Sample Calculation

The results of step-height calculations for various pattern geometries based on the discrete pad model are listed in Table 2.3. Table 2.4 and Fig. 2.7 show the time evolution of pressure, material removal rate, height at high and low features and the step-height. In this calculation, $(t_2 - t_1)/\tau_2 = 11$ and thus the step-height reaches almost zero before it enters Stage 3. Additionally, $(t_{ep} - t_2)/\tau_3$ is in the range between 4 and 10 as listed in Table 2.3. Therefore, the final step-height at the polishing endpoint is close to $h_s(\infty)$, and the material removal rates at both the high and low features are close to MRR_{∞}.

Table 2.3 Step-height calculation based on the discrete pad model: w = 50 µm and $\alpha = 1$.

w/λ	δ (nm)	δ_o (nm)	t_1 (s)	τ_2 (s)	t_2 (s)	τ_3 (s)	t_2^*	t_o^*	$h_s(t_2)$ (nm)	$h_s(\infty)$ (nm)	$h_s(t_{ep})$ (nm)	$h_h(t_{ep})$ (nm)	$h_l(t_{ep})$ (nm)
0.1	67	67	180	15	340	16	11.0	10.0	0	62	62	-58	-120
0.2	76	76	158	15	318	18	11.0	10.2	0	69	69	-73	-143
0.3	87	87	137	15	297	20	11.0	10.1	0	78	78	-88	-166
0.4	101	101	115	15	275	23	11.0	9.7	0	90	90	-111	-201
0.5	121	121	93	15	253	27	11.0	9.1	0	105	105	-137	-242
0.6	152	152	72	15	232	33	11.0	8.2	0	127	127	-173	-300
0.7	202	202	50	15	210	42	11.0	7.0	0	161	161	-230	-391
0.8	303	303	29	15	189	57	11.0	5.5	0	220	219	-316	-535
0.9	607	607	7	15	167	89	11.0	3.7	0	344	336	-476	-812

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}	E_p = 300 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}	H_o = 1.3 mm	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1		t_{ep} = 480 s
	p_{av} = 14 kPa		
	v_R = 0.9 m/s		

Table 2.4 Evolution of step-height based on the discrete pad model: w = 50 μm, λ = 100 μm and α = 1.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\bar{p}_h$ (kPa)	$\bar{p}_l$ (kPa)	$\lvert dh_h / dt \rvert$ (nm/min)	$\lvert dh_l / dt \rvert$ (nm/min)
0	900	1500	600	28	0	500	0
30	650	1250	600	28	0	500	0
60	400	1000	600	28	0	500	0
90	150	750	600	28	0	500	0
120	18	559	541	16	12	288	212
150	2	426	424	14	14	254	246
180	0	300	300	14	14	251	249
210	0	175	175	14	14	250	250
240	0	50	50	14	14	250	250
270	47	-14	-61	19	9	25	152
300	86	-28	-114	24	4	30	72
330	99	-44	-143	25	3	32	46
360	103	-60	-163	26	2	33	37
390	105	-76	-181	26	2	33	34
420	105	-93	-198	26	2	33	34
450	105	-110	-215	26	2	33	33
480	105	-126	-231	26	2	33	33

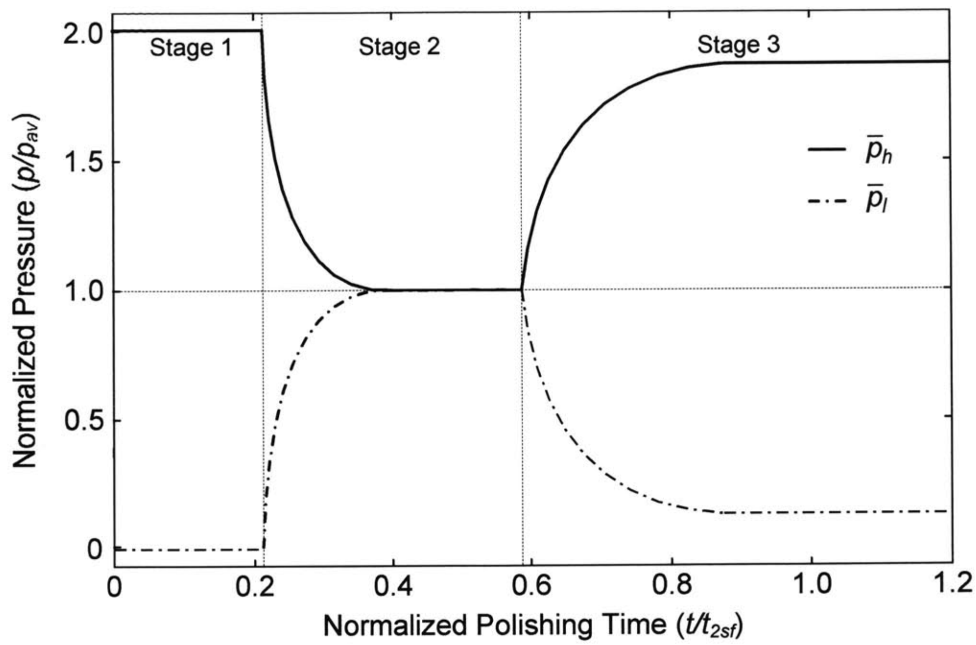
* Parameters include:

h_{si} = 900 nm
h_{Cu} = 1500 nm
h_I = 1000 nm

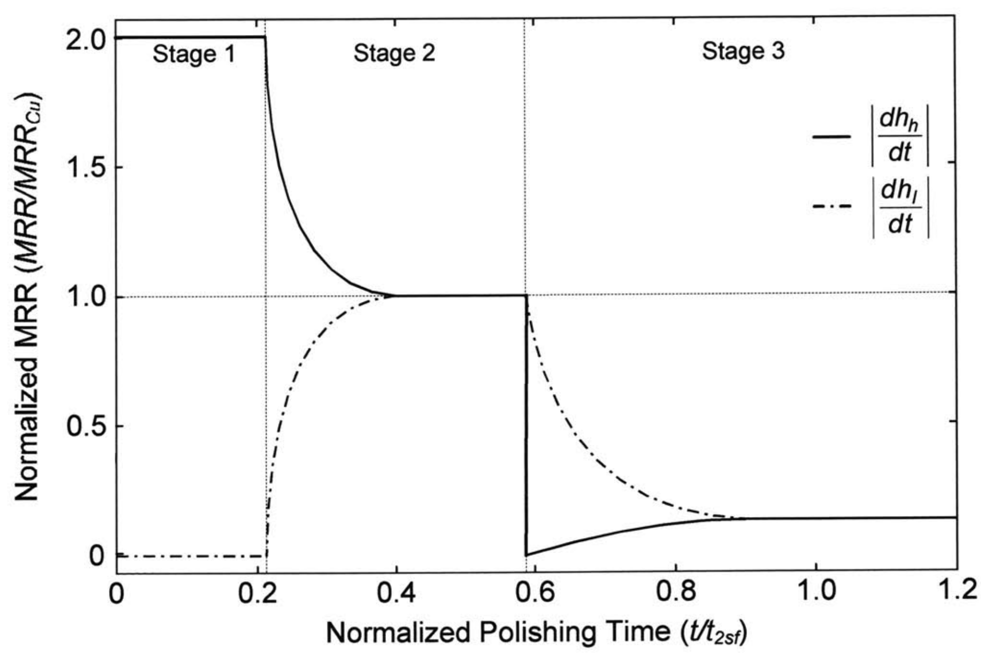
$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}
$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}
$S_{Cu/ox}$ = 14.1
p_{av} = 14 kPa
v_R = 0.9 m/s

E_p = 300 MPa
H_o = 1.3 mm

t_{2f} = 360 s
t_{2sf} = 435 s
t_{ep} = 480 s

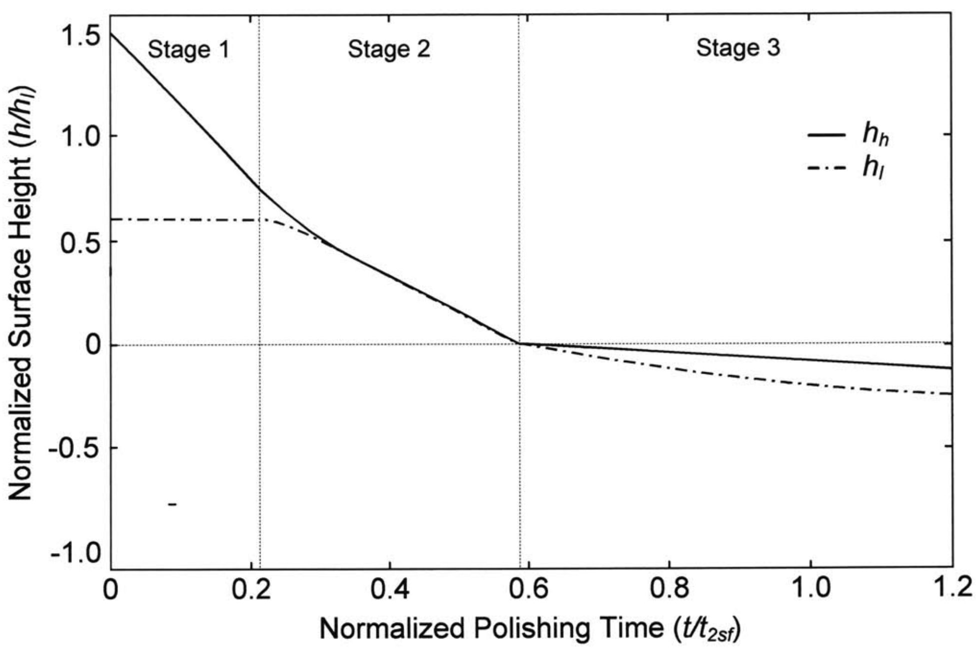


(a)

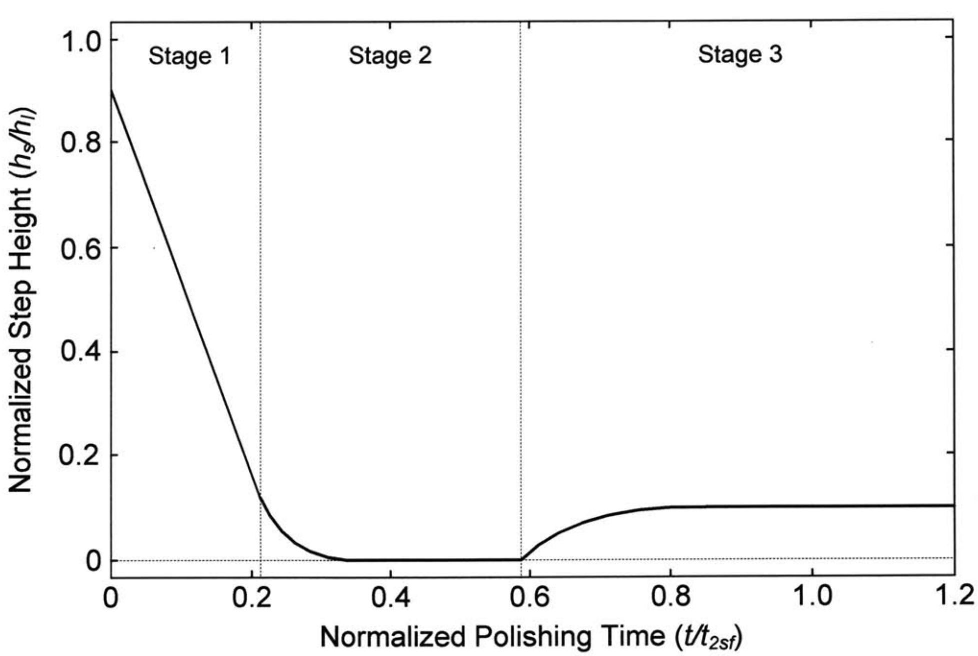


(b)

Figure 2.7 Time evolution of various parameters in the discrete pad model: (a) pressure, (b) material removal rate (c) polished surface height at the high- and low-feature and (d) step-height (continued).



(c)



(d)

Figure 2.7 Time evolution of various parameters in the discrete pad model: (a) pressure, (b) material removal rate (c) polished surface height at the high- and low-feature and (d) step-height.

2.5 The Simplified Rough Pad Model: Uniform Asperity Height Distribution

In the smooth and discrete pad models, sections 2.3 and 2.4, the pad is assumed to have perfectly smooth contact surfaces, and thus the true contact areas are the same as the areas of the high and the low features. However, the real polishing pad is rough and the wafer is supported by many pad asperities. Therefore, the true contact area is much smaller than the nominal area.

In this section and the following section, we consider the contact between pad asperities and the wafer based on the measured pad surface profile, which will determine the local pressure and material removal rate realistically. It is assumed that the slurry distribution is uniform across the wafer, there are enough abrasive particles in the contact area, and each asperity carries at least one abrasive particle. The mechanism of abrasive entrapment by the pad asperities, however, requires further study [Luo and Dornfeld, 2003].

The wafer-scale nominal pressure, p_{av}, is defined as the ratio of the normal load to the projected area of the wafer. The true contact pressure, or stress, however, cannot be determined from the average pressure unless of course the asperities on the pad are flattened elastically or plastically, so that the real contact area is close to the nominal area. In CMP practice, for example, the pad surface contains many pores and the ratio of the average pressure to the Young's modulus of the pad is about 10^{-4}. Therefore, the deformation at each pad asperity is small and the real contact area is much smaller than the nominal area.

Many of the rough pad models in CMP adopt the Greenwood-Williamson approach to calculate the contact area and the spacing of pad asperities in contact with the wafer [Greenwood and Williamson, 1966; Johnson, 1985]. The effect of pad roughness on the material removal rate in CMP was considered by Yu *et al.* although the area fraction effect was not considered [Yu *et al.*, 1993]. Vlassak proposed a feature-scale wafer surface evolution model based on the contact mechanics between a rough pad and a patterned wafer [Vlassak, 2001, 2004]. The pad asperity contact is modeled based on the Greenwood-Willamson approach with exponential asperity-height distribution. Then, the average pressure and the pad deformation profile was expressed by integral form based on the smooth contact mechanics analysis, which was used by Chekina *et al.* The pad deformation and average pressure profile was calculated iteratively by assuming an initial pressure distribution. This model requires a full profile of pad deformation in the nominal area to calculate average pressure profile, and thus the calculation is cumbersome. Borucki presented a mathematical analysis of the evolution of blanket wafer material removal rate based on the rough pad contact [Borucki, 2002]. Nguyen derived a dishing model based on the rough pad surface with Gaussian pad asperity-height distribution [Nguyen *et al.*, 2003]. This model has two different contact modes. The first is the rough pad asperity contact when Cu linewidth is larger than diameter of the asperity contact. In this mode, the Greenwood-Williamson approach has been adopted to investigate the contact pressure of pad asperities under a given load. The other is when the pad asperity contact diameter is smaller than the Cu linewidth. In this mode, the maximum pad asperity deformation is assumed to be proportional to Cu linewidth and the pressure is expressed as a first-order ordinary differential equation. In this model, however, erosion is neglected, therefore, dishing rate is assumed to be directly proportional to the pressure at Cu area only, which is valid only when a high-selectivity slurry is used. Furthermore, the model considers only the overpolishing stage and does not explain the evolution of step-height. Finally, it assumes the mean contact pressure is the same regardless of the area fraction of Cu interconnects.

Seok *et al.* [2003] presented a multiscale model for material removal in CMP. Abrasive particle, asperity and wafer-scale were considered based on the deformation of hyper-elastic asperities attached to a linear elastic pad. The asperity deformation behavior was combined with the Greenwood-Williamson rough pad analysis to calculate the blanket wafer polishing behavior. Then, the time evolution of a feature is analyzed by using an iterative numerical analysis from the complete wafer surface profile. This model, however, relies heavily on the finite element analysis and thus requires intense computation. Additionally, how to determine each model parameter and the physical meaning of each step are not fully described.

In this section, a simplified rough surface model is developed for step-height evolution in patterned wafer polishing. The pad topography is idealized as uniform with all asperities having the same height, radius of curvature, and of a certain spatial density. This idealization is valid for two reasons. First, when the mean deformation of the pad is so small that only the high end of the distribution is actually in contact with the wafer, variation in the number of asperities in contact is relatively small during polishing. Second, the pad/wafer contact is not static but dynamic. Therefore, the effect of actual contact pressure on any feature of the wafer during polishing can be represented as an average value. That is, even though the heights of pad asperities in contact at any given moment are different between two features, if the number of pad asperities that pass over a feature during the total polishing time is large, the effect of pad asperity height can be averaged as long as the spatial distribution and the radius of curvature of asperities are the same.

As shown in Fig. 2.8(a), let R_a be the radius, λ_a the spacing of pad asperities, r_{a_h} and r_{a_l} the radii of the asperity contact areas, $\overline{p}_{a_h}$ and $\overline{p}_{a_l}$ the mean pad asperity contact pressures at the high and low features, respectively.

2.5.1 Elastic Contact

Contact between the pad asperities and the wafer is assumed to follow the Hertz theory of elastic contacts. The key assumptions of the model are:

- The pad is isotropic, linear elastic material.
- The wafer surface remains horizontal.
- Each pad asperity behaves as an elastic half-space if $R_a \ll \lambda_a$.
- The strain of pad asperity is very small: $r_{a_h}, r_{a_l} \ll R_a$.
- The mean contact pressure is much less than the fully plastic stress: $\overline{p}_{a_h}, \overline{p}_{a_l} \ll 3Y_p$.
- Friction between the pad and the wafer is negligible: $\mu \sim 0.1$.

Let η_a be the number of pad asperities per unit area. Then, the mean spacing of the pad asperities, λ_a, is:

$$\lambda_a = \left(\frac{1}{\eta_a}\right)^{1/2} \tag{2.80}$$

Since pad asperities can be considered as moving bodies on the wafer, the mean pressure, $\overline{p}$, in the nominal area, λ_a^2, can be expressed by the pad asperity contact pressure, p_a, as:

$$\overline{p} \equiv p_a\left(\frac{\pi r_a^2}{\lambda_a^2}\right) \tag{2.81}$$

In this case, the radius of the asperity contact area, r_a, the deformation, δ_a, and the asperity contact pressure, p_a, at each asperity contact can be expressed as:

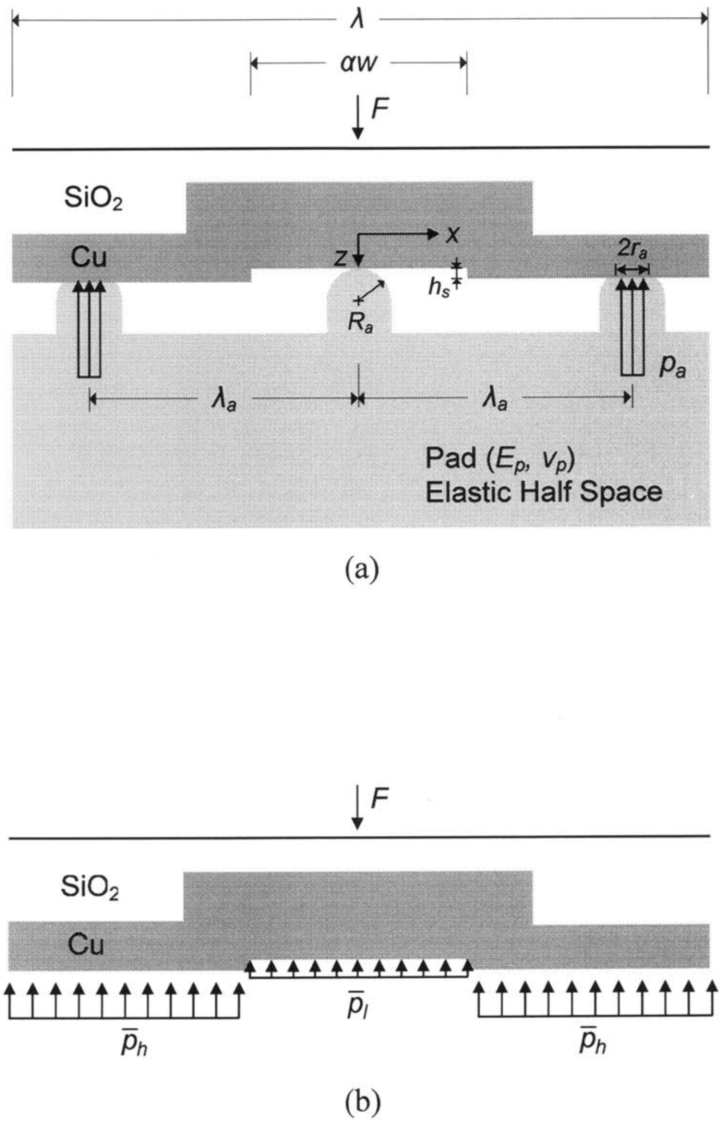


Figure 2.8 Schematics of the pattern/simplified rough pad contact interface: (a) pad asperity contact pressure and radius of contact and (b) average pressure at the high- and low-feature.

$$r_a = \left(\frac{3}{4}\lambda_a^2 R_a\right)^{1/3}\left(\frac{\overline{p}}{E^*}\right)^{1/3} \tag{2.82}$$

$$\delta_a = \left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3}\left(\frac{\overline{p}}{E^*}\right)^{2/3} \tag{2.83}$$

$$p_a = \frac{2}{3}\left(\frac{6\lambda_a^2}{\pi^3 R_a^2}\right)^{1/3}\left(\overline{p}E^{*2}\right)^{1/3} \tag{2.84}$$

If the pad asperities are assumed to deform independently, the deformation of pad asperities at the high and the low features can be expressed as:

$$\begin{aligned} \delta_h &= \left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3}\left(\frac{\overline{p}_h}{E^*}\right)^{2/3} \\ \delta_l &= \left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3}\left(\frac{\overline{p}_l}{E^*}\right)^{2/3} \end{aligned} \tag{2.85}$$

Stage 1

The pad deformation, δ, is represented as the relative deformation of the pad asperities at the high feature with the condition $\overline{p}_l = 0$ as:

$$\delta = \left(\frac{\overline{p}_h}{E^*}\right)^{2/3}\left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3} \tag{2.86}$$

From force equilibrium, Eq. (2.86) can be rewritten as:

$$\delta = \left(\frac{1}{1-\alpha w/\lambda}\right)^{2/3}\left(\frac{p_{av}}{E^*}\right)^{2/3}\left(\frac{9\lambda_a^4}{16R}\right)^{1/3} \tag{2.87}$$

If $\delta \geq h_{si}$, the low feature is also in contact and thus the polishing starts with Stage 2. If the polishing starts with Stage 1, $\delta < h_{si}$, then the general procedure in section 2.2 can be applied.

Stage 2

Since both the high and low features are in contact with the pad now, the step-height is expressed as the difference of the pad asperity deformation at the high and low features as:

$$h_s = \left(\frac{\overline{p}_h}{E^*}\right)^{2/3}\left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3} - \left(\frac{\overline{p}_l}{E^*}\right)^{2/3}\left(\frac{9\lambda_a^4}{16R_a}\right)^{1/3} \tag{2.88}$$

By combining Eqs. (2.85)and (2.88), the force equilibrium in Stage 2 is expressed as:

$$\delta_h^{3/2}(1-\alpha w/\lambda)+(\delta_h-h_s)^{3/2}(\alpha w/\lambda)=\left(\frac{p_{av}}{E^*}\right)\left(\frac{9\lambda_a^4}{16R_a}\right)^{1/2} \tag{2.89}$$

This equation needs to be numerically solved for δ_h at a given h_s. Therefore, the procedure to determine the evolution of step-height is iterative. Once δ_h and δ_l are determined at a given step-height, the pressure, material removal rates at the high and low features and step-height reduction rate can be calculated as:

$$\begin{aligned}\overline{p}_h &= E^*\left(\frac{16R_a\delta_h^3}{9\lambda_a^4}\right)^{1/2}\\ \overline{p}_l &= E^*\left(\frac{16R_a\delta_l^3}{9\lambda_a^4}\right)^{1/2}\end{aligned} \tag{2.90}$$

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_{Cu}}p_{av}v_R\left[\left(\frac{E^*}{p_{av}}\right)\left(\frac{16R_a\delta_h^3}{9\lambda_a^4}\right)^{1/2}\right]\\ \frac{dh_l}{dt} &= -k_{p_{Cu}}p_{av}v_R\left[\left(\frac{E^*}{p_{av}}\right)\left(\frac{16R_a\delta_l^3}{9\lambda_a^4}\right)^{1/2}\right]\end{aligned} \tag{2.91}$$

$$\frac{dh_s}{dt} = -k_{p_{Cu}}p_{av}v_R\left(\frac{E^*}{p_{av}}\right)\left(\frac{16R_a}{9\lambda_a^4}\right)^{1/2}\left(\delta_h^{3/2}-\delta_l^{3/2}\right) \tag{2.92}$$

The pressure and material removal rate at the high and low features, and the step-height reduction rate at any given time t can be calculated recursively. The end of Stage 2 is marked by the polishing time, t_2, when the pad surface at high feature reaches the oxide surface, that is, $h_h(t_2)=0$.

Stage 3

In this stage, the materials being polished at the high and low features are oxide and Cu, respectively: $k_{p_h}=k_{p_{ox}}$ and $k_{p_l}=k_{p_{Cu}}$. The new area fraction in Eq. (2.12) is calculated based on the designed Cu linewidth, w, instead of αw. Although the pressure at the high and the low features in Stage 3 are the same with those in Stage 2, the material removal rates in the oxide and Cu regions in Stage 3 are different due to the different Preston constants. Thus,

$$
\begin{aligned}
\frac{dh_h}{dt} &= -k_{p_{ox}} p_{av} v_R \left[\left(\frac{E^*}{p_{av}} \right) \left(\frac{16 R_a \delta_h^3}{9 \lambda_a^4} \right)^{1/2} \right] \\
\frac{dh_l}{dt} &= -k_{p_{Cu}} p_{av} v_R \left[\left(\frac{E^*}{p_{av}} \right) \left(\frac{16 R_a \delta_l^3}{9 \lambda_a^4} \right)^{1/2} \right]
\end{aligned} \tag{2.93}
$$

$$
\frac{dh_s}{dt} = -k_{p_{Cu}} p_{av} v_R \left(\frac{E^*}{p_{av}} \right) \left(\frac{16 R_a}{9 \lambda_a^4} \right)^{1/2} \left(\frac{1}{S_{Cu/ox}} \delta_h^{3/2} - \delta_l^{3/2} \right) \tag{2.94}
$$

At the onset of Stage 3, $t=t_2$, there is an initial step-height, $h_s(t_2)$. The step-height in Stage 3 can also be solved numerically by the same procedure as that of Stage 2.

2.5.2 Fully Plastic Contact

In the fully plastic contact mode, the deformation of pad asperity and the mean pressure at the high and low features has linear relationship as [Johnson, 1985]:

$$
\begin{aligned}
\delta_h &= \left(\frac{\overline{p}_h}{Y_p} \right) \left(\frac{\lambda_a^2}{6 \pi R_a} \right) \\
\delta_l &= \left(\frac{\overline{p}_l}{Y_p} \right) \left(\frac{\lambda_a^2}{6 \pi R_a} \right)
\end{aligned} \tag{2.95}
$$

Therefore, the step-height analysis procedure of this subsection is exactly the same as the discrete pad model by replacing the initial pad thickness H_o with $\lambda_a^2 / 6 \pi R_a$ and the Young's modulus of the pad, E_p, with the yield strength of the pad, Y_p.

Stage 1

The initial relative pad deformation, δ, is the same as the pad displacement at the high feature:

$$
\delta = \left(\frac{\overline{p}_h}{Y_p} \right) \left(\frac{\lambda_a^2}{6 \pi R_a} \right) \tag{2.96}
$$

Combined with the force balance equation and the condition $\overline{p}_l = 0$:

$$
\delta = \left(\frac{1}{1 - \alpha w / \lambda} \right) \left(\frac{p_{av}}{Y_p} \right) \left(\frac{\lambda_a^2}{6 \pi R_a} \right) \tag{2.97}
$$

The final heights at the high and low features and the step-height, and the time at the end of Stage 1 can be obtained by subsituting δ in Eq. (2.97) into Eqs. (2.20) – (2.22), respectively.

Stage 2

The pressure and the material removal rates at the high and low features at any given time t can be expressed as a function of the step-height:

$$\bar{p}_h(t) = p_{av}\left[1 + (\alpha w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right]$$
$$\bar{p}_l(t) = p_{av}\left[1 - (1 - \alpha w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right] \tag{2.98}$$

$$\frac{dh_h}{dt} = -k_{p_{Cu}} p_{av} v_R\left[1 + (\alpha w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right]$$
$$\frac{dh_l}{dt} = -k_{p_{Cu}} p_{av} v_R\left[1 - (1 - \alpha w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right] \tag{2.99}$$

The step-height, h_s, is expressed by the first-order ordinary differential equation as:

$$\frac{dh_s}{dt} + \left[k_{p_{Cu}} Y_p v_R\left(\frac{6\pi R_a}{\lambda_a^2}\right)\right]h_s = 0 \tag{2.100}$$

At the onset of Stage 2, $t = t_1$, the low feature barely contacts the pad: $h_s(t_1) = \delta$. Therefore, the general solution for step-height in Stage 2, $t_1 < t < t_2$, is expressed as:

$$h_s(t) = \delta exp\left(-\frac{t - t_1}{\tau_2}\right) \tag{2.101}$$

$$\tau_2 = \left(\frac{1}{k_{p_{Cu}} Y_p v_R}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right) \tag{2.102}$$

The analysis is the exactly the same as that of the discrete pad model. Thus, once the δ, t_1, and τ_2 are solved, the final heights, h_h and h_l, step-height, h_s, polishing time, t_2, at the end of Stage 2 can be obtained by replacing H_o by $\lambda_a^2 / 6\pi R_a$ and E_p by Y_p in the discrete pad model.

Stage 3

The pressure and the material removal rate in the oxide and Cu regions in Stage 3 can be represented by the step-height in the same way as in Eqs. (2.71) and (2.72).

$$\bar{p}_h(t) = p_{av}\left[1 + (w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right]$$
$$\bar{p}_l(t) = p_{av}\left[1 - (1 - w / \lambda)\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)h_s(t)\right] \tag{2.103}$$

and,

$$\frac{dh_h}{dt} = -k_{p_{Cu}} p_{av} v_R \left[1 + (w/\lambda) \left(\frac{Y_p}{p_{av}} \right) \left(\frac{6\pi R_a}{\lambda_a^2} \right) h_s(t) \right]$$
$$\frac{dh_l}{dt} = -k_{p_{Cu}} p_{av} v_R \left[1 - (1 - w/\lambda) \left(\frac{Y_p}{p_{av}} \right) \left(\frac{6\pi R_a}{\lambda_a^2} \right) h_s(t) \right] \quad (2.104)$$

The step-height, h_s, in Stage 3 is expressed by the first-order ordinary differential equation:

$$\frac{dh_s}{dt} + \left\{ \left[(1 - w/\lambda) k_{p_{Cu}} + (w/\lambda) k_{p_{ox}} \right] p_{av} v_R \left(\frac{Y_p}{p_{av}} \right) \left(\frac{6\pi R_a}{\lambda_a^2} \right) \right\} h_s = \left[k_{p_{Cu}} - k_{p_{ox}} \right] p_{av} v_R \quad (2.105)$$

At the onset of Stage 3, $t = t_2$, there is an initial step-height, $h_s(t_2)$. Therefore, the general solution for step-height in stage 3, $t \geq t_2$:

$$h_s(t) = h_s(t_2) + \left[h_s(\infty) - h_s(t_2) \right] \left\{ 1 - exp\left[-\left(\frac{t - t_2}{\tau_3} \right) \right] \right\} \quad (2.106)$$

$$h_s(\infty) \equiv \left[\frac{S_{Cu/ox} - 1}{(1 - w/\lambda) S_{Cu/ox} + w/\lambda} \right] \left(\frac{p_{av}}{Y_p} \right) \left(\frac{\lambda_a^2}{6\pi R_a} \right) \quad (2.107)$$

$$\tau_3 \equiv \left[\frac{S_{Cu/ox}}{(1 - w/\lambda) S_{Cu/ox} + w/\lambda} \right] \left(\frac{1}{k_{p_{Cu}} Y_p v_R} \right) \left(\frac{\lambda_a^2}{6\pi R_a} \right) \quad (2.108)$$

where $h_s(\infty)$ is the asymptotic step-height as $t \to \infty$ and τ_3 the time constant of stage 3.

2.5.3 Sample Calculation

Results of step-height calculation for various pattern geometry based on the simplified rough pad model are listed in Table 2.5 and 2.7. Tables 2.6 and 2.8, and Fig. 2.9 show the time evolution of pressure, material removal rate, height at high and low feature and the step-height based on the simplified rough pad model for both elastic and plastic contact conditions.

The initial relative pad deformation, δ, in the simplified rough pad model is greater than that of the discrete model. However, the ratio of the polishing interval to the time constant, t_2^* and t_o^*, is still greater than 4 but it is smaller than that of the discrete pad model. Thus, the step-height decreases in Stage 2 and increases in Stage 3 slower than that in the discrete case. The step-height increases until the polishing endpoint and the material removal rates at both the high and low features also approach to the asymptotic material removal rate, MRR_∞.

Table 2.5 Step-height based on the simplified rough pad model (elastic contact): w = 50 μm and α = 1.

w/λ	δ (nm)	t_1 (s)	t_2 (s)	$h_s(t_2)$ (nm)	$h_s(\infty)$ (nm)	$h_s(t_{ep})$ (nm)	$h_h(t_{ep})$ (nm)	$h_l(t_{ep})$ (nm)
0.1	402	107	339	11	331	262	-51	-313
0.2	435	89	318	12	356	289	-63	-352
0.3	475	71	297	13	386	320	-78	-397
0.4	527	53	276	15	423	355	-93	-448
0.5	595	37	254	17	471	398	-110	-508
0.6	690	20	234	19	535	451	-139	-590
0.7	836	5	213	22	626	522	-172	-694
0.8	1096	–	193	28	769	620	-222	-843
0.9	1739	–	174	34	1037	771	-302	-1073

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}}$ = 3.31x10^{-13} Pa^{-1}	E_p = 300 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}}$ = 0.23x10^{-13} Pa^{-1}	ν_p = 0.3	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1	R_a = 6 μm	t_{ep} = 480 s
	p_{av} = 14 kPa	λ_a = 100 μm	
	v_R = 0.9 m/s		

Table 2.6 Evolution of step-height based on the simplified rough pad model (elastic contact): w = 50 µm, λ = 100 µm and α = 1.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\overline{p}_h$ (kPa)	$\overline{p}_l$ (kPa)	$\lvert dh_h/dt \rvert$ (nm/min)	$\lvert dh_l/dt \rvert$ (nm/min)	r_{a_h} (µm)	r_{a_l} (µm)	p_{a_h} (MPa)	p_{a_l} (MPa)
0	900	1500	600	28	0	500	0	1.54	0	54	0
30	650	1250	600	28	0	500	0	1.54	0	54	0
60	418	1009	591	25	3	447	53	1.49	0.73	52	26
90	257	803	547	21	7	376	124	1.40	0.97	49	34
120	156	628	472	18	10	328	172	1.34	1.08	47	38
150	94	472	378	17	11	297	203	1.30	1.14	45	40
180	57	329	271	16	12	279	221	1.27	1.18	44	41
210	34	192	158	15	13	267	233	1.25	1.20	44	42
240	21	60	40	15	13	260	240	1.24	1.21	43	42
270	72	-5	-77	16	12	20	214	1.28	1.16	45	41
300	158	-16	-174	18	10	23	172	1.34	1.08	47	38
330	223	-28	-252	20	8	26	140	1.38	1.01	48	35
360	274	-42	-316	21	7	27	116	1.41	0.95	49	33
390	313	-55	-369	23	5	29	98	1.43	0.90	50	31
420	345	-70	-415	23	5	29	84	1.45	0.85	51	30
450	369	-85	-454	24	4	30	74	1.46	0.81	51	28
480	388	-100	-489	24	4	31	65	1.47	0.78	51	27

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}	E_p = 300 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}	ν_p = 0.3	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1	R_a = 6 µm	t_{ep} = 480 s
	p_{av} = 14 kPa	λ_a = 100 µm	
	v_R = 0.9 m/s		

Table 2.7 Step-height based on the simplified rough pad model (plastic contact): w = 50 μm and $\alpha = 1$.

w/λ	δ (nm)	δ_o (nm)	t_1 (s)	τ_2 (s)	t_2 (s)	τ_3 (s)	t_2^*	t_o^*	$h_s(t_2)$ (nm)	$h_s(\infty)$ (nm)	$h_s(t_{ep})$ (nm)	$h_h(t_{ep})$ (nm)	$h_l(t_{ep})$ (nm)
0.1	149	149	162	32	340	35	5.5	4.5	1	137	135	-61	-197
0.2	167	167	141	32	319	39	5.5	4.6	1	153	151	-72	-223
0.3	191	191	119	32	297	44	5.5	4.6	1	172	170	-89	-259
0.4	223	223	98	32	276	51	5.5	4.4	1	198	195	-106	-301
0.5	267	267	76	32	254	60	5.5	4.1	1	232	228	-128	-356
0.6	334	334	54	32	232	73	5.5	3.7	1	281	274	-162	-435
0.7	446	446	33	32	211	92	5.5	3.2	2	355	340	-203	-544
0.8	668	668	11	32	189	125	5.5	2.5	2	484	444	-268	-712
0.9	1337	1337	0	32	179	196	5.6	1.6	3	758	611	-395	-1007

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}} = 3.31\text{x}10^{-13}$ Pa^{-1}	Y_p = 20 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}} = 0.23\text{x}10^{-13}$ Pa^{-1}	R_a = 6 μm	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1	λ_a = 100 μm	t_{ep} = 480 s
	p_{av} = 14 kPa		
	v_R = 0.9 m/s		

Table 2.8 Evolution of step-height based on the simplified rough pad model (plastic contact): w = 50 μm and λ = 100 μm and α = 1.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\overline{p}_h$ (kPa)	$\overline{p}_l$ (kPa)	$\lvert dh_h / dt \rvert$ (nm/min)	$\lvert dh_l / dt \rvert$ (nm/min)	r_{a_h} (μm)	r_{a_l} (μm)	p_{a_h} (MPa)	p_{a_l} (MPa)
0	1500	600	900	28	0	500	0	1.46	0	60	0
30	1250	600	650	28	0	500	0	1.46	0	60	0
60	1000	600	400	28	0	500	0	1.46	0	60	0
90	761	589	171	23	5	410	90	1.32	0.62	60	60
120	583	517	66	17	11	312	188	1.16	0.90	60	60
150	438	412	26	15	13	274	226	1.08	0.98	60	60
180	305	295	10	15	13	259	241	1.05	1.01	60	60
210	177	173	4	14	14	254	246	1.04	1.03	60	60
240	51	49	1	14	14	251	249	1.04	1.03	60	60
270	-13	-69	55	17	11	21	198	1.14	0.92	60	60
300	-25	-150	125	21	7	26	133	1.25	0.75	60	60
330	-39	-206	168	23	5	29	93	1.32	0.63	60	60
360	-54	-247	193	24	4	31	69	1.36	0.54	60	60
390	-69	-278	209	25	3	32	55	1.38	0.49	60	60
420	-85	-303	218	25	3	32	46	1.39	0.45	60	60
450	-101	-325	223	26	2	33	41	1.40	0.42	60	60
480	-118	-344	227	26	2	33	38	1.41	0.40	60	60

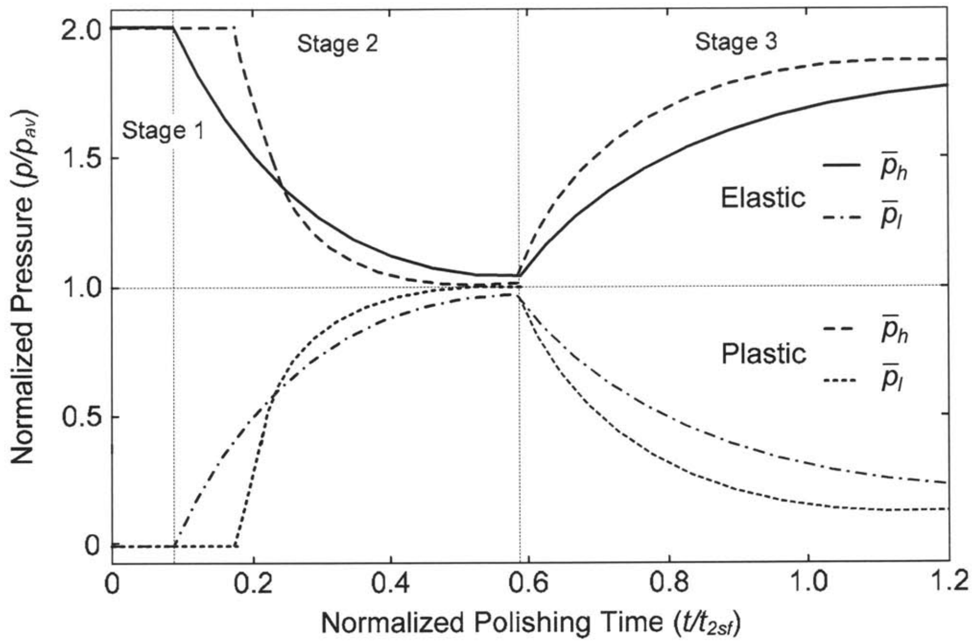
* Parameters include:

h_{si} = 900 nm
h_{Cu} = 1500 nm
h_I = 1000 nm

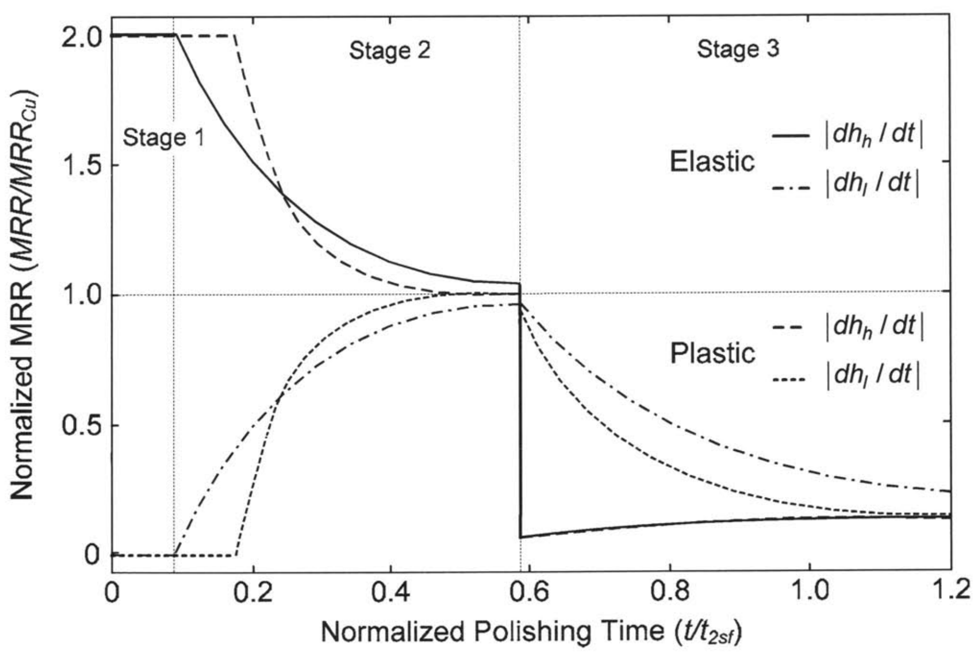
$k_{p_{Cu}} = 3.31\text{x}10^{-13}\ \text{Pa}^{-1}$
$k_{p_{ox}} = 0.23\text{x}10^{-13}\ \text{Pa}^{-1}$
$S_{Cu/ox}$ = 14.1
p_{av} = 14 kPa
v_R = 0.9 m/s

Y_p = 20 MPa
R_a = 6 μm
λ_a = 100 μm

t_{2f} = 360 s
t_{2sf} = 435 s
t_{ep} = 480 s

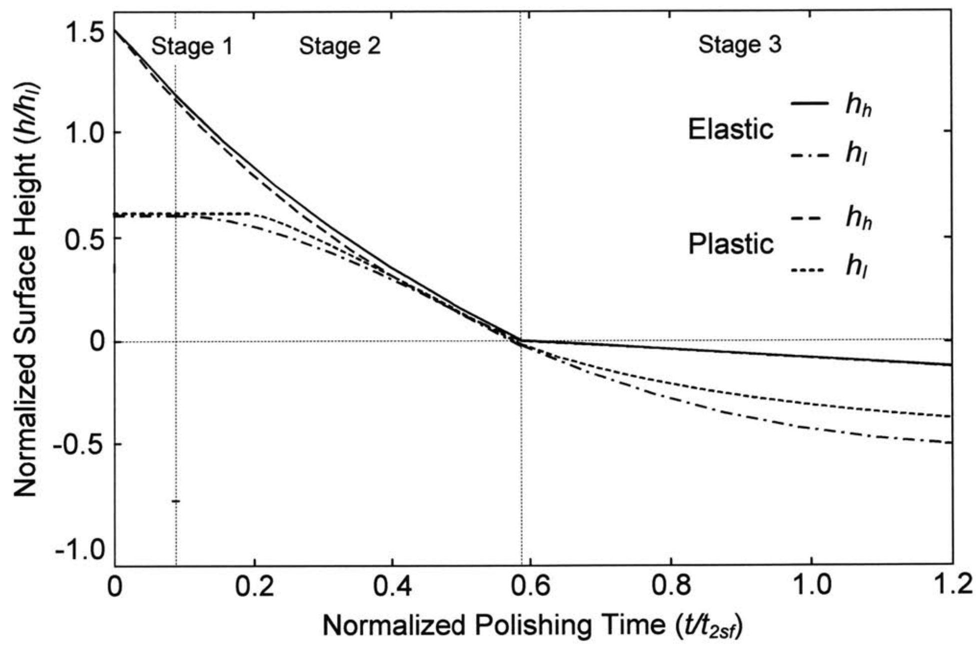


(a)

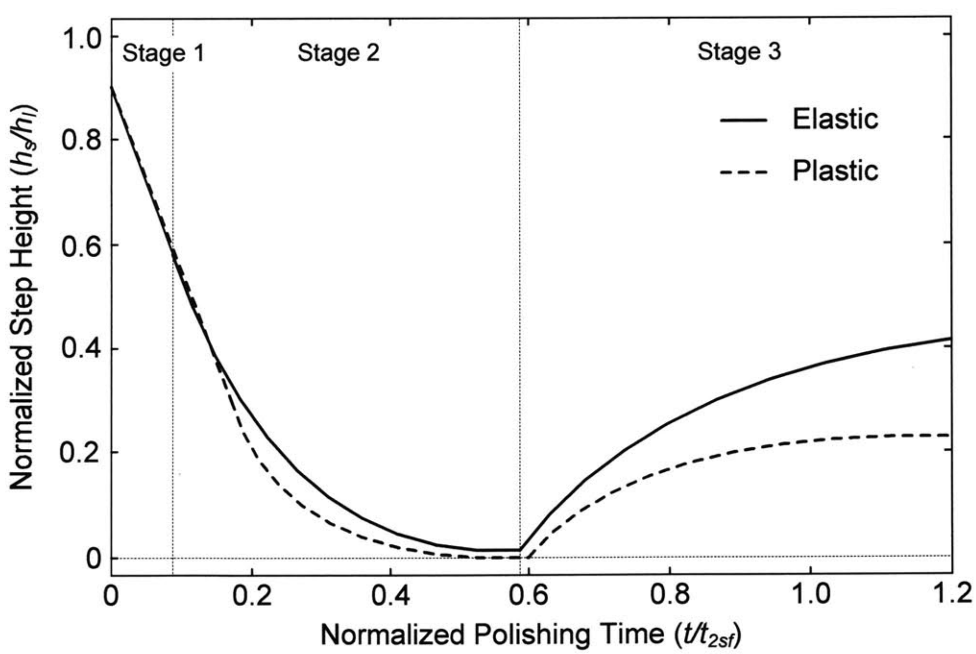


(b)

Figure 2.9 Time evolution of various parameters in the simplified rough pad model: (a) pressure, (b) material removal rate, (c) polished surface height at the high- and low-feature and (d) step-height (continued).



(c)



(d)

Figure 2.9 Time evolution of various parameters in the simplified rough pad model: (a) pressure, (b) material removal rate, (c) polished surface height at the high- and low-feature and (d) step-height.

2.6 The General Rough Pad Model

In this section, a general rough surface model is developed for step-height evolution in patterned wafer polishing. First, a rough surface contact on the blanket wafer is analyzed and a set of model parameters is determined. By using these parameters, the surface contact on the patterned wafer at a given step-height is analyzed.

2.6.1 Rough Surface Contact with a Blanket Wafer

Figure 2.10 shows the sample surface profile, asperity height distribution, and cumulative height distribution of a commercial pad. Although the polishing process wears out the pad asperities, the "continuously conditioned" pad profile is used in the analysis by assuming continuous conditioning of the pad during polishing.

From the profilometer measurements, the average surface roughness, R_{avg}, and the root-mean-square surface roughness, R_{RMS}, can be expressed as:

$$R_{avg} = \frac{1}{n}\sum_{i=1}^{n}|z_i| \tag{2.109}$$

$$R_{RMS} = \left(\frac{1}{n}\sum_{i=1}^{n} z_i^2\right)^{1/2} \tag{2.110}$$

where n is the total number of data and z_i the ith height from the mean line of the profile.

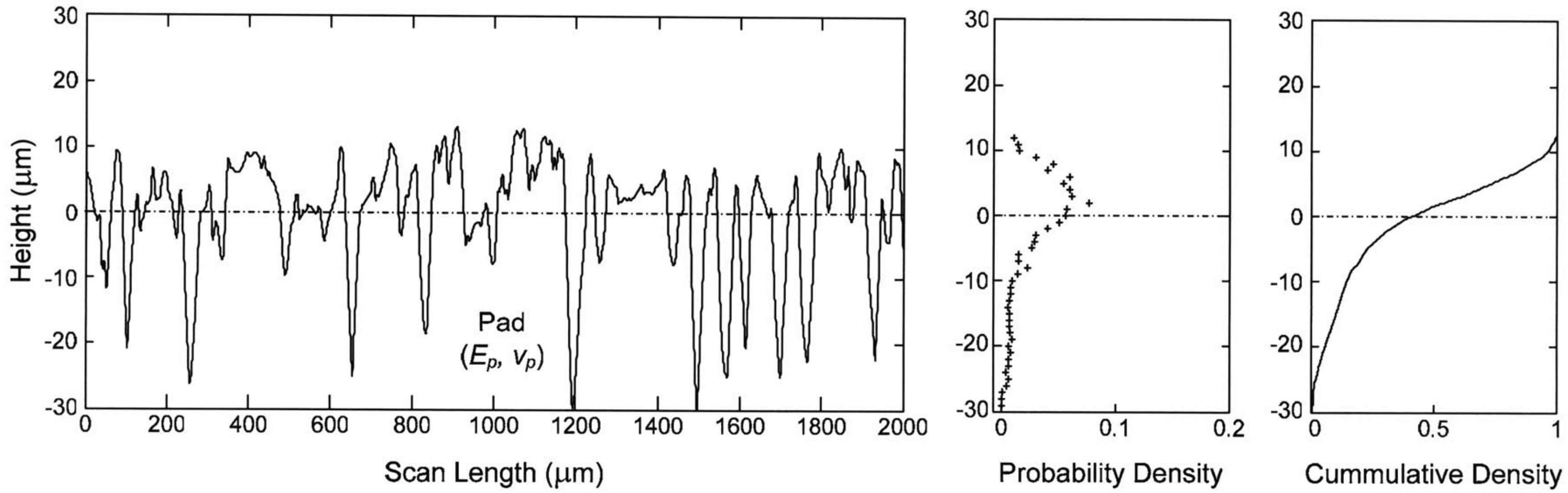
Throughout the analysis of the surface profilometer measurement, peak, summit and asperity are separately defined as follows.

- **Peak:** Points of convexity throughout a two-dimensional surface profile in the (x, z) plane, subscripted p.
- **Summit:** Points of convexity by extending two-dimensional surface profile to the three-dimensional (x, y, z) plane, subscripted s.
- **Asperity:** Summits actually in contact with the wafer, subscripted a.

The slope, m, and the curvature, κ, of the peaks are defined as:

$$m_i \equiv \frac{z_{i+1} - z_i}{\Delta x} \tag{2.111}$$

$$\kappa_i \equiv \frac{z_{i+1} - 2z_i + z_{i-1}}{\Delta x^2} \tag{2.112}$$



(a) (b) (c)

Figure 2.10 Topography of a commercial polishing pad (Rohm and Haas, IC1400): (a) surface profile, (b) asperity height distribution and (c) cumulative height distribution (continuously conditioned).

Peaks are decided by selecting data points where z_i is greater than both z_{i+1} and z_{i-1}. Figure 2.11 shows spatial distribution, peak height distribution, and cumulative height distribution of a commercial pad from the profile in Fig. 2.10. The distributions of surface height and pad peak heights are nearly Gaussian as shown in Fig. 2.12. In this case, the mean radius of curvature of summits and pad asperites, R_s and R_a, can be approximated as [Johnson, 1985]:

$$R_s = R_a \approx 1/\sigma_\kappa \tag{2.113}$$

where σ_κ is the standard deviation of the curvature of the peak defined by Eq. (2.112).

In the Gaussian distribution, the number of summits per unit area, η_s, can be estimated by the number of peaks per unit area, η_p, as:

$$\eta_s \approx 1.8\eta_p^2 \tag{2.114}$$

The number of summits, n_s, in the nominal area, A_o, are calculated as:

$$n_s = \eta_s A_o \tag{2.115}$$

Figure 2.13 shows a schematic of a general rough pad and the blanket wafer contact interface. Under no load, there will be only one contact between the pad and the wafer with a zero contact area. In this case, the separation between the mean line of the pad surface and the blanket wafer, d, is the maximum summit height, $z_{s_{\max}}$. As the load is increased, d decreases and the summits whose height z_s is greater than d start supporting the load. In this case, the total number of pad asperities in contact with the wafer in the nominal area A_o, n_a, can be calculated as:

$$n_a = n_s \int_d^{\infty} \phi(z_s)\, dz_s \tag{2.116}$$

Thus, the average spacing of the pad asperities in contact with the wafer, $\bar{\lambda}_a$, is defined as:

$$\bar{\lambda}_a = \left(\frac{1}{\eta_a}\right)^{1/2} \tag{2.117}$$

where η_a is the number of pad asperities per unit area, which is defined as $\eta_a = n_a / A_o$.

Each asperity is assumed to have a spherical shape. Let the deformation of the pad asperity be δ_a ,the radius of contact r_a. The asperity contact area A_a, and the load by each pad asperity F_a is a function of δ_a.

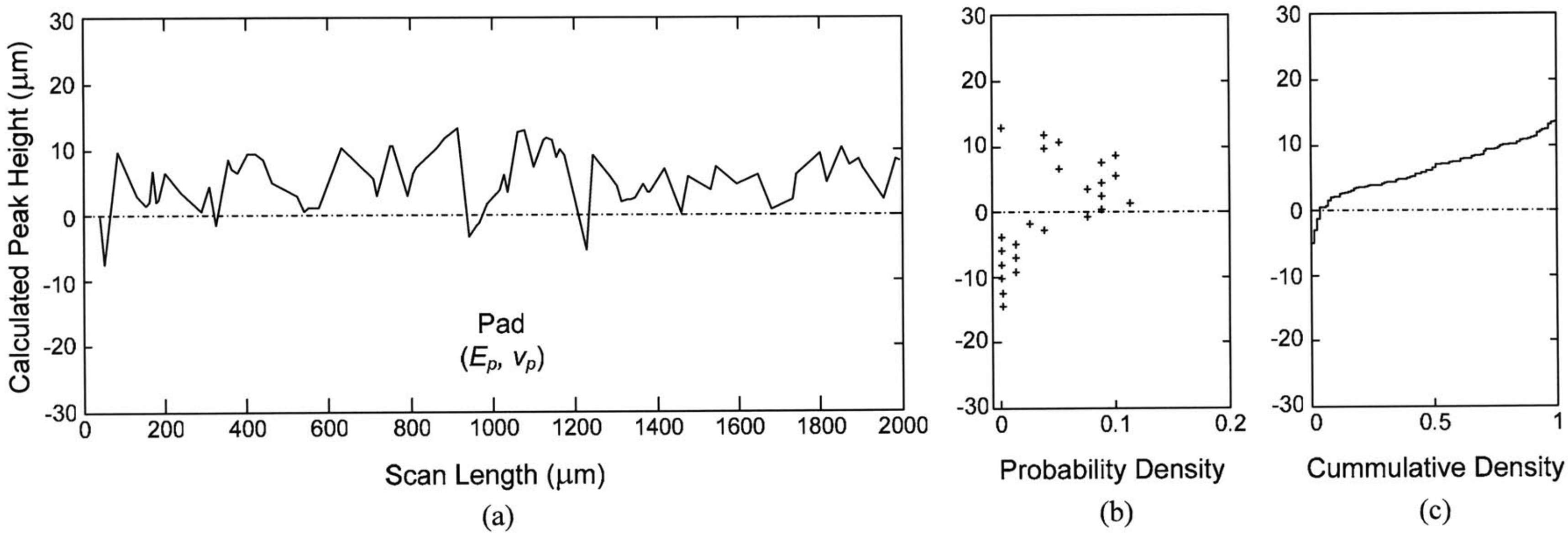


Figure 2.11 Plots of a commercial polishing pad (Rohm and Haas, IC1400): (a) calculated peak profile, (b) height distribution and (c) its cumulative height distribution (continuously conditioned).

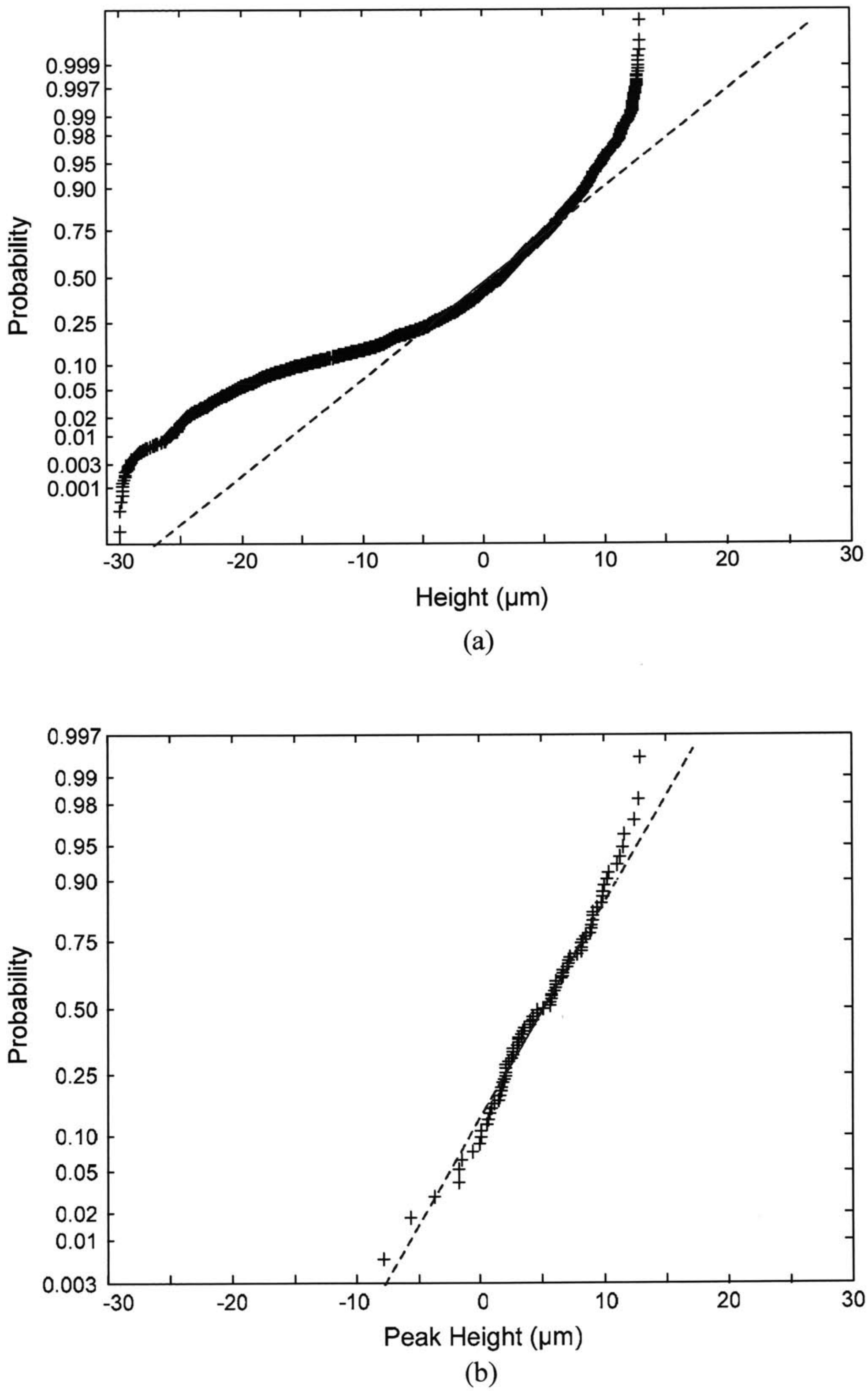


Figure 2.12 Normality plots of (a) the surface profile and (b) calculated peak profile of a commercial polishing pad (Rohm and Haas, IC1400).

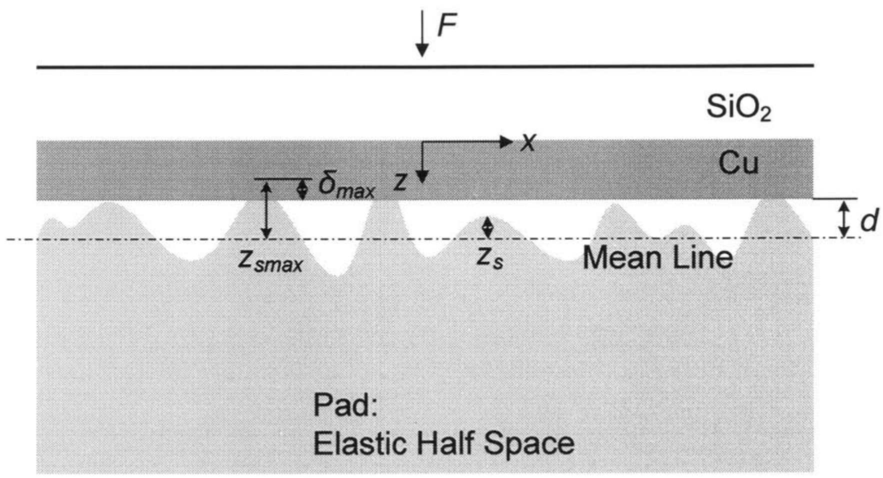


Figure 2.13 Schematic of the blanket wafer/rough pad contact interface. The vertical scale of the pad surface is greatly amplified for clarity.

$$A_a = \pi r_a^2 = f(\delta_a) \tag{2.118}$$

$$F_a = g(\delta_a) \tag{2.119}$$

The total area of contact, A_c, and the total applied load, F, when the separation between the mean line of the pad surface and the wafer is d can be expressed as:

$$A_c = n_s \int_d^{\infty} f(z_s - d) \cdot \phi(z_s) dz_s \tag{2.120}$$

$$F = \overline{p} A_o = n_s \int_d^{\infty} g(z_s - d) \cdot \phi(z_s) dz_s \tag{2.121}$$

where $\overline{p}$ is the average pressure in the nominal area A_o.

The mean pad asperity contact area, $\overline{A}_a$, and the mean applied load, $\overline{F}_a$, for each asperity in contact, are defined in indicial I_f and I_g, respectively.

$$I_f \equiv \overline{A}_a = \frac{A_c}{n_a} = \frac{\int_d^{\infty} f(z_s - d) \cdot \phi(z_s) dz_s}{\int_d^{\infty} \phi(z_s) dz_s} \tag{2.122}$$

$$I_g \equiv \overline{F}_a = \frac{F}{n_a} = \frac{\int_d^{\infty} g(z_s - d) \cdot \phi(z_s) dz_s}{\int_d^{\infty} \phi(z_s) dz_s} \tag{2.123}$$

Therefore, the mean contact radius, $\overline{r}_a$, and the mean contact pressure, $\overline{p}_a$, per each asperity in contact, can be expressed as:

$$\overline{r}_a = \left(\frac{I_f}{\pi} \right)^{1/2} \tag{2.124}$$

$$\overline{p}_a = \frac{I_g}{I_f} = \frac{\int_d^{\infty} g(z_s - d) \cdot \phi(z_s) dz_s}{\int_d^{\infty} f(z_s - d) \cdot \phi(z_s) dz_s} \tag{2.125}$$

The ratio of the total true contact area and the nominal area, A_c / A_o, is calculated by:

$$\frac{A_c}{A_o} = \frac{I_f}{I_g} \overline{p} \tag{2.126}$$

In the rough pad analysis, the index parameter I_f and I_g are determined by the asperity contact area, A_a, and the load by each pad asperity, F_a dependency on the pad deformation, which categorized either elastic or plastic contact.

Elastic contact

Assuming that the asperities are spherical and the contact is the Hertzian, the contact area and load per asperity are given as a function of the asperity deformation δ_a as:

$$A_a = \pi R_a \delta_a \tag{2.127}$$

$$F_a = \left(\frac{4}{3} R_a^{1/2} \delta_a^{3/2}\right) E^* \tag{2.128}$$

Although the complete surface profile data should be used in the full analysis, it is common to approximate the surface profile distribution either as Gaussian or as exponential as shown in Fig. 2.14. In the case of a Gaussian distribution,

$$\phi(z_s) = \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left(-\frac{z_s^2}{2\sigma_s^2}\right): \quad -\infty \leq z_s \leq \infty \tag{2.129}$$

In the exponential distribution case,

$$\phi(z_s) = \frac{1}{\sigma_s} \exp\left(-\frac{z_s}{\sigma_s}\right): \quad z_s > 0 \tag{2.130}$$

where σ_s is the standard deviation of the summit heights.

Figure 2.15 shows the summit distribution by the Gaussian assumption. Since only small portion of asperities whose height is large are actually in contact, the difference between the Gaussian and exponential distribution is small in the area of interest. If the asperity distribution is assumed to be exponential, the number of pad asperities can be expressed as:

$$n_a = \frac{n_s}{2} \exp\left(-\frac{d}{\sigma_s}\right) \tag{2.131}$$

Furthermore, the mean asperity contact area and the mean load per asperity are:

$$\overline{A}_a = I_f = \pi R_a \sigma_s \tag{2.132}$$

$$\overline{F}_a = I_g = \frac{4}{3} E^* R_a^{1/2} \sigma_s^{3/2} \tag{2.133}$$

The mean contact pressure and the ratio of the total contact area to the nominal area are:

$$\overline{p}_a = \frac{I_g}{I_f} = \left(\frac{\sigma_s}{\pi R_a}\right)^{1/2} E^* \tag{2.134}$$

$$\frac{A_c}{A_o} = \left(\frac{\pi R_a}{\sigma_s}\right)^{1/2} \left(\frac{\overline{p}}{E^*}\right) \tag{2.135}$$

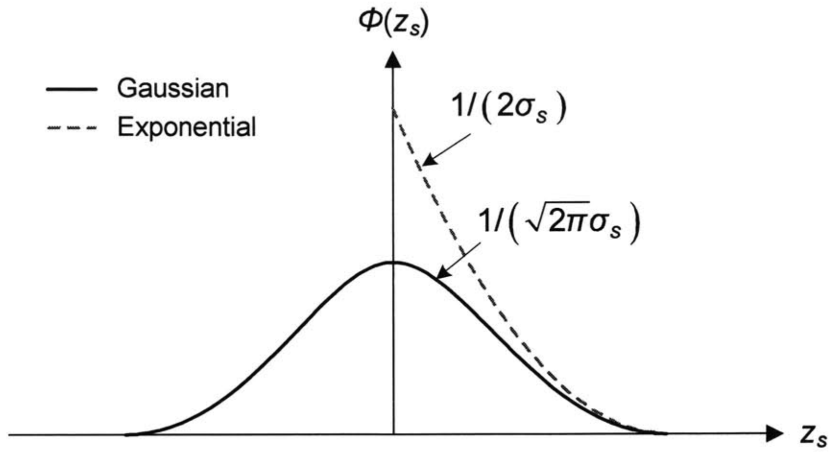


Figure 2.14 Schematics of Gaussian and exponential distribution of pad asperties.

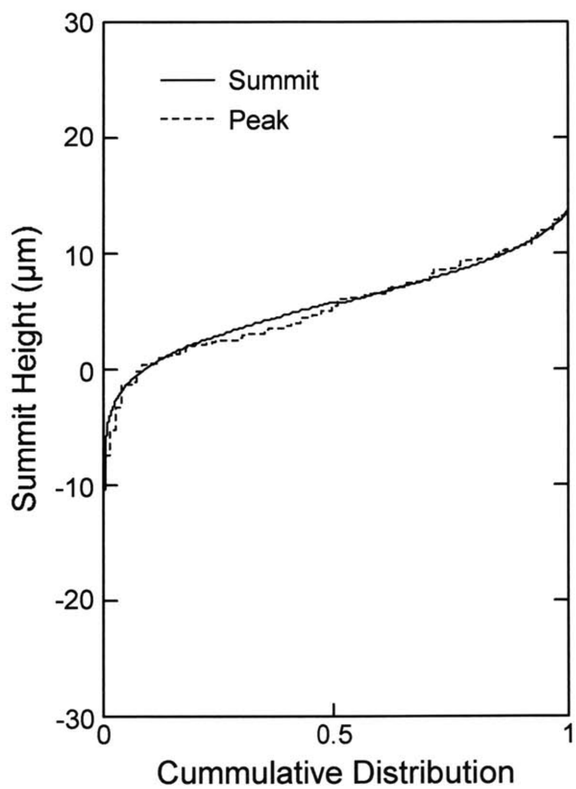


Figure 2.15 Pad asperity (summit) height distribution based on the same Gaussian distribution as the measured peak height distribution.

Fully Plastic contact

As the applied load increases, pad asperity deformation increases and the contact mode changes from elastic to plastic. Assuming that asperities are in perfectly plastic contact mode, the contact area and the load per asperity are calculated as:

$$A_a = 2\pi R_a \delta_a \tag{2.136}$$

$$F_a = (6\pi R_a \delta_a) Y_p \tag{2.137}$$

If pad asperity distribution is assumed exponential, the mean asperity contact area and the mean applied load per asperity in contact is expressed as:

$$\overline{A}_a = I_f = 2\pi R_a \sigma_s \tag{2.138}$$

$$\overline{F}_a = I_g = (6\pi R_a \sigma_s) Y_p \tag{2.139}$$

where Y_p is the yield strength of the asperity. The mean contact pressure and the ratio of the real contact area to the nominal contact area are:

$$\overline{p}_a = \frac{I_g}{I_f} = 3Y_p \tag{2.140}$$

$$\frac{A_c}{A_o} = \frac{I_f}{I_g}\overline{p} = \frac{\overline{p}}{3Y_p} \tag{2.141}$$

Here, the plastic contact pressure, $\overline{p}_a = 3Y_p$, is the hardness of the asperity, H_p.

To delineate the actual contact mode, the plasticity index, ψ, is defined as the mean contact pressure ratio between the elastic and fully plastic cases for the exponential distribution.

$$\psi \equiv \frac{\overline{p}_a)_{Elastic}}{\overline{p}_a)_{Plastic}} = \left(\frac{\sigma_s}{\pi R_a}\right)^{1/2} \frac{E^*}{3Y_p} \approx \left(\frac{\sigma_s}{\pi R_a}\right)^{1/2} \frac{E^*}{H_p} \tag{2.142}$$

The interpretation of the plasticity index in the rough surface contact is described by Greenwood and Williamson [Greenwood and Williamson, 1967]. If $\psi < 0.3$, the contact can be assumed to be in the elastic mode and if $\psi > 1$ in the full plastic mode. However, most polishing contacts are neither elastic nor fully plastic. Some portion of the pad asperities, with large summit heights, are in the plastic mode while others are in the elastic mode.

2.6.2 Rough Surface Contact with a Patterned Wafer

When a patterned wafer is in contact with the pad as shown in Fig. 2.16, the high and low features can be considered separately. The number of summits at the high and low features, n_{s_h} and n_{s_l}, respectively, can be expressed as:

$$\begin{aligned} n_{s_h} &= n_s\left(1 - A_f\right) \\ n_{s_l} &= n_s A_f \end{aligned} \tag{2.143}$$

The number of pad asperties, n_a, and contact area and contact force indices, I_f and I_g, in Eqs. (2.116), (2.122) and (2.123) will be represented as a function of the gap between the pad and the wafer, d: $n_a = n_a(d)$, $I_f = I_f(d)$ and $I_g = I_g(d)$. Assuming that the summits at the high and low features have the same statistical properties, the number of pad asperities in contact with the high and low features for a step-height, h_s, can be expressed as:

$$\begin{aligned} n_{a_h} &= n_{s_h}\int_d^{\infty}\phi(z_s)\,dz_s = \left(1 - A_f\right)\cdot n_a(d) \\ n_{a_l} &= n_{s_l}\int_{d+h_s}^{\infty}\phi(z_s)\,dz_s = A_f\cdot n_a(d+h_s) \end{aligned} \tag{2.144}$$

The asperity contact area at the high and low features are calculated as:

$$\begin{aligned} A_{c_h} &= n_{s_h}\int_d^{\infty} f(z_s - d)\cdot\phi(z_s)\,dz_s = \left(1 - A_f\right)\cdot n_a(d)\cdot I_f(d) \\ A_{c_l} &= n_{s_l}\int_{d+h_s}^{\infty} f(z_s - d)\cdot\phi(z_s)\,dz_s = A_f\cdot n_a(d+h_s)\cdot I_f(d+h_s) \end{aligned} \tag{2.145}$$

The total applied load at the high and low features are expressed as:

$$\begin{aligned} F_h &= n_{s_h}\int_d^{\infty} g(z_s - d)\cdot\phi(z_s)\,dz_s = \left(1 - A_f\right)\cdot n_a(d)\cdot I_g(d) \\ F_l &= n_{s_l}\int_{d+h_s}^{\infty} g(z_s - d)\cdot\phi(z_s)\,dz_s = A_f\cdot n_a(d+h_s)\cdot I_g(d+h_s) \end{aligned} \tag{2.146}$$

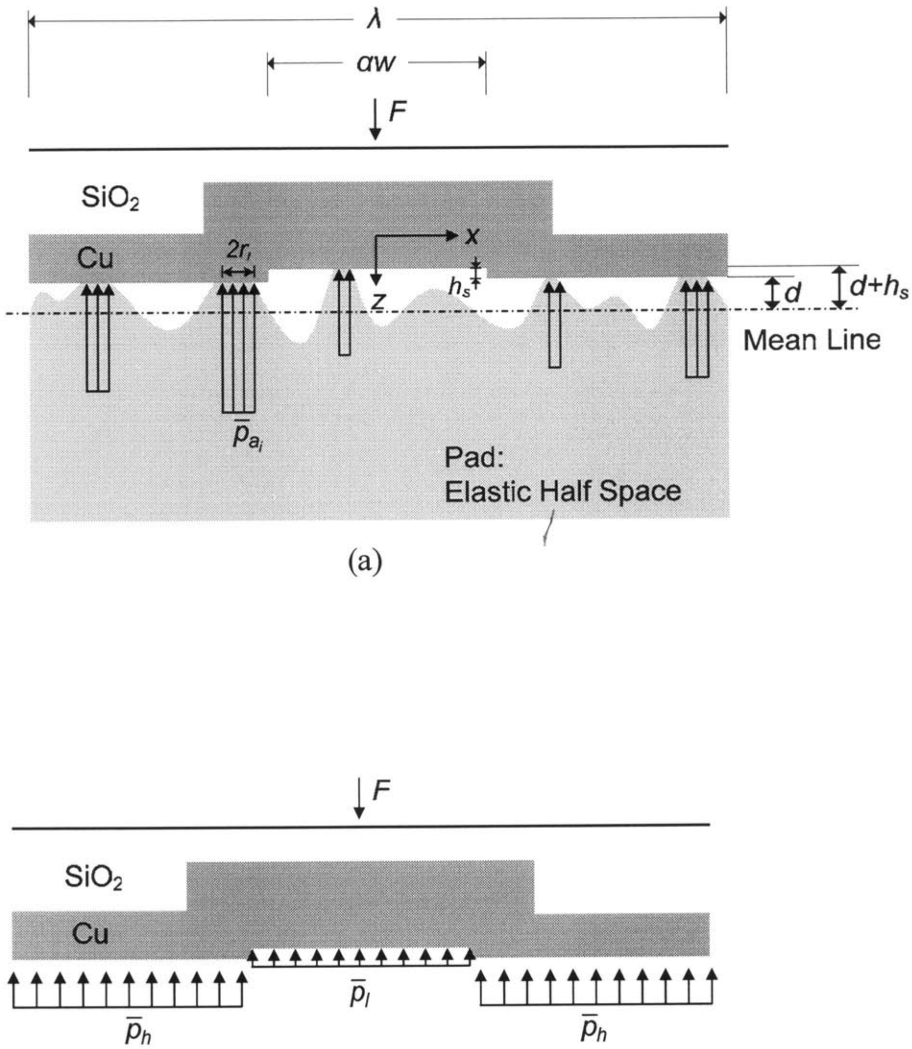
Therefore, the total applied load at the nominal area is represented as:

$$F = \left(1 - A_f\right)\cdot n_a(d)\cdot I_g(d) + A_f\cdot n_a(d+h_s)\cdot I_g(d+h_s) \tag{2.147}$$

The mean pressures at the high and low features are expressed as:

$$\begin{aligned} \overline{p}_h &= \frac{F_h}{A_h} = \frac{n_a(d)I_g(d)}{A_o} \\ \overline{p}_l &= \frac{F_l}{A_l} = \frac{n_a(d+h_s)I_g(d+h_s)}{A_o} \end{aligned} \tag{2.148}$$

where A_o is the nominal area, typically $A_o = \lambda\cdot L$.



(a)

(b)

Figure 2.16 Schematics of the pattern/rough pad contact interface: (a) asperity contact pressure and radius of contact and (b) average pressure at the high and low features.

Finally, the material removal rates at the high and low features can be expressed by the Preston equation:

$$\begin{aligned}\frac{dh_h}{dt} &= -k_{p_h} p_{av} v_R \left[\frac{n(d) I_g(d)}{A_o p_{av}}\right] \\ \frac{dh_l}{dt} &= -k_{p_l} p_{av} v_R \left[\frac{n(d+h_s) I_g(d+h_s)}{A_o p_{av}}\right]\end{aligned} \tag{2.149}$$

where h_h and h_l are the polishing surface heights of the high and low features relative to the top of the oxide, k_{p_h} and k_{p_h} are the Preston constants, and $\overline{p}_h$ and $\overline{p}_l$ are the average pressures at the high and low features as shown in Fig. 2.16(b). The step-height, $h_s(t)$, at any given time t is the height difference between the high and low features and can be solved numerically.

2.6.3 Sample Calculation

Table 2.9 shows the statistical topography parameters of the Rohm and Haas IC1400 pad before polishing. Tencor P10 profilometer was used to measure the surface profile; scan lengh was 2 mm. From the analysis, the average radius of curvature of the summit is 6.0 μm. Based on the parameters in Table 2.9, the contact between the blanket wafer and the pad was analyzed as in Table 2.10 and Fig. 2.17. The applied pressure, mean asperity contact pressure, mean asperity deformation, mean contact radius, mean asperity spacing and plasticity index are plotted in terms of the maximum asperity deformation. The maximum asperity deformation can be calculated for a known average applied pressure. For example, for p_{av} = 14 kPa (2 psi), the maximum asperity deformation in case of elastic contact is 0.85 μm. Once the maximum deformation of the contact is determined, the other parameters listed in Fig. 2.17 can be determined. In this case, $\overline{p}_a$ = 40 MPa, $\overline{\delta}_a$ = 0.45 μm, $\overline{r}_a$ = 1.15 μm, $\overline{\lambda}_a$ = 120 μm. The plasticity index ψ is about 0.7, which implies that there are both elastic and plastic contacts between the pad asperities and the wafer.

Given the relationship between the maximum pad deformation and the average applied pressure, the relationship between the step-height and the maximum asperity deformation, and the mean pressures at the high and the low feature can be expressed as shown in Fig. 2.18. Thus, at a given step-height, the mean pressure and thus material removal rate can be calculated, Tables 2.11 and 2.12.

Figure 2.19 shows the time evolution of pressure, material removal rate, height of the high

Table 2.9 Surface topography parameters of the IC1400 pad.

Property	Value
Scan Length, L (μm)	2,000
Average Surface Roughness, R_{avg} (μm)	6.7
RMS Surface Roughness, R_{RMS} (μm)	8.8
Maximum Surface Roughness, $R_{\max}$ (μm)	43
Number of Peaks, n_p	87
Maximum Peak Height, $z_{p_{\max}}$ (μm)	13.56
Number of Summits, n_s	14,000
Average Summit Radius of Curvature, R_s (μm)	6.0

Table 2.10 Contact area and pressure calculation based on the rough pad model on the blanket wafer.

Contact Type	$\delta_{\max}$ (μm)	d (μm)	$\overline{p}$ (kPa)	$\overline{p}_a$ (MPa)	$\overline{\delta}$ (μm)	n_a	A_c / A_o (x10^{-3})	$\overline{r}_a$ (μm)	$\overline{\lambda}_a$ (μm)
Elastic	0	13.6	0	0	0	0	0	0	–
	0.5	13.1	3	32	0.24	102	0.08	0.9	198
	1.0	12.6	19	43	0.44	253	0.34	1.2	126
	1.5	12.1	58	52	0.64	446	0.88	1.5	95
	2.0	11.6	129	60	0.84	655	1.71	1.7	78
	2.5	11.1	244	67	1.04	906	2.92	1.9	66
	3.0	10.6	415	72	1.21	1215	4.56	2.0	57
	3.5	10.1	655	77	1.39	1559	6.71	2.2	51
	4.0	9.6	977	81	1.56	1944	9.42	2.3	45
	4.5	9.1	1397	86	1.73	2386	12.78	2.4	41
	5.0	8.6	1934	89	1.87	2908	16.88	2.5	37
Plastic	0	13.6	0	0	0	0	0	0	–
	0.5	13.1	3	60	0.24	102	0.15	1.3	198
	1.0	12.6	19	60	0.44	253	0.69	1.7	126
	1.5	12.1	58	60	0.64	446	1.76	2.1	95
	2.0	11.6	129	60	0.84	655	3.43	2.4	78
	2.5	11.1	244	60	1.04	906	5.84	2.7	66
	3.0	10.6	415	60	1.21	1215	9.13	2.9	57
	3.5	10.1	655	60	1.39	1559	13.41	3.1	51
	4.0	9.6	977	60	1.56	1944	18.84	3.3	45
	4.5	9.1	1397	60	1.73	2386	25.56	3.4	41
	5.0	8.6	1934	60	1.87	2908	33.77	3.5	37

* Parameters include: L = 2 mm, E_p = 300 MPa, ν_p = 0.3 and Y_p = 20 MPa.

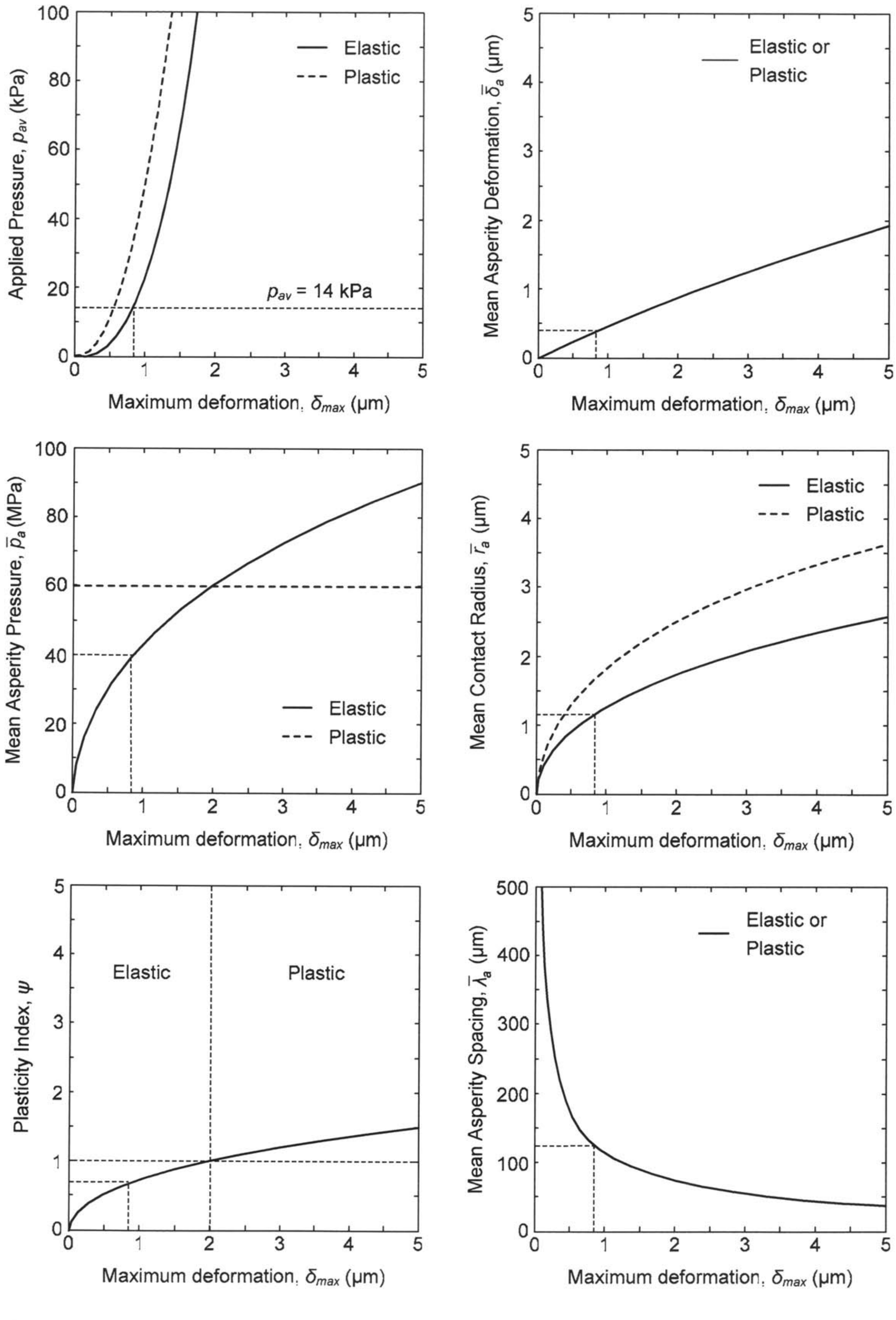


Figure 2.17 Results of rough pad contact model on the blanket wafer. Parameters include: L = 2 mm, $E_p = 300$ MPa, $\nu_p = 0.3$ and $Y_p = 20$ MPa.

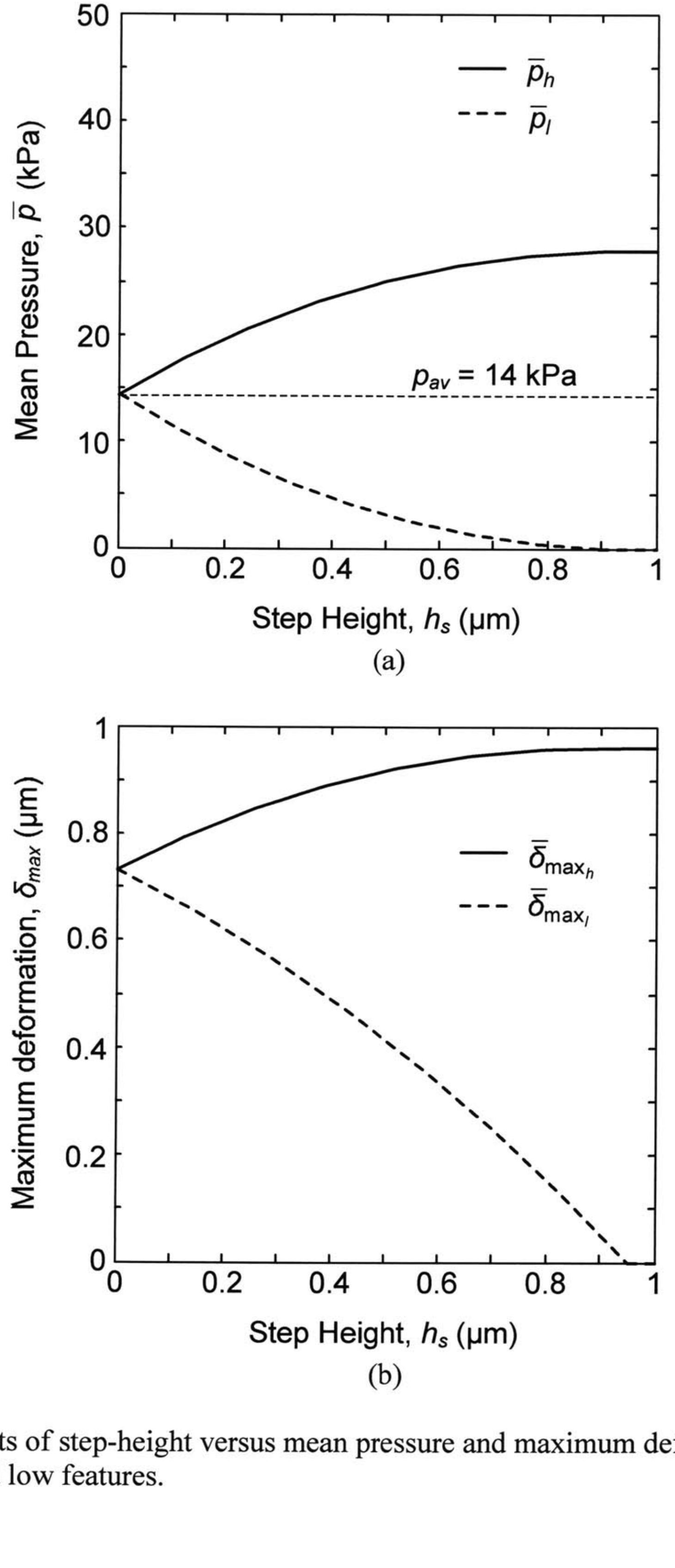


Figure 2.18 Plots of step-height versus mean pressure and maximum deformation at the high and low features.

Table 2.11 Step-height calculation based on the general rough pad model: w =50 μm and α = 1.

Contact Mode	w/λ	δ (nm)	t_2 (s)	$h_s(t_2)$ (nm)	$h_s(\infty)$ (nm)	$h_s(t_{ep})$ (nm)	$h_h(t_{ep})$ (nm)	$h_l(t_{ep})$ (nm)
	0.1	770	339	24	490	309	-53	-362
	0.2	800	319	25	520	338	-63	-401
	0.3	850	297	26	550	373	-73	-446
	0.4	900	276	27	580	410	-90	-500
Elastic	0.5	960	256	30	610	450	-108	-558
	0.6	1050	235	32	650	498	-133	-632
	0.7	1160	215	35	710	556	-169	-726
	0.8	1350	194	39	790	634	-221	-856
	0.9	1730	175	43	920	745	-314	-1061
	0.1	500	338	12	360	267	-51	-318
	0.2	530	317	13	380	291	-63	-354
	0.3	560	296	13	400	320	-75	-395
	0.4	600	276	14	430	351	-91	-442
Plastic	0.5	660	254	15	470	388	-112	-499
	0.6	730	234	18	510	432	-140	-572
	0.7	840	213	19	560	489	-177	-665
	0.8	1010	192	23	660	566	-238	-804
	0.9	1370	171	26	800	683	-338	-1021

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}}$ = 3.31x10^{-13} Pa^{-1}	E_p = 300 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}}$ = 0.23x10^{-13} Pa^{-1}	ν_p = 0.3	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1	Y_p = 20 MPa	t_{ep} = 480 s
	p_{av} = 14 kPa		
	v_R = 0.9 m/s		

Table 2.12 Evolution of step-height based on the general rough pad model (elastic contact): w = 50 μm, λ = 100 μm and α = 1.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\bar{p}_h$ (kPa)	$\bar{p}_l$ (kPa)	$\lvert dh_h / dt \rvert$ (nm/min)	$\lvert dh_l / dt \rvert$ (nm/min)
0	900	1500	600	28	0	497	0
30	658	1254	596	26	1	470	22
60	454	1027	573	24	4	420	72
90	303	827	524	21	7	373	125
120	198	649	452	19	9	340	168
150	129	490	361	17	11	310	197
180	83	342	259	16	12	281	212
210	54	204	150	15	13	272	228
240	34	68	33	15	13	263	237
270	76	-7	-83	16	12	20	212
300	164	-17	-181	18	10	23	182
330	235	-29	-264	20	9	25	155
360	292	-42	-334	21	7	26	130
390	338	-55	-393	22	6	27	108
420	376	-69	-445	23	6	29	98
450	408	-84	-492	23	5	29	85
480	435	-99	-533	24	4	30	80

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}}$ = 3.31x10^{-13} Pa^{-1}	E_p = 300 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}}$ = 0.23x10^{-13} Pa^{-1}	ν_p = 0.3	t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1		t_{ep} = 480 s
	p_{av} = 14 kPa		
	v_R = 0.9 m/s		

Table 2.13 Evolution of step-height based on the general rough pad model (plastic contact): w = 50 μm, λ = 100 μm and α = 1.

Time (s)	h_s (nm)	h_h (nm)	h_l (nm)	$\bar{p}_h$ (kPa)	$\bar{p}_l$ (kPa)	$\lvert dh_h / dt \rvert$ (nm/min)	$\lvert dh_l / dt \rvert$ (nm/min)
0	900	1500	600	28	0	501	0
30	650	1250	600	28	0	501	0
60	417	1008	591	25	3	454	49
90	252	801	549	21	7	383	123
120	149	625	476	18	10	330	172
150	88	469	381	16	11	294	200
180	52	327	275	15	12	271	219
210	31	192	161	15	13	271	239
240	18	60	42	15	13	260	239
270	70	-6	-76	16	12	20	209
300	154	-17	-171	18	10	23	172
330	218	-29	-247	21	8	26	139
360	267	-43	-309	21	6	27	109
390	304	-56	-361	22	5	28	96
420	334	-71	-405	23	5	29	83
450	359	-86	-445	24	4	30	71
480	377	-101	-478	25	4	31	65

* Parameters include:

h_{si} = 900 nm	$k_{p_{Cu}}$ = 3.31x10^{-13} Pa^{-1}	Y_p = 20 MPa	t_{2f} = 360 s
h_{Cu} = 1500 nm	$k_{p_{ox}}$ = 0.23x10^{-13} Pa^{-1}		t_{2sf} = 435 s
h_I = 1000 nm	$S_{Cu/ox}$ = 14.1		t_{ep} = 480 s
	p_{av} = 14 kPa		
	v_R = 0.9 m/s		

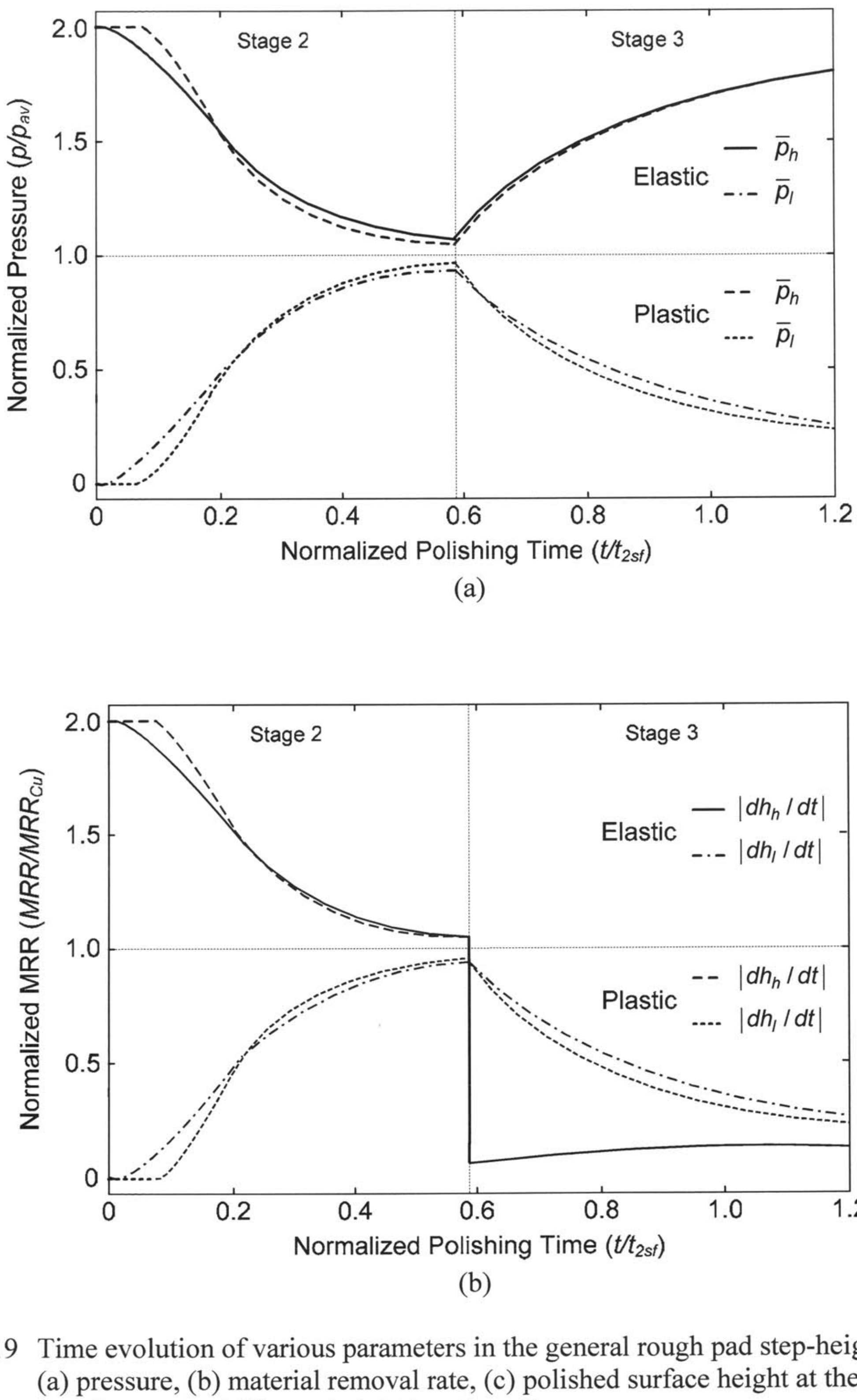
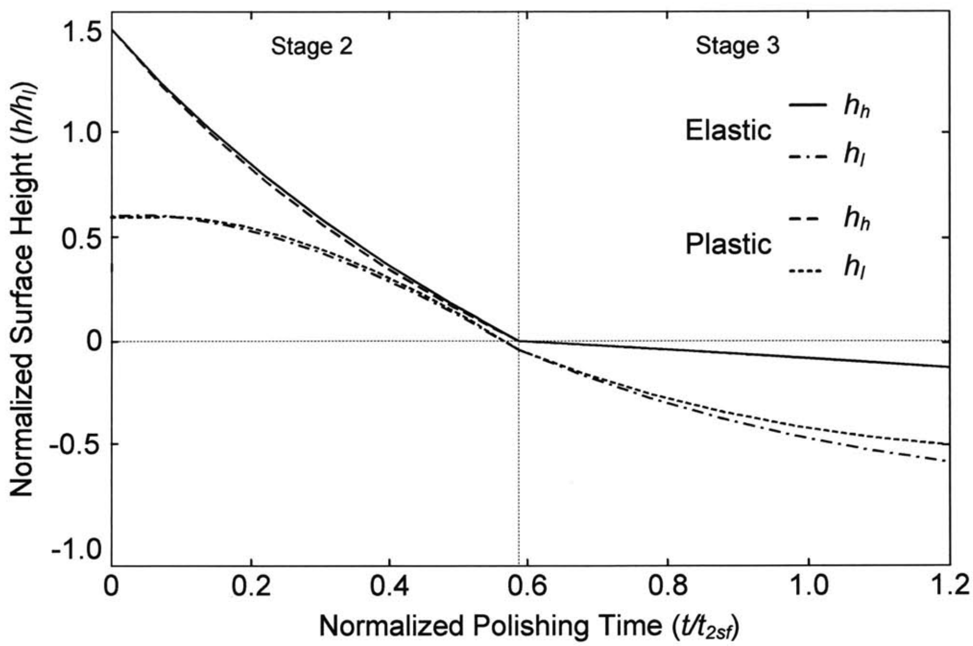
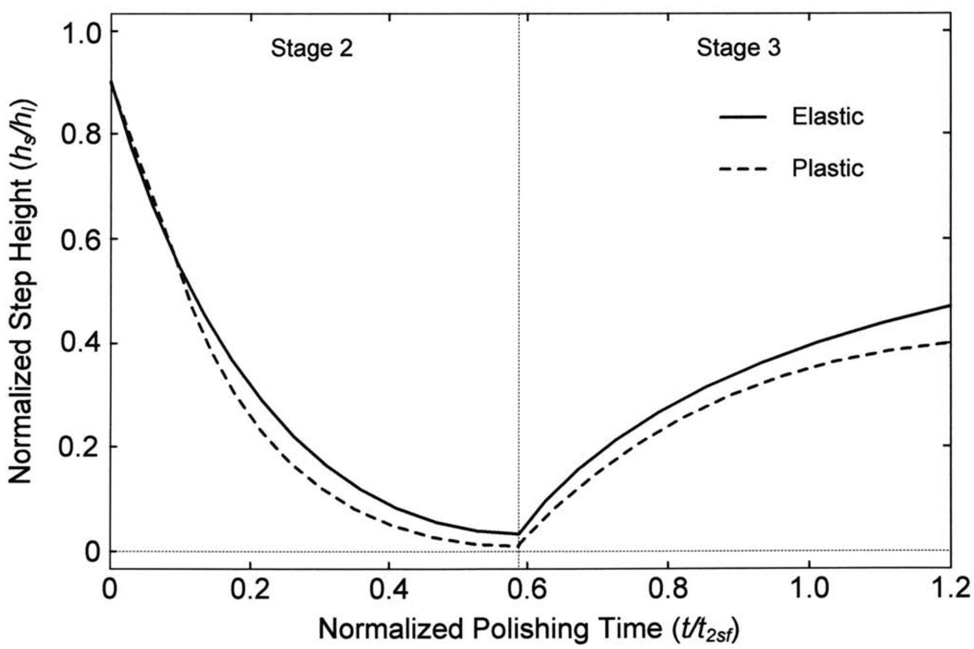


Figure 2.19 Time evolution of various parameters in the general rough pad step-height model: (a) pressure, (b) material removal rate, (c) polished surface height at the high and low features and (d) step-height (continued).



(c)



(d)

Figure 2.19 Time evolution of various parameters in the general rough pad step-height model: (a) pressure, (b) material removal rate, (c) polished surface height at the high and low features and (d) step-height.

and low features and the step-height based on the general rough pad model. If the initial relative pad deformation is higher than the initial step-height, $\delta > h_{si}$, the process starts from Stage 2. As polishing progresses, the step-height decreases. The step-height at the end of Stage 2, $h_s(t_2)$, is greater than that in the smooth-pad or discrete-pad model. Furthermore, the step-height increases much faster than in the smooth-pad and the discrete-pad cases, and the asymptotic value of the step-height in Stage 3 is also greater than the other two cases.

2.6.4 Comparison with Previous Models

The effect of pad roughness on the material removal rate in CMP has been addressed by several researchers [Yu *et al.*, 1993; Vlassak, 2001 & 2004; Nguyen *et al.*, 2003; Seok *et al.*, 2003]. In these models, an in the present model, the pad/wafer contact is analyzed by the Greenwood and Williamson approach.

Nguyen *et al.* derived a dishing model based on the rough pad surface with Gaussian pad asperity distribution. In their model, however, evolution of high-feature was neglected and thus step-height reduction rate was assumed to be directly proportional to the pressure on Cu area only. As a result, this model is valid only a high-selectivity slurry is used. Moreover, the model considers only the overpolishing stage and does not address the evolution of step-height. Furthermore, it assumes the mean contact pressure is the same, regardless of the area fraction of Cu interconnects.

Vlassak's model requires a full profile of pad deformation in the nominal area to calculated average pressure profile and thus calculation is intensive. Additionally, due to the smooth pad contact mechanics model for the average pressure, dishing is almost linearly proportional to the interconnect linewidth, which is valid only in the intermediate-wiring level. In Cu CMP, for example, in the global wiring level where linewidth is larger than 100 μm, Cu dishing remains almost constant regardless of the linewidth and mainly depends on the amount of overpolishing.

In the model by Seok *et al.*, the time evolution of a feature is also analyzed by using the iterative numerical analysis from the complete wafer surface profile. This model heavily relies on the finite element analysis and thus requires intense computation. How to determine each model parameter and its physical meaning are not fully discussed, however.

The present rough pad model uses the blanket wafer analysis in the patterned wafer analysis. The mean pressures at the high and low features are solved by combining the force balance equation and index functions from the blanket wafer analysis. This model is expected to reduce the computational time dramatically compared with the other models, since it only needs a blanket wafer analysis and simple time-step iterative calculation at the high and low features, not based on the full profile of wafer surface.

2.7 Summary

In this chapter, feature-scale step-height models based on various pad/wafer contact conditions are developed. First, contact between the wafer and the pad is analyzed on the assumption that the polishing pad is elastic and smooth. This assumption is valid only in the submicron device level where Cu linewidth is smaller than the diameter of asperity contact. This model explains the small step-height and "steady-state" material removal behavior in the submicron level. Second, the pad is assumed to deform as discrete blocks. This model explains the step-height reduction behavior in the global wiring level fairly well, but relies on the finite thickness and the physically inadmissible discontinuous deformation of the pad. Third, a simplified rough pad model, which idealizes that the pad asperities are of uniform height and of the same radius of curvature, is developed. The mean asperity contact radius and the asperity spacing are estimated by the analysis based on the general random rough surface of the pad. Both elastic and plastic deformations of pad asperities are considered and the evolution of the step-height during polishing is followed. In the case of a plastic contact, the simplified rough pad model establishes the step-height reduction model in a simple algebraic form. Finally, the rough pad/wafer contact is analyzed by adapting the classical analysis by Greenwood and Williamson. Throughout the general rough pad model analysis based on a given pad surface profile, surface topography parameters for the simplified rough pad model are obtained.

For all the contact conditions listed earlier, the general procedure to calculate the step-height at the given polishing time is the same. First, the contact pressure distribution at a given step height is related either to the elastic deformation of the pad surface, in smooth pad assumption, or to the elastic/plastic deformation of pad asperities, in the rough pad analysis. Then, the evolution of step-height in a patterned wafer is calculated by integrating the effects of pad profile and the geometry of Cu interconnects at each polishing stage.

From the developed step-height models, integrated dielectric erosion and Cu dishing models will be presented in terms of geometric, material and process parameters in the next chapter.

Nomenclature

A_a = contact area of each pad asperity (m^2)

A_{Cu} = total Cu interconnect area in a characteristic area (m^2)

A_c = total pad/wafer contact area (m^2)

A_o = nominal area (m^2)

A_f = area fraction of Cu interconnects

$A_1, A_{hh}, A_{hl}, A_{lh}, A_{ll}$ = dimensionless geometric coefficients

$B_{11}, B_{12}, B_{21}, B_{22}$ = dimensionless geometric coefficients

E_p, E^* = Young's modulus and effective Young's modulus of a pad (N/m^2)

F = applied normal force (N)

$f(\delta), g(\delta)$ = functions of contact area and normal load per pad asperity

H_o = undeformed pad thickness (m)

h = film thickness (m)

h_{Cu}, h_b, h_I = film thickness of Cu, barrier layer and interconnect (m)

h_h, h_l = height of the high and the low features (m)

$h_s, h_{si}, h_s(\infty)$ = step-height, initial step-height and asymptotic step-height (m)

I_f, I_g = integral functions of contact area and normal load of asperities

k_p = Preston constant (m^2/N)

k_{p_h}, k_{p_l} = Preston constant of high and low features (m^2/N)

$k_{p_{Cu}}, k_{p_{ox}}$ = Preston constant of Cu and oxide (m^2/N)

MRR = material removal rate (m/s)

MRR_∞ = asymptotic material removal rate (m/s)

m = slope of peaks in pad profile

n_a, n_p, n_s = number of asperities, peaks and summits in a nominal area

p = pressure (N/m^2)

p_{av} = average pressure (N/m^2)

$\overline{p}_h, \overline{p}_l$ = mean pressure at high and low features (N/m^2)

$\overline{p}_{a_h}, \overline{p}_{a_l}$ = mean contact pressure at high and low features (N/m^2)

R_a, R_s = radius of curvature of asperity, summit (m)

R_{avg}, R_{RMS} = average and RMS surface roughnesses (m)

r_a, r_{a_h}, r_{a_l} = radius of pad asperity contact (m)

$S_{Cu/ox}, S_{b/ox}$ = Cu-to-oxide and barrier-to-oxide selectivities

t = polishing time (s)

t_{ep} = process endpoint (s)

t_1, t_2, t_o = polishing time at the end of Stage 1, Stage 2 and overpolishing time (s)

t_2^*, t_o^* = dimensionless polishing time of Stage 2 and Stage 3

$\overline{u}_z, \overline{u}_z'$ = displacement of pad surface (m)

v_R = magnitude of relative velocity (m/s)

w, w_s = designed Cu interconnect linewidth and "surface linewidth" (m)

x, y, z = Cartesian coordinates

Y_p = yield strength of pad material (N/m^2)

$\Delta h_{ff}, \Delta h_{sf}$ = material removed in the fastest and slowest fields (m)

Δh_o = amount of overpolishing at the wafe reference point (m)

α = Cu deposition factor

δ, δ_o = maximum deformation of pad (asperity) (m)

δ_a = deformation of each pad asperity (m)

δ_h, δ_l = deformation of pad (asperity) at the high and the low features (m)

$\phi(z_s)$ = frequency function of summit distribution

η_a, η_p, η_s = number of asperities, peaks, and summits per unit area (m^{-2})

κ = curvature of peaks (m^{-1})

λ = pitch of Cu interconnect lines (m)

λ_a = spacing of pad asperities (m)

μ_h = mean of the amount of material removed in a wafer (m)

ν_p = Poisson's ratio of pad material

σ_s = standard deviation of summit distribution (m)

$\sigma_x, \tau_{xy}, \tau_{xz}$ = normal and shear stress (N/m^2)

$\sigma_{\Delta h}$ = standard deviation of the amount of material removed in a wafer

τ_2, τ_3 = time constant of Stage 2 & 3 (s)

ψ = plasticity index

References

Ahmadi, G. and Xia, X., 2001, "A Model for Mechanical Wear and Abrasive Particle Adhesion During the Chemical Mechanical Polishing Process," *J. Electrochemical Soc.*, vol. 148, pp. 99-109.

Borucki, L., 2002, "Mathematical Modeling of Polish-Rate Decay in Chemical-Mechanical Polishing," *J. Engineering Mathematics*, vol. 43, pp. 105-114.

Chekina, O. G., Keer, L. M. and Liang, H., 1998, "Wear-Contact Problems and Modeling of Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 145, pp. 2100-2106.

Chen, D.-Z. and Lee B.-S., 1999, "Parameter Analysis of Chemical Mechanical Polishing: An Investigation Based on the Pattern Planarization Model," *J. Electrochemical Soc.*, vol. 146, pp. 3420-3424.

Chen, D.-Z. and Lee B.-S., 1999, "Pattern Planarization Model of Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 146, pp. 744-748.

Elbel, N., Neureither, B., Ebersberger, B. and Lahnor, P., 1998, "Tungsten Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 145, pp. 1659-1664.

Fu, G. and Chandra, A., 2003, "An Analytical Dishing and Step Height Reduction Model for Chemical Mechanical Planarization (CMP)," *IEEE Trans. on Semiconductor Manufacturing*, vol. 16, pp. 477-485.

Fu, G., Chandra, A., Guha, S. and Subhash, G., 2001, "A Plasticity-Based Model of Material Removal in Chemical-Mechanical Polishing (CMP)," *IEEE Trans. on Semiconductor Manufacturing*, vol. 14, pp. 406-417.

Greenwood, J. A. and Williamson, J. B. P., 1966, "Contact of Nominally Flat Surfaces," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 295, pp. 300-319.

Grillaert, J., Meuris, M., Heylen, N., Devriendt, K., Vrancken, E. and Heyns, M., 1998, "Modeling Step Height Reduction and Local Removal Rates Based on Pad-substrate Interactions," *Proc. CMP-MIC Conf.*, pp. 79-86.

Guo, Y., Chandra, A. and Bastawros, A.-F., 2004, "Analytical Dishing and Step Height Model for CMP With a Viscoelestic Pad," *J. Electrochemical Soc.*, vol. 151, pp. G583-G589.

Jindal, A. and Babu, S. V., 2004, "Effect of pH on CMP of Copper and Tantalum," *J. Electrochemical Soc.*, vol. 151, pp. G709-G716.

Johnson, K. L., 1985, **Contact Mechanics**, Cambridge University Press, Cambridge, UK.

Lai, J.-Y., Saka, N. and Chun, J.-H., 2002, "Evolution of Coppe-Oxide Damascene Structures in Chemical Mechanical Polishing, I - Contact Mechanics Modeling," *J. Electrochemical Soc.*, vol. 149, pp. G31-G40.

Lai, J.-Y., Saka, N. and Chun, J.-H., 2002, "Evolution of Coppe-Oxide Damascene Structures in Chemical Mechanical Polishing, II - Copper Dishing and Oxide Erosion," *J. Electrochemical Soc.*, vol. 149, pp. G41-G50.

Lai, J.-Y., 2001, "Mechanics, Mechanisms, and Modeling of the Chemical Mechanical Polishing Process," Ph.D. thesis, Department of Mechanical Engineering, MIT, Cambridge, MA.

Liu, C.-W., Dai B.-T., Tseng, W.-T. and Yeh, C.-F., 1996, "Modeling of the Wear Mechanism during Chemical-Mechanical Polishing," *J. Electrochemical Soc.*, vol. 143, pp. 716-721.

Luo, J. and Dornfeld, D. A., 2003, "Material Removal Regions in Chemical Mechanical Planarization for Submicron Integrated Circuit Fabrication: Coupling Effects of Slurry Chemicals, Abrasive Size Distribution, and Wafer-Pad Contact Area," *IEEE Trans. on Semiconductor Manufacturing*, vol. 16, pp. 45-56.

Luo, J. and Dornfeld, D. A., 2003, "Effect of Abrasive Size Distribution in Chemical Mechanical Planarization: Modeling and Verification," *IEEE Trans. on Semiconductor Manufacturing*, vol. 16, pp. 469-476.

Moinpour, M., Tregub, A., Oehler, A. and Cadien, K., 2002, "Advances in Chracterization of CMP Consumables," *MRS Bulletin*, vol. 27, pp. 766-771.

Nguyen, V. H., Daamen, R., Kranenburg, H., Velden, P. and Woerlee, P. H., 2003, "A Physical Model for Dishing during Metal CMP," *J. Electrochemical Soc.*, 150, pp. G689-G693.

Noh, K., Kopanski, K., Saka, N. and Chun, J.-H., 2005, "The Effect of Pad Topography on Surface Non-uniformity in Cu CMP," *Proc. CMP-MIC Conf.*, pp. 443-451.

Noh, K., Saka, N. and Chun, J.-H., 2004, "Effect of Slurry Selectivity on Dielectric Erosion and Cu dishing in Copper Chemical Mechanical Polishing," *Annals of the CIRP*, vol. 53, pp. 463-466.

Park, T., Tugbawa, T. and Boning, D., 2000, "Overview of Methods for Characterization of Pattern Dependencies in Copper CMP," *Proc. CMP-MIC Conf.*, pp. 196-205.

Preston, F. W., 1927, "The Theory and Design of Plate Glass Polishing Machines," *J. Soc Glass Technology*, vol. 11, pp. 214-256.

Runnels, S. R., 1994, "Feature-Scale Fluid-Based Erosion Modeling for Chemical-Mechanical Polishing," *J. Electrochemical Soc.*, vol. 141, pp. 1900-1904.

Runnels, S. R., Kim, I., Schleuter, J., Karlsrud, C. and Desai, M., 1998, "A Modeling Tool for Chemical-Mechanical Polishing Design and Evaluation," *IEEE Trans. on Semiconductor Manufacturing*, vol. 11, pp. 501-510.

Runnels, S. R., Miceli, F. and Kim, I., 1999, "Validation of a Large Area Three Dimensional Erosion Simulator for Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 146, pp. 4619-4625.

Saka, N., Lai, J.-Y., Chun, J.-H. and Suh, N. P., 2001, "Mechanisms of the Chemical Mechanical Polishing (CMP) Process in Integrated Circuit Fabrication," *Annals of the CIRP*, vol. 50, pp. 233-238.

Seok, J., Sukam, C. P., Kim, A. T., Tichy, J. A. and Cale, T.S., 2004, "Material Removal Model for Chemical-Mechanical Polishing Considering Wafer Flexibility and Edge Effects," *Wear*, vol. 257, pp. 496-508.

Seok, J., Sukam, C. P., Kim, A. T., Tichy, J. A. and Cale, T.S., 2003, "Multiscale Material Removal Modeling of Chemical Mechanical Polishing," *Wear*, vol. 254, pp. 307-320.

Stavreva, Z., Zeidler, D., Pltner, M. and Drescher, K., 1995, "Chemical Mechanical Polishing of Copper for Multilevel Metallization," *Appl. Surface Sci.*, vol. 91, pp. 192-196.

Stavreva, Z., Zeidler, D., Pltner, M., Grasshoff, G. and Drescher, K., 1997, "Chemical-Mechanical Polishing of Copper for Interconnect Formation," *Microelectronic Engineering*, vol. 33, pp. 249-257.

Steigerwald, J. M., Zirpoli, R., Murarka, S. P., Price, D. and Gutmann, R. J., 1994, "Pattern Geometry Effects in the Chemical-Mechanical Polishing of Inlaid Copper Structures," *J. Electrochemical Soc.*, vol. 141, pp. 2842-2848.

Stine, B., Ouma, D., Divecha, R., Boning, D., Chung, J., Hetherington, D. L., Ali, I., Shinn, G., Clark, J., Nakagawa, O. S., and Oh, S.-Y., 1997, "A Closed-Form Analytic Model for ILD Thickness Variation in CMP Process," *Proc. CMP-MIC Conf.*, pp. 266-273.

Tugbawa, T., Park, T. and Boning, D., 2002, "Integrated Chip-Scale Simulation of Pattern Dependencies in Copper Electroplating and Copper Chemical Mechanical Polishing Processes," *International Interconnect Technology Conference*, San Francisco, CA.

Vlassak, J. J., 2001, "A Contact-Mechanics Based Model for Dishing and Erosion in Chemical-Mechanical Polishing," *Proc. Mat. Res. Soc. Symp.*, vol. 671, pp. M4.61-M4.66.

Vlassak, J. J., 2004, "A Model for Chemical-Mechanical Polishing of a Material Surface Based on Contact Mechanics," *J. Mechanics and Physics of Solids*, vol. 52, pp. 847-873.

Yang, L., 2000, "Modeling CMP for Copper Dual Damascene Interconnects," *Solid State Technology*, vol. 10, pp. 111-124.

Yu, T.-K., Yu, C. C. and Orlowski, M., 1993, "A Statistical Polishing Pad Model for Chemical-Mechanical Polishing," *International Electron Devices Meeting*, pp. 35.4.1-35.4.4.

CHAPTER 3

THE INTEGRATED NON-UNIFORMITY MODELS

3.1 Introduction

In Chapter 2, the feature-scale polishing behavior based on local pressure distribution at the pad/wafer interface was studied and the evolution of step-height and wafer surface height was characterized. The pad/wafer interaction regimes include contact of smooth, discrete and rough pads with patterned wafers.

The primary problem in Cu CMP, however, is that the material removal rate across the wafer is non-uniform due to non-uniform pressure and velocity distributions, and even the Preston constant. Figure 3.1 shows the definition of various length scales across the wafer: wafer-, die-, subdie- and feature-scales. In patterned wafer polishing, there is not only inter-die variation depending on the location of dies on a wafer, but also inter-subdie or feature-scale variations depending on the pattern layout of subdies in a die. Therefore, it is necessary to extend feature-scale models to a model that include the die- and wafer-scale variations.

In this chapter, the possible causes of non-uniformities at the wafer-, die- and feature-scales are identified in terms of geometric, material and process parameters. Then, integrated dielectric erosion and Cu dishing models are developed by combining wafer-, die- and feature-scale variations with feature-scale step-height models. Finally, the effects of model parameters on dishing and erosion are discussed.

3.2 Surface Non-uniformities at Various Scales

3.2.1 Feature-scale Non-uniformity

Feature-scale variation is mostly represented by pattern geometry. From the early days of CMP, pattern geometry dependency has been experimentally demonstrated. Warnock proposed a phenomenological model in ILD CMP to predict material removal rates of arrays of features for various pattern geometries quantitatively [Warnock, 1991]. In this model, he described the

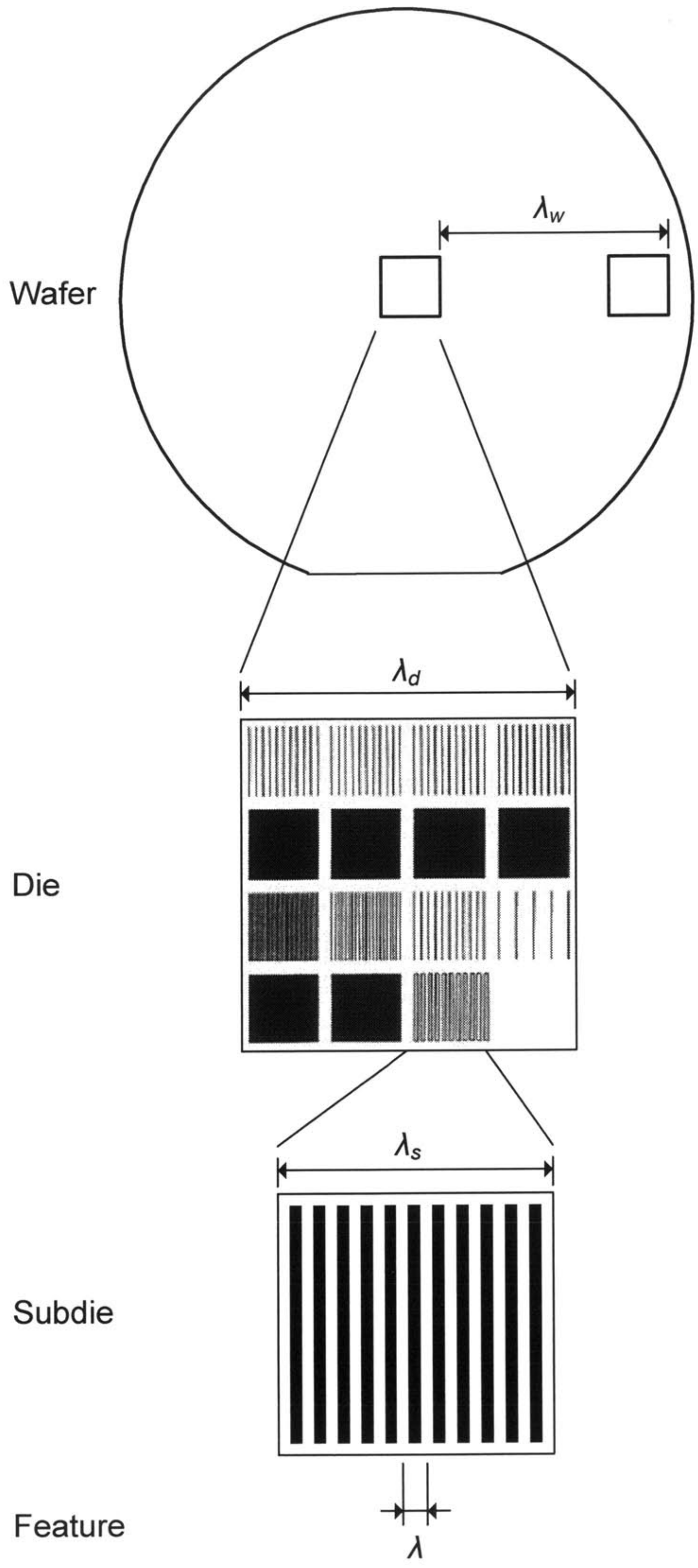


Figure 3.1 Definition of length scales in CMP: wafer-, die-, subdie- and feature-scales.

material removal rate as a microscopic mathematical model based on the behavior observed during polishing, such as the effect of extrusion and protrusion of a feature and the neighboring points. In this model, however, the physical meaning of the correlation between pattern geometry and material removal rate mechanism was not addressed. The effect of pattern geometry on erosion and dishing in Cu CMP was experimentally studied by Steigerwald *et al.* They proposed that dielectric erosion depends primarily on the area fraction, and Cu dishing on the linewidth of Cu interconnects. The effect of overpolishing on step-height reduction was also observed by Stavreva *et al.* [Steigerwald *et al.*, 1994; Stavreva *et al.*, 1997].

Feature-scale pattern geometry of the wafer surface prior to CMP is represented by the linewidth and the step height. Due to the different characteristics of the Cu deposition processes, such as PVD and electroplating, the surface Cu topography is generally different from the underlying trench pattern. The effect of Cu deposition on the initial pattern geometry was recognized by Lee [Lee, 2002]. He defined three different types of initial patterns: positive deposition, zero deposition, and negative deposition biases as shown in Fig. 3.2. Later, Park *et al.* presented a methodology for characterizing and modeling of the Cu electroplated wafer surface topography [Park *et al.*, 2004]. Semiempirical response surface models have been generated and model parameters extracted from conventional and superfill plating processes. In the present models, however, h_{Cu} is assumed to be uniform, i.e., zero deposition bias as shown in Fig. 3.2(b), regardless of the underlying pattern geometries since the experimental wafers are deposited by the PVD process.

Since h_{Cu} is assumed to be uniform in this thesis, surface profile can be expressed by the "surface linewidth," w_s, and the initial step height, h_{si} as shown in Fig. 3.3. The "surface linewidth," w_s, and the initial step height, h_{si}, may be smaller than the underlying Cu linewidth, w and the interconnect thickness, h_I. The Cu deposition factor α is defined as:

$$\alpha \equiv \frac{w_s}{w} \quad (0 \leq \alpha \leq 1) \tag{3.1}$$

If $\alpha = 0$, the initial Cu surface topography is flat regardless of the underlying pattern geometry as in Fig. 3.4(a), and if $\alpha = 1$ and $h_{si} = 1$, the initial surface topography is a true replica of the underlying trench pattern as in Fig. 3.4(c).

Even though h_{Cu} is assumed uniform, the general surface of deposited Cu can be expressed by setting appropriate h_{Cu}, α and h_{si} values depending on the Cu deposition methods.

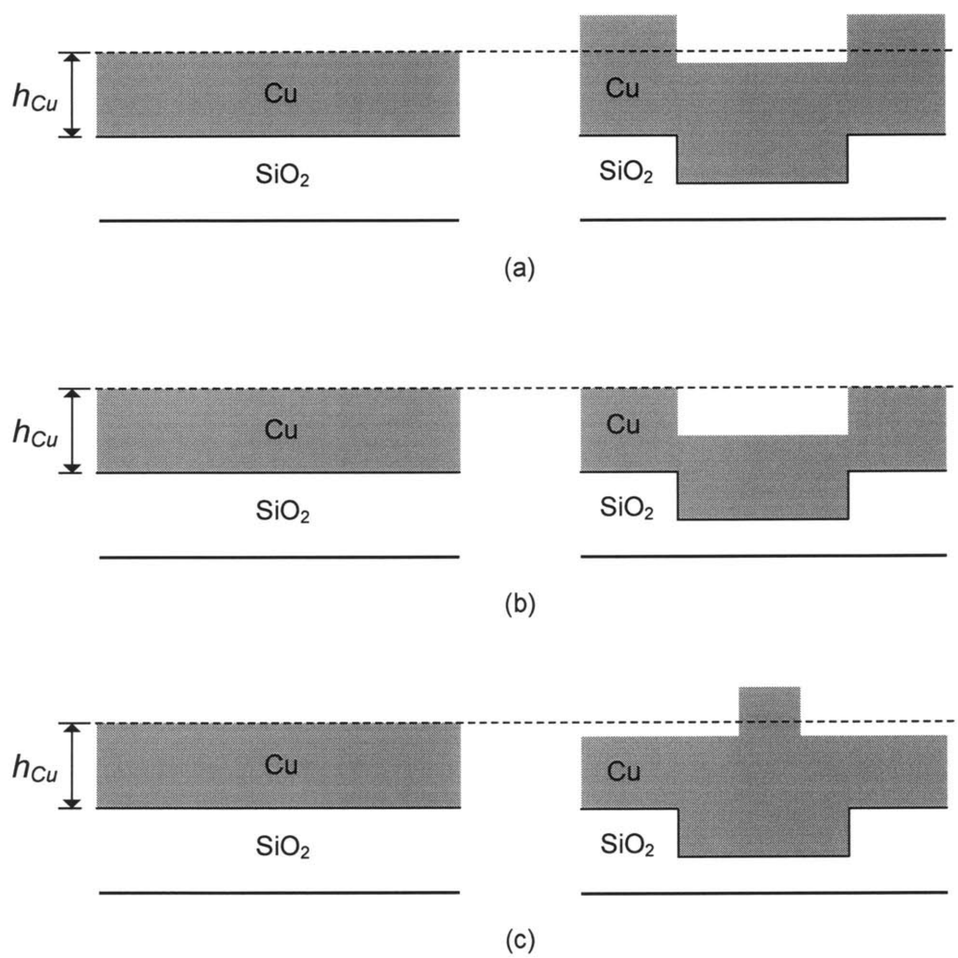


Figure 3.2 Cu deposition effect on the initial pattern geometry: (a) Positive deposition bias, (b) Zero deposition bias and (c) Negative deposition bias.

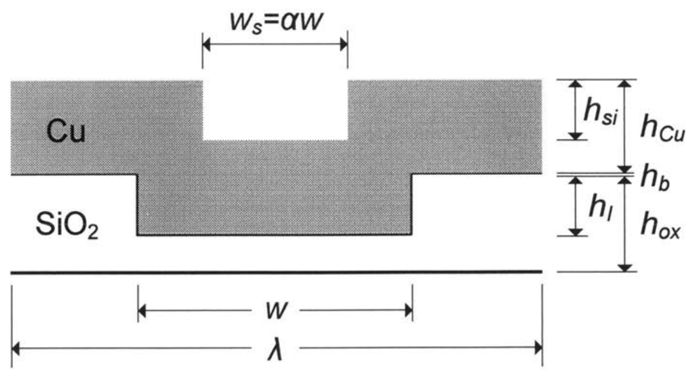


Figure 3.3 Definition of pattern geometry in Cu damascene structure based on the Cu interconnect linewidth, w, pitch, λ, Cu deposition factor α and the initial step-height, h_{si}.

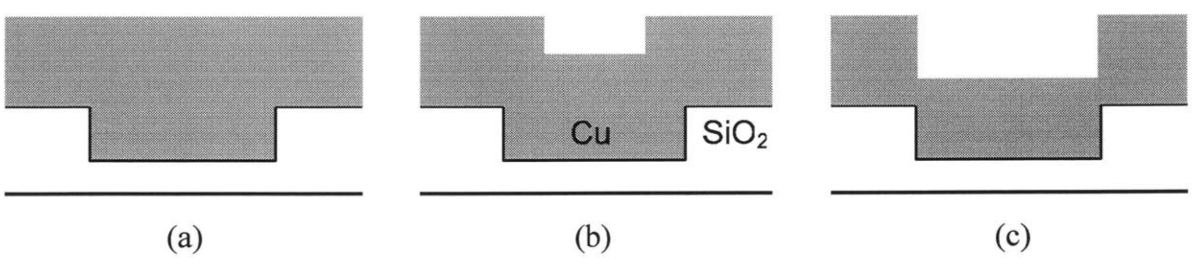


Figure 3.4 Effect of Cu deposition factor, α, and initial step-height, h_{si}, when (a) $\alpha = 0$ or $h_{si} = 0$, (b) $0 < \alpha < 1$ or $0 < h_{si} < 1$ and (c) $\alpha = 1$ or $h_{si} = h_I$.

3.2.2 Die-scale (Inter-subdie) Non-uniformity

In addition to the feature-scale non-uniformity, there are die-scale non-uniformities in each die on the wafer. The die-scale variation mainly depends on the pattern geometry and the materials in contact with the pad at each polishing height.

There have been several researches that try to simulate die or chip scale variations depending on the pattern layout. First approach was to develop a chip-scale step-height reduction simulator based on the planarization length. In this simulation, the average pressure at each pattern density unit is calculated based on the smooth contact mechanics model. The effect of pattern geometry on the material removal rate was investigated and a pattern density model was defined based on the planarization length concept in ILD CMP [Stine *et al.*, 1998]. In this pattern density model, the material removal rate of high-feature was assumed the same as the blanket wafer material removal rate divided by the pattern density. Later, a chip scale step-height reduction model was proposed based on the assumption that pressure at the high-feature decreases and pressure at the low-feature increases linearly as step-height decreases, if the step-height is less than a critical step-height [Tugbawa *et al.*, 2002]. Another approach was to develop a chip scale erosion simulator based on the feature-scale fluid based erosion model [Runnels et al, 1999].

To express pattern geometry at die-scale of complex chip-style patterns, it is required to define a characteristic area that can be considered as a separate region in a die in terms of an average pressure. The average pressure in the characteristic area mainly depends on the area fraction, A_f, of Cu in the underlying pattern. This area fraction can be expressed as the ratio of the Cu interconnects area, A_{Cu}, to the total characteristic area, A_{total}.

$$A_f \equiv \frac{A_{Cu}}{A_{total}} \quad (0 \leq A_f \leq 1) \tag{3.2}$$

If $A_f = 0$, the region is an oxide field region and there are no Cu lines within the area. On the other hand, if $A_f = 1$, the entire area is monolithic Cu. In this study, each die in a patterned wafer consists of 16 separate subdies with various periodic patterns as shown in Fig. 3.5(a) and (b), and thus each subdie will be assumed as the characteristic area. Thus, the area fraction A_f of a subdie for linear features may be expressed the same as the area fraction of each feature as:

$$A_f = \frac{w}{\lambda} \quad (0 \leq A_f \leq 1) \tag{3.3}$$

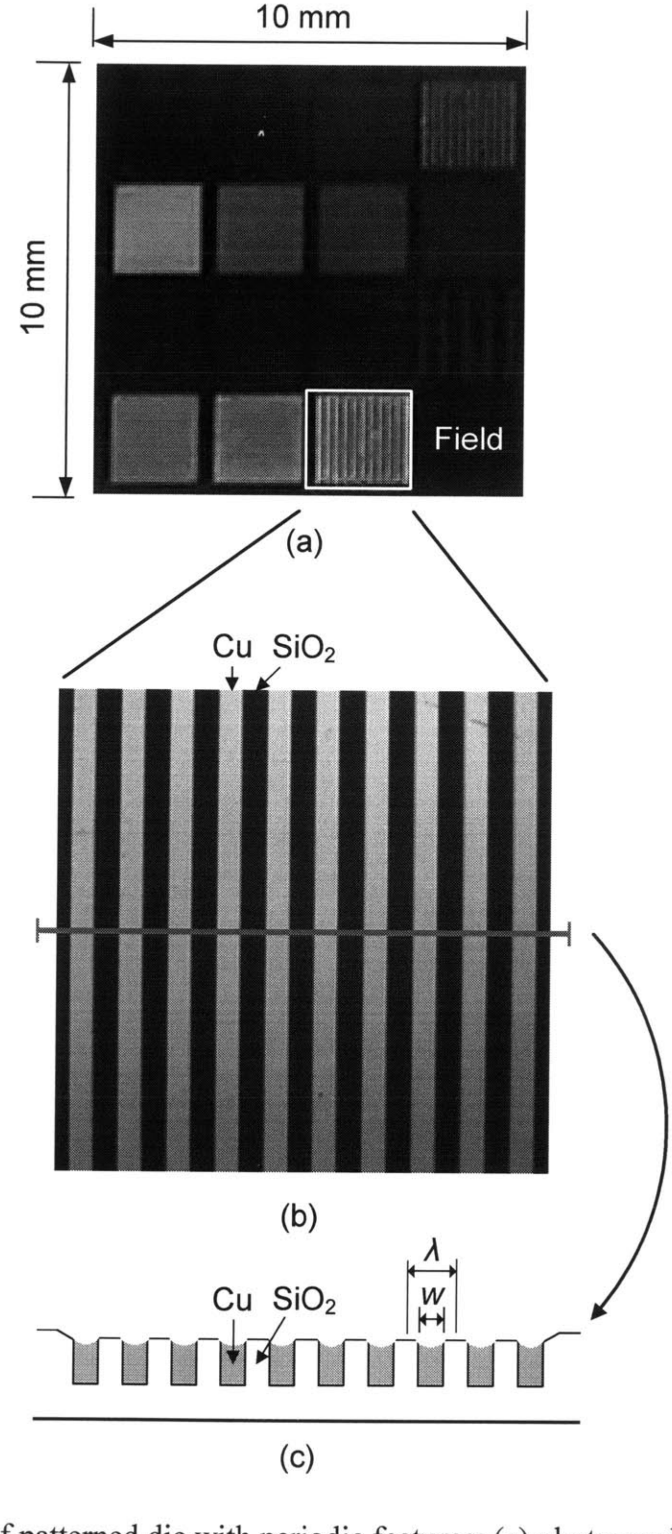


Figure 3.5 Examples of patterned die with periodic features: (a) photograph of a die, (b) micrograph of a subdie and (c) schematic cross-section of a subdie after CMP.

where w is the Cu linewidth and λ the pitch of the underlying pattern geometry in a Cu damascene structure as shown in Fig. 3.5(c).

Therefore, die-scale (inter-subdie) variation in this thesis is expressed by feature-scale variation due to its periodic pattern geometry.

3.2.3 Wafer-scale (Inter-die) Non-uniformity

Wafer-scale non-uniformity is due to non-uniform pressure and velocity distributions, and non-uniform Preston constant due to uneven slurry distribution.

Runnels *et al.* proposed a wafer-scale phenomenological model to represent the wafer-scale material removal rate [Runnels *et al.*, 1998]. In this model, they analyzed the kinematics of rotary-type CMP process and considered a wafer-centered quadratic pressure distribution. Then, interpolation formulae were developed to determine the relationship between the material removal rate and relative velocity and pressure distribution. Tichy *et al.* presented a model to predict the wafer-scale pressure distribution during CMP by using one-dimensional contact mechanics and hydrodynamic lubrication theory [Tichy *et al.*, 1999]. Fluid pressure across the wafer in one direction and an effective gap between the rough pad and the blanket wafer were measured. Fu and Chandra presented an analytical expression for wafer-scale pressure distribution at the pad/wafer interface for a viscoelastic pad [Fu and Chandra, 2002]. They showed that the material removal rate decreases as polishing progresses due to the pressure decay except in the edge area of a wafer. Seok *et al.* modeled the wafer-scale pressure based on the assumption that the wafer acts as a thin beam resting on an elastic foundation [Seok *et al.*, 2004]. The simulations show high contact pressure at the edge of the wafer.

A complete analysis of wafer-scale variations is complicated since not only the individual contributions require thorough investigation, but also the effect of one parameter on another needs to be considered.

In CMP, the wafer-scale non-uniformity is expressed as within-wafer non-uniformity (*WIWNU*), defined as:

$$WIWNU \equiv \frac{\sigma_h}{\mu_h} \tag{3.4}$$

where σ_h is the standard deviation and μ_h is the mean of the amount of material removed in a given time at the sampled points.

The *WIWNU* estimation requires statistical analysis. To compare non-uniformity at any two points, however, the ratio of material removed by polishing in a given time is more useful. In this thesis, therefore, the maximum wafer-scale non-uniformity is focused. Although the reference point in a die can be any pattern geometry so long as the subdies with the same pattern geometry in the slowest and the fastest die are compared, the field region is the best to compare with in part due to the ease of measurement. Thus, the wafer-scale uniformity factor, β, is defined as the ratio of the material removal rates at the slowest and fastest field regions. Thus,

$$\beta \equiv \frac{\Delta h_{sf}}{\Delta h_{ff}} \quad (0 < \beta \leq 1) \tag{3.5}$$

where Δh_{sf} and Δh_{ff} are the thicknesses of material removed in the slowest and the fastest field regions, respectively, in a given time.

As shown in Fig. 3.6, the edge region of the wafer polishes faster and hence erosion and dishing are the greatest in the outer dies in many of the conventional face-down CMP tools. In a blanket wafer, every point in a wafer can be considered as a field region as shown in Fig. 3.6(a). Therefore, β expresses the maximum material removal ratio in a wafer. If $\beta = 1$, the wafer is polished uniformly across the entire area; if $\beta < 1$, the wafer is polished nonuniformly. In a patterned wafer, on the other hand, it is important to differentiate between the wafer-scale non-uniformity and the die-scale non-uniformity. The simplest way to do this is to compare a subdie of the same pattern geometry in each die. The fastest or the slowest polishing region in a die, in some cases, may not be the field region, however. Nevertheless, it is convenient to select the field region in each die as a reference point as shown in Fig. 3.6(b). Even when $\beta = 1$, there could be die-scale non-uniformity. This means that a subdie at the same relative position in each die on a wafer will have the same non-uniformity. If $\beta < 1$, of course, subdies in each die will have different non-uniformities.

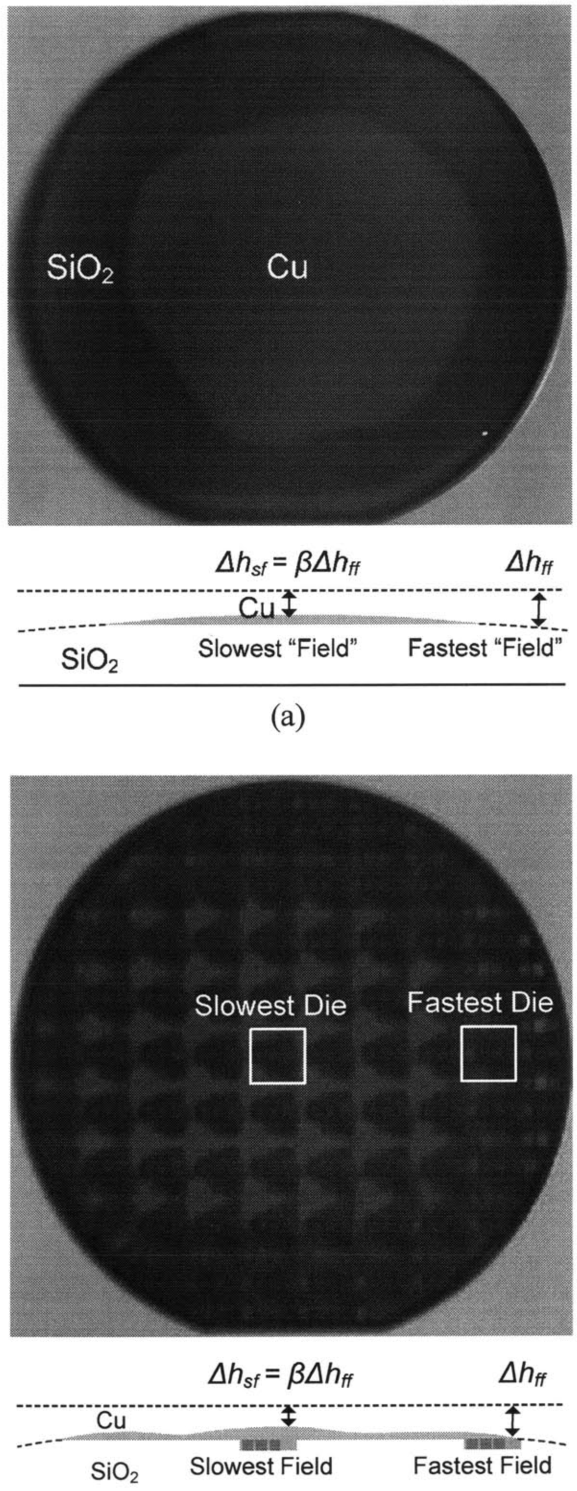


Figure 3.6 Definition of wafer-scale uniformity factor β : (a) blanket wafer and (b) patterned wafer.

3.3 Cu Dishing and Dielectric Erosion

3.3.1 General Approach

The first approach to integrate wafer- and die-scale variations into the feature-scale Cu dishing model was addressed by Yang [Yang, 2000]. He introduced the concept of dishing susceptibility, χ, to quantify the dishing dependence of over-polishing in Cu CMP. In this model, χ represents the slope of the line between two points connecting final and initial dishing when the x-axis is the amount of corresponding overpolishing at field region. Then, he expressed the overpolishing time at each feature in terms of the standard deviation of the wafer-scale non-uniformity. Although this is conceptually similar to the current model in this chapter, there are several limitations in Yang's model. First, there is no physical model describing feature-scale dishing and it assumes the final dishing is linearly proportional to the overpolishing time. Second, the dishing susceptibility χ is a lumped parameter which depends not only on the feature-scale initial pattern geometry but also on the final dishing and the process end-point at wafer-scale. Thus χ requires new measurements whenever the process end-point varies.

In this section, integrated dielectric erosion and Cu dishing models are developed by combining wafer-, die- and feature-scale variations into the feature-scale step-height calculation models. Figure 3.7 shows the evolution of the pad/wafer contact surface based on the feature-scale step-height models in Chapter 2. In these models, the polishing surface in the general Cu CMP is represented in three stages and end of each stage is separated as polishing time:

Stage 1: Initially, the pad contacts high-feature only if the relative deformation of the pad between the high- and low-feature, δ, under the given load is smaller than the initial step-height h_{si}: $\delta < h_{si}$. Materials being polished at high- and low-feature are the same: Cu. End of the Stage 1 is designated as t_1: $0 \leq t \leq t_1$.

Stage 2: The pad starts contacting both high- and low-feature. Materials being polished at high- and low-feature are the same: Cu. End of the Stage 2 is designated as t_2: $t_1 \leq t \leq t_2$.

Stage 3: The pad contacts both high- and low-feature. Materials being polishing at high- and low-feature, however, are different: oxide at the high-feature and Cu at the low-feature. End of the Stage 3 is designated as the process end-point t_{ep}: $t_2 \leq t \leq t_{ep}$.

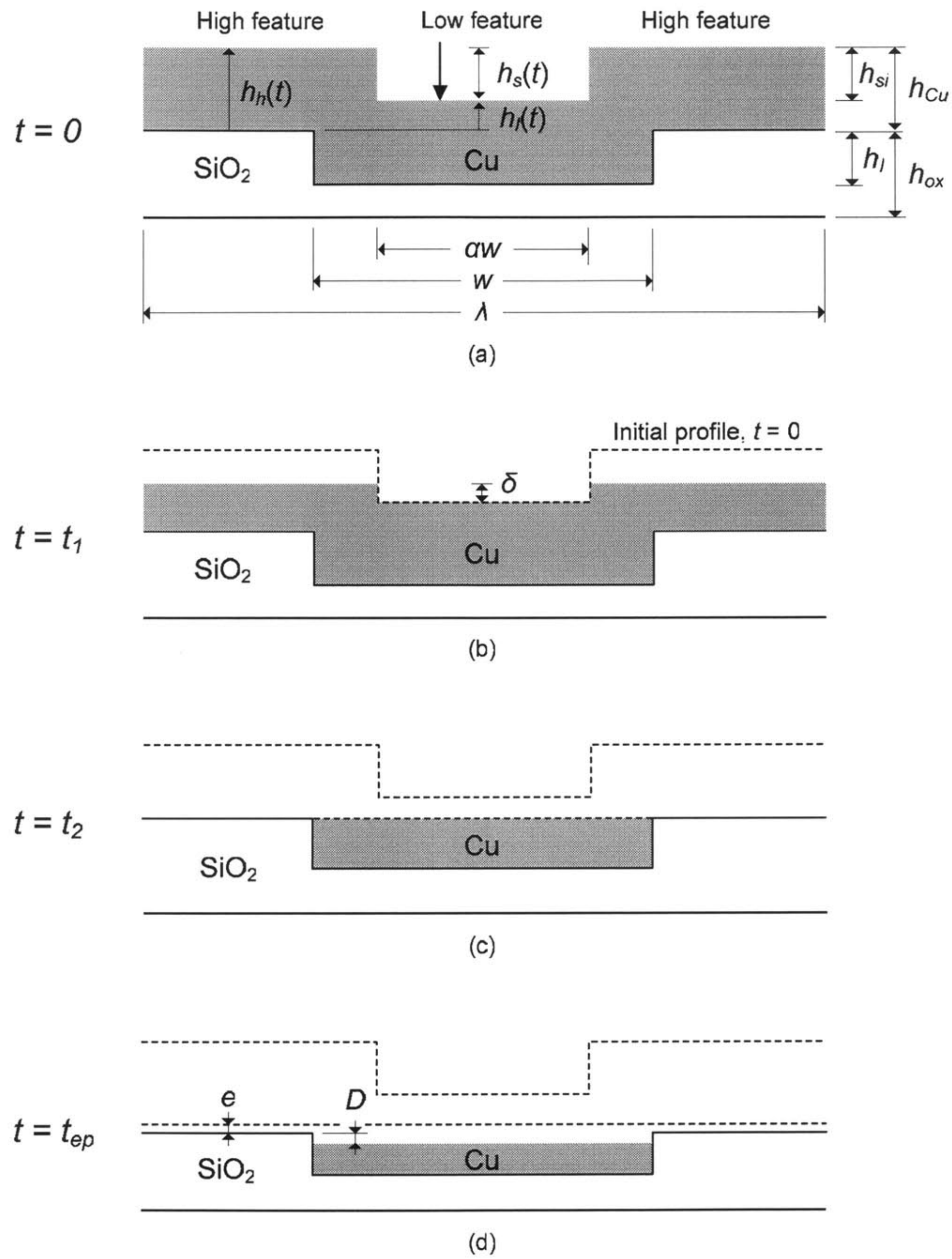


Figure 3.7 Schematics of the evolution of the polishing surface and non-uniformity definition at the feature-scale for $0 \leq t \leq t_{ep}$. Polishing time: (a) $t = 0$, (b) $t = t_1$, (c) $t = t_2$ and (d) $t = t_{cp}$.

Cu dishing, D, and dielectric erosion, e, are defined as the step-height and the height at high-feature at each pattern geometry when the polishing process ends, $t = t_{ep}$, as shown in Fig. 3.7(d). Thus,

$$D \equiv h_s(t_{ep}) \tag{3.6}$$

$$e \equiv h_h(t_2) - h_h(t_{ep}) \tag{3.7}$$

As far as the single feature is concerned, it is ideal to end the polishing process when the polishing surface reaches the top of the oxide, i.e., dielectric layer, $t_{ep} = t_2$. Cu dishing at this moment can be is defined as an initial dishing, D_i, and there is no erosion. Thus,

$$D_i \equiv h_s(t_2) \tag{3.8}$$

$$e = 0 \tag{3.9}$$

Based on the step-height model analysis in Chapter 2, the step-height decreases as polishing progresses in Stage 2 and thus, the initial dishing is almost zero. If the thickness Cu deposition is relatively smaller than the initial step-height, however, there may not be enough time to reduce the step-height and there can be a significant initial dishing. In any case, the initial dishing, D_i, is the minimum dishing at the feature during polishing process if the selectivity is greater than one.

The polishing times at the end of Stages 1 and 2, t_1 and t_2, are not global variables across a wafer. They depend on the non-uniformity variations at wafer-, die-, and feature-scale as defined in Section 3.2. In the case of the smooth and perfectly flat mode, for example, t_1 and t_2 in the fastest die are expressed as:

$$t_1 = \frac{h_{si}(1 - \alpha w / \lambda)}{k_{p_{Cu}} p_{av} v_R} \tag{3.10}$$

$$t_2 = \frac{h_{Cu} - (\alpha w / \lambda) h_{si}}{k_{p_{Cu}} p_{av} v_R} \tag{3.11}$$

In this case, t_2 increases as Cu deposition factor α, initial step-height h_{si}, and the area fraction w / λ decreases. This is reasonable since the total amount of Cu to be polished until Stage 2 increases as α, h_{si} and w / λ decreases.

Although it is desirable to end the polishing process at $t = t_2$ at each feature, in the conventional face-down CMP setup, it is impossible to setup different endpoints on a wafer since the whole area of the wafer is always in contact with the polishing pad. There is only one process endpoint across entire wafer. Most of previous feature- or at die-scale dishing and erosion models [Runnels, 1999; Park et al, 2000 and 2004; Vlassak, 2001 and 2004; Fu and Chandra, 2003], a process end-point in CMP is independently expressed at a time-scale and determined by an end-point setting by a process operator. However, in this approach, it is difficult to characterize effect of wafer-scale variation on erosion and dishing at each feature in a wafer.

The key approach to integrate wafer-scale variation into the feature-scale erosion and dishing is to express time-scale variables in terms of model parameters. For example, required polishing time of a wafer can be expressed by blanket material removal rate, pattern geometry and wafer-scale non-uniformities. It is expressesd as the time when the excess Cu in the field region of the slowest die, i.e., wafer reference point, is completely removed, $t = t_{2sf}$ as shown in Fig. 3.8. That is,

$$t_{2sf} = \frac{h_{Cu}}{\beta k_{p_{Cu}} p_{av} v_R} \tag{3.12}$$

In polishing practices, however, the wafer is slightly overpolished to ensure that there is no excess Cu or barrier material on the entire wafer, the so-called "overpolishing". The amount of overpolishing, Δh_o, is defined as the amount of oxide polished in the field region of the slowest die. Therefore, the process end-point, t_{ep}, is expressed as:

$$t_{ep} = \frac{h_{Cu}}{\beta k_{p_{Cu}} p_{av} v_R} + \frac{\Delta h_o}{\beta k_{p_{ox}} p_{av} v_R} \tag{3.13}$$

The amount of erosion and dishing are closely related to the amount of time the pad spends after it reaches the top of oxide at the feature. Accordingly, the overpolishing time at each feature, t_o, is defined as the additional time spent after the polishing time reaches the top of oxide at the general feature in the fastest die. Thus,

$$t_o \equiv t_{ep} - t_2 = \left(\frac{1}{k_{p_{Cu}} p_{av} v_R}\right)\left[\left(\frac{1}{\beta} - 1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \tag{3.14}$$

Since t_2 depends on pattern geometry and non-uniformity parameters, the overpolishing time,

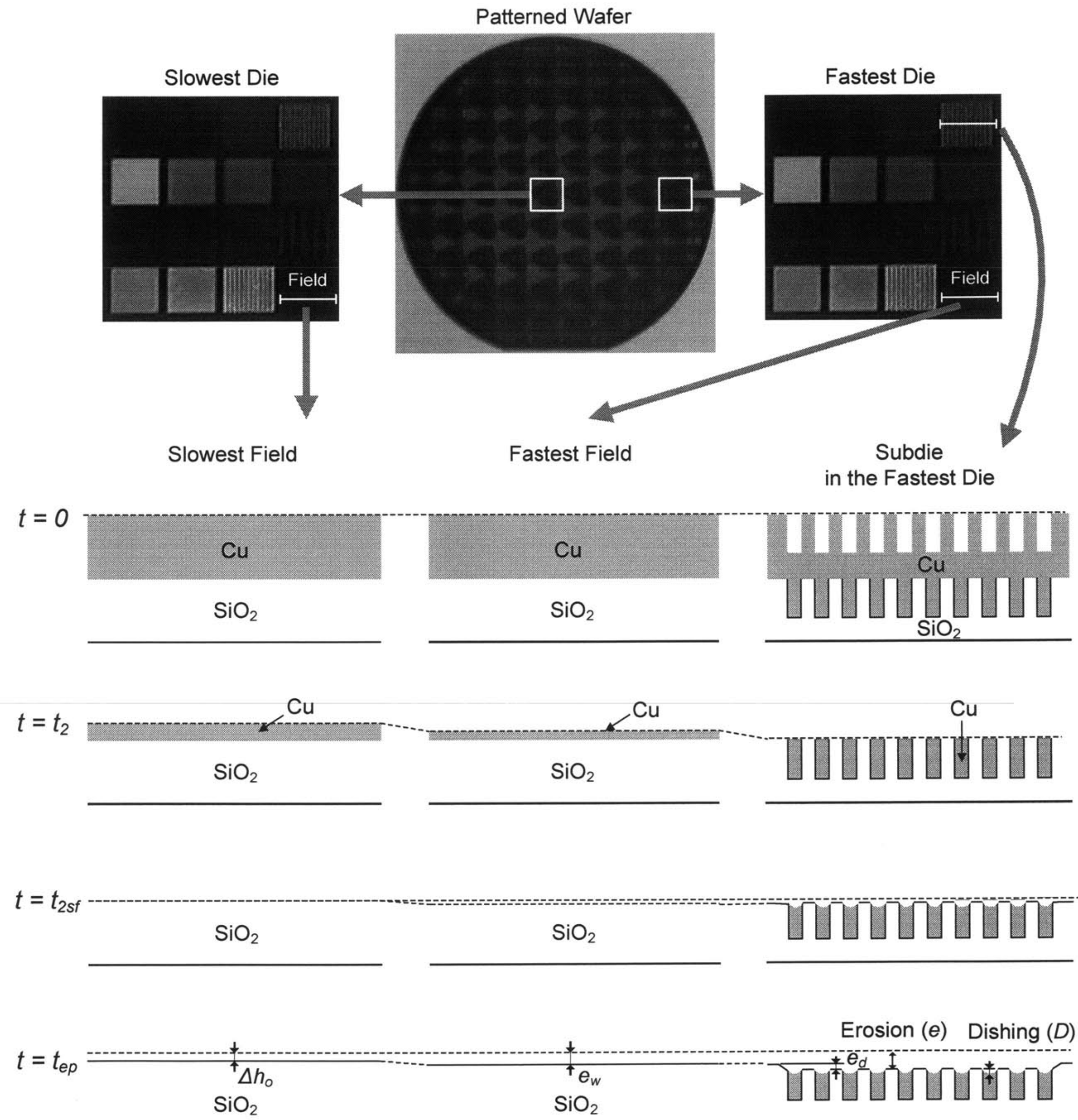


Figure 3.8 Evolution of the slowest and fastest polishing field regions, and a subdie in the fasted polishing die in a wafer ($0 \leq t \leq t_{ep}$).

t_o, also varies from feature to feature in a die, and from die to die in a wafer. Equation (3.14) shows three terms that affect the overpolishing time, t_o. Each terms represents lenth-scale variation parameters instead of time-scale variation. First, if $\beta = 1$, there is no wafer-scale variation and thus each die in a wafer will have exactly same die-scale variation regardless of its location in a wafer. Second, if $\alpha = 0$ or $h_{si} = 0$, the initial feature surface is perfectly flat and thus material removal rate is the same as that in the field region. In this case, there is no subdie-to-subdie variation or feature-to-feature variation in a die and the wafer surface height decreases at a rate of blanket wafer material removal rate until the polishing surface reaches the top of oxide layer. Third, if $\Delta h_o = 0$, there is no erosion and dishing developed in the wafer reference point. Additionally, if there is no wafer- and feature-scale variation, $\beta = 1$ and $\alpha = 0$ or $h_{si} = 0$, the overpolishing time at each feature, t_o, becomes zero and thus, there is no erosion and dishing across a wafer. If any of these three conditions are not satisfied, there is always certain amount of overpolishing at each feature and so are erosion and dishing.

The amount of oxide overpolishing at the field region at the slowest die, Δh_o, can be calculated based on the material removal rate of oxide and the overpolishing time as:

$$\Delta h_o = \beta \cdot k_{p_{ox}} p_{av} v_R \cdot (t_{ep} - t_{2sf}) \tag{3.15}$$

As the overpolishing time at each feature, t_o, increases, erosion and dishing increase. Cu dishing, however, does not increase linearly and may approach an asymptotic value, since the maximum step-height is limited by the oxide area surrounding Cu interconnects as observed in Chapter 2. This asymptotic value is represented by D_∞. Cu dishing D, D_i and D_∞ are expressed in dimensionless forms D^*, D_i^* and D_∞^*, respectively, the ratio of dishing to the nominal Cu interconnect thickness, h_I.

$$D^* \equiv \frac{D}{h_I},\ D_i^* \equiv \frac{D_i}{h_I},\ D_\infty^* \equiv \frac{D_\infty}{h_I} \tag{3.16}$$

Dielectric erosion, e, is defined as thickness of oxide worn out as expressed in (3.7). In the integrated erosion model, wafer-scale and die-scale erosions are decoupled to explain the effects of non-uniformity parameters. Wafer-scale dielectric erosion e_w is defined as the amount of overpolished oxide in the fastest die field region by the time polishing reaches the end-point, t_{ep}, as shown in Fig. 3.8. Furthermore, die-scale dielectric erosion e_d is defined as the amount of overpolished oxide in each suddie relative to the field region in the same die. Total dielectric erosion, e, is defined as the amount of overpolished oxide at the surrounding oxide and also the sum of the wafer-scale erosion, e_w, and the die-scale erosion, e_d.

$$e = e_w + e_d \tag{3.17}$$

Dielectric erosion, e, e_w and e_d are also expressed in dimensionless form e^*, e_w^* and e_d^* as:

$$e^* \equiv \frac{e}{h_I},\ e_w^* \equiv \frac{e_w}{h_I},\ e_d^* \equiv \frac{e_d}{h_I} \tag{3.18}$$

Integrated erosion and dishing models are developed by combining various step-height models described in Chapter 2 and the non-uniformity parameters defined in section 3.2.

Feature-scale step-height models in Chapter 3 can be expressed as three different types in terms of calculation method: iterative, numerical and analytic.

- **Iterative solution:** General rough pad model with elastic and plastic contacts
- **Numerical solution:** Simplified rough pad model with elastic contact.
- **Analytic solution:** Smooth pad, discrete pad model and simplified rough pad model with plastic contact.

Although it is possible to develop integrated erosion and dishing model in iterative, numerical solutions, only analytic solutions are considered in this chapter to develop integrated erosion and dishing model in simple symbolic forms. In this section, therefore, three feature-scale step-height models are considered: smooth, discrete and simplified rough pad with fully plastic contact models. Figure 3.9 shows schematics of each feature-scale pad/wafer contact models and Table 3.1 lists the final form of the step-height models.

3.3.2 The Smooth Pad Model

In the smooth pad step-height model, the pad is assumed as a homogeneous, monolithic elastic semi-infinite body as shown in Fig. 3.9(a). Although most of commercial polishing pad have rough surfaces and there are many asperity/wafer contacts during polishing, the smooth pad model is valid if the size of asperity contact area is sufficiently smaller than the width of Cu interconnects. In the submicron device level, for example, Cu interconnect linewidth is from 65

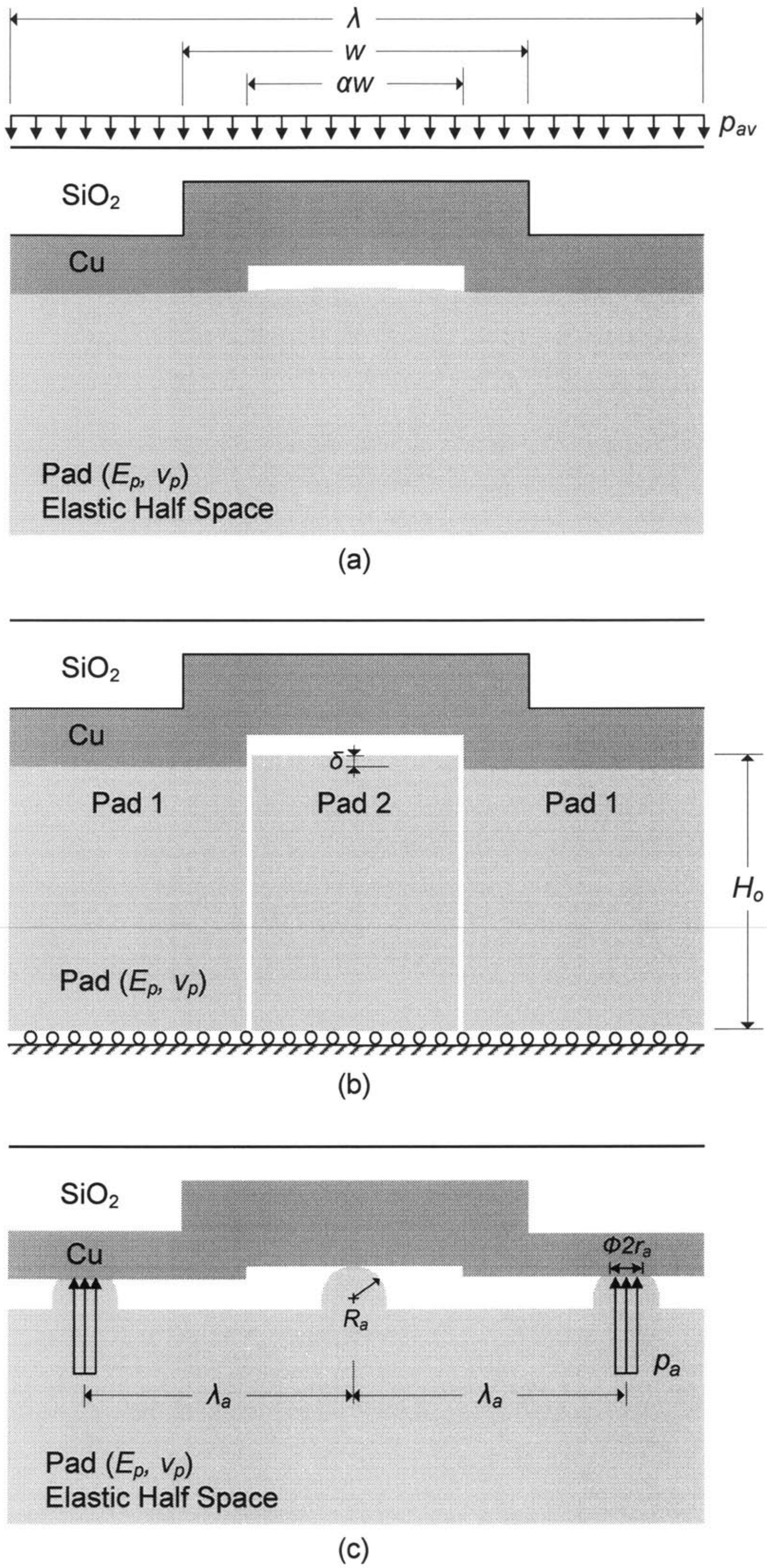


Figure 3.9 Schematics of feature-scale step-height for three contact modes: (a) smooth pad, (b) discrete pad and (c) simplified rough pad.

Table 3.1 Comparison of feature-scale step-height models.

Model	Initial Pad Deformation, Step-height, Height at the High Feature Values
Smooth Pad	$\delta = \frac{A_1\alpha}{(1-\alpha w/\lambda)}\left(\frac{p_{av}}{E^*}\right)w$ $h_s(t) = \left(\frac{B_{21}S_{Cu/ox} - B_{11}}{B_{12} - B_{22}S_{Cu/ox}}\right)\left(\frac{p_{av}}{E^*}\right)w\left[1 - exp\left(-\frac{t-t_2}{\tau_3}\right)\right]$ $\tau_3 = \left(\frac{S_{Cu/ox}}{B_{12} - B_{22}S_{Cu/ox}}\right)\left(\frac{1}{k_{p_{Cu}}E^* v_R}\right)w$ $h_h(t) = -MRR_\infty(t - t_2)$ where $MRR_\infty \equiv k_{p_{Cu}} p_{av} v_R \left[\frac{1}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]$
Discrete Pad	$\delta = \left(\frac{1}{1-\alpha w/\lambda}\right)\left(\frac{p_{av}}{E_p}\right)H_o$ $h_s(t) = \left[\frac{S_{Cu/ox} - 1}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{p_{av}}{E_p}\right)H_o\left\{1 - exp\left[-\left(\frac{t-t_2}{\tau_3}\right)\right]\right\}$ $\tau_3 \equiv \left[\frac{S_{Cu/ox}}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{1}{k_{p_{Cu}}E_p v_R}\right)H_o$ $h_h(t) = -MRR_\infty(t - t_2) + \left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]h_s(t)$
Rough Pad - Simplified, Elastic	$\delta = \left(\frac{1}{1-\alpha w/\lambda}\right)^{2/3}\left(\frac{p_{av}}{E^*}\right)^{2/3}\left(\frac{9\lambda_a^4}{16R}\right)^{1/3}$
Rough Pad - Simplified, Fully plastic	$\delta = \left(\frac{1}{1-\alpha w/\lambda}\right)\left(\frac{p_{av}}{Y_p}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right)$ $h_s(t) = \left[\frac{S_{Cu/ox} - 1}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{p_{av}}{Y_p}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right)\left\{1 - exp\left[-\left(\frac{t-t_2}{\tau_3}\right)\right]\right\}$ $\tau_3 \equiv \left[\frac{S_{Cu/ox}}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{1}{k_{p_{Cu}}Y_p v_R}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right)$ $h_h(t) = -MRR_\infty(t - t_2) + \left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]h_s(t)$

* Simplified rough pad (elastic contact) and general rough pad models are calculated iteratively.

** $h_s(t_2)$ is assumed zero for all models.

nm to 500 nm and the radius of pad asperity contact area is about 1~2 μm when the applied pressure is 14 kPa (2 psi).

Additionally, the Cu deposition factor, α, and the initial step-height, h_{si}, at the sub-micron device level are much smaller than those at the global wiring level. For example, in PVD Cu coating, $\alpha = 0.1$ and $h_{si} / h_I = 0.1$ when Cu linewidth is 0.5 μm. Thus, the pad can be assumed to be perfectly smooth and flat during Stage 1 since the initial relative deformation in the smooth pad model is much smaller than the initial step-height: $\delta / h_{si} < 10^{-3}$. Based on this assumption, the polishing time t_1 and t_2 in the smooth pad model are the same as in (3.10) and (3.11), respectively. Furthermore, at the end of Stage 2, the step-height is zero, $h_s(t_2) = 0$. Therefore, the step-height can be expressed as listed in Table 3.1.

As the polishing time increases, material removal rates at both high- and low-features approach the same value. This asymptotic material removal rate at the high- and low-features was defined as MRR_∞, and is useful to compare material removal rates of oxide of each step-height models.

In Stage 3, material removal rates in the oxide and Cu region remain almost the same since the Stage 3 time constant of the smooth pad step-height model is much smaller than the overpolishing time: $t_o \gg \tau_3$. Thus, $h_h(t)$ decreases at the rate of MRR_∞.

Cu Dishing

In the conventional CMP practice, the time constant of the smooth pad model, τ_3, listed in Table 3.1 is much smaller than overpolishing time at each feature, $t_o / \tau_3 \gg 10$. Thus the Cu dishing can be considered time-independent. From the smooth pad step-height in Table 3.1,

$$D = \left(\frac{B_{21} S_{Cu/ox} - B_{11}}{B_{12} - B_{22} S_{Cu/ox}} \right) \left(\frac{p_{av}}{E_p^*} \right) w \tag{3.19}$$

In this model, dishing is proportional to the Cu linewidth. However, this model is valid in the region that Cu linewidth is smaller that the radius of the pad asperity contact area. Furthermore, in the conventional polishing practices, $p_{av} / E^* \sim 10^{-4}$, and thus the amount of dishing is negligible in the submicron device level.

Dielectric Erosion

From Eq. (3.7), dielectric erosion in the smooth pad model can be expressed as:

$$e = MRR_{\infty}\left(t_{ep} - t_2\right) \tag{3.20}$$

Therefore, dielectric erosion increases linearly as the polishing time increases and is independent of Cu dishing. Combined with MRR_{∞} and the overpolishing time, t_o, in Eq. (3.14), erosion can be rewritten as:

$$e = \left[\frac{1}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left[\left(\frac{1}{\beta} - 1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \tag{3.21}$$

Wafer-scale erosion can be simply obtained by substituting $w = 0$ in Eq. (3.21). Thus

$$e_w = \frac{1}{S_{Cu/ox}}\left(\frac{1}{\beta} - 1\right)h_{Cu} + \frac{1}{\beta}\Delta h_o \tag{3.22}$$

Die-level erosion is calculated by subtracting the wafer-level erosion, e_w, from the total erosion, e, as:

$$\begin{aligned} e_d = &\left[\frac{1}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left[\left(\frac{1}{\beta} - 1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \\ &- \frac{1}{S_{Cu/ox}}\left(\frac{1}{\beta} - 1\right)h_{Cu} - \frac{1}{\beta}\Delta h_o \end{aligned} \tag{3.23}$$

3.3.3 The Discrete Pad Model

In this model, the pad is assumed to be discrete, uniaxially loaded blocks as shown in Fig. 3.9(b). Although the model relies on the finite pad thickness and physically inadmissible discontinuous boundary conditions, it reflects the evolution of the step-height in the simplest form and fairly well.

Stage 2 ends when the pad reaches the top of oxide at the high feature. The end of Stage 2, t_2, of the discrete pad model approaches to that of the smooth pad model, Eq. (3.11), and $h_s(t_2)$ approaches zero, as the ratio $(t_2 - t_1)/\tau_2$ increases. As described in the sample calculation in Chapter 2, $(t_2 - t_1)/\tau_2 > 4$ in conventional Cu CMP. Therefore, in this section, $h_s(t_2)$ is assumed zero and thus the final $h_s(t)$ and $h_h(t)$ are listed in Table 3.1.

Cu dishing

Ideally, polishing should end when the excess Cu at the high-feature is completely removed:

$t = t_2$. Cu dishing at this moment is defined as the initial dishing, D_i. However, since $h_s(t_2)$ is assumed zero, the initial dishing is also zero.

Therefore, Cu dishing can be rewritten as:

$$D = \left[\frac{S_{Cu/ox} - 1}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{p_{av}}{E_p}\right)H_o\left[1 - exp\left(-\frac{t_{ep} - t_2}{\tau_3}\right)\right] \tag{3.24}$$

To evaluate the amount of exponential term in Eq. (3.24), it is useful define a dimensionless overpolishing time, t_o^*, as:

$$t_o^* \equiv \frac{t_{ep} - t_2}{\tau_3} \tag{3.25}$$

The dimensionless overpolishing time, t_o^* is an index of how close Cu dishing is to the asymptotic value, D_∞. For example, if $t_o^* > 4$, the Cu dishing, D approaches D_∞ within 2%. Therefore, the final form of Cu dishing of discrete pad is:

$$D = \left[\frac{S_{Cu/ox} - 1}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left(\frac{p_{av}}{E_p}\right)H_o\left[1 - exp\left(-t_o^*\right)\right] \tag{3.26}$$

where,

$$t_o^* = \left[\frac{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}{S_{Cu/ox}}\right]\left(\frac{E_p}{p_{av}}\right)\left(\frac{1}{H_o}\right)\left[\left(\frac{1}{\beta} - 1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \tag{3.27}$$

Dielectric erosion

Based on the step-height model in Table 3.1, dielectric erosion in the discrete pad model is represented as a similar form to that in the smooth pad model:

$$e = MRR_\infty(t_{ep} - t_2) - \left[\frac{w/\lambda}{(1 - w/\lambda)S_{Cu/ox} + w/\lambda}\right]D \tag{3.28}$$

The first term in (3.28) is the same as the dielectric erosion in the smooth pad model. The second term reflects the effect of Cu dishing on erosion. In the smooth pad model, the pressure at the high-feature is assumed to have the steady-state value during Stage 3. In the discrete model, however, the pressure at the high-feature increases from the pressure at the field area to the steady state value and the time constant to approach the steady state value is not negligible.

Therefore, the erosion in the discrete model, if the overpolishing time is the same, is smaller than that in the smooth pad model. Additionally, the effect of dishing on the erosion increases as area fraction of Cu interconnect increases. If $w/\lambda = 0$, reduction in erosion due to Cu dishing is $D/S_{Cu/ox}$. As w/λ approaches unity, reduction is erosion due to Cu dishing increases to D.

As the overpolishing time t_o increases, erosion increases almost linearly with MRR_{ss} since the dishing approaches an asymptotic value D_∞. Combining with the steady state material removal rate in (3.28) and the overpolishing time in (3.14), erosion can be rewritten as:

$$\begin{aligned} e = &\left[\frac{1}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left[\left(\frac{1}{\beta}-1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \\ &-\left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]D \end{aligned} \tag{3.29}$$

Since the wafer-scale erosion is defined in the field region, there is no dishing. Thus, the wafer-scale erosion in the discrete pad model is the same as that in the smooth pad model as in (3.22). Die-scale erosion, e_d, is calculated by subtracting the wafer-scale erosion, e_w, from the total erosion, e, as:

$$\begin{aligned} e_d = &\left[\frac{1}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]\left[\left(\frac{1}{\beta}-1\right)h_{Cu} + \alpha(w/\lambda)h_{si} + S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \\ &-\left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox} + w/\lambda}\right]D - \frac{1}{S_{Cu/ox}}\left(\frac{1}{\beta}-1\right)h_{Cu} - \frac{1}{\beta}\Delta h_o \end{aligned} \tag{3.30}$$

3.3.4 The Simplified Rough Pad Model with Plastic Contact

In the simplified rough pad model, the pad topography is idealized to have pad asperities with a uniform height and the same radius of curvature as shown in Fig. 3.9(c). Contact between the pad asperities and the wafer are assumed as either fully elastic or fully plastic.

In the elastic pad asperity contact condition, the relationship between the contact pressure and the pad asperity deformation follows the Hertzian contact theory. Unlike the discrete pad model, therefore, the pressure and the material removal rate at the high- and low-features are not linear with the step-height. Thus, the pad deformation at the high-feature δ_h at a given step-height h_s needs to be solved numerically. Therefore, in this section, the deformation of each pad asperity is assumed to be fully plastic.

In the plastic contact, the deformation of pad asperity and the mean pressures at the high- and low-features are linear. Therefore, the step-height analysis procedure is exactly the same as in the discrete pad model, replacing H_o by $\lambda_a^2 / 6\pi R_a$ and E_p by Y_p as listed in Table 3.1.

Cu dishing

Similar to the discrete pad analysis, the initial dishing in this case is also much smaller than the interconnect thickness, thus can be neglected. Therefore Cu dishing in the simplified rough pad model with plastic contact mode is solved as:

$$D = \left[\frac{S_{Cu/ox}-1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left(\frac{p_{av}}{Y_p}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right)\left[1-exp\left(-t_o^*\right)\right] \tag{3.31}$$

where,

$$t_o^* = \left[\frac{(1-w/\lambda)S_{Cu/ox}+w/\lambda}{S_{Cu/ox}}\right]\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right] \tag{3.32}$$

Dielectric erosion

The relationship between the dielectric erosion and Cu dishing in this model is the exactly the same with that in the discrete model, described in Eqs. (3.29) – (3.30). It is noticeable that the effect of material properties and the surface topography of polishing pad on dielectric erosion are marginal as long as Cu-to-oxide selectivity is high. However, if slurry selectivity approaches to unity, the amount of dishing should be considered to calculate erosion in the global wiring feature.

3.4 Process Parameter Sensitivity Analysis

In the previous section, integrate erosion and dishing models are developed based on the feature-scale step-height models for various pad/wafer contact regimes. Table 3.2 lists the summary of the developed integrated models.

In this section, the effects of model parameters on dishing and erosion are discussed. As described earlier, Cu dishing is dominant at the global-wiring level and dielectric erosion is

Table 3.2 Summary of developed integrated erosion and dishing models.

Model	Cu Dishing and Dielectric Erosion Values
Smooth Pad	$D=\left(\frac{B_{21}S_{Cu/ox}-B_{11}}{B_{12}-B_{22}S_{Cu/ox}}\right)\left(\frac{p_{av}}{E^{*}}\right)w$
	$e=\left[\frac{1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right]$
Discrete Pad	$D=\left[\frac{S_{Cu/ox}-1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left(\frac{p_{av}}{E_p}\right)H_o\left[1-exp\left(-t_o^{*}\right)\right]$
	$t_o^{*}=\left[\frac{(1-w/\lambda)S_{Cu/ox}+w/\lambda}{S_{Cu/ox}}\right]\left(\frac{E_p}{p_{av}}\right)\left(\frac{1}{H_o}\right)\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right]$
	$e=\left[\frac{1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right]-\left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]D$
Rough Pad - Simplified, Fully plastic	$D=\left[\frac{S_{Cu/ox}-1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left(\frac{p_{av}}{Y_p}\right)\left(\frac{\lambda_a^2}{6\pi R_a}\right)\left[1-exp\left(-t_o^{*}\right)\right]$
	$t_o^{*}=\left[\frac{(1-w/\lambda)S_{Cu/ox}+w/\lambda}{S_{Cu/ox}}\right]\left(\frac{Y_p}{p_{av}}\right)\left(\frac{6\pi R_a}{\lambda_a^2}\right)\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right]$
	$e=\left[\frac{1}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]\left[\left(\frac{1}{\beta}-1\right)h_{Cu}+\alpha(w/\lambda)h_{si}+S_{Cu/ox}\frac{1}{\beta}\Delta h_o\right]-\left[\frac{w/\lambda}{(1-w/\lambda)S_{Cu/ox}+w/\lambda}\right]D$

* Simplified rough pad (elastic contact) and general rough pad models are calculated iteratively.

dominant at the submicron device level in Cu CMP. Therefore, erosion of the smooth pad model, listed in Table 3.2, is considered to express erosion at the submicron level. Since the extent of dishing at the submicron device level is already below the industrial specification, the focus will be on the minimization of dielectric erosion by adopting the smooth pad model. At the global wiring level, on the other hand, rough pad models are applicable to express erosion and dishing. It is well known that die-scale erosion at the global wiring level is much less than the industrial specification. Since the wafer-scale erosion in a die depends on the wafer-scale uniformity factor and the amount of overpolishing only, it is considered in the erosion at the submicron level. Cu dishing, therefore, is mainly discussed based on the simplified rough-pad model, listed in Table 3.2, at the global wiring level.

3.4.1 Parameter Analysis

All the model parameter values for parameter sensitivity analysis in this chapter are listed in Table 3.3. The smooth pad model is used for the analysis of dielectric erosion and the rough pad model with fully plastic contact is considered for the analysis of Cu dishing.

Figure 3.10 shows the effect of slurry selectivity ($S_{Cu/ox}$) on dishing and erosion. When Cu-to-oxide selectivity is unity, the pad polishes Cu and SiO_2 at the same rates, and the step-height decreases with polishing time even after the pad contacts the dielectric. From the simplified rough pad model analysis, the step-height becomes much less than 5% of Cu thickness before the pad contacts the top of the oxide layer regardless of the location of die or pattern geometry. Cu dishing in this case is much less than 5% of Cu interconnect thickness.

As slurry selectivity is increased, dishing increases since Cu in the interconnect region polishes faster than the surrounding dielectric. However, the amount of Cu dishing depends on the amount of overpolishing time at a specific feature, which is determined by the initial pattern geometry, wafer-scale uniformity factor, and the extent of overpolishing at the global reference point. By contrast, dielectric erosion decreases as the selectivity is increased if the material removal rate of Cu is fixed. If the selectivity is unity, both SiO_2 and Cu are polished at the same rate as a blanket SiO_2 or Cu wafer. If the selectivity is very large, either the dielectric is completely rigid or the Cu is polished very fast. Since a fixed Cu material removal rate is considered, the latter case is applied, and there is no erosion regardless of the model parameters and the pattern geometries. With a fixed selectivity larger than one, erosion depends on the

Table 3.3 Model parameters for parameter sensitivity analysis.

Property	Value
Thickness of Cu Interconnect, h_I (nm)	1,000
Area Fraction of Cu Interconnect, w / λ	0.5
Cu-to-oxide Selectivity, $S_{Cu/ox}$	14.1
Wafer-scale Uniformity Factor, β	0.83
Thickness of Cu Deposition, h_{Cu} (nm)	1,500
Cu Deposition Factor, α, for Erosion Model	0.1
Cu Deposition Factor, α, for Dishing Model	1
Initial Step-height, h_{si} (nm)	900
Amount of Overpolishing, Δh_o (nm)	10
Applied Pressure, p_{av} (kPa)	14
Young's Modulus of Pad, E_p (MPa)	300
Poisson's Ratio of Pad, ν_p	0.3
Yield Strength of Pad, Y_p (MPa)	20
Radius of Pad Asperity, R_a (μm)	6.0
Spacing between Pad Asperities, λ_a (μm)	120

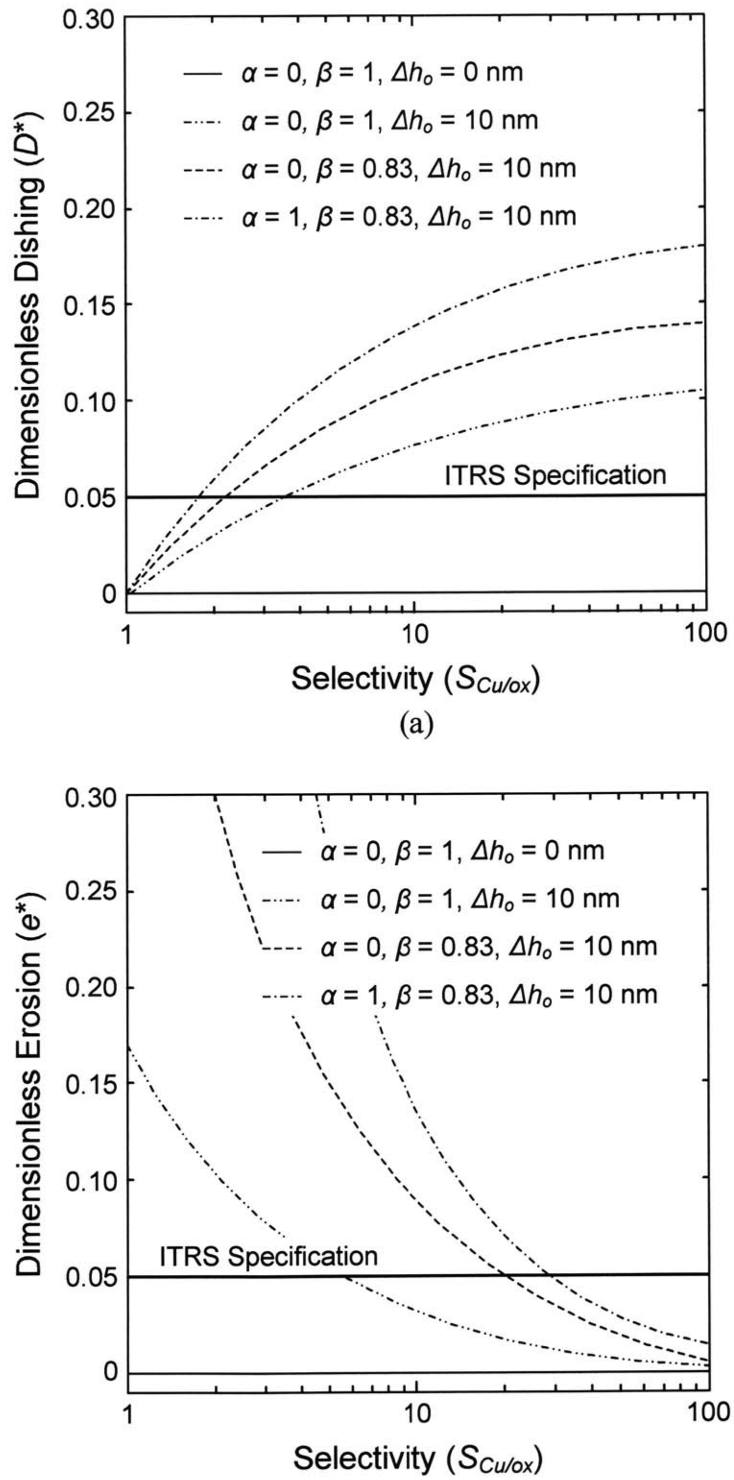


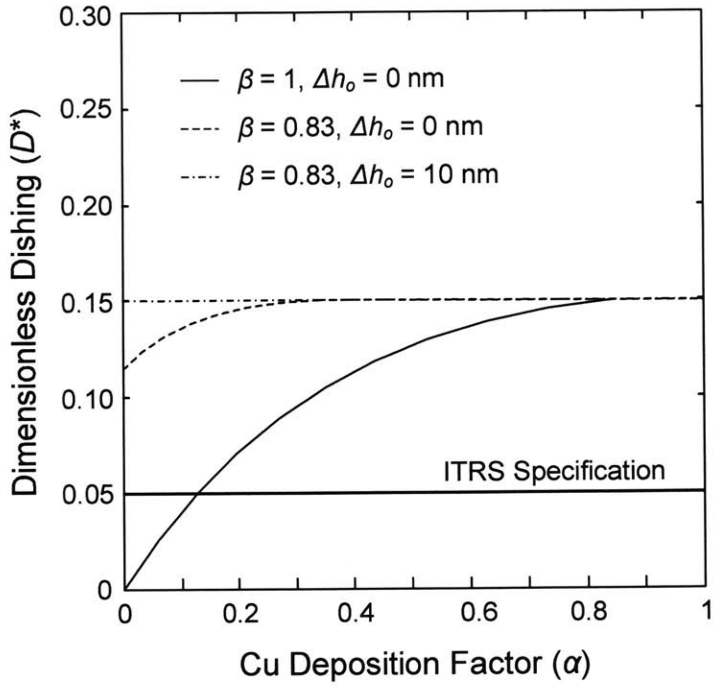
Figure 3.10 Effect of the slurry selectivity on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).

amount of overpolishing time at any specific feature, which is determined by the initial pattern geometry, wafer-scale uniformity factor, and the amount of overpolishing at the global reference point.

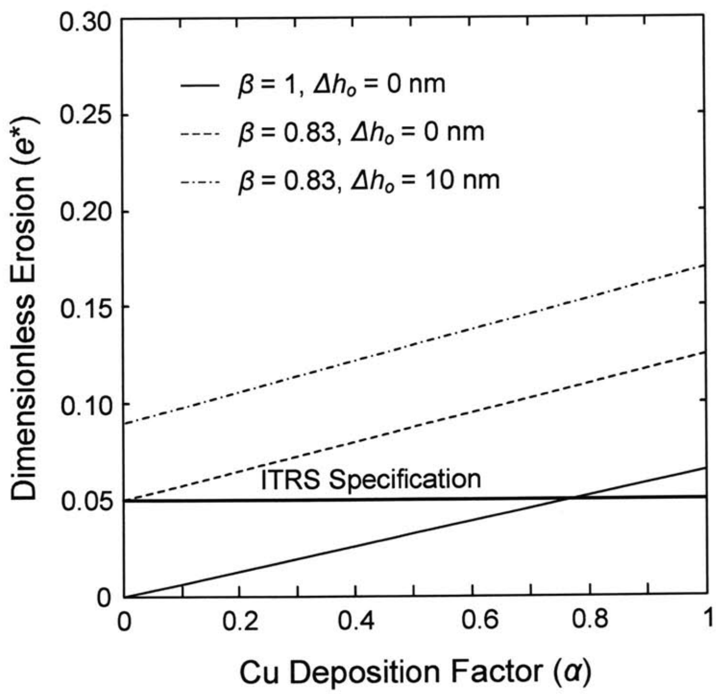
Figure 3.11 shows the effect of Cu deposition factor, α, on dishing and erosion. If α is zero, the initial surface can be considered as a field region and thus, there will be no die-scale variation until the polishing surface reaches the top of oxide layer. Both dishing and erosion developed up to this moment are zero, so are the final erosion and dishing if there is no wafer-scale variation and the global overpolishing: $\beta = 1$ and $\Delta h_o = 0$. If $\beta < 1$ or $\Delta h_o > 0$, however, the interconnect pattern affects erosion and dishing. As α increases, the material removal rate at the feature increases and thus this feature will reach the top of oxide layer earlier than the in the field region. Combined with the effect of β and Δh_o, α also affects the overpolishing time, and dishing and erosion increase.

Figure 3.12 shows the effect of wafer-scale uniformity factor, β, on dishing and erosion. If $\beta = 1$, there is no wafer-scale variation. Additionally, if α is zero, there is no die-scale variation before the polishing surface reaches the oxide layer, and the dishing and erosion can be maintained as zero if polishing is stopped at the exact end-point, i.e., $\Delta h_o = 0$. If $\beta < 1$, the feature in the fastest die reaches the oxide earlier than that in the slowest die, and thus the amount of overpolishing in the fastest die increases. Although the commercial equipment claims that within-wafer non-uniformity (WIWNU) less than 5%, based on the standard deviation of the material removal rate distribution across the wafer, standard deviation is not a good measure of wafer-scale uniformity since it may not have information about the worst case in a wafer. Therefore, β represents the maximum value of the wafer-scale material removal rate variation to focus on the maximum dishing and erosion in a wafer.

Figure 3.13 shows the effect of the amount of overpolishing, Δh_o, at the global reference point on dishing and erosion. If $\beta = 1$ and $\alpha = 0$, there is no wafer- and die-scale variations in a wafer, and thus dishing and erosion can be zero if $\Delta h_o = 0$. As Δh_o increases, however, the amount of overpolishing at the fastest die increases and thus, dishing and erosion increase. In the polishing process, Δh_o is determined by various end-point detection techniques and maintained as the certain value to ensure there is no excess Cu, or barrier layer residue, on the oxide. This is because first, the slowest region in the slowest die may or may not be the field region depending

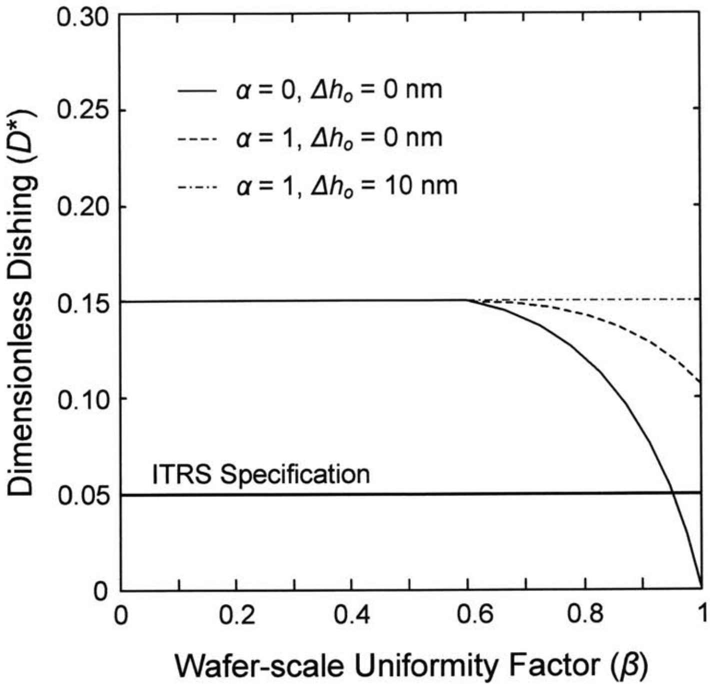


(a)

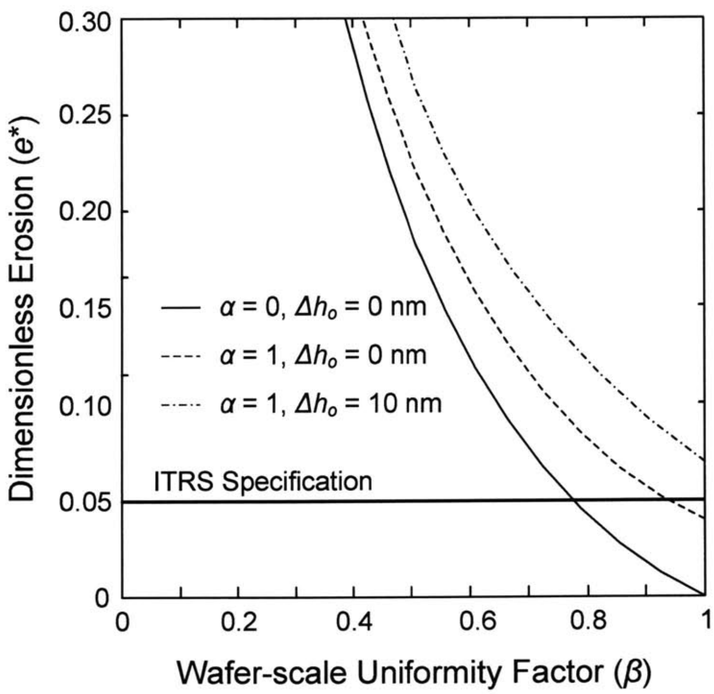


(b)

Figure 3.11 Effect of the Cu deposition factor, α, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).

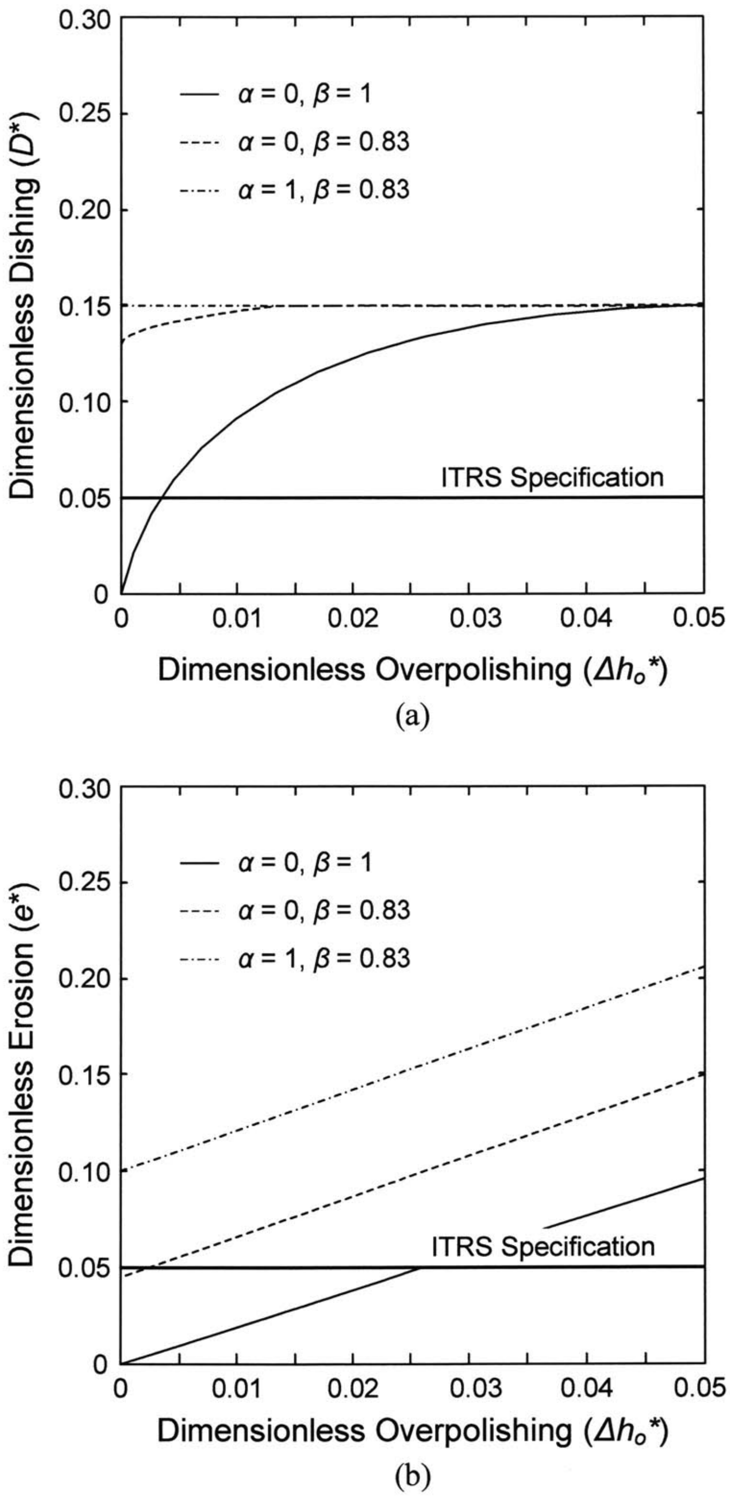


(a)



(b)

Figure 3.12 Effect of the wafer-scale uniformity factor, β, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).



(a)

(b)

Figure 3.13 Effect of non-dimensional overpolishing, Δh_o^*, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).

on the layout of subdies and, second, the time to polish the barrier layer may not be significantly low if the hardness of the barrier layer is harder than that of Cu.

Figure 3.14 shows the effect of applied pressure and the yield strength of the polishing pad on dishing and erosion. Cu dishing is proportional to the amount of maximum pad deformation if the pattern geometry of the feature and the slurry of selectivity is fixed. As p_{av} / Y_p decreases, i.e., the nominal pressure decreases or the yield strength of the pad increases, dishing decreases.

In the general rough pad contact, deformations of pad asperities are not fully elastic or plastic. Only some of the pad asperities are in the plastic contact mode and rest in the elastic mode. The ratio of elastic and plastic deformation depends on the surface topography and the applied load. In the case of elastic contact mode, the yield strength of the pad in Fig. 3.14 is replaced by the Young's modulus of the pad with corresponding range of parameter, p_{av} / E_p. The pad stiffness or yield strength does not significantly affect dielectric erosion in the submicron level based on the smooth pad erosion model. However, the decrease of yield strength or Young's modulus of pad decreases erosion at the global wiring level, since erosion depends on dishing. As the applied pressure increases, material removal rates of both Cu and SiO_2 increase at the same rate. Accordingly, in the smooth pad model, erosion is the time independent value if Δh_o is fixed instead of overpolishing time and does not depend on the applied pressure.

Figures 3.15 and 3.16 show the effect of pad topography, expressed by the spacing, λ_a, and the radius curvature, R_a, of pad asperities, on erosion and dishing. These parameters represent how much of the pressure is localized on the pad asperities. With the same radius of curvature and height of the pad asperities, the normal load on each pad asperities decreases as λ_a decreases since the number of the pad asperities per unit area increases. Therefore, the deformation of pad asperities decreases and thus Cu dishing decreases. On the other hand, if λ_a is fixed, Cu dishing decreases as R_a increases. Additionally, as λ_a increases or R_a decreases, the contact regime can change from the elastic to the plastic contact, which requires the investigation of pad wear. Furthermore, as the pad gets smoother, slurry flow and the number of abrasive particles trapped in the each pad asperities may decrease.

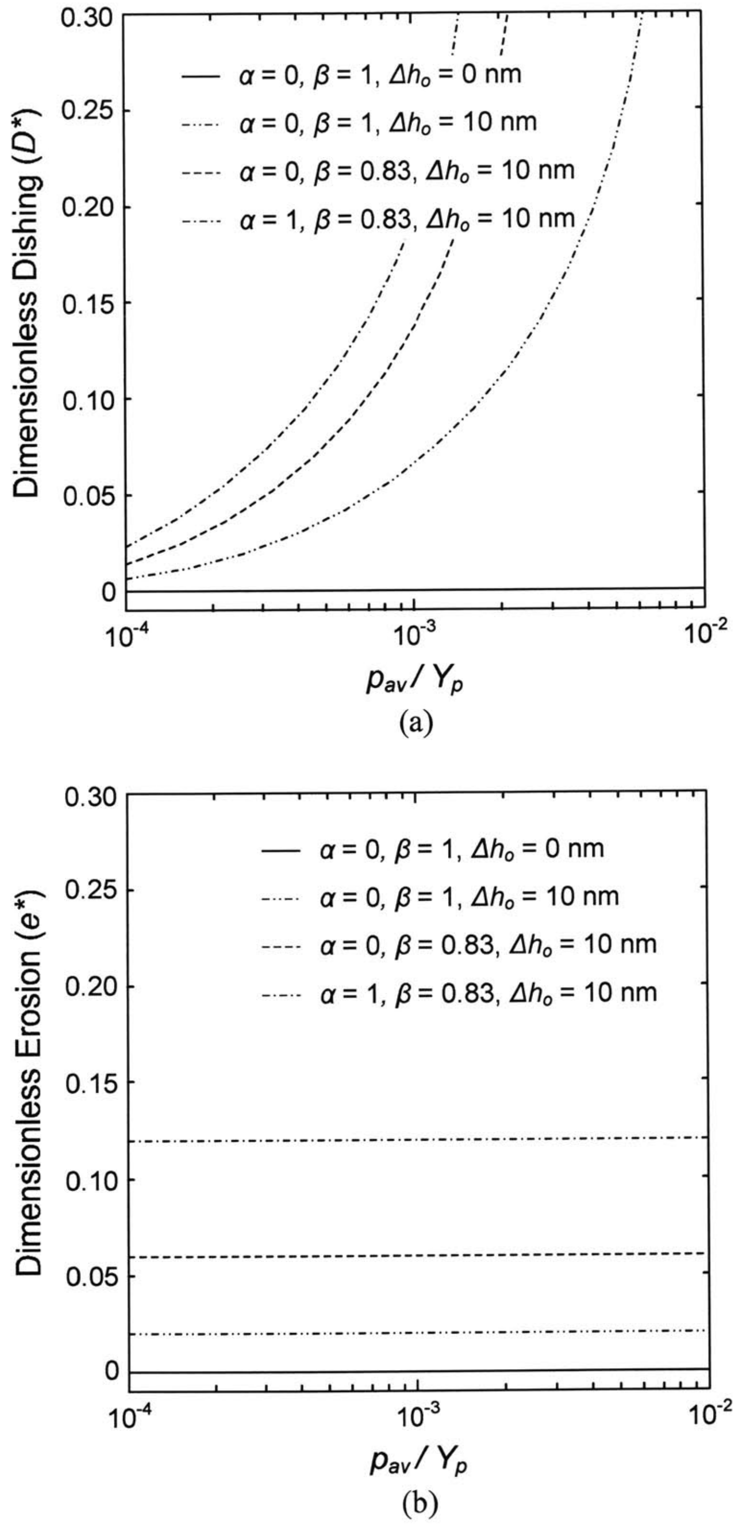


Figure 3.14 Effect of the applied pressure and yield strength ratio, p_{av} / Y_p, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).

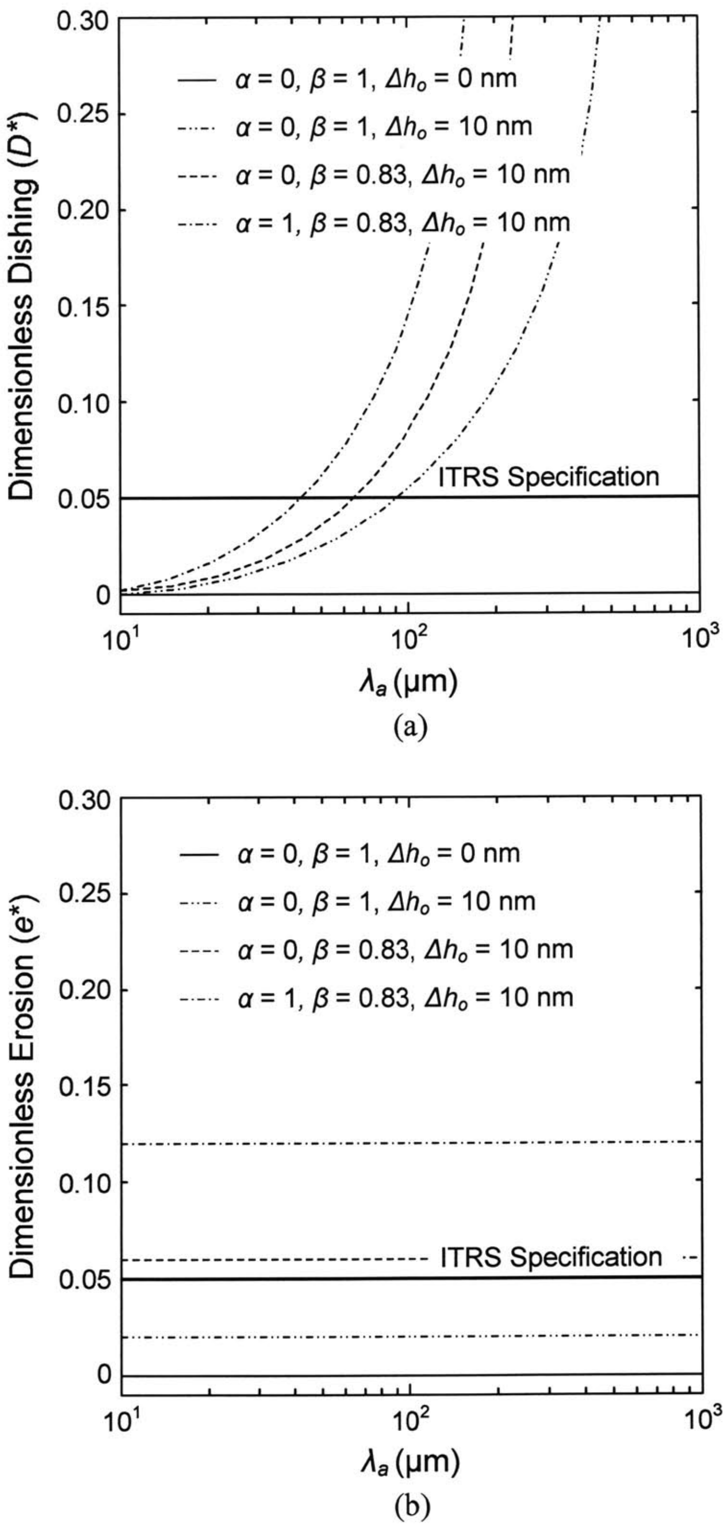
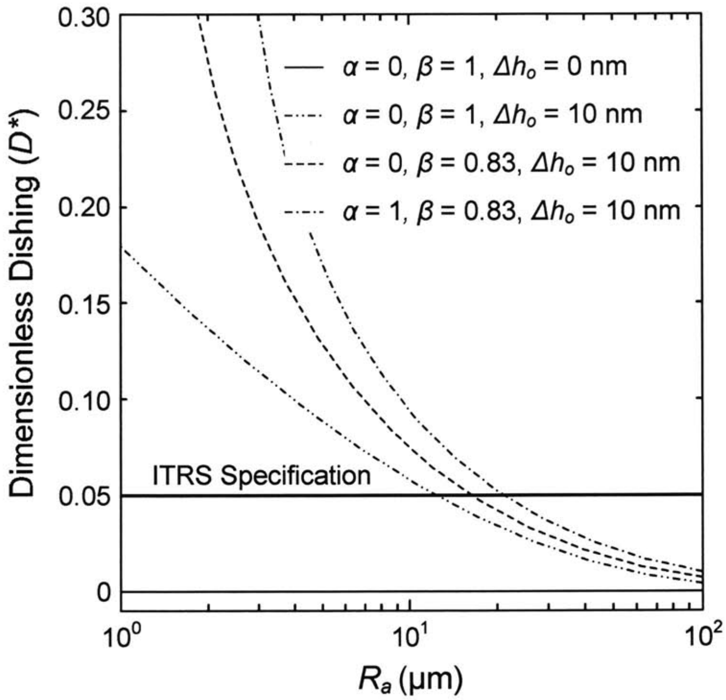
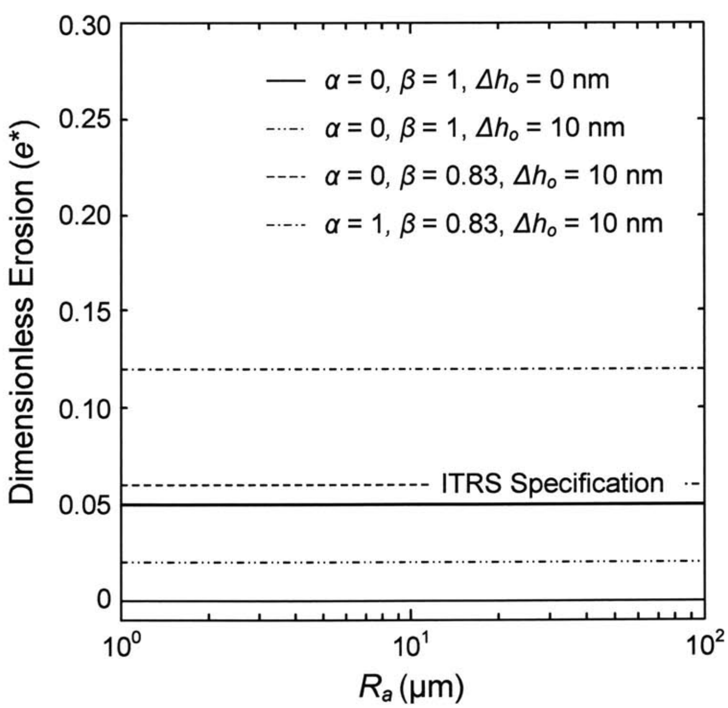


Figure 3.15 Effect of the pad asperity spacing, λ_a, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).



(a)



(b)

Figure 3.16 Effect of the radius of curvature of pad asperities, R_a, on (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model).

3.4.2 Process Optimization

From the model analysis in the previous subsection, the model parameters needs to be optimized to minimize dishing and erosion as follows.

To reduce Cu dishing,

- Use slurry with a low selectivity: $S_{Cu/ox} \rightarrow 1$
- Decrease the initial pattern geometry variation: $\alpha \rightarrow 0$
- Increase wafer-scale uniformity: $\beta \rightarrow 1$
- Minimize the amount of overpolishing: $\Delta h_o \rightarrow 0$
- Use a low pressure and/or a stiff pad: decrease p_{av} / E_p or p_{av} / Y_p
- Use a smooth pad: decrease λ_a and increase R_a

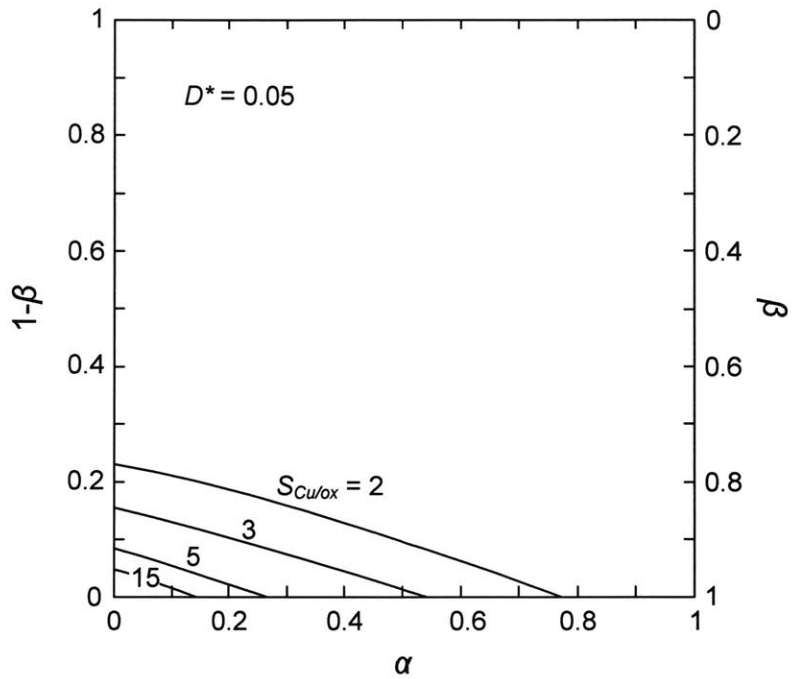
To reduce dielectric erosion,

- Use slurry with a high selectivity: $S_{Cu/ox} \gg 1$
- Decrease the initial pattern geometry variation: $\alpha \rightarrow 0$
- Increase wafer-scale uniformity: $\beta \rightarrow 1$
- Minimize the amount of overpolishing: $\Delta h_o \rightarrow 0$

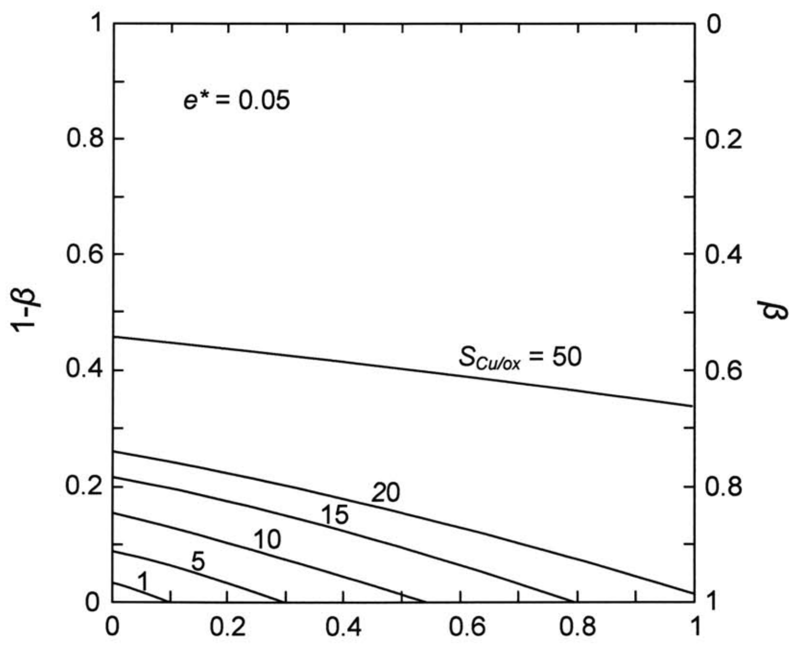
The amount of overpolishing depends primarily on the end-point technique during polishing, which can be controlled separately. Thus in this section, it is assumed that there is no overpolishing, $\Delta h_o = 0$. Although the low applied pressure is beneficial in reducing dishing, it also reduces the material removal rate, which is directly related to the throughput. The polishing pad requires a higher Young's modulus and a smooth surface. The effect of pad properties on the slurry and abrasive distribution is beyond the scope of this work and requires further investigation.

Thus the Cu deposition factor α, the wafer-scale unformity factor β, and the slurry selectivity $S_{Cu/ox}$ are chosen as optimization parameters. Slurry selectivity $S_{Cu/ox}$ conflicts with minimum dishing and erosion requirements in single-step polishing. Nevertheless, by decreasing α and increasing β, the single step-polishing can be made to satisfy a given specification.

In Figs 3.17 and 3.18, requirements for Cu deposition factor (α), wafer-scale uniformity factor (β) for various selectivites ($S_{Cu/ox}$), to satisfy the industrial specifications (5%) of



(a)



(b)

Figure 3.17 Requirement of the wafer-scale uniformity factor, β, and Cu deposition factor, α, to meet the industrial specification (5%) for (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model) with various slurry selectivities.

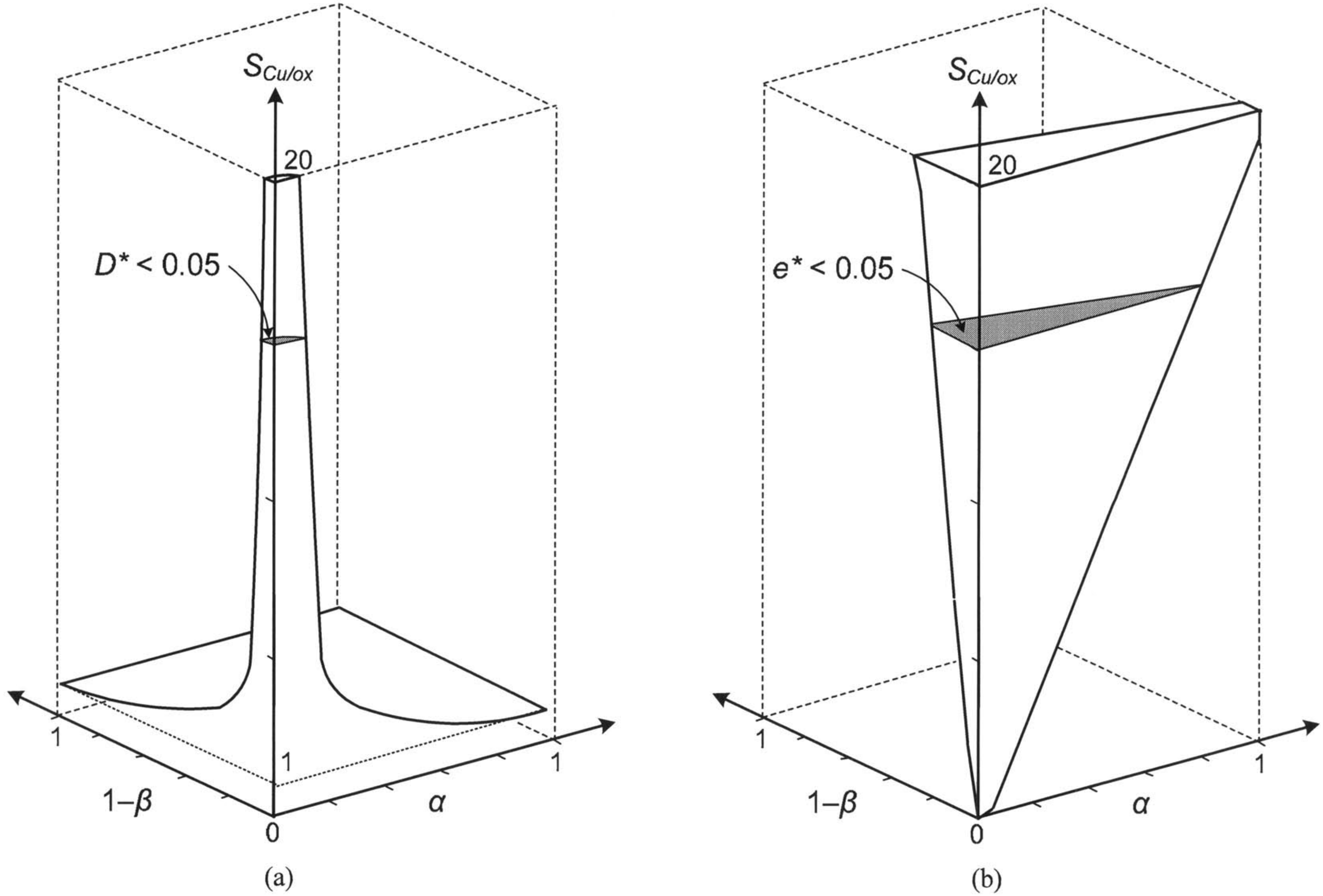


Figure 3.18 Three dimensional plot of requirements of the wafer-scale uniformity factor, β, and Cu deposition factor, α, to meet the industrial specification (5%) for (a) Cu dishing (simplified rough pad model, fully plastic) and (b) dielectric erosion (smooth pad model) with various slurry selectivities.

dielectric erosion and Cu dishing, are presented. As the selectivity of the slurry increases, the requirements of α and β for erosion becomes more flexible but for dishing becomes tighter. In single-step polishing, for example, Cu deposition factor should be less than 0.1 and the wafer-scale uniformity factor needs to be greater than 0.95 to maintain both erosion and dishing within 5% of the interconnect thickness across the wafer if the polishing slurry has a selectivity of 15.

3.5 Summary

In this chapter, integrated erosion and dishing models have been developed by combining wafer-, die- and feature-scale non-uniformities with feature-scale step-height models. First, non-uniformities in Cu CMP at various scales are defined. The plausible causes of erosion and dishing at wafer-, die- and feature-scales are identified in terms of the geometric and physical parameters. Such parameters include: Cu interconnect deposition factor, α , wafer-scale uniformity factor, β, and Cu-to-oxide selectivity, $S_{Cu/ox}$. To model wafer-, die- and feature-scale non-uniformities, it is required to consider three separate points on the wafer. First, to calculate the wafer-scale non-uniformity, field regions in the slowest and the fastest dies are considered. These two field regions are defined as the wafer and die reference points, respectively. Additionally, to calculate die-scale non-uniformities, the general feature in the fastest die, which is the same die with the local reference point, is considered. Feature-scale non-uniformity is characterized as Cu dishing and dielectric erosion.

Second, an overpolishing time at each feature is expressed in terms of non-uniformiy parameters at wafer-, die- and feature-scales, and erosion and dishing is calculated from the step-height model at the process end-point. In the smooth pad model, the amount of dishing is neglected due to its small pad deformation and erosion increases at the asymptotic material removal rate, MRR_{∞}. In the discrete pad model, the amount of dishing is related to the pattern geometry, slurry selectivity, Young's modulus of pad, initial thickness of pad. In the simplified rough pad model with plastic asperity contact condition, the dishing and erosion analysis is exactly the same as that in the discrete pad model by replacing H_o by $\lambda_a^2 / 6\pi R_a$ and E_p by Y_p. In both discrete and simplified rough pad model, erosion increase at the asymptotic material removal rate as polishing progresses. However, the actual amount of erosion is less than the smooth pad model due to the time required to develop a dishing.

Finally, based on the parameter analysis, the following conclusions are drawn to reduce dishing and erosion.

- Use slurry with a low selectivity to decrease dishing or a high selectivity to decrease erosion,
- Decrease the initial pattern geometry variation,
- Increase wafer-scale uniformity factor,
- Minimize the amount of overpolishing, and
- Use a low pressure and a stiff, smooth pad.

The requirements of wafer-scale uniformity and the initial deposited pattern geometry to satisfy both dishing and erosion specification for a given slurry selectivity were obtained. In single-step polishing, for example, Cu deposition factor should be less than 0.1 and the wafer-scale uniformity factor greater than 0.95 to maintain both erosion and dishing within 5% of the interconnect thickness across the wafer if the slurry selectivity is 15.

Nomenclature

A_{Cu} = total Cu interconnect area in a characteristic area (m^2)
A_o = nominal area (m^2)
A_f = area fraction of Cu interconnects
$B_{11}, B_{12}, B_{21}, B_{22}$ = dimensionless geometric coefficients
D, D_{∞}, D_i = total, asymptotic and initial Cu dishing (m)
D^*, D^*_{∞}, D^*_i = dimensionless total, asymptotic and initial Cu dishing
E_p, E^* = Young's modulus and effective Young's modulus of pad (N/m^2)
F = applied normal force (N)
H_o = undeformed pad thickness (m)
MRR = material removal rate (m/s)
MRR_{∞} = asymptotic material removal rate (m/s)
R_a = radius of curvature of asperity (m)
$S_{Cu/ox}, S_{b/ox}$ = Cu-to-oxide and barrier-to-oxide selectivities
WIWNU = within-wafer non-uniformity
Y_p = yield strength of pad material (N/m^2)
e, e_d, e_w = total, die-scale and wafer-scale dielectric erosion (m)
e^*, e^*_d, e^*_w = dimensionless total, die-scale and wafer-scale dielectric erosion
h = film thickness (m)
h_{Cu}, h_b, h_I = film thickness of Cu, barrier layer and interconnect (m)
h_h, h_l = surface height at high and low features (m)
$h_s, h_{si}, h_s(\infty)$ = step-height, initial step-height and asymptotic step-height (m)
k_p = Preston constant (m^2/N)
$k_{p_{Cu}}, k_{p_{ox}}$ = Preston constant of Cu and oxide (m^2/N)
p = pressure (N/m^2)
p_{av} = average pressure (N/m^2)
t = polishing time (s)
t_{ep} = process end-point (s)
t_1, t_2, t_o = polishing time at the end of Stage 1 & 2, and overpolishing time (s)
t^*_2, t^*_o = dimensionless polishing time at Stage 2 & 3
v_R = magnitude of relative velocity (m/s)

x, y, z = Cartesian coordinates

$\Delta h_{ff}, \Delta h_{sf}$ = material removed in the fastest and the slowest regions (m)

Δh_o = amount of overpolishing (m)

α = Cu deposition factor

β = wafer-scale uniformity factor

δ, δ_o = maximum deformation of pad (asperity) (m)

λ_a = spacing of pad asperities (m)

μ_h = mean of the amount of material removed in a wafer (m)

ν_p = Poisson's ratio of pad material

σ_h = standard deviation of the amount of material removed in a wafer

τ_2, τ_3 = time constants of Stage 2 & 3 (s)

References

Bullen, D., Scarfo, A., Koch, A., Bramano, D., Coppeta and J., Racz, L., 2000, "In Situ Technique for Dynamic Fluid Film Pressure Measurement during Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 147, pp. 2741-2748.

Chekina, O. G., Keer, L. M. and Liang, H., 1998, "Wear-Contact Problems and Modeling of Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 145, pp. 2100-2106.

Coppeta, J., Rogers, C., Racz, L., Philipossian, A. and Kaufman, F. B., 2000, "Investigating Slurry Transport Beneath a Wafer during Chemical Mechanical Polishing Processes," *J. Electrochemical Soc.*, vol. 147, pp. 1903-1909.

Fu, G. and Chandra, A., 2003, "An Analytical Dishing and Step Height Reduction Model for Chemical Mechanical Planarization (CMP)," *IEEE Trans. on Semiconductor Manufacturing*, vol. 16, pp. 477-485.

Fu, G. and Chandra, A., 2002, "A Model for Wafer Scale Variation of Material Removal Rate in Chemical Mechanical Polishing Based on Viscoelastic Pad Deformation," *J. Electronic Materials*, vol. 31, pp. 1066-1073.

Lee, B., 2002, "Modeling of Chemical Mechanical Polishing for Shallow Trench Isolation," Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA.

Park, T., Tugbawa, T. and Boning, D., 2000, "Overview of Methods for Characterization of Pattern Dependencies in Copper CMP," *Proc. CMP-MIC Conf.*, pp. 196-205.

Park, T., Tugbawa, T., Boning, D., Chidambaram, C., Borst, C. and Shin, G., 2004, "Chip-Scale Modeling of Electroplated Copper Surface Profile," *J. Electrochemical Soc.*, vol. 151, pp. C418-C430.

Runnels, S. R., Kim, I., Schleuter, J., Karlsrud, C. and Desai, M., 1998, "A Modeling Tool for Chemical-Mechanical Polishing Design and Evaluation," *IEEE Trans. on Semiconductor Manufacturing*, vol. 11, pp. 501-510.

Runnels, S. R., Miceli, F. and Kim, I., 1999, "Validation of a Large Area Three Dimensional Erosion Simulator for Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 146, pp. 4619-4625.

Seok, J., Sukam, C. P., Kim, A. T., Tichy, J. A. and Cale, T.S., 2004, "Material Removal Model for Chemical-Mechanical Polishing Considering Wafer Flexibility and Edge Effects," *Wear*, vol. 257, pp. 496-508.

Stavreva, Z., Zeidler, D., Pltner, M., Grasshoff, G. and Drescher, K., 1997, "Chemical-Mechanical Polishing of Copper for Interconnect Formation," *Microelectronic Engineering*, vol. 33, pp. 249-257.

Steigerwald, J. M., Zirpoli, R., Murarka, S. P., Price, D. and Gutmann, R. J., 1994, "Pattern Geometry Effects in the Chemical-Mechanical Polishing of Inlaid Copper Structures," *J. Electrochemical Soc.*, vol. 141, pp. 2842-2848.

Stine, B. E., Ouma, D. O., Divecha, R. R. and Boning, D.S., 1998, "Rapid characterization and Modeling of Pattern-Dependent Variation in Chemical Mechanical Polishing," *IEEE Trans. on Semiconductor Manufacturing*, vol. 11, pp. 129-140.

Tichy, J., Levert, J. A., Shan, L. and Danyluk, S., 1999, "Contact Mechanics and Lubrication Hydrodynamics of Chemical Mechanical Polishing," *J. Electrochemical Soc.*, vol. 146, pp. 1523-1528.

Tugbawa, T., Park, T. and Boning, D., 2002, "Integrated Chip-Scale Simulation of Pattern Dependencies in Copper Electroplating and Copper Chemical Mechanical Polishing Processes," *International Interconnect Technology Conference*, San Francisco, CA.

Vlassak, J. J., 2001, "A Contact-Mechanics Based Model for Dishing and Erosion in Chemical-Mechanical Polishing," *Proc. Mat. Res. Soc. Symp.*, vol. 671, pp. M4.61-M4.66.

Vlassak, J. J., 2004, "A Model for Chemical-Mechanical Polishing of a Material Surface Based on Contact Mechanics," *J. Mechanics and Physics of Solids*, vol. 52, pp. 847-873.

Warnock, J., 1991, "A Two-Dimensional Process Model for Chemimechanical Polishing Planarization," *J. Electrochemical Soc.*, vol. 138, pp. 2398-2402.

Yang, K., 2000, "General Analysis of Dual-Step CMP for Copper Damascene Process," *Proc. CMP-MIC Conf.*, pp. 23-32.

Zabasajja, J., Merchant, T., Ng, B., Banerjee, S., Green, D., Lawing, S. and Kura, H., 2001, "Modeling and Characterization of Tungsten Chemical and Mechanical Polishing Processes," *J. Electrochemical Soc.*, vol. 148, pp. 73-77.

CHAPTER 4

EXPERIMENTAL VALIDATION

4.1 Introduction

In this chapter, experimental means of determining the model parameters are outlined. The model parameters include: blanket wafer material removal rate, selectivity, wafer-scale uniformity factor, Cu deposition factor and the initial step-height. The parameters are obtained by polishing experiments on blanket and patterned wafers followed by profilometry and scanning electron microscopy (SEM). The feature-scale step-height models in Chapter 2 are validated by 100 mm patterned wafer polishing experiments and measurement of step-heights at various polishing times. Polishing rates of the slowest and the fastest dies are compared to determine the effect of the wafer-scale uniformity factor. The effect of pattern geometry is explained by measuring each subdie in the slowest and the fastest dies. Finally, Cu dishing and dielectric erosion at each subdie in the fastest die is measured and compared with the integrated erosion and dishing models developed in Chapter 3.

4.2 Parameter Determination: Experiments and Measurement

4.2.1 Consumables and Equipment

Commercial pads, Rohm and Hass IC1400, with K-type grooves were used in the experiments. The groove geometry and the thickness of the stack are shown in Fig. 4.1, and the material properties are listed in Table 4.1. The top layer is polyurethane with high Young's modulus and pores on the surface. The bottom layer is more compliant than the top layer to increase conformity between the wafer and the pad. The concentric K-type grooves are for enhancing slurry distribution during polishing. The sample surface profile of polishing pad is shown previously in Chapter 2, Fig. 2.10. The Young's modulus of the polishing pad is 300 MPa at room temperature and the yield strength is assumed to be 20 MPa.

A commercial slurry with alumina abrasive particles, Cabot Microelectronics iCue5001, was

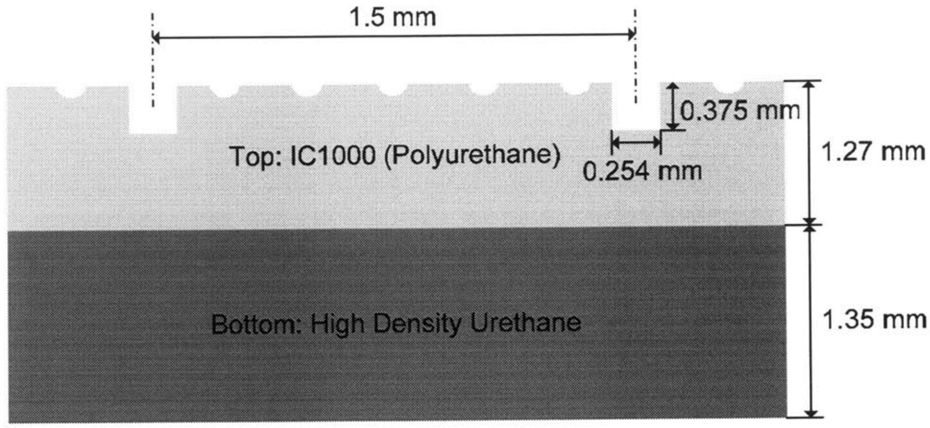


Figure 4.1 Schematic of Rohm and Haas IC1400 pad geometry.

Table 4.1 Properties of Rohm and Haas IC1400 pad.

Property	Top Layer	Composite
Designation	IC1000	IC1400
Material	Polyurethane	Polyurethane + Urethane
Diameter (mm)	300	300
Thickness (mm)	1.27	2.62
Young's Modulus (MPa) at 24 °C	300	56
Poisson's Ratio	0.3	–
Hardness (Shore D)	57	–
Yield Strength (MPa)*	20	–
Type of Groove	Concentric K	–
Size of Groove, (mm x mm)	0.254x0.375	–
Pitch of Groove (mm)	1.5	–
Size of Pore (μm)	20–60 (isolated)	–

– Rohm and Haas Corporation [http://www.rohmhaas.com].
* Polyurethane Elastomer [http://www.matweb.com].

used in both blanket and the patterned wafer polishing experiments. The particle diameter is between 0.8 μm to 11 μm as shown in Fig. 4.2 and the mean particle size is 3.6 μm. The physical properties of the slurry are listed in Table 4.2. Hydrogen peroxide was added during the slurry preparation process by the amount of 2.5% of total slurry volume and mixed before polishing for at least 30 min. It is well known that hydrogen peroxide reacts with Cu to form a thin soft Cu layer and thus increases Cu material removal rate.

All polishing experiments were conducted on a face-down, 100 mm CMP tool shown in Fig. 4.3. The wafer was mounted underneath the wafer carrier and pushed against the polishing pad with a certain normal load and both were rotated in the same direction. The experimental conditions for both blanket and patterned wafer are listed in Table 4.3. During both blanket- and patterned-wafer experiments, the average pressure was either 14 kPa (2 psi) or 28 kPa (4 psi) and the rotational speeds of wafer and pad were the same. Slurry was fed from the edge of the wafer and was carried by the pad through the pad/wafer interface. Slurry flow rate was 50 ml/min.

4.2.2 Material Removal Rate and Selectivity

To characterize selectivity in the Cu CMP process, the material removal rate in a blanket wafer for each substrate was determined. Although it is required to achieve material removal rates of each layer in substrate stacks as shown in Fig. 4.4, it is convenient to use separately coated blanket wafers to determine material removal rate of each substrate material. Thus, 100 mm blanket wafers with various coatings, Cu, SiO_2, Ta, and TaN, were polished. The properties of these coatings are listed in Table 4.4. The total polishing time was limited to 1–2 min to ensure that only the coatings were polished across the wafer. The loss of weight, Δm, was measured to a resolution of 0.01 mg and converted into thickness removed, Δh, using the material density listed in Table 4.4.

Tables 4.5 and 4.6 show results of polishing experiments on blanket wafers. The Preston constant of Cu is between 2.5 and 3.5×10^{-13} Pa^{-1}, similar to that of the industrial value [Park *et al.*, 2000; ThomasWest Inc., 2005]. Although the material removal rate increases as the applied pressure increases, the nominal pressure in the patterned wafer polishing experiments was set at 14kPa (2psi) for better observation of surface profile evolution as polishing progresses. The current set of consumables shows the Cu-to-SiO_2 selectivity to be 14.1 and 11.1, depending on the process parameter sets listed in Table 4.6. The Ta-to-SiO_2 selectivity is 1.9, which is much

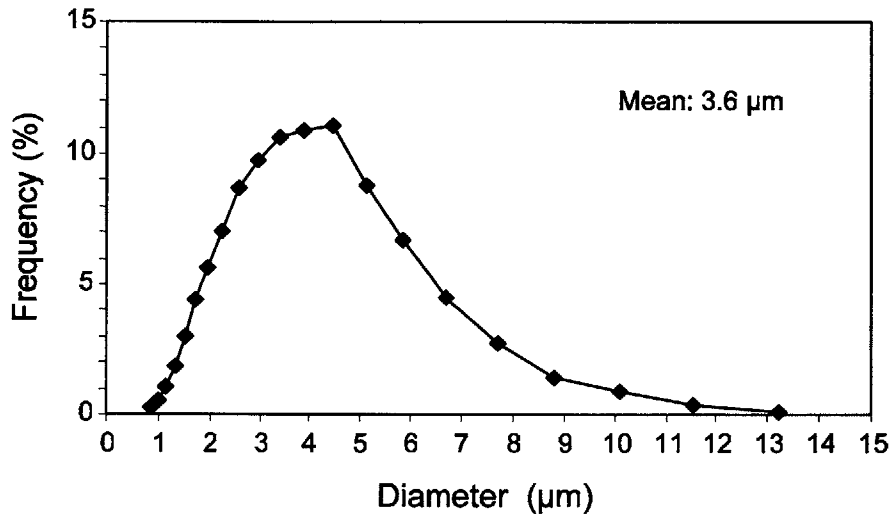


Figure 4.2 Abrasive particle size distribution in Cabot iCue5001 slurry: Cabot Microelectronic Corporation [http://www.cabotcmp.com].

Table 4.2 Properties of Cabot iCue5001 slurry.

Property	Value
Abrasive Material	Al_2O_3
Mean Particle Size (µm)	3.2
Hardness of Abrasive (MPa)	29,400
Volume Percent of Abrasive	5
pH	7.5
Viscosity (Pa·s)	0.0082
Additive	H_2O_2
Volume Percent of H_2O_2	2.5

– Cabot Microelectronic Corporation [http://www.cabotcmp.com].

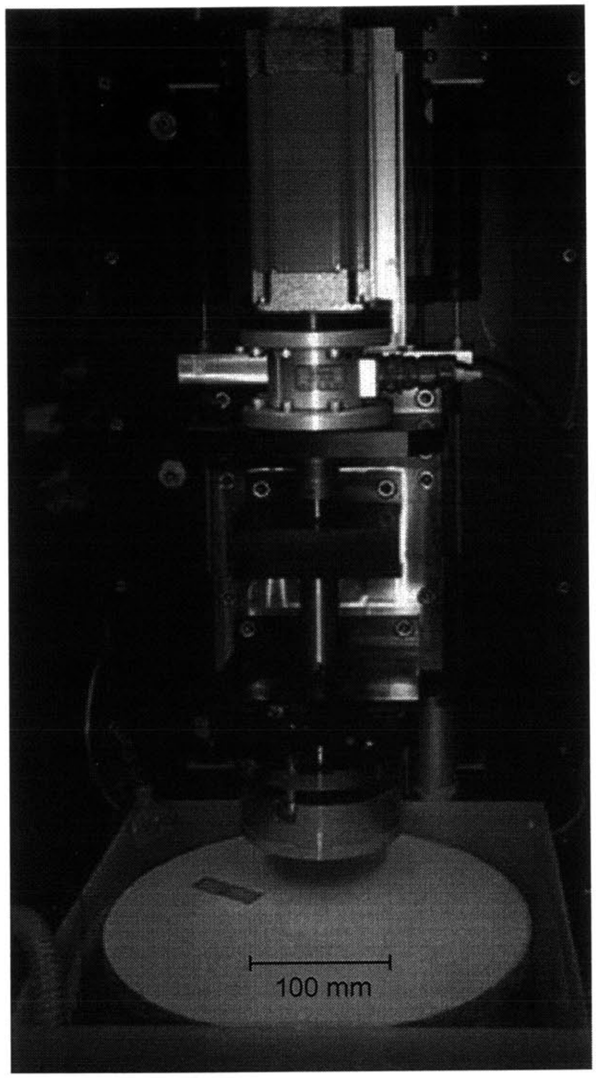
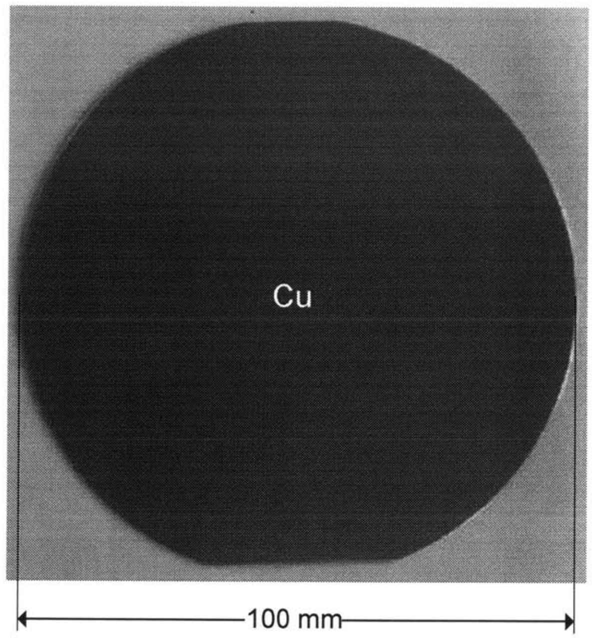


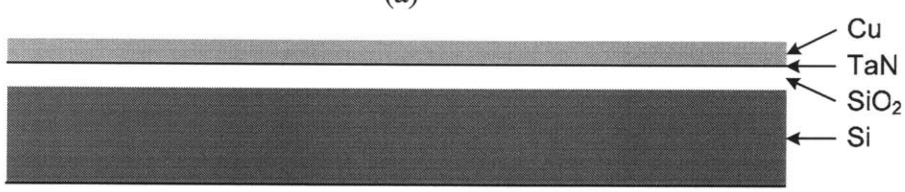
Figure 4.3 Photograph of the face-down, 100 mm CMP tool.

Table 4.3 Experimental conditions

Parameter	Present Work	Industrial CMP
Diameter of Wafer (mm)	100	200 – 300
Diameter of Pad (mm)	300	600 – 1000
Normal Load (N)	110 – 220	440 – 880
Normal Pressure (kPa)	14 – 28	14 – 28
Rotational Speed of Wafer (rpm)	75 – 100	50 – 120
Rotational Speed of Pad (rpm)	75 – 100	50 – 120
Center-to-Center Distance (mm)	85	200 – 300
Relative Velocity (m/s)	0.67 – 0.90	0.5 – 3.0
Slurry Flow Rate (ml/min)	20 – 100	50 – 200
Polishing Time (s)	60 – 480	60 – 300



(a)



(b)

Figure 4.4 The 100 mm blanket Cu wafer: (a) photograph and (b) schematic of substrate stacks.

Table 4.4 Mechanical and physical material properties of coatings and the substrate.

Property	Cu	TaN	SiO_2	Si
Thickness, h (nm)	1,000	20	1,500	0.5×10^6
Deposition Process	PVD	CVD	TEOS	–
Young's Modulus, E (GPa)	128	180	73	112
Hardness, Brinell (MPa)	686	10,800	7,840	24,500
Density, ρ (kg/m^3)	8,941	13,800	2,500	2,330
Electrical Resistivity, ρ ($\mu\Omega\cdot$cm)	1.741	130	1.0×10^{20}	1.0×10^5
Reflectivity (%)	90	78	–	28

– CRC Materials Science & Engineering Handbook, CRC Press Inc.
– ASM Metals Reference Book, American Society for Metals.
– Matweb Inc., [http://www.matweb.com].

Table 4.5 Results of polishing experiments on blanket wafers with various coatings.

Experimental Conditions	Coating	Δt (min)	m_i (mg)	m_f (mg)	Δm (mg)	Δh (nm)	MRR (nm/min)	k_p (x10^{-13} Pa^{-1})
p = 14 kPa, v_R = 0.9 m/s, Q = 50 ml/min	Cu	2	9031.98	9000.72	31.26	446.6	223	2.94
		2	9000.72	8972.08	28.64	409.1	204	2.70
		2	8972.08	8942.96	29.12	416.0	207	2.74
		2	9046.26	9018.30	27.96	399.4	199	2.63
		2	9018.30	8987.02	31.28	446.9	223	2.95
		2	8987.02	8957.44	29.58	422.6	211	2.79
	SiO_2	2	8694.48	8693.94	0.54	7.7	14	0.18
		2	8693.94	8693.23	0.71	10.1	18	0.24
		2	8693.23	8692.71	0.52	7.4	13	0.18
		2	8694.09	8693.43	0.66	9.4	17	0.22
		2	8693.43	8692.85	0.58	8.3	15	0.20
		2	8692.85	8692.24	0.61	8.7	16	0.21
p = 28 kPa, v_R = 0.67 m/s, Q = 150 ml/min	Cu	1	11262.15	11239.42	22.73	324.7	324	2.88
		1	11269.46	11244.01	25.45	363.5	362	3.22
		1	11294.71	11263.48	31.23	446.2	445	3.95
		1	11276.45	11246.66	29.79	425.6	424	3.77
		1	11270.34	11243.70	26.64	380.6	379	3.37
		1	11275.67	11247.66	28.01	400.1	399	3.54
	SiO_2	2	8693.72	8692.50	1.22	17.5	31	0.28
		2	8692.50	8690.98	1.52	21.7	39	0.34
		2	8690.98	8689.64	1.34	19.1	34	0.30
	Ta	2	9418.60	9411.11	7.49	107.0	62	0.55
		2	9418.11	9411.36	6.74	96.3	56	0.50
		2	9414.55	9405.47	9.08	129.7	76	0.67
	TaN	2	9149.18	9140.76	8.42	120.3	39	0.35
		2	9158.70	9151.14	7.56	108.0	35	0.31
		2	9057.86	9048.34	9.52	135.9	44	0.39

Table 4.6 Experimental results on blanket wafers: MRR, Preston constant and selectivities.

Experimental Conditions	Parameter	Coating	Present Work	Industrial Value*
p = 14 kPa v_R = 0.9 m/s Q = 50 ml/min	MRR (nm/min)	Cu	211	–
		SiO_2	15	–
	k_p ($\times 10^{-13}$ Pa^{-1})	Cu	2.79	–
		SiO_2	0.20	–
	Selectivity	$S_{Cu/ox}$	14.1	–
p = 28 kPa v_R = 0.67 m/s Q = 150 ml/min	MRR (nm/min)	Cu	389	800
		Ta	65	60
		SiO_2	35	20
	k_p ($\times 10^{-13}$ Pa^{-1})	Cu	3.45	3.80
		Ta	0.58	0.29
		SiO_2	0.20	0.10
	Selectivity	$S_{Cu/ox}$	11.1	40.0
		$S_{b/ox}$	1.9	3.0

* Industrial values are based on the conditions that p_{av} = 28 kPa, v_R =1.25 m/s.

lower than that of Cu-to-SiO_2. Although the hardness of Ta is close to the hardness of Cu, the material removal rate of Ta is much less than that of Cu.

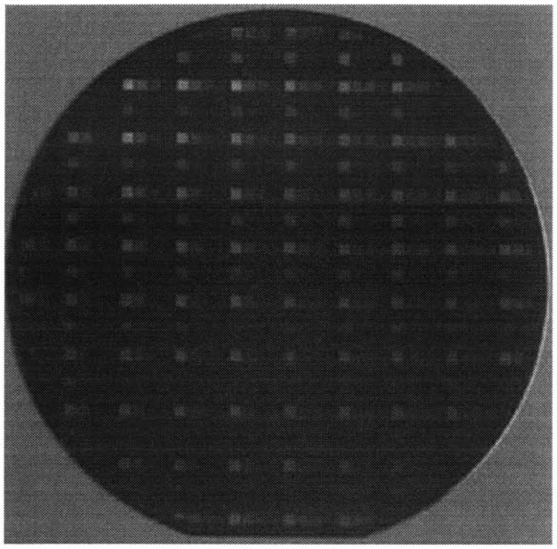
4.2.3 Cu Deposition Factor and Initial Step-Height

Figures 4.5 – 4.8 show schematics and photographs of the dies and subdies in the patterned masks. Tables 4.7 – 4.10 list the mechanical and physical properties of film substrates on the patterned wafers and the pattern layout of each mask. The linewidths of Cu interconnect in Mask 1 are mostly smaller than 5 μm, except two 25 μm and one 100 μm subdies. On the other hand, Mask 2 has global wiring level interconnects, in the range 10 – 100 μm.

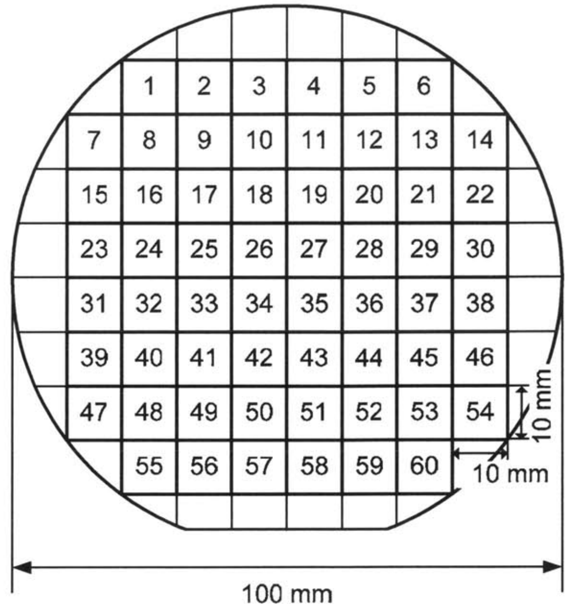
In the intermediate and global wiring level, both Cu deposition factor, α, and initial the step-height, h_{si}, can be easily measured by the surface profilometer as shown in Fig. 4.9. As linewidth decreases, however, the proflometer reaches its measurement limit. The Tencor P-10 surface profilometer used in this investigation can only measure features separated by 5 μm or more. When the linewidth is smaller than 5 μm, SEM was used to measure α and h_{si}. Figure 4.10 shows SEM micrographs of each subdie in Mask 1. Table 4.11, Figs 4.11 and 4.12 show the α and h_{si} values for both masks. Both α and h_{si} decrease as the interconnect linewidth decreases. However, if the linewidth is greater than 25 μm, which lies in the global wiring level, α is close to one and h_{si} is close to the thickness of interconnects. That is, the underlying trench pattern is closely reproduced by Cu deposition. In the submicron linewidth level, α becomes less than 0.1 and h_{si} decreases to about 10% of interconnect thickness, which makes the subdie resemble to the field area.

4.2.4 Wafer-scale Uniformity Factor

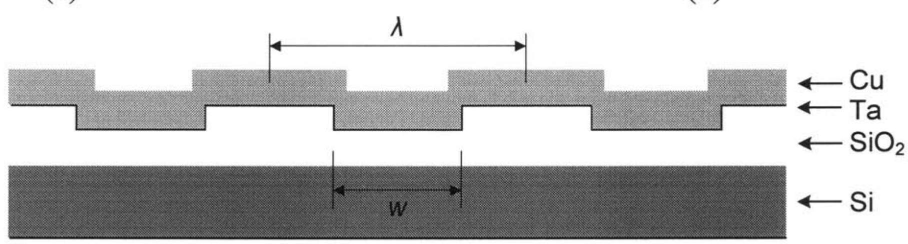
The wafer-scale uniformity factor β was obtained by blanket and patterned wafer polishing experiments by comparing the polishing time when the excess Cu in the field subdie at the fastest and the slowest die. The material removal rate of field region is the same as that of blanket wafer and the excess Cu thickness, h_{Cu}, across the wafer is assumed to be the same in PVD Cu deposition. Therefore, the wafer-scale uniformity factor, β, the ratio of material removed at the slowest and the fastest fields, can be rewritten as the ratio of the polishing times as:



(a)



(b)



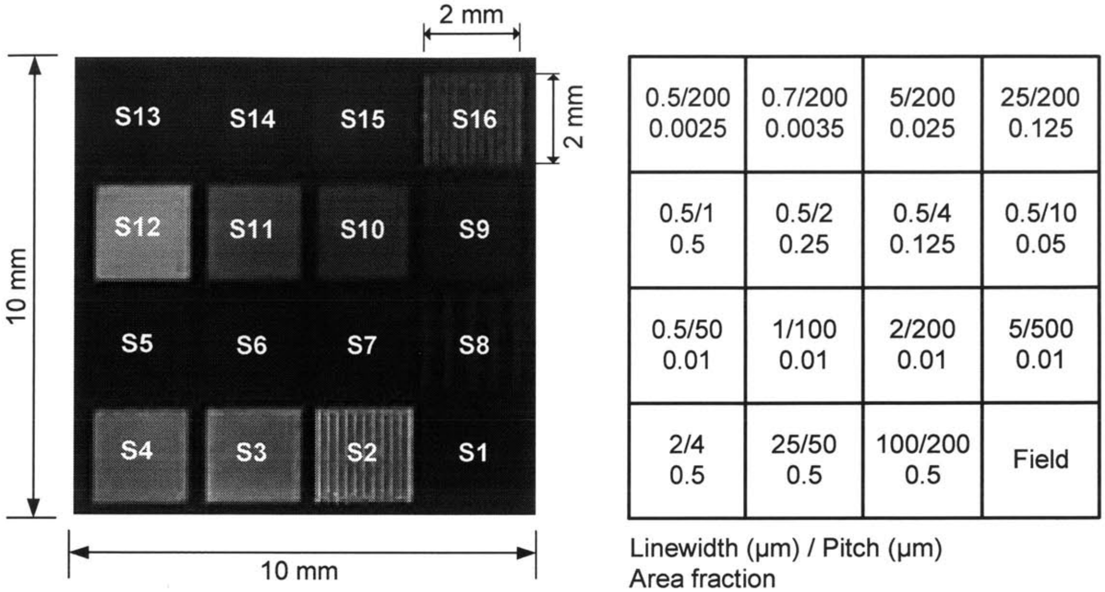
(c)

Figure 4.5 The 100 mm wafer patterned with Mask 1: (a) Photograph, (b) schematics of die layout and (c) schematic of cross-section.

Table 4.7 Mechanical and physical material properties of Cu and substrates (Mask 1).

Property	Cu	Ta	SiO_2	Si
Thickness, h (nm)	1,500	20	1,500	0.5×10^6
Deposition Process	PVD	CVD	TEOS	–
Young' Modulus, E (GPa)	128	180	73	112*
Hardness, Brinell (MPa)	686	980	7,840	24,500
Density, ρ (kg/m^3)	8,941	16,656	2,500	2,330
Electrical Resistivity, ρ (μΩ·cm)	1.741	12.5	1.0×10^{20}	1.0×10^5
Reflectivity (%)	90	78	–	28

– CRC Materials Science & Engineering Handbook, CRC Press Inc.
– ASM Metals Reference Book, American Society for Metals.
– Matweb Inc., [http://www.matweb.com].



(a) (b)

Figure 4.6 Schematics of Mask 1: (a) die layout and (b) feature geometries in each die.

Table 4.8 Pattern layout of Mask 1.

Subdie No.	Cu linewidth, w (μm)	Pitch, λ (μm)	Area Fraction $A_f = w/\lambda$
1	Field	–	–
2	100	200	0.5
3	25	50	0.5
4	2	4	0.5
5	0.5	50	0.01
6	1	100	0.01
7	2	200	0.01
8	5	500	0.01
9	0.5	10	0.05
10	0.5	4	0.125
11	0.5	2	0.25
12	0.5	1	0.5
13	0.5	200	0.0025
14	0.7	200	0.0035
15	5	200	0.025
16	25	200	0.125

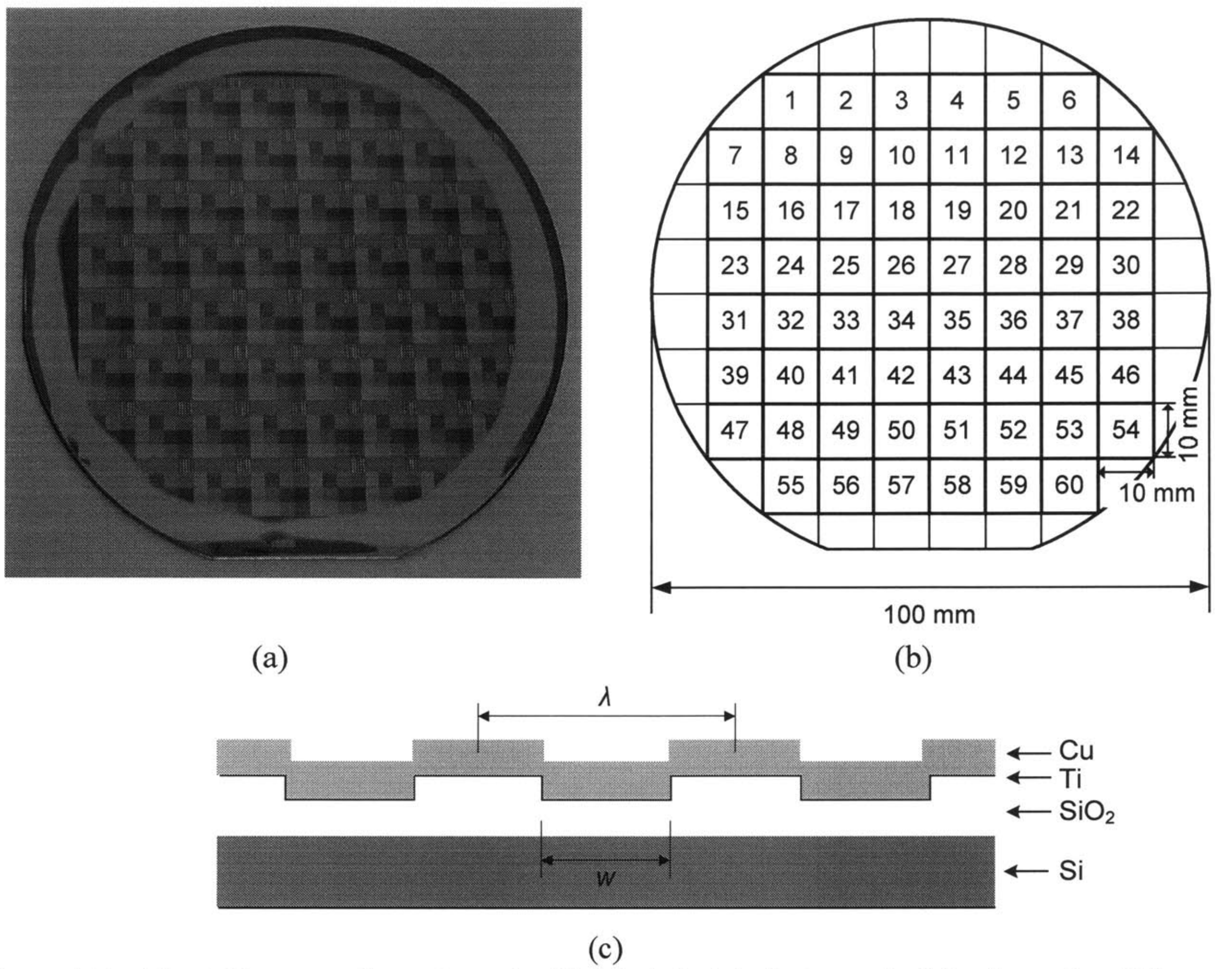
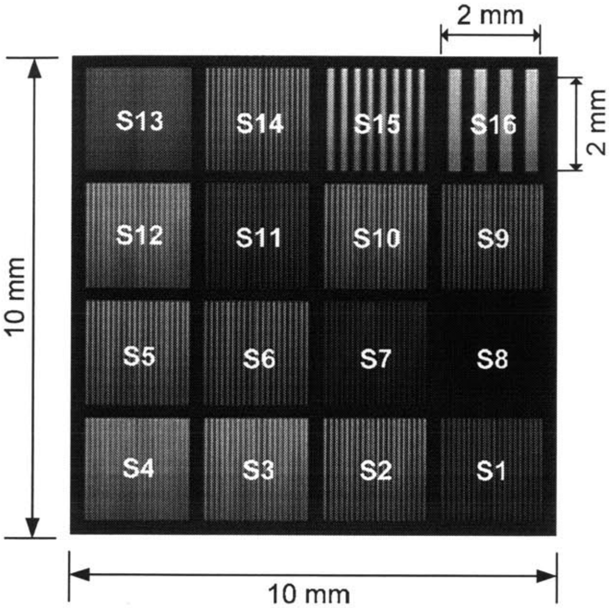


Figure 4.7 The 100 mm wafer patterned with Mask 2: (a) photograph, (b) schematic of die layout and (c) schematic of cross-section.

Table 4.9 Mechanical and physical material properties of Cu and substrates (Mask 2).

Property	Cu	Ti	SiO_2	Si
Thickness, h (nm)	1500	20	1500	0.5×10^6
Deposition Process	PVD	CVD	TEOS	–
Young' Modulus, E (GPa)	128	186	73	112*
Hardness, Brinell (MPa)	686	950	7840	24500
Density, ρ (kg/m^3)	8941	16650	2500	2330
Electrical Resistivity, ρ (μΩ·cm)	1.741	12.5	1.0×10^{20}	1.0×10^5
Reflectivity (%)	90	–	–	28

– CRC Materials Science & Engineering Handbook, CRC Press Inc.
– ASM Metals Reference Book, American Society for Metals.
– Matweb Inc., [http://www.matweb.com].



10/20 0.5	50/100 0.5	125/250 0.5	250/500 0.5
75/100 0.75	25/100 0.25	62.5/100 0.625	37.5/100 0.375
62.5/100 0.625	37.5/100 0.375	12.5/100 0.125	Field
87.5/100 0.875 May, 1998 MIT	75/100 0.75	50/100 0.5	25/100 0.25

Linewidth (μm) / Pitch (μm)
Area fraction

(a) (b)

Figure 4.8 Schematics of Mask 2: (a) die layout and (b) feature geometries in each die.

Table 4.10 Pattern layout of Mask 2.

Subdie No.	Cu linewidth, w (μm)	Pitch, λ (μm)	Area Fraction $A_f = w / \lambda$
1	25	100	0.25
2	50	100	0.5
3	75	100	0.75
4	87.5	100	0.875
5	62.5	100	0.625
6	37.5	100	0.375
7	12.5	100	0.125
8	Field	–	–
9	37.5	100	0.375
10	62.5	100	0.625
11	25	100	0.25
12	75	100	0.75
13	10	20	0.5
14	50	100	0.5
15	125	250	0.5
16	250	500	0.5

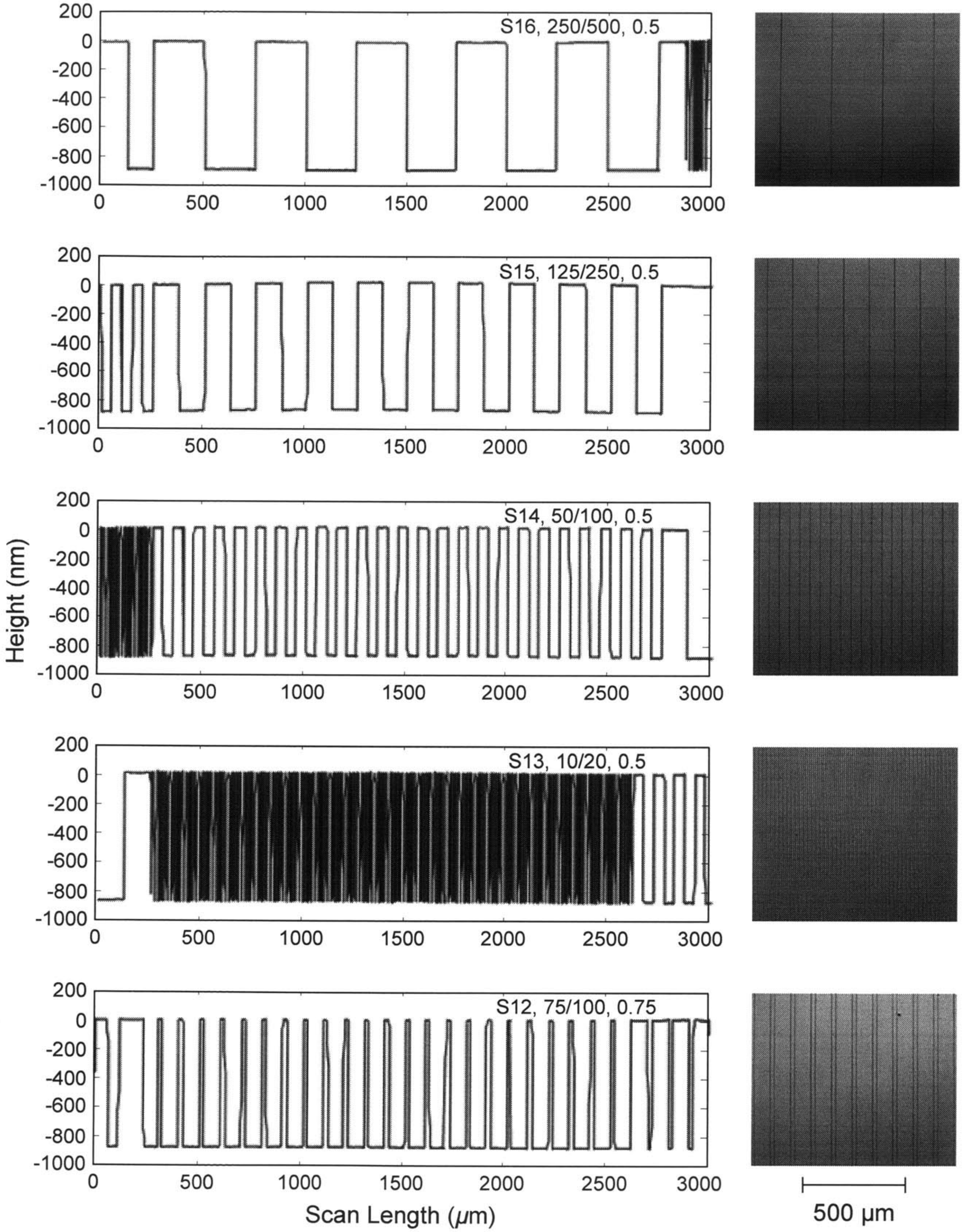


Figure 4.9 Surface profiles of subdies (Mask 2): Subdie No., Linewidth/Pitch, Area Fraction.

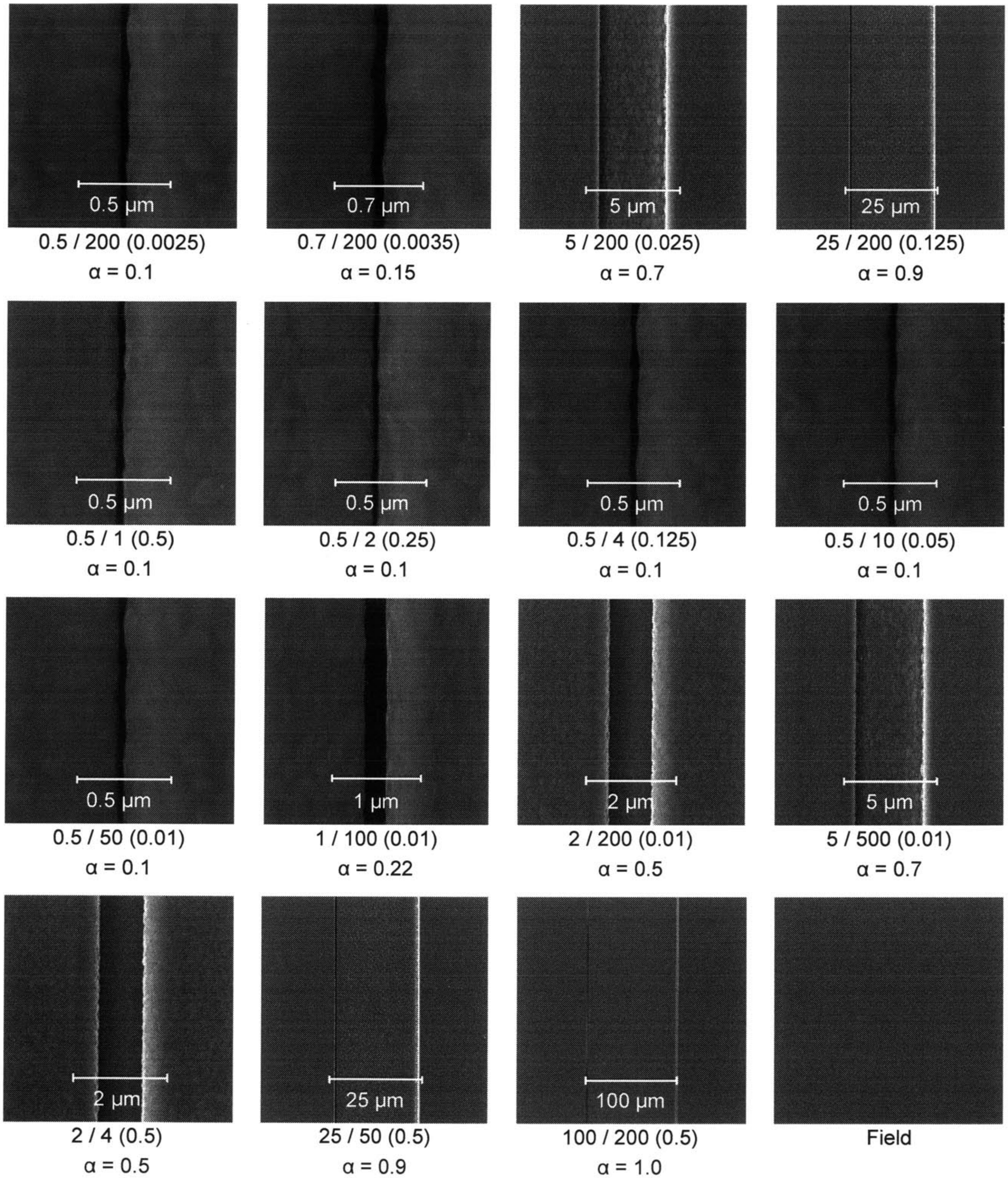


Figure 4.10 Measurement of Cu deposition factor, α, by SEM micrographs of each subdie (Mask 1): Cu linewidth, w / Pitch, λ (Area fraction, w / λ)

Table 4.11 Cu deposition factor, α, and initial step-height, h_{si}, of each subdie in the patterned Cu wafers.

Subdie No.	Mask 1					Mask 2				
	w (μm)	λ (μm)	w/λ	α	h_{si} (μm)	w (μm)	λ (μm)	w/λ	α	h_{si} (μm)
1	Field	–	–	–	–	25	100	0.25	0.97	0.87
2	100	200	0.5	1.0	0.95	50	100	0.5	1.0	0.90
3	25	50	0.5	0.9	0.95	75	100	0.75	1.0	0.90
4	2	4	0.5	0.5	0.12	87.5	100	0.875	1.0	0.90
5	0.5	50	0.01	0.1	0.10	62.5	100	0.625	1.0	0.90
6	1	100	0.01	0.22	0.10	37.5	100	0.375	1.0	0.90
7	2	200	0.01	0.5	0.12	12.5	100	0.125	0.95	0.87
8	5	500	0.01	0.7	0.75	Field	–	–	–	–
9	0.5	10	0.05	0.1	0.10	37.5	100	0.375	1.0	0.90
10	0.5	4	0.125	0.1	0.10	62.5	100	0.625	1.0	0.90
11	0.5	2	0.25	0.1	0.10	25	100	0.25	0.97	0.87
12	0.5	1	0.5	0.1	0.10	75	100	0.75	1.0	0.90
13	0.5	200	0.0025	0.1	0.10	10	20	0.5	0.9	0.85
14	0.7	200	0.0035	0.15	0.10	50	100	0.5	1.0	0.90
15	5	200	0.025	0.7	0.75	125	250	0.5	1.0	0.90
16	25	200	0.125	0.9	0.80	250	500	0.5	1.0	0.90

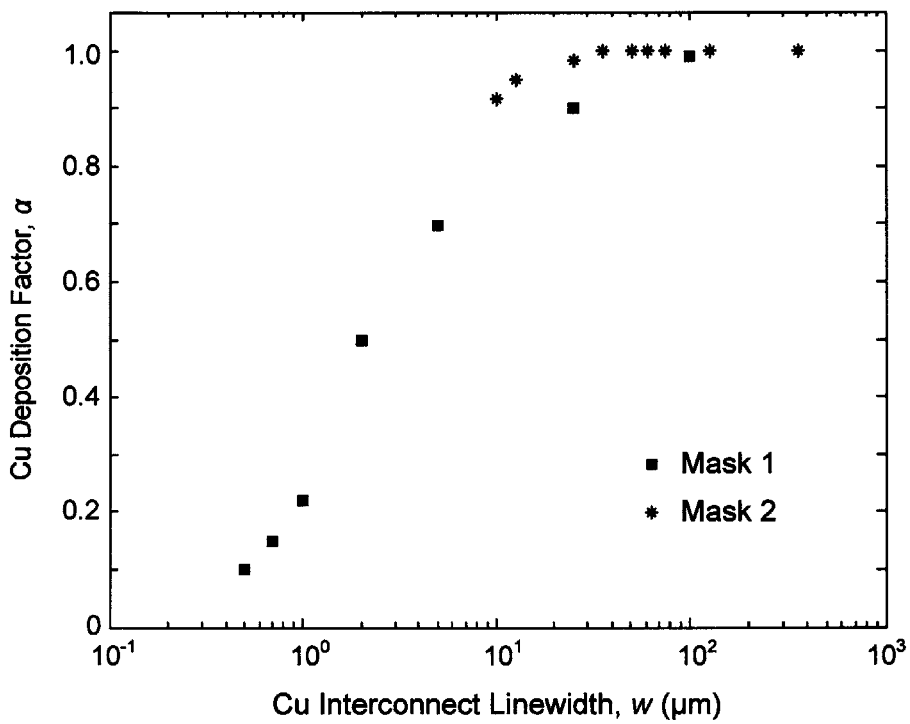


Figure 4.11 Plot of Cu deposition factor, α, of wafers patterned with Masks 1 & 2 versus Cu interconnect linewidth, w.

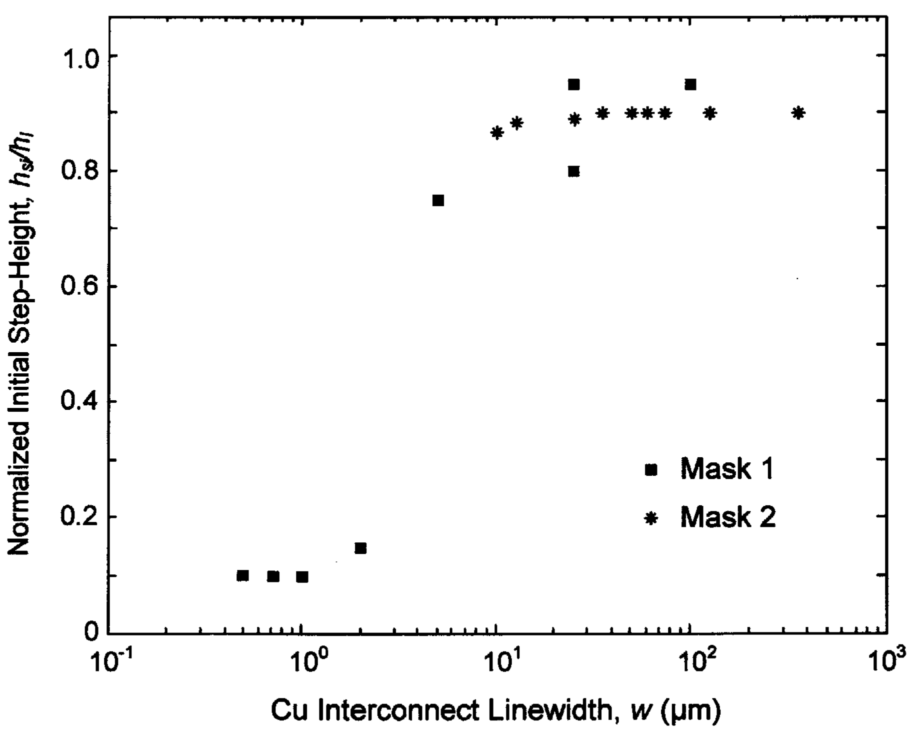


Figure 4.12 Plot for initial step-height, h_{si}, of wafers patterned with Masks 1 & 2 versus Cu interconnect linewidth, w.

$$\beta = \frac{\Delta h_{sf}}{\Delta h_{ff}} = \frac{t_{2ff}}{t_{2sf}} \tag{4.1}$$

where t_{2sf} and t_{2ff} is the time to remove the excess Cu at the slowest and the fastest fields, respectively.

Blanket and patterned wafers were polished under conditions listed in Table 4.3. Figs 4.13 and 4.14 show the evolution of the wafer-scale non-uniformity during polishing. In blanket wafer polishing experiments, the edge of the wafer is polished faster than the center of the wafer and thus the Cu/SiO_2 boundary propagates from the edge to the center as polishing progresses. In patterned wafer polishing, the wafer-scale non-uniformity needs to be combined with the die-scale non-uniformity. However, by determining the polishing time of the field regions, the boundary between Cu and SiO_2 can be defined in the same way as the blanket wafer. In patterned wafer with Mask 2, the region in the edge of the wafer with width 10 mm does not have any die. Therefore, the material removal rate of this area is lower than that of the other regions.

Table 4.12 shows wafer-scale uniformity. The wafer-scale uniformity factor of patterned wafers is slightly higher than that of the blanket wafers. This may be due to better slurry distribution in patterned wafers. In the current setup, β values are in the range 0.67 – 0.83. The wafer-scale uniformity factor, β , depends not only on process parameters but also on consumables, patterns on the wafer, and especially the polishing tool. Therefore, it is important to determine β under exactly the same set of experimental conditions before characterizing step-height, dielectric erosion and Cu dishing.

The amount of overpolishing of the slowest field, Δh_o, can be calculated from the wafer-scale unformity factor and the oxide material removal rate as:

$$\Delta h_o = \beta \cdot k_{p_{ox}} p_{av} v_R \cdot \left(t_{ep} - t_{2sf} \right) \tag{4.2}$$

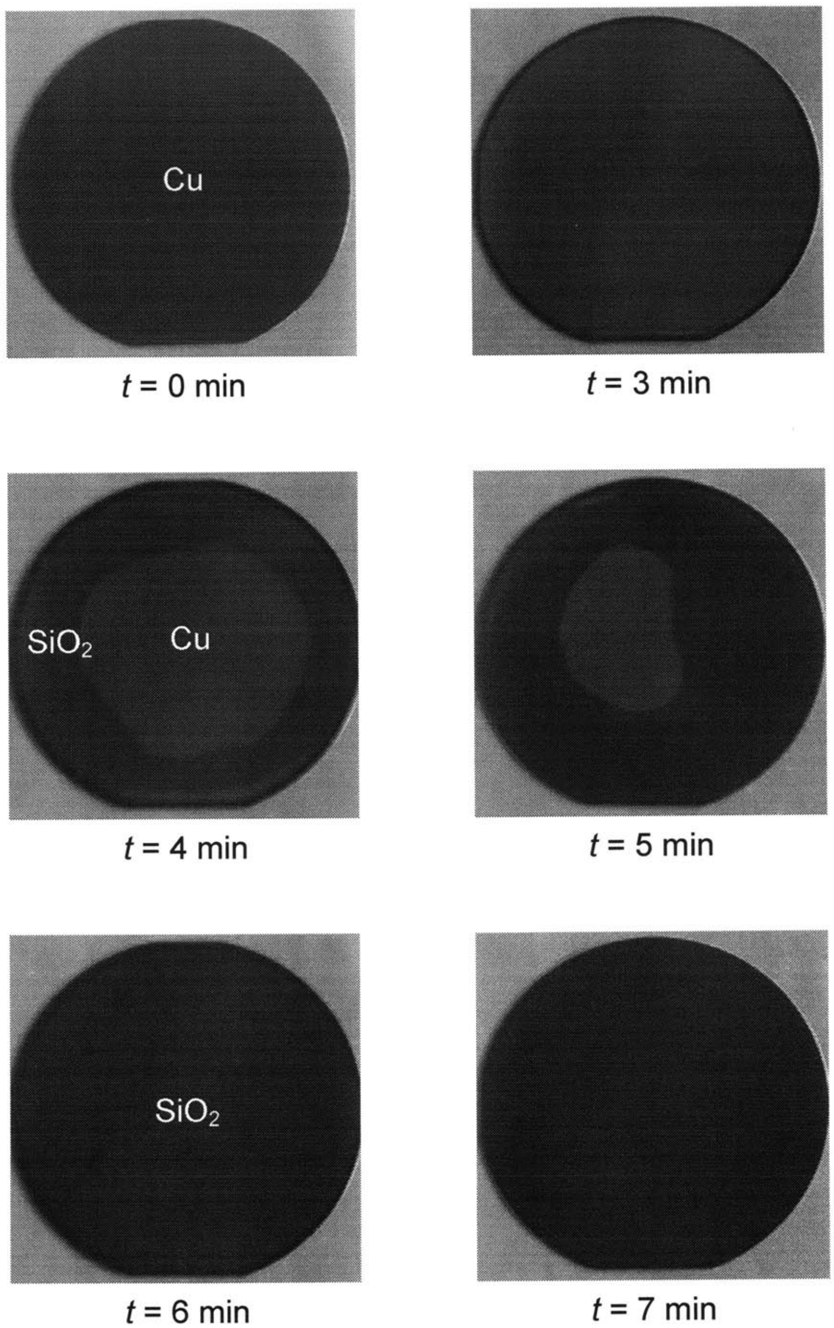


Figure 4.13 Wafer-scale non-uniformity in blanket wafer polishing.

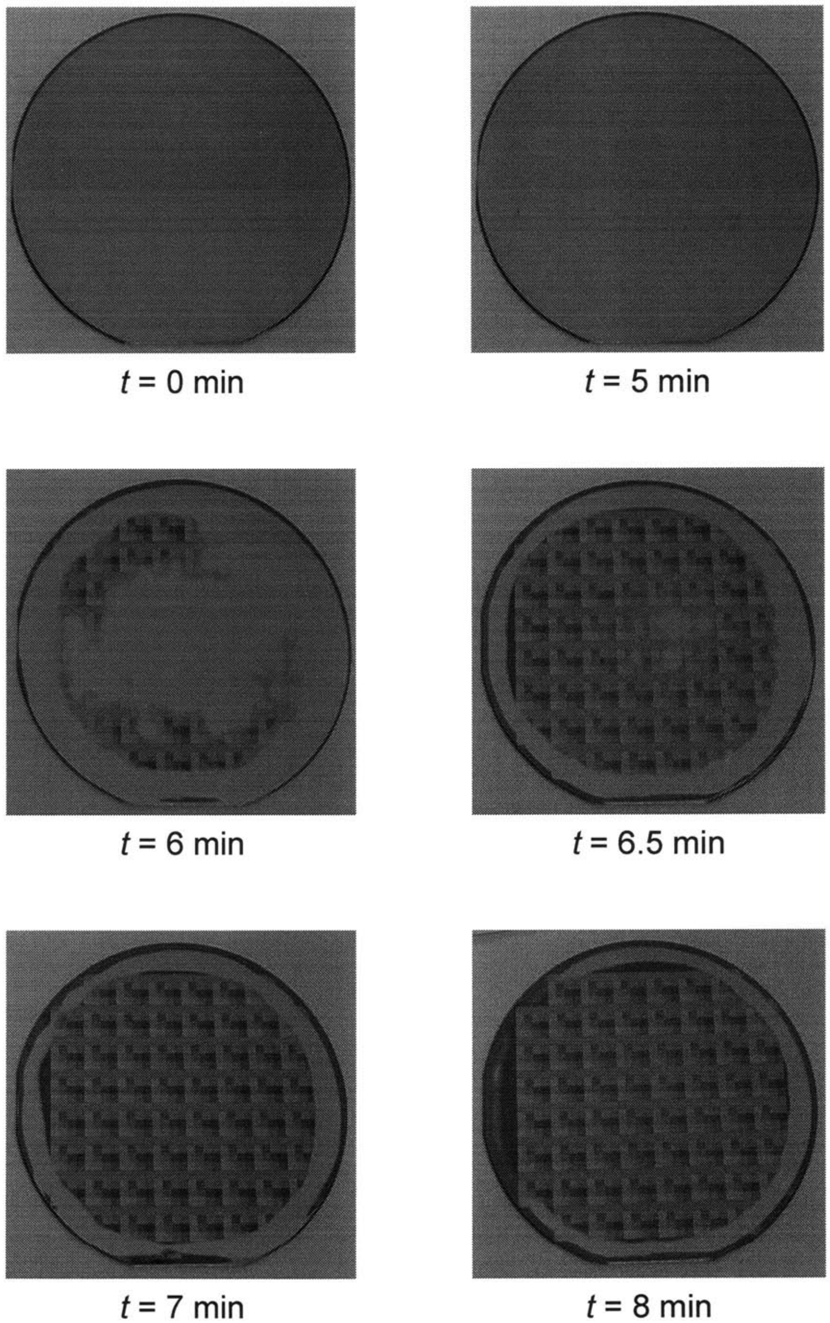


Figure 4.14 Wafer-scale non-uniformity in patterned wafer polishing (Mask 2).

Table 4.12 Wafer-scale uniformity factor, β, and the amount of overpolishing, Δh_o.

Experiment	t_{2ff} (s)	t_{2sf} (s)	t_{ep} (s)	Δt_o (s)	β	MRR_{ox} (nm/min)	Δh_o (nm)
Blanket Wafer Polishing	240	360	420	60	0.67	18	12
Patterned Wafer Polishing (Mask 1)	330	420	480	60	0.80	15	14
Patterned Wafer Polishing (Mask 2)	360	435	480	45	0.83	15	11

4.3 Step-height

Wafers patterned by Mask 2, Fig. 4.7, were used to investigate the evolution of step-height due to its large linewidth. The mechanical and physical properties of patterned wafer substrates are listed in Table 4.9. Each die has 16 subdies with various pattern geometries as shown in Fig. 4.8 and listed in Table 4.10. Most of the lines are in the range of global wiring level, and the area fraction is between 0.125 and 0.875. The polishing conditions for patterned wafers are listed in Table 4.3. The average pressure was 14kPa (2psi). The slurry was fed at the circumference of the wafer at a flow rate of 50 ml/min. Eight wafers were polished to observe the time evolution of the step-height. Each wafer was polished from 1 min to 8 min and the polished wafer surface topography was measured by Tencor P10 surface profilometer. Table 4.13 lists the material removal rate, Preston constant, and the observations by naked eye. The oxide in the field region in the fastest die was exposed after 6 min, and that in the slowest die shortly after 7 min of polishing time. Polishing was continued up to 8 min to ensure that the excess Cu on the entire wafer is removed.

Tables 4.14 and 4.15 show the step-heights in the slowest and the fastest dies, respectively. Figs. 4.15 and 4.16 show the profilometer data and light micrographs as polishing progresses. The initial step-height decreases and becomes less than 50 nm at 6 min, and it is very hard to catch the interconnect region by profilometer. However, as polishing progresses further, the Cu region recesses more than the oxide region, i.e., Cu dishing, profilometer measurement becomes viable again.

Figure 4.17 shows the evolution of step-height in Cu CMP for various area fractions when the pitch is fixed at 100 μm. The data in the slowest die in Fig. 4.17(a) show almost the same behavior with the fasted die in Fig. 4.17(b) by stretching the time scale by a factor of $1/\beta$. This is reasonable since the wafer-scale uniformity factor, β, is defined based on the field regions and thus represents the material removal rate ratio between the slowest and the fastests. As the area fraction increases, the step-height decreases faster. The end of Stage 2, t_2, however, shows the almost the same value, which represents that the die-level variation is relatively small. This can be explained as follows. Although the Cu deposition factor in the Mask 2 patterned wafer is almost uniform, $\alpha = 1$, the effect of pattern geometry is less because subdies with large and the small area fractions are located alternately. This is especially observed at the subdie with 0.625

Table 4.13 Experimental results of patterned wafer (Mask 2) polishing.

Δt (min)	m_i (mg)	m_f (mg)	Δm (mg)	Δh (nm)	MRR (nm/min)	k_p ($\times 10^{-13}$ Pa^{-1})	Coating Exposed
1	9761.24	9747.95	13.29	190	189	2.50	Cu
2	9616.68	9589.83	26.85	384	191	2.53	Cu
3	9747.85	9706.54	41.31	590	196	2.59	Cu
4	9598.12	9544.08	54.04	772	192	2.54	Cu
5	9624.84	9559.75	65.09	930	185	2.45	Cu
6	9591.85	9519.44	72.41	1,034	172	2.27	Patterns partially exposed.
7	9609.95	9530.94	79.01	1,129	161	2.13	Patterns partially exposed.
8	9595.33	9508.45	86.88	1,241	155	2.05	Patterns fully exposed.

Table 4.14 Step-height in the various subdies of the slowest die during patterned wafer polishing (Mask 2).

Subdie No.	w (μm)	λ (μm)	w/λ	Polishing Time (min) 0	1	2	3	4	5	6	7	8
4	87.5	100	0.875	900	425	175	150	65	35	25	75	130
7	12.5	100	0.125	870	750	600	325	250	135	5	10	75
9	37.5	100	0.375	900	650	450	250	160	70	10	30	75
10	62.5	100	0.625	900	650	450	225	130	30	5	30	100
11	25	100	0.25	870	700	475	300	150	80	5	30	85
12	75	100	0.75	900	500	225	175	75	20	10	45	135
13	10	20	0.5	850	550	290	200	90	10	5	25	85
14	50	100	0.5	900	600	325	225	125	30	5	50	110
15	125	250	0.5	900	625	380	250	140	30	5	50	125
16	250	500	0.5	900	600	400	270	175	30	15	50	125

Table 4.15 Step-height in the various subdies of the fastest die during patterned wafer polishing (Mask 2).

Subdie No.	w (μm)	λ (μm)	w/λ	Polishing Time (min) 0	1	2	3	4	5	6	7	8
4	87.5	100	0.875	900	340	140	96	52	20	75	130	150
7	12.5	100	0.125	870	600	440	260	140	30	10	75	100
9	37.5	100	0.375	900	536	360	196	128	35	30	75	130
10	62.5	100	0.625	900	496	360	180	120	20	30	100	125
11	25	100	0.25	870	560	368	240	120	20	30	85	120
12	75	100	0.75	900	392	180	140	60	10	50	135	150
13	10	20	0.5	850	440	220	180	60	10	25	85	100
14	50	100	0.5	900	480	260	180	96	20	45	110	130
15	125	250	0.5	900	500	304	200	112	20	50	125	150
16	250	500	0.5	900	480	320	200	140	20	50	125	150

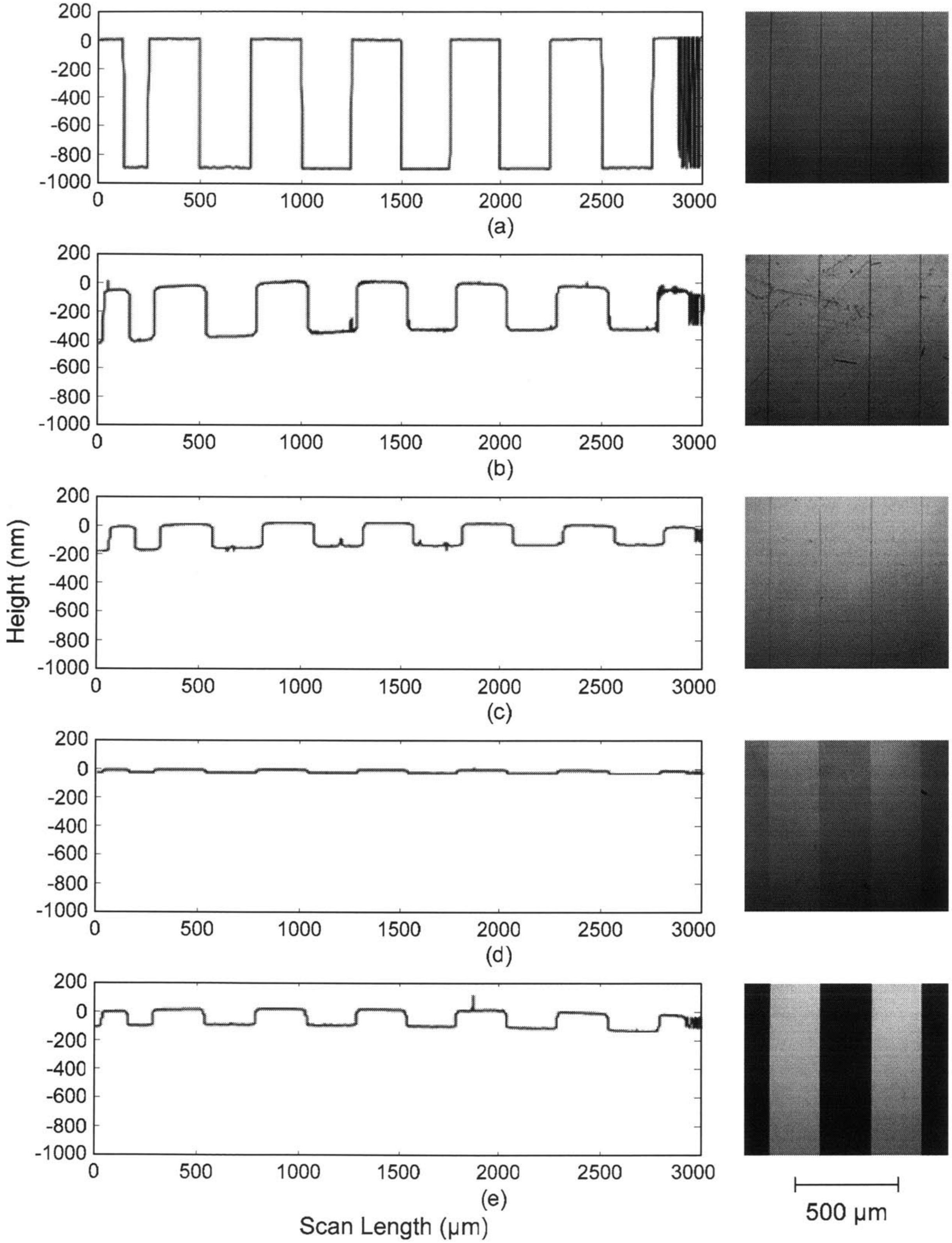


Figure 4.15 Evolution of the surface profile in Cu CMP for various times: (a) t = 0 min, (b) 2 min, (c) 4 min, (d) 6 min and (e) 8min. $w = 250$ μm and $\lambda = 500$ μm.

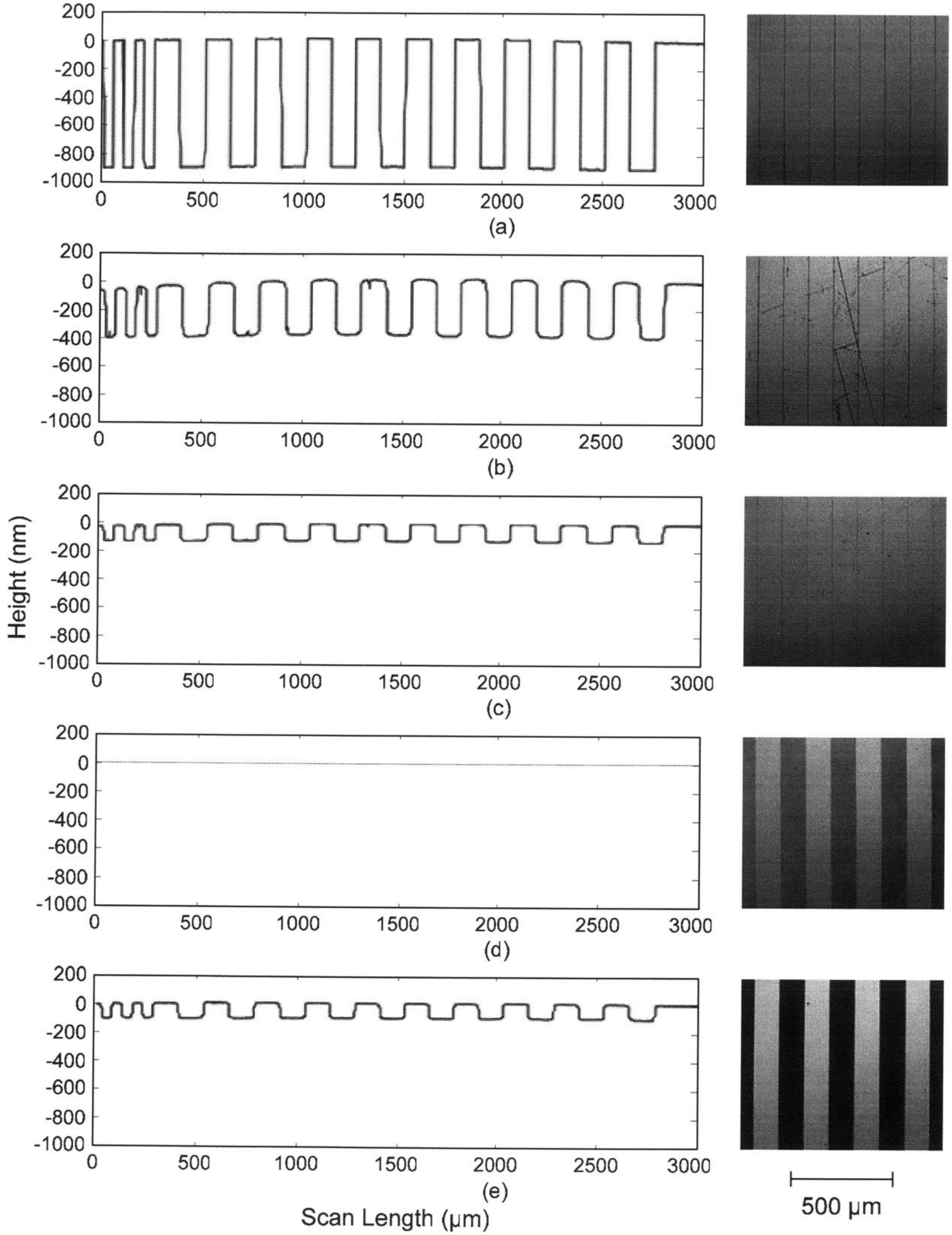
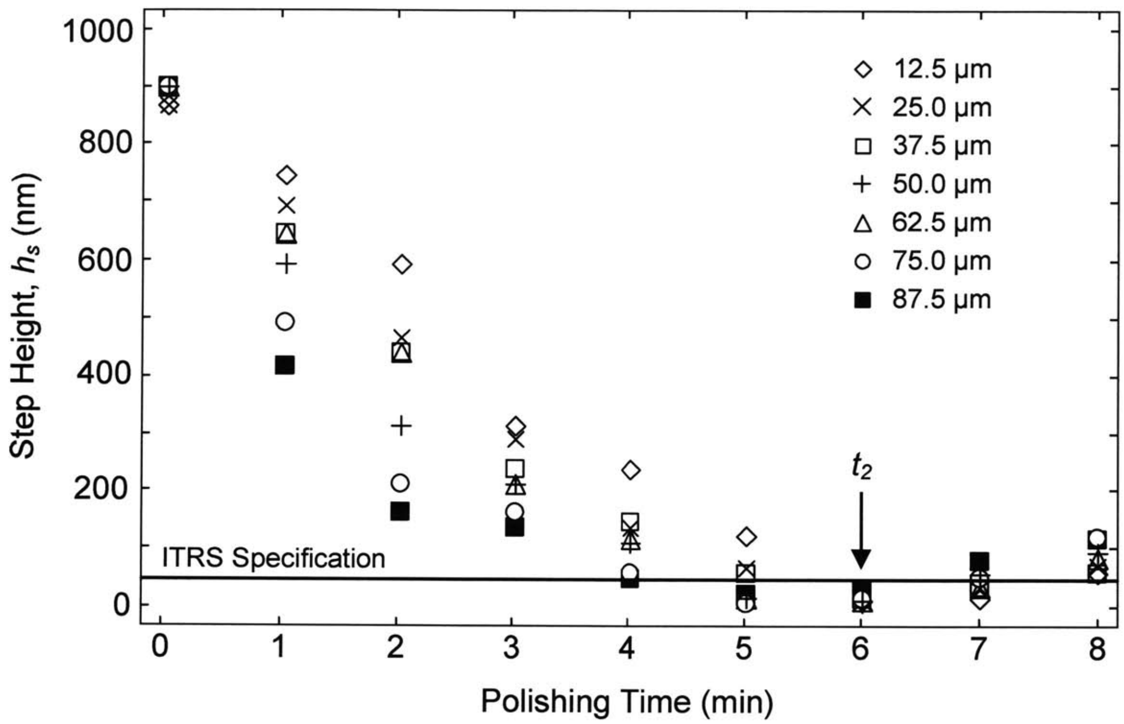
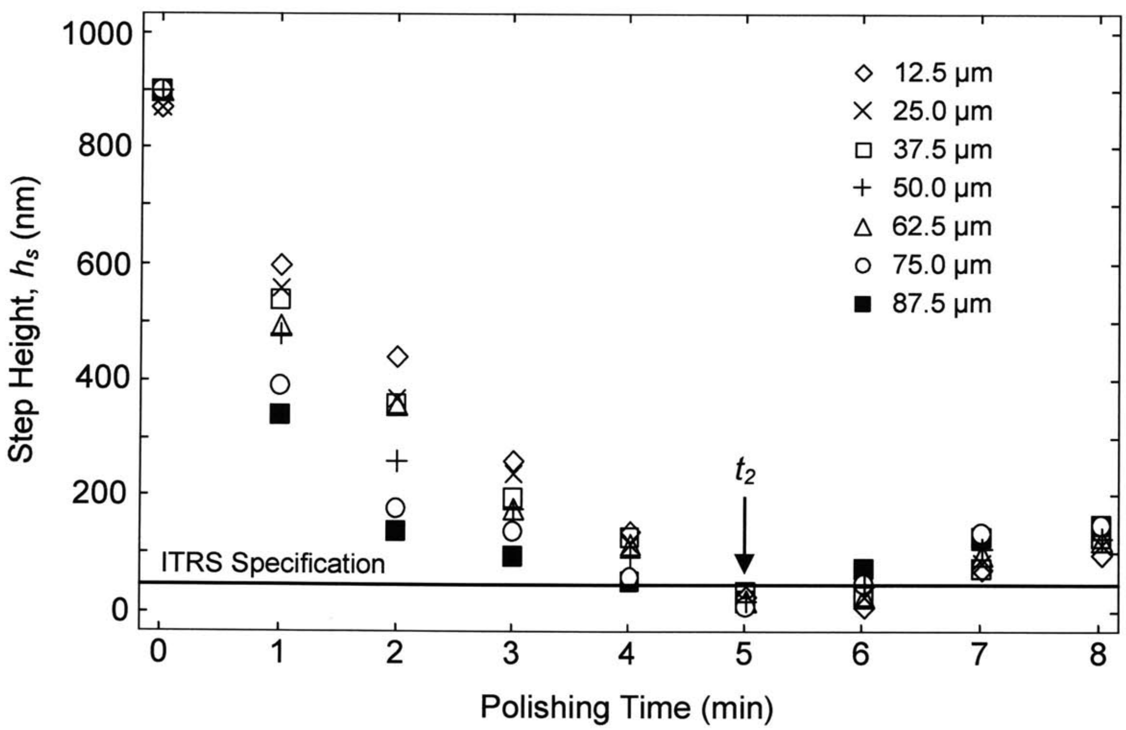


Figure 4.16 Evolution of the surface profile in Cu CMP for various times: (a) t = 0 min, (b) 2 min, (c) 4 min, (d) 6 min and (e) 8min. w = 125 μm and λ = 250 μm.



(a)



(b)

Figure 4.17 Evolution of the step-height in Cu CMP for various area fractions and pitch $\lambda = 100$ μm: (a) slowest die and (b) fastest die.

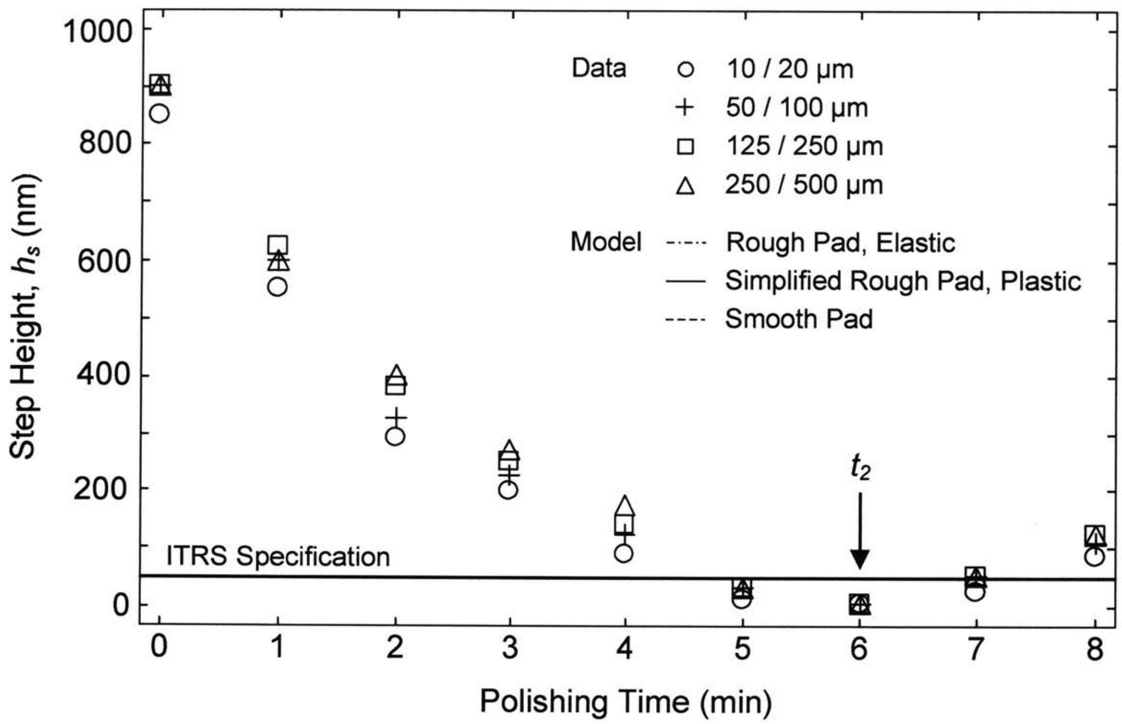
area fraction. Since this subdie is located between subdies with 0.325 area fraction and the field region, the actual pattern geometry effect is diminished. This effect is closely related to the so-called "planarization length", which will not be addressed in this study.

Figure 4.18 shows the effect of linewidth. Since the range of Cu linewidth is in the global wiring level, the effect of linewidth is relatively small. This explains that the smooth pad model cannot be used in the global wiring level because the wafer is in contact with the pad asperities, not the complete nominal area. Therefore, as far as the Cu linewidth is greater than the asperity contact size, the effect of linewidth on the step-height can be ignored.

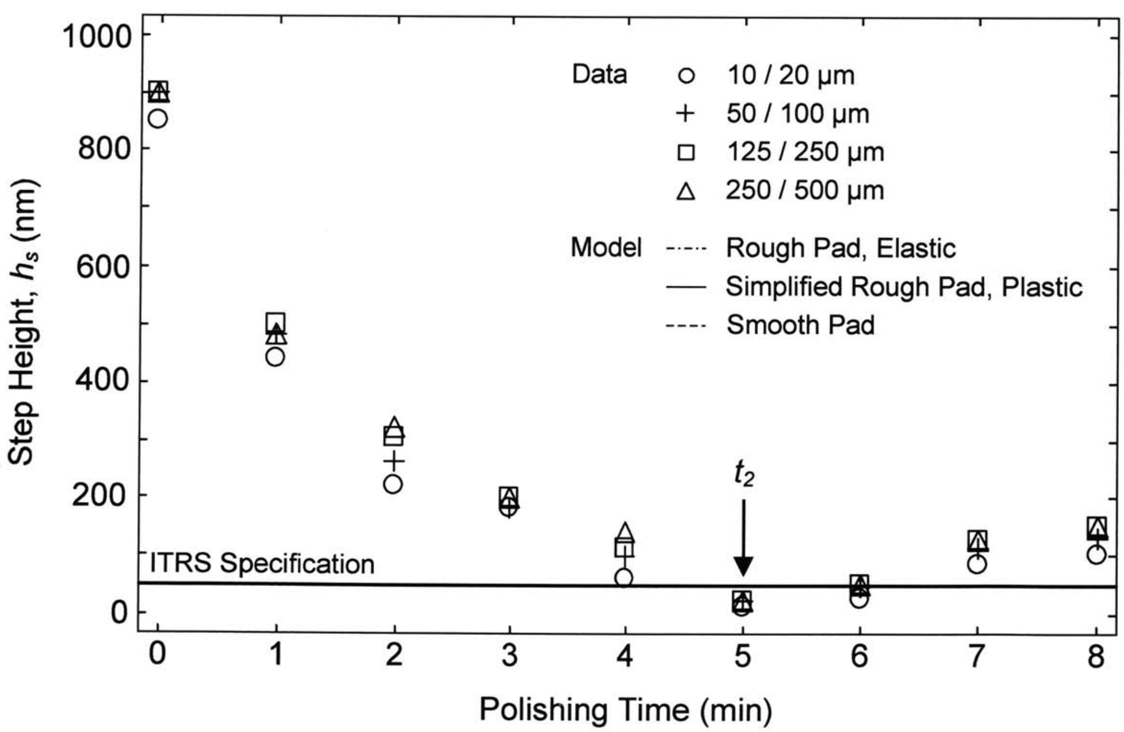
Finally, the developed models are compared with the experimental data as shown in Fig. 4.19. Again, the difference between the models in the slowest die and the fasted die is the wafer-scale uniformity factor, β, and thus it can just stretch at the time scale. The duration of Stage 1 is short, which explains why the smooth pad model is not appropriate to explain the step-height in the global wiring level. However, the smooth pad model is useful in the submicron region since the asperity contact diameter is greater than the interconnect linewidth. Among developed models, the general random rough pad model gives the best fit in the step-height reduction regime, but it overestimates dishing in the overpolishing regime. The simplified rouph pad model and the discrete pad explain the evolution of step-height best in the overpolishing regime.

Although the general random rough pad model may be the best way to explain the feature-scale contact between the pad and the wafer, the simplified rough pad model may be the best choice for developing an integrated non-uniformity model and to avoid the tedious iterative statistical calculations. Since our focus is on the step-height behavior during overpolishing, the simplified rough pad model with plastic contact is selected to describe dishing and erosion.

The discrete pad model assumes that the real contact area is the same as the nominal area, which in general is not true. Additionally, the discontinuous boundary condition and the fact that the local stress is affected by the initial pad thickness make the discrete pad model difficult to interpret the actual contact behavior. Therefore, the discrete pad model will not be considered further to describe erosion and dishing.

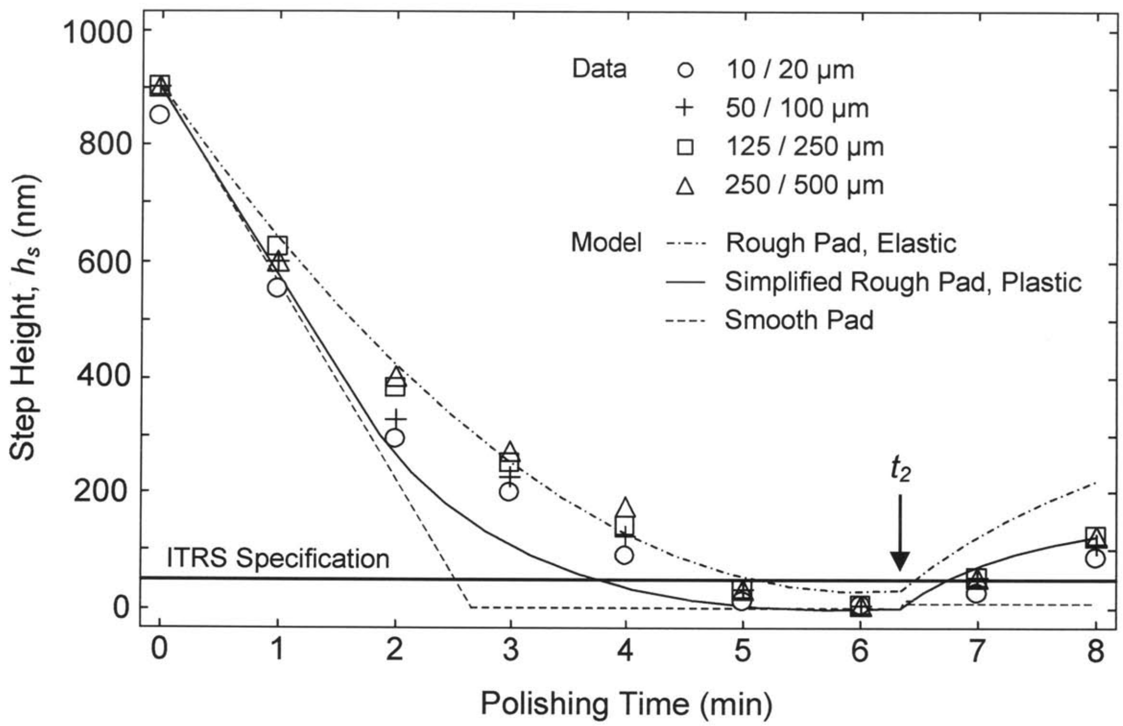


(a)

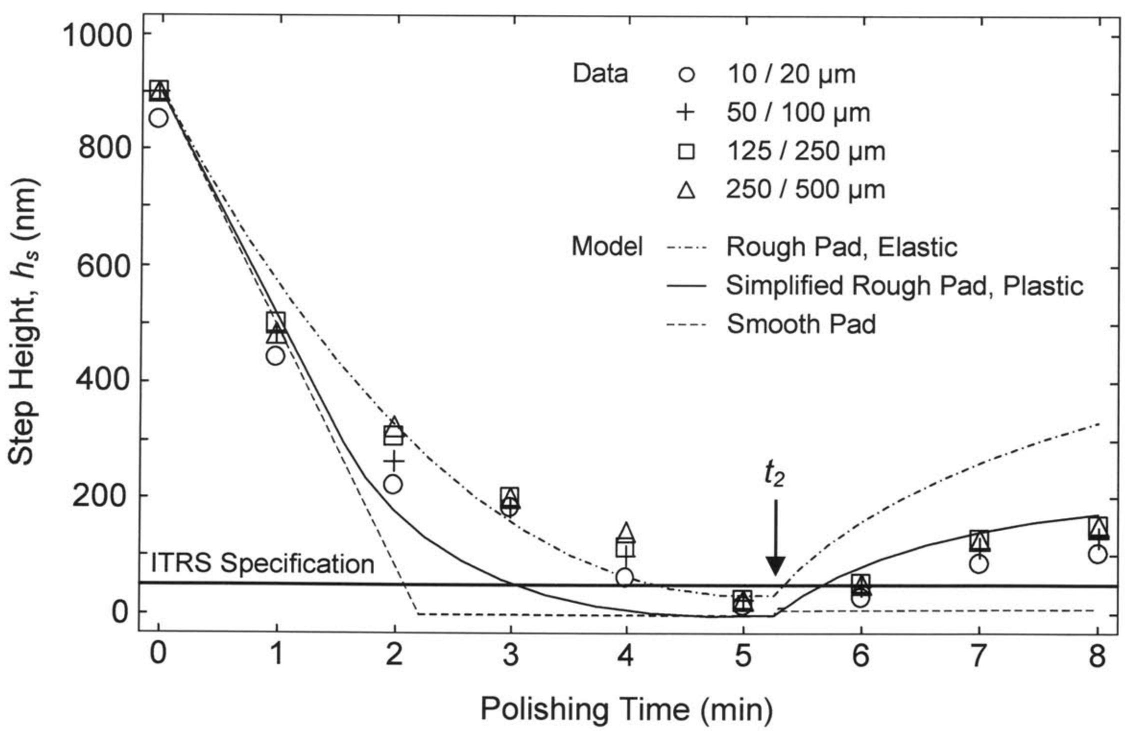


(b)

Figure 4.18 Evolution of the step-height in Cu CMP for various Cu linewidths and $w / \lambda = 0.5$: (a) slowest die and (b) fastest die.



(a)



(b)

Figure 4.19 Data and model comparison for various contact conditions. Data are for $w = 50$ μm $\lambda = 100$ μm. Models include: smooth pad, simplified rough pad, random rough pad contact conditions.

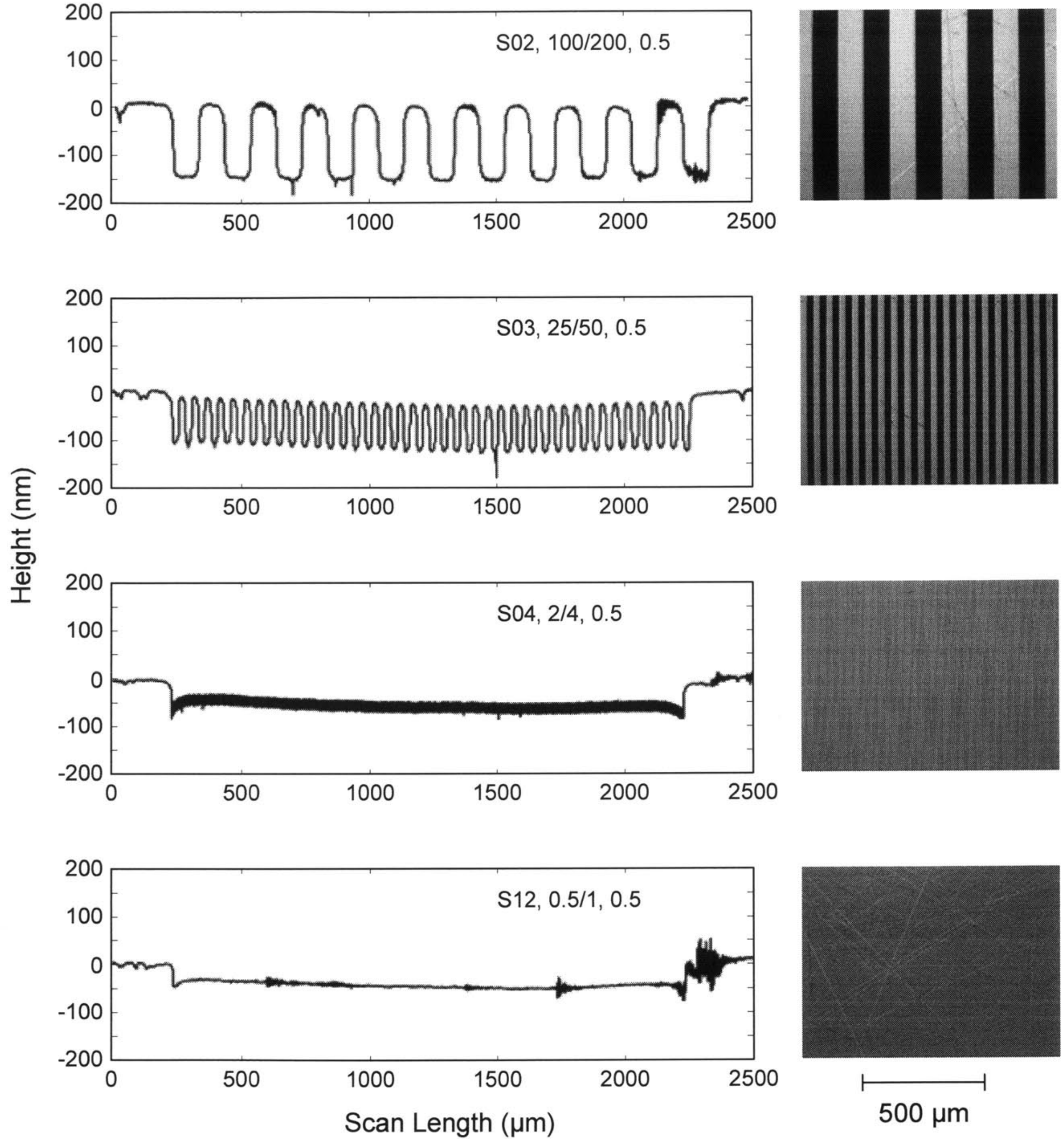
4.4 Dishing and Erosion

To verify dielectric erosion and Cu dishing in the both sub-micron and the global-wiring levels, wafers patterned by both masks were used. Each die in both masks has 16 subdies with various pattern geometries. In Mask 1, the Cu linewidth varies from 0.5 μm to 100 μm with area fractions between 0.0025 and 0.5. There is a field subdie and each subdies are surrounded by field region with 500 μm width. This sidewall will be considered as a field region and assumed to have the same thickness as the field subdie after polishing to compare the dielectric erosion. In Mask 2, Cu interconnect lines are between 10 μm and 250 μm and the area fraction is between 0.125 and 0.875. Wafers patterned by Mask 2 are used to express Cu dishing in the global wiring level. Although Mask 2 also has a field subdie, it is hard to calculate erosion at each subdie since there is not enough sidewall field area between subdies.

Figure 4.20(a) shows surface profiles for various linewidths with a fixed area fraction, $w/\lambda = 0.5$. At the subdie where Cu linewidth is 100 μm, most of non-uniformity is expressed by Cu dishing and the amount of Cu dishing is 150 nm. In the 0.5 μm linewidth subdie, by contrast, dielectric erosion is responsible for most of the non-uniformity and Cu dishing is less than 20 nm, which is about 10% of that in the 100 μm subdie. Figure 4.20(b) shows the effect of area fraction in the sub-micron subdies. As the area fraction increases, erosion increases while Cu dishing is still less than 20 nm.

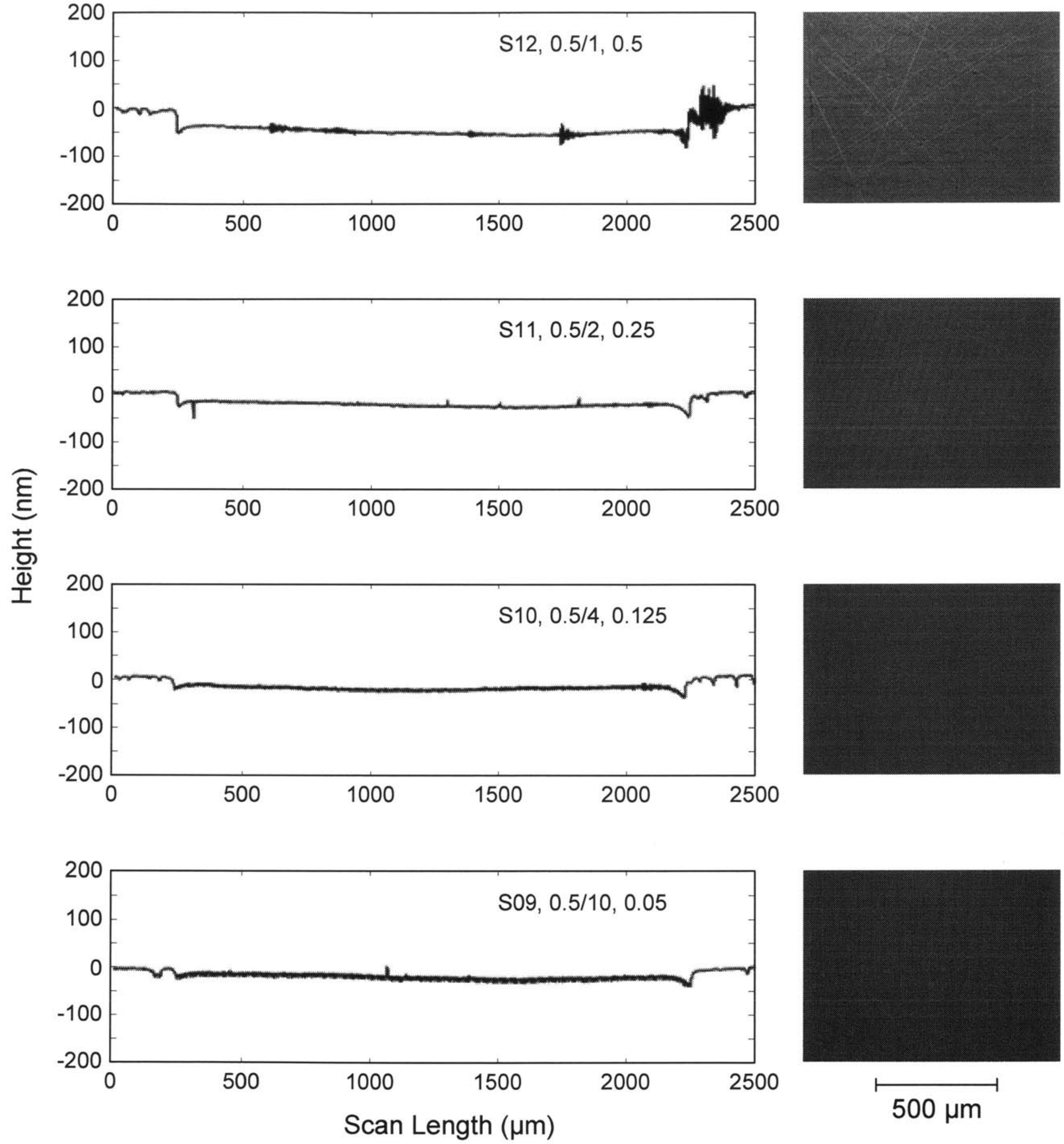
Figure 4.21(a) shows the surface profiles as the Cu linewidth varies. Cu dishing in the subdie where the linewidth is 50, 125, 250 μm is 135 – 150 nm and does not vary significantly with Cu linewidth. In the 10 μm linewidth region, Cu dishing is less about 100 nm and erosion starts to increase. The effects of area fraction and linewidth on Cu dishing are shown in Fig. 4.21(b). Although, to observe the effect of area fraction, it is required to fix the linewidth and vary pitch between interconnect lines only, in Mask 2 the pitch is fixed as 100 μm and Cu linewidth varies to set area fractions in the rage of 0.125 – 0.875. In this case, Cu dishing increased from 100 nm to 150 nm as area fraction increases from 0.125 to 0.875.

Dielectric erosion and Cu dishing data from the experiments on wafers patterned by both masks versus Cu interconnect area fraction and linewidth are plotted in Figs. 4.22 and 4.23. Both erosion and dishing are expressed in dimensionless forms, the ratio of erosion or dishing to the oxide trench depth or the nominal Cu interconnect thickness, h_I. In the submicron area, the



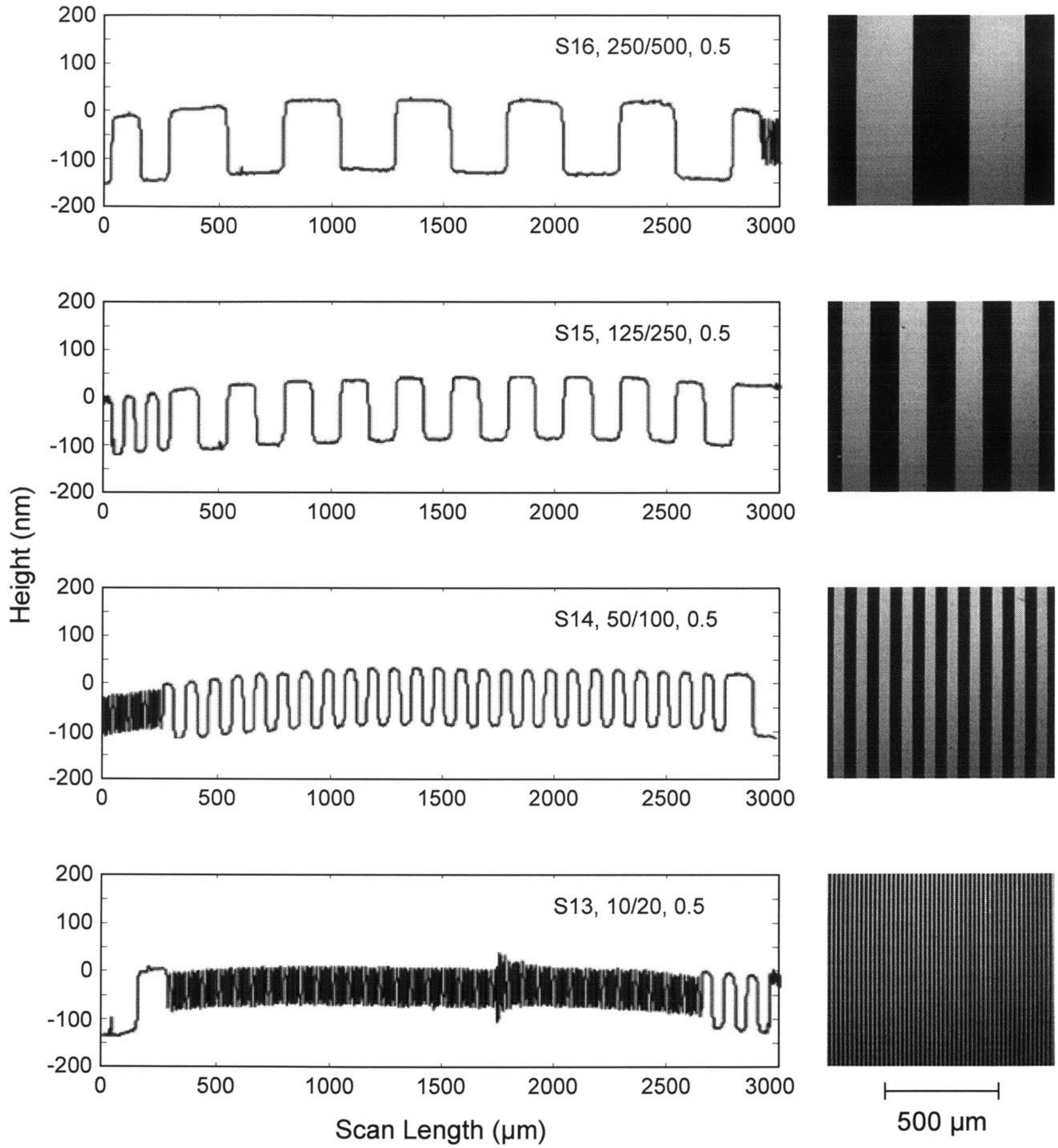
(a)

Figure 4.20 Surface profile and micrograph of Mask 1 after Cu CMP of (a) various linewidths with an area fraction of 0.5 and, (b) various area fractions with a linewidth of 0.5 μm (continued). $p = 14$ kPa, $v_R = 0.9$ m/s, $t = 8$ min.



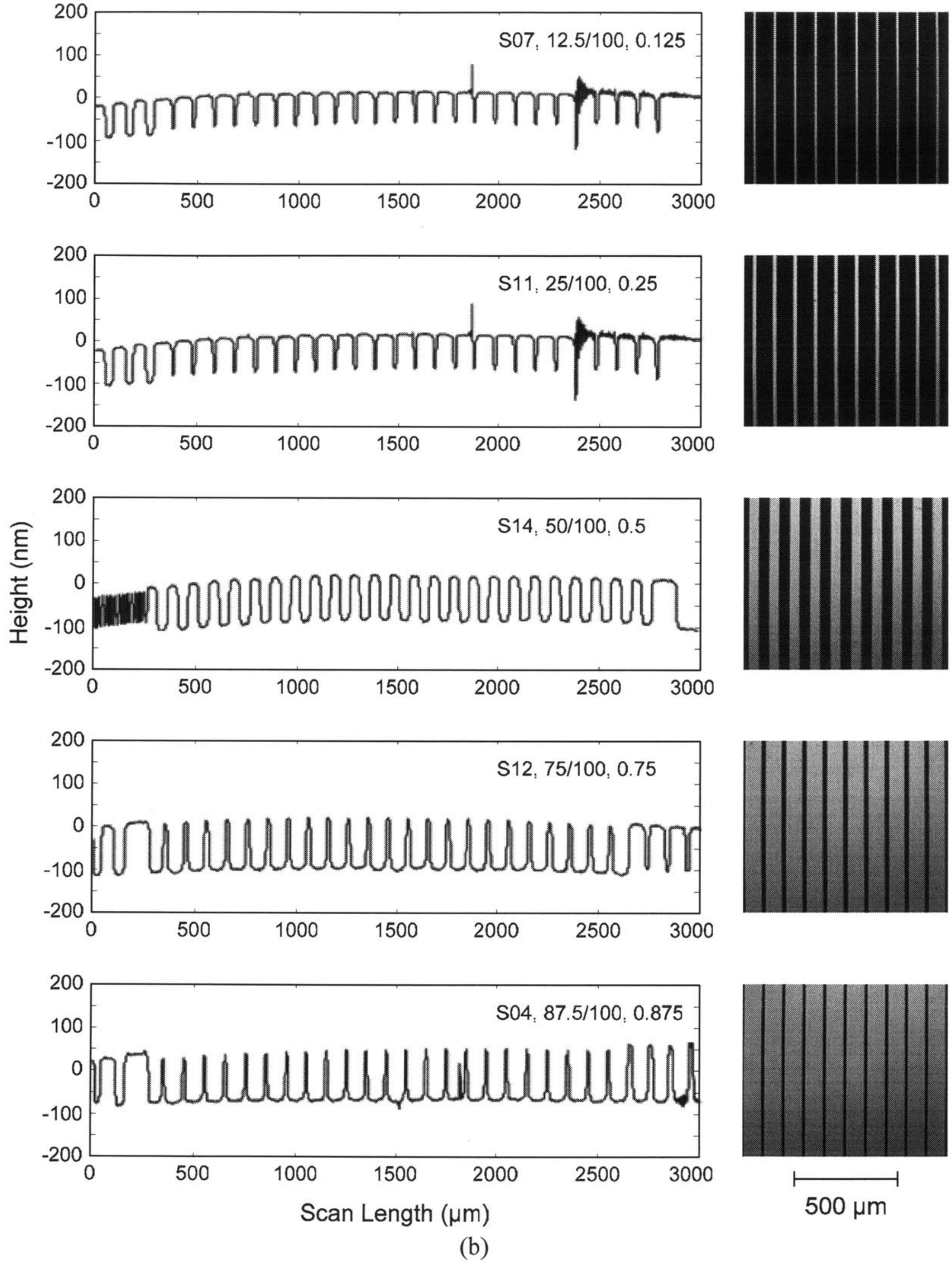
(b)

Figure 4.20 Surface profile and micrograph of Mask 1 after Cu CMP of (a) various linewidths with an area fraction of 0.5 and (b) various area fractions with a linewidth of 0.5 μm. p = 14 kPa, v_R = 0.9 m/s, t = 8 min.



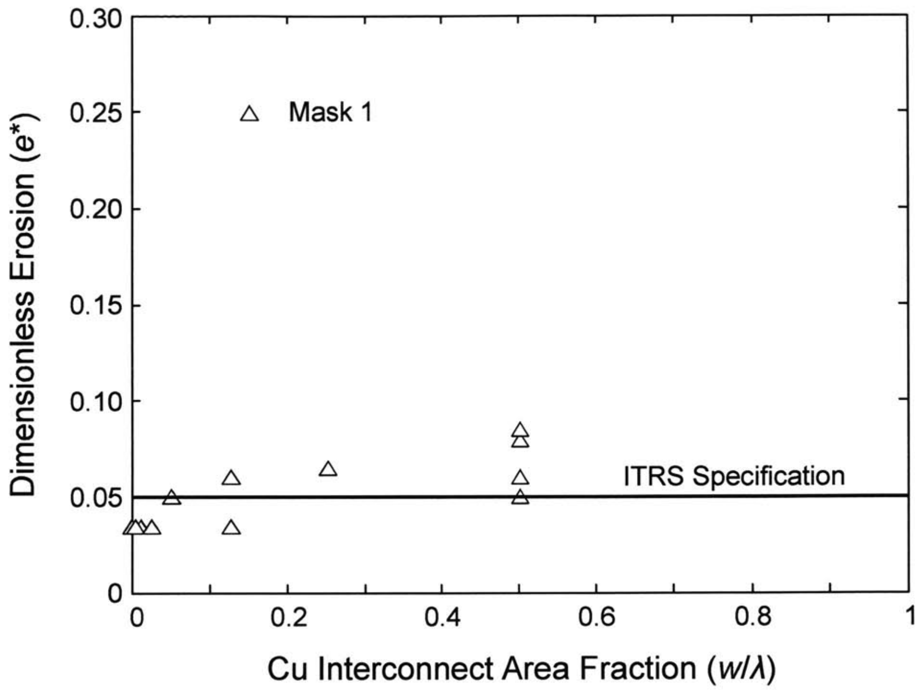
(a)

Figure 4.21 Surface profile and micrograph of Mask 2 after Cu CMP of (a) various linewidths with an area fraction of 0.5 and (b) various area fractions with a pitch of 100 μm (continued). p = 14 kPa, v_R = 0.9 m/s, t = 8 min.

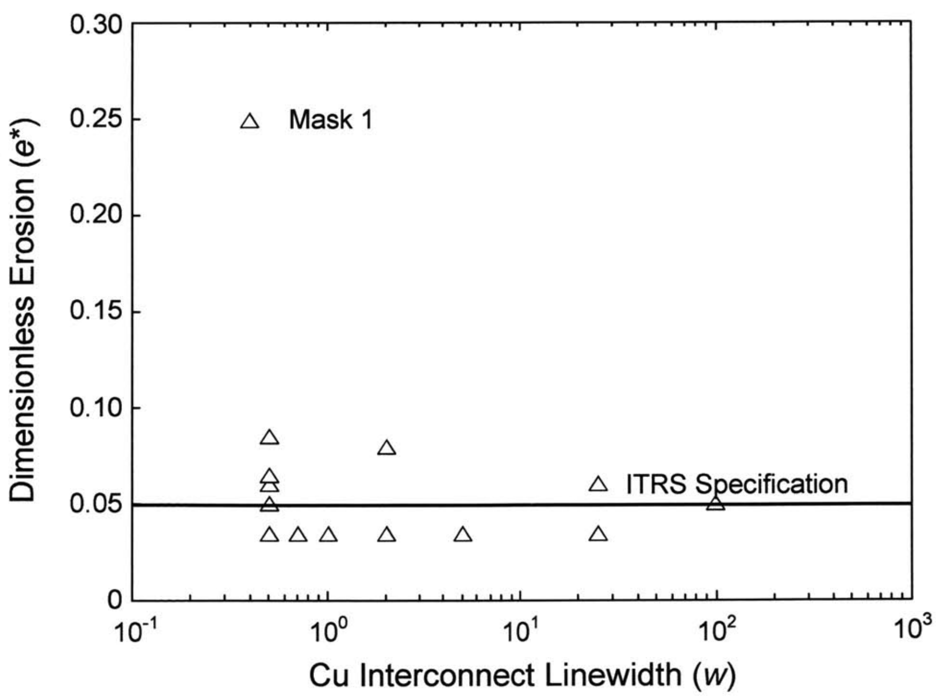


(b)

Figure 4.21 Surface profile and micrograph of Mask 2 after Cu CMP of (a) various linewidths with an area fraction of 0.5 and (b) various area fractions with a pitch of 100 μm. $p = 14$ kPa, $v_R = 0.9$ m/s, $t = 8$ min.

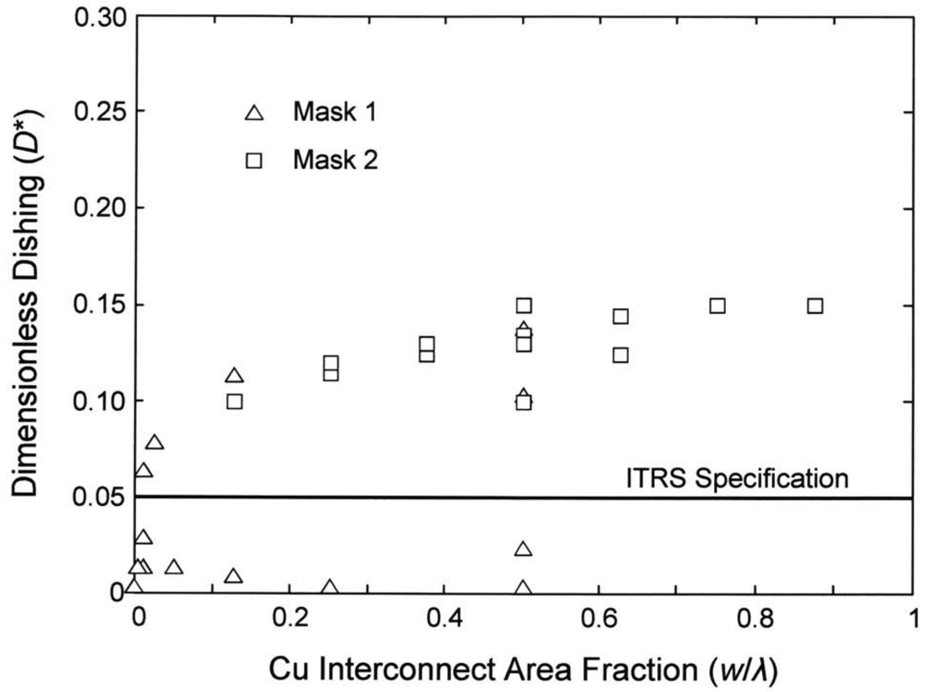


(a)

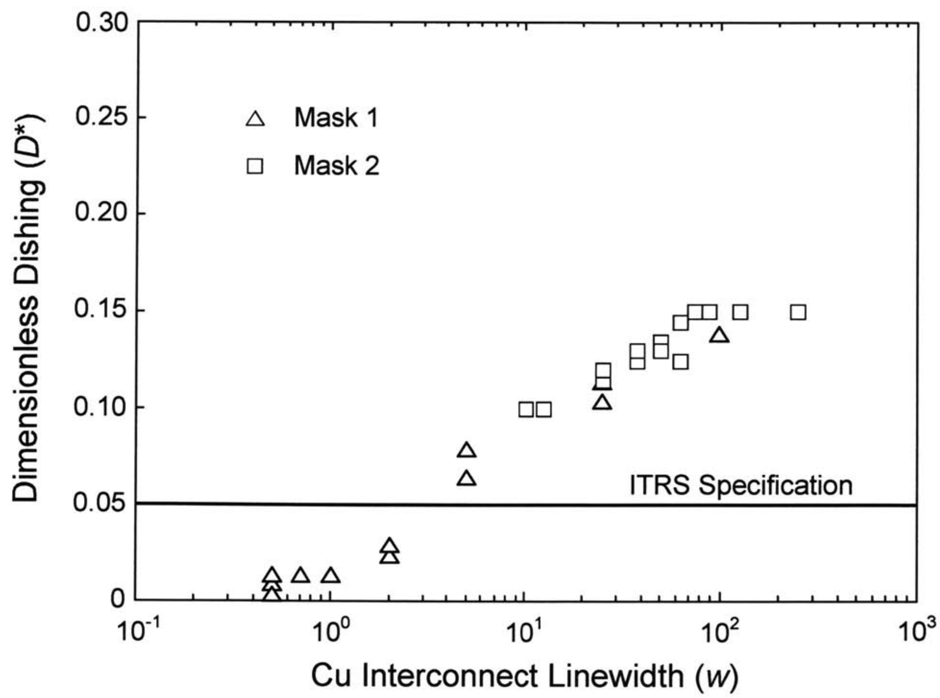


(b)

Figure 4.22 Dielectric erosion versus (a) area fraction and (b) linewidth.



(a)



(b)

Figure 4.23 Cu dishing versus (a) area fraction and (b) linewidth.

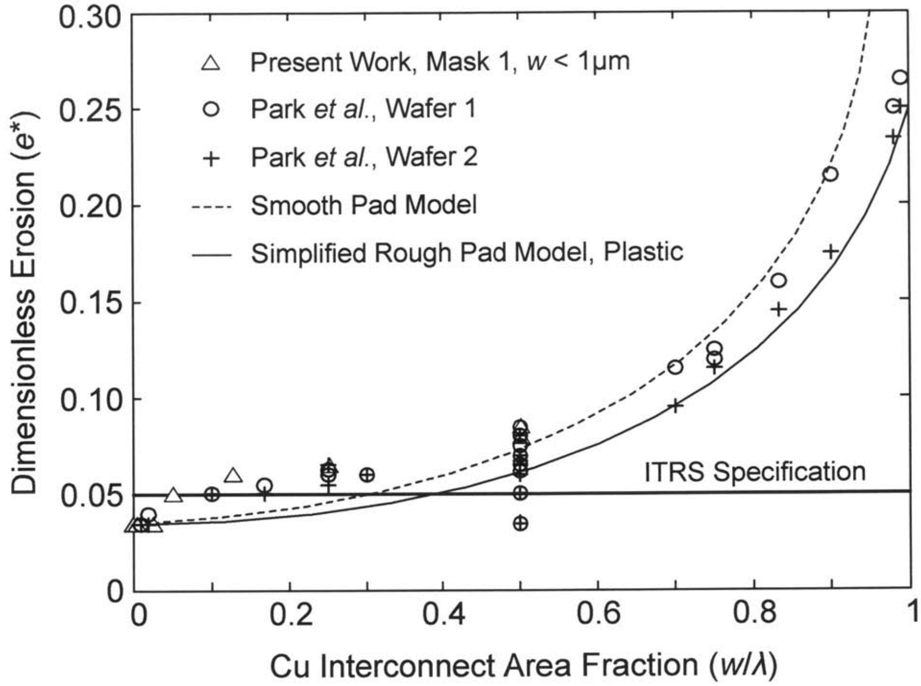
dielectric erosion is responsible for most of the non-uniformity and it mainly depends on the area fraction of Cu interconnects – erosion increases as the area fraction increases. In the global wiring level where Cu linewidth is greater than 50 μm, the die-scale erosion is relatively smaller than the wafer-scale erosion and thus the total erosion is close to that in the field region. Dielectric erosion at the global wiring level is less than that in the submicron level when the area fraction is fixed. This is explained by the discrete or rough pad model since the time evolution of erosion is affected by the amount of dishing.

Cu dishing mainly depends on Cu interconnect linewidth as shown in Fig. 4.23(b). If the Cu linewidth is submicron, the amount of dishing is well under the specification, 5% of interconnect thickness [ITRS – Interconnect, 2003]. If Cu linewidth is greater than 50 μm, Cu dishing is about 15% of the interconnect thickness and does not vary significantly with Cu linewidth.

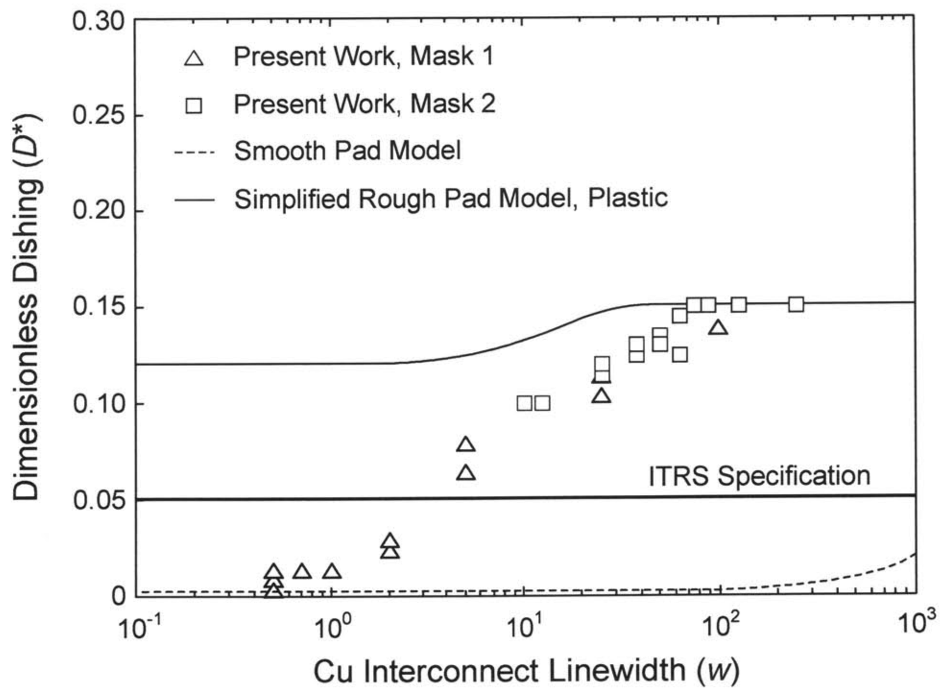
Figure 4.24 compares experimental data with two different models: smooth pad and simplified rough pad model. From the previous section, the simplified rough pad model is chosen to describe the feature-scale contact behavior at the global-wiring level. The smooth pad model is added to compare contact behavior at the submicron linewidth level.

The dielectric erosion dependency on the area fraction of Cu interconnects are well explained by both smooth and simplified rough pad models. Although erosion in the rough pad model is slightly less than erosion in the smooth pad model as shown in Fig. 4.24(a), the difference between two models is marginal unless the area fraction is close to one. Therefore, the smooth pad model is selected to characterize erosion and parameter analysis.

Cu dishing, however, is more complicated. At the submicron device level, the amount of dishing is well below 50 nm and does not vary with Cu linewidth. In this region, the smooth pad model needs to be applied since diameter of pad asperity contact is greater than linewidth. In the global wiring level, by contrast, rough pad surface should be considered. From the step-height model comparison, it was shown that the simplified rough pad model is the best fit. If $w > 100$ μm, amount of dishing does not depend on the interconnect linewidth and the simplified rough pad model reflects it. As interconnect linewidth decreases, Cu dishing slightly decreases even in the simplified rough pad model. This is due to the initial pattern variation in a die. Although the present model does not explain the linewidth dependency in the intermediate level, our main concern is on the maximum dishing at the global wiring level and the model follows the data reasonably well.



(a)



(b)

Figure 4.24 Comparison of the models with the experimental data from the present work and literature [Park et al., 2002]: (a) dielectric erosion and (b) Cu dishing.

4.5 Summary

In this chapter, experimental and analytical means of determining the model parameters, are outlined. Feature-scale step-height models are validated by 100 mm patterned wafer polishing experiments and measurement of step-heights at various polishing times. Cu dishing and dielectric erosion in each subdie of the fastest polishing die is measured, and compared with integrated erosion and dishing models. Based on the experiments, the following conclusions are drawn.

The general random rough pad model explains the step-height reduction data fairly well, and the simplified rough pad model explains best in the overpolishing regime. Since our focus is on the step-height behavior during overpolishing, the simplified rough pad model with plastic contact is the best choice to describe dishing and erosion at the global wiring level.

Dielectric erosion at the submicron device level is well explained by the smooth pad model. Cu dishing in the global wiring level was characterized by the simplified rough pad model and the model follows the data reasonably well.

Nomenclature

A_f = area fraction of Cu interconnects

D = Cu dishing (m)

D^* = dimensionless Cu dishing

e = dielectric erosion (m)

e^* = dimensionless dielectric erosion

h = film thickness (m)

h_{Cu}, h_b, h_I = film thickness of Cu, barrier layer and interconnect (m)

$h_s, h_{si}, h_s(\infty)$ = step-height, initial step-height and asymptotic step-height (m)

k_p = Preston constant (m^2/N)

MRR = material removal rate (m/s)

MRR_{Cu}, MRR_{ox} = material removal rate of Cu and SiO_2 (m/s)

$m_i, m_f, \Delta m$ = initial and final mass, mass change in polishing (kg)

p = pressure (N/m^2)

Q = slurry feed rate (m^3/s)

$S_{Cu/ox}, S_{b/ox}$ = Cu-to-oxide and barrier-to-oxide selectivities

t = polishing time (s)

t_{ep} = process end-point (s)

t_{2ff}, t_{2sf} = polishing time at the end of Stage 2 of fastest and slowest dies (s)

v_R = magnitude of relative velocity (m/s)

w = linewidth of Cu interconnect (m)

$\Delta h_{ff}, \Delta h_{sf}$ = material removed in the fastest and the slowest regions (m)

Δh_o = amount of overpolishing (m)

α = Cu deposition factor

β = wafer-scale uniformity factor

References

Americal Society of Metals, 1993, **ASM Metals Reference Book**, ASM International, Metals Park, OH.

Cabot Microelectrnic Corporation, 2005, [Online] http://www.cabotcmp.com.

International Technology Roadmap for Semiconductors (ITRS) – Interconnect, 2003, [Online] http://www.itrs.net.

Park, T., Tugbawa, T. and Boning, D., 2000, "Overview of Methods for Characterization of Pattern Dependencies in Copper CMP," *Proc. CMP-MIC Conf.*, pp. 196-205.

Matweb Inc., 2005, [Online] http://www.matweb.com.

Rohm and Haas Corporation, 2005, [Online] http://www.rohmhaas.com.

Shackelford, J. F., and Alexander, W., 2000, **CRC Materials Science & Engineering Handbook**, CRC Press Inc., Cleveland, OH.

Thomas West Inc., 2005, [Online] http://www.thomaswest.com.

CHAPTER 5

CONCLUSION

5.1 Summary

In this thesis, dielectric erosion and Cu dishing models that integrate non-uniformities at the wafer-, die- and feature-scales have been developed by computing the step-height evolution during CMP. Based on the developed models, practical solutions to reduce erosion and dishing are suggested.

Chapter 2 presents feature-scale step-height models for various pad/wafer contact conditions. First, the contact between the wafer and the pad was analyzed on the assumption that the polishing pad is perfectly elastic and smooth. This model is applicable in the submicron device level of where the linewidth is much smaller than the size of the pad asperities. Second, the pad is assumed to have flat surface and deforms as discrete blocks. This model accounted for step-height fairly well, but relied on the finite thickness and the physically inadmissible discontinuous deformation of the pad. Third, a simplified rough pad model, which idealizes that pad asperities have uniform height and the same radius of curvature, was developed. The mean asperity contact radius and the asperity spacing were estimated by the analysis based on the general random rough surface of the pad. Finally, a general random rough surface contact between the pad asperities and the patterned wafer during polishing was considered. In these models, the asperity height distribution was assumed to be either exponential or Gaussian. Both elastic and plastic deformations of pad asperities were considered, and their effect on the evolution of step-height during polishing was compared.

Chapter 3 proposes integrated, dielectric erosion and Cu dishing models by combining wafer-, die- and feature-scale wafer surface non-uniformity variations with feature-scale step-height models of Chapter 2. The plausible causes of erosion and dishing at wafer-, die- and feature-scales were identified in terms of the geometric and physical parameters. Such parameters include: Cu interconnect deposition factor, α, wafer-scale uniformity factor, β, and Cu-to-oxide slurry selectivity, $S_{Cu/ox}$. To model wafer-, die- and feature-scale non-uniformities, it was required to consider three separate points on a wafer. First, to calculate the wafer-scale uniformity, field regions in the slowest and the fastest die were considered. These two field regions were defined as the global reference point and local reference point, respectively. Additionally, to calculate die-scale non-uniformities, the general feature in the fastest die, which is the same die with the local reference point, was considered. The polishing time at various polishing stages were expressed by predefined parameters at the wafer-, die- and feature-scales, and integrated into the feature-scale step-height models.

Based on the developed erosion and dishing models, the effects of model parameters on the wafer-surface non-uniformity in Cu CMP are discussed. To reduce dishing and erosion,

- Use slurry with a low selectivity to decrease dishing or a high selectivity to decrease erosion,
- Decrease the initial pattern geometry variation,
- Increase wafer-scale uniformity,
- Minimize the amount of overpolishing, and
- Use a low pressure and a stiff and smooth pad.

Requirements of the wafer-scale uniformity and the initial deposited pattern geometry to satisfy both dishing and erosion specification with a given slurry selectivity were obtained. In single-step polishing, for example, the Cu deposition factor should be less than 0.1 and the wafer-scale uniformity factor needs to be greater than 0.95 to maintain both erosion and dishing within 5% of interconnect thickness across the wafer if the polishing slurry has a selectivity of 15.

Chapter 4 comprises the general procedure for experimental validation. Experimental and analytical means of determining the model parameters were outlined. The chemical and chemo-mechanical effects were included as slurry selectivities and obtained by the average material removal rates from the 100 mm blanket Cu, barrier (Ta) and SiO_2 wafer polishing experiments. The interconnect deposition factor, α, of patterned wafers was obtained by profilometry and SEM. The initial step-height, h_{si}, was measured by a surface profilometer. In PVD Cu patterned wafer, the results showed that the surface profile is almost the same as the underlying pattern geometry: $\alpha = 1$ and $h_{si} \approx h_I$. On the other hand, in the submicron region, the surface profile was close to the field area: $\alpha \approx 0$ and $h_{si} \ll h_I$. The wafer-scale uniformity factor β in a patterned wafer was obtained by comparing the polishing times between two selected points, the field regions in the fastest and slowest dies on the same wafer. Results of polishing experiments on 100 mm patterned Cu wafers validated both the step-height models and the integrated non-uniformity model. Based on the present models, erosion and dishing across the wafer was bounded by predefined parameters. Additionally, as expected in the model, it was observed that the step-heights of the slowest die and the fastest die evolve at the ratio of the wafer-scale uniformity factor.

5.2 Suggestions for Future Work

Based on the integrated erosion and dishing models, further research in the following areas is recommended to reduce erosion and dishing in the CMP process: wafer, equipment, and consumables.

Wafer: The two critical parameters to control erosion and dishing are the initial pattern geometry variation and the wafer-scale uniformity factor. Initial pattern geometry can be controlled by Cu deposition method or additional processing before CMP, e.g., electropolishing.

Equipment: To enhance wafer-scale uniformity, a face-up type CMP equipment is recommended. A kinematic wafer-scale uniformity control scheme is proposed in this thesis, but not validated. Thus, SiO_2 blanket wafer experiments, due to the capability of film thickness measurement, will be useful to verify the theory. Furthermore, a wafer-scale end-point detection mechanism should be incorporated to minimize the overpolishing time of each feature of a wafer in face-up CMP.

Slurry: In this thesis, it is assumed that the material removal is done by pad asperity and wafer contacts. Each pad asperity is assumed to have at least one abrasive particle. The effect of abrasive particles, however, needs more careful investigation, especially on the actual abrasive entrapment mechanism between pad asperity/abrasive/wafer contact interface to increase material removal rate in CMP.

Pad: Based on the proposed integrated erosion and dishing model, the stiffness, yield strength and surface topography of polishing pad play important roles to reduce erosion and dishing. Thus, further investigations employing pads of different characteristics are suggested.

APPENDIX A

FACE-UP CMP: KINEMATICS AND MATERIAL REMOVAL RATE

A.1 Introduction

The primary concern of this thesis is to model and minimize dielectric erosion and Cu dishing across the wafer. The developed erosion and dishing models suggest that the initial Cu surface profile and the wafer-scale uniformity, in addition to selectivity, are the key factors that affect surface polishing non-uniformity. In single-step polishing, the Cu deposition factor should be less than 0.1 and the wafer-scale uniformity factor greater than 0.95 to maintain both erosion and dishing within 5% of the interconnect thickness across the wafer if polishing slurry has a selectivity of 15.

At present, 300 mm wafers are the industry standard. When these wafers are polished by the conventional *face-down* rotary CMP tools, shown in Fig. A.1, the edge polishes faster than the center. An example of wafer-scale non-uniform polishing of 100 mm wafer is shown in Figs. 4.13 and 4.14. This results in both dielectric erosion and the dishing of Cu interconnects. From the Preston equation, non-uniform polishing at the wafer-scale is due to variation in relative velocity, pressure, or the Preston constant. Even if the relative velocity can be precisely set and controlled, and the pressure, too, maintained fairly uniform over the wafer, wafer-scale non-uniformity still cannot be eliminated. In some of the commercial polishers, the wafer carrier is so designed that different pressures are applied in different annuli of the wafer, the so-called zone-pressure-control. While this method is promising, accomplishing uniform wafer-scale polishing is still a formidable task.

The architectures, i.e., geometry and kinematics, of the current commercial CMP tools are compared in Fig. A.1(a)-(d). In the *face-down* CMP, it is difficult to maintain uniform slurry distribution for achieving uniform material removal across the standard 300 mm wafer. Even in the polishing of 100 mm wafer, the problem is acute. Accordingly, a new CMP tool

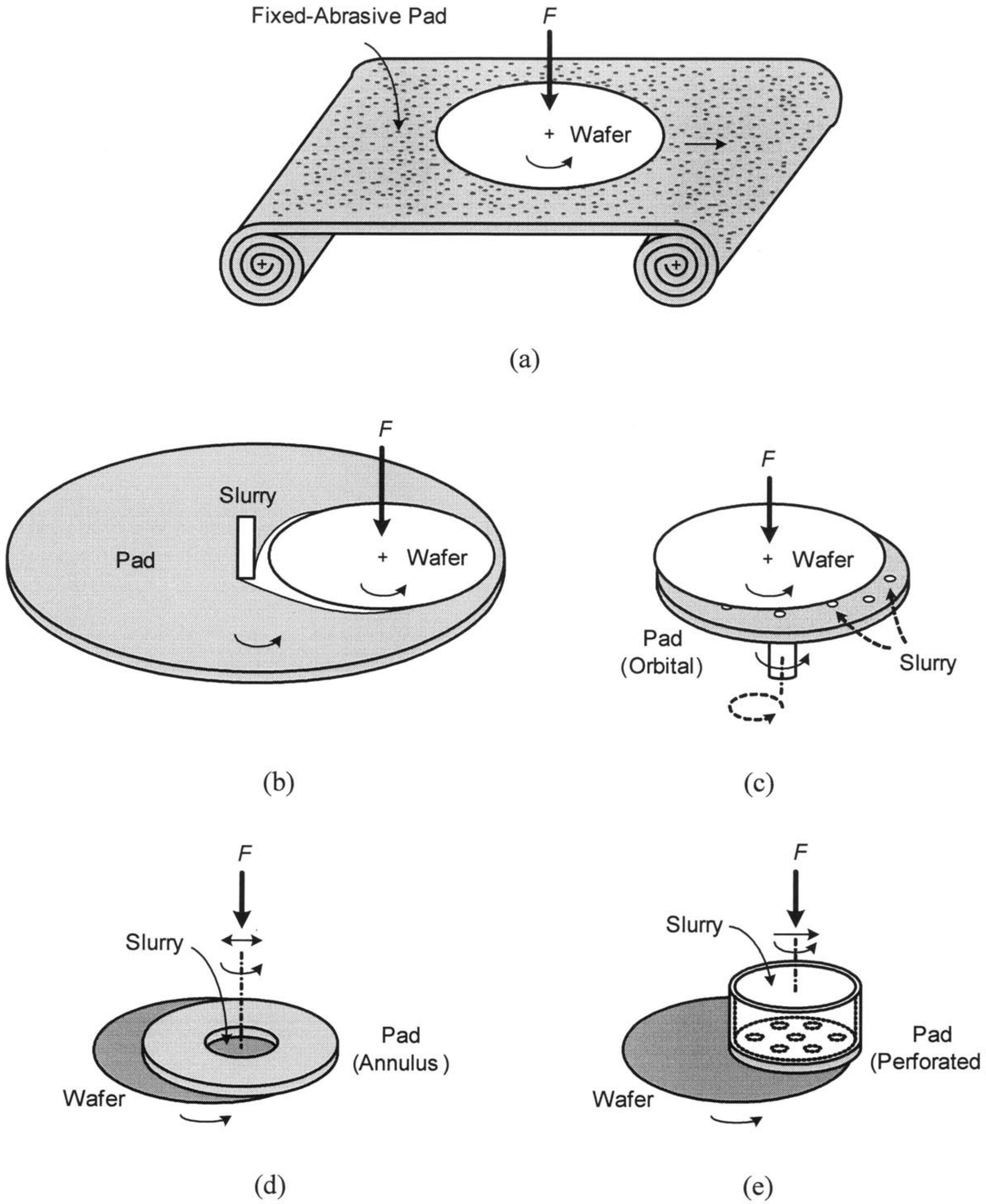


Figure A.1 Geometry and kinematics of the various CMP tools: (a) Face-down, linear; (b) Face-down, rotary; (c) Face-down, orbital; (d) Face-up, rotary; (e) Face-up, rotary.

"architecture" is proposed to control wafer-scale non-uniformity as shown in Fig. A.1(e). In this configuration, the wafer is polished *face up,* and the pad diameter is about the wafer radius. Slurry is fed uniformly in the contact region through perforations in the pad.

A detailed kinematic and kinetic analyses for the conventional *face-down* polisher were conducted by Lai [Lai, 2001]. But because the *face-up* architecture is an inverted version of the *face-down,* the kinematics, material removal rate, etc., need to be reworked in detail. In this appendix, therefore, the kinematic and kinetic analysis are presented for the *face-up* polishing scheme.

A.2 Theory of Face-up Polishing

A.2.1 Geometry

Since the wafer is not completely covered by the pad in the *face-up* scheme, as shown in Fig. A.2, the pad "contact angle" θ_c directly affects on the material removal rate at any radius, r, of the wafer. From the triangle $O_w O_p P$,

$$r_p^2 = r^2 + r_{cc}^2 - 2rr_{cc}\cos\theta_c \tag{A.1}$$

where r_p is the radius of the pad and r_{cc} the distance between the centers of the wafer and the pad. The expression for θ_c has different forms, depending on the center-to-center distance, r_{cc}, wafer radius, r_w, and pad radius, r_p, as shown in Figs. A.2 and A.3.

$$0 \le r_{cc} < r_p: \quad \theta_c(r) = \begin{cases} \pi & r < (r_p - r_{cc}) \\ \cos^{-1}\left(\dfrac{r^2 + r_{cc}^2 - r_p^2}{2rr_{cc}}\right) & (r_p - r_{cc}) \le r \le (r_{cc} + r_p) \\ 0 & r > (r_{cc} + r_p) \end{cases}$$

$$r_{cc} = r_p: \qquad \theta_c(r) = \begin{cases} \cos^{-1}\left(\dfrac{r}{2r_{cc}}\right) & 0 \le r \le (r_{cc} + r_p) \\ 0 & r > (r_{cc} + r_p) \end{cases} \tag{A.2}$$

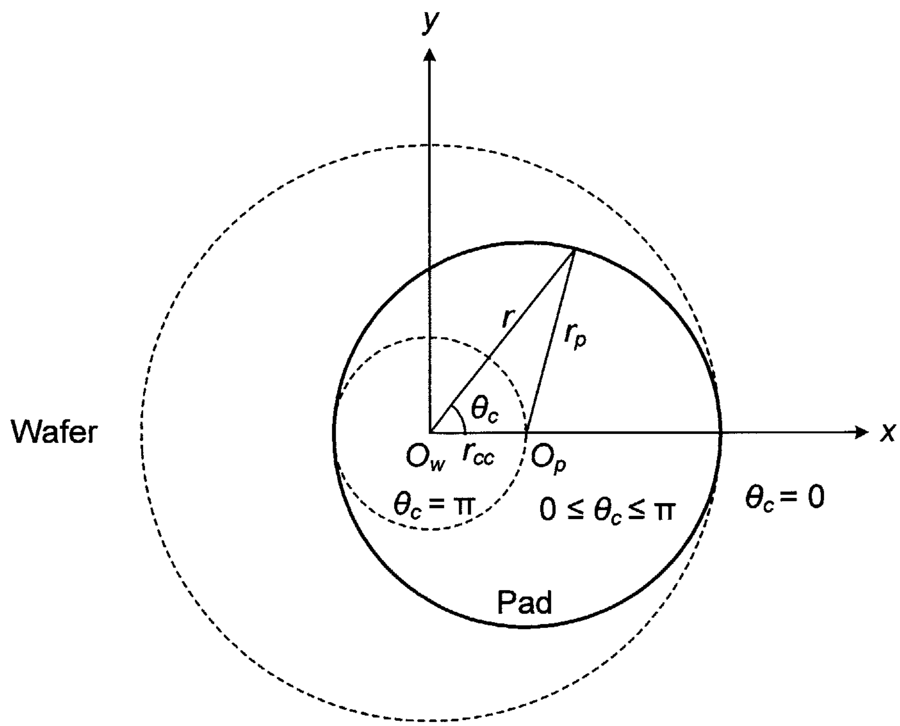


Figure A.2 Definition of pad contact angle in the face-up CMP tool.

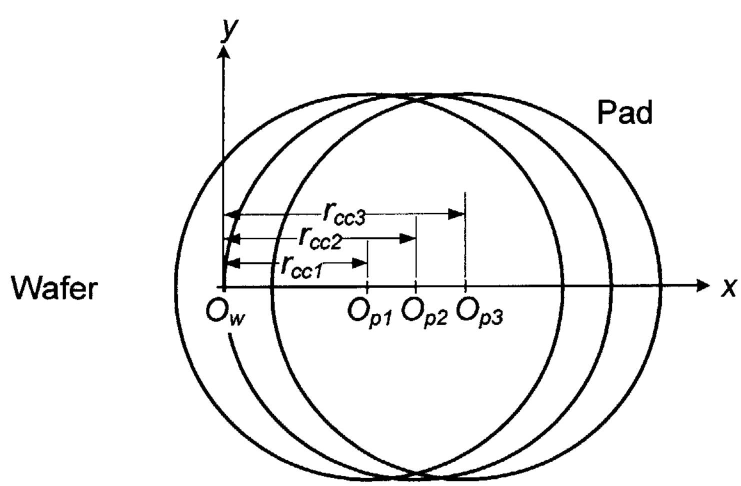


Figure A.3 Three different pad positions in the face-up CMP tool.

$$r_{cc} > r_p: \qquad \theta_c(r) = \begin{cases} 0 & 0 \le r < (r_{cc} - r_p) \\ \cos^{-1}\left(\dfrac{r^2 + r_{cc}^2 - r_p^2}{2rr_{cc}}\right) & (r_{cc} - r_p) \le r \le (r_{cc} + r_p) \\ 0 & r > (r_{cc} + r_p) \end{cases}$$

The maximum contact angle is calculated from the condition $d\theta_c / dr = 0$:

$$\frac{d\theta_c}{dr} = \frac{\dfrac{\left[r^2 - \left(r_{cc}^2 - r_p^2\right)\right]}{2r_{cc}r^2}}{\sqrt{1 - \left[\dfrac{r^2 + \left(r_{cc}^2 - r_p^2\right)}{2rr_{cc}}\right]^2}} = 0 \tag{A.3}$$

Therefore, $\theta_c(r)$ is maximum when $r = \sqrt{r_{cc}^2 - r_p^2}$ and

$$\theta_{c_{\max}} = \cos^{-1}\left[\frac{2\left(r_{cc}^2 - r_p^2\right)}{2\sqrt{r_{cc}^2 - r_p^2} \cdot r_{cc}}\right] = \cos^{-1}\sqrt{1 - \left(\frac{r_p}{r_{cc}}\right)^2} \tag{A.4}$$

It is useful to express the pad contact angle in Eq. (A.2) by the dimensionless variable r / r_w as:

$$0 \le \frac{r_{cc}}{r_w} < \frac{r_p}{r_w}: \qquad \theta_c(r / r_w) = \begin{cases} \pi \\ \cos^{-1}\left[\dfrac{(r / r_w)^2 + (r_{cc} / r_w)^2 - (r_p / r_w)^2}{2\left(rr_{cc} / r_w\right)^2}\right] \\ 0 \end{cases}$$

$$\frac{r_{cc}}{r_w} = \frac{r_p}{r_w}: \qquad \theta_c(r / r_w) = \begin{cases} \cos^{-1}\left[\dfrac{r / r_w}{2\left(r_{cc} / r_w\right)}\right] \\ 0 \end{cases} \tag{A.5}$$

$$\frac{r_{cc}}{r_w} > \frac{r_p}{r_w}: \qquad \theta_c(r / r_w) = \begin{cases} 0 \\ \cos^{-1}\left[\dfrac{(r / r_w)^2 + (r_{cc} / r_w)^2 - (r_p / r_w)^2}{2\left(rr_{cc} / r_w^{\,2}\right)}\right] \\ 0 \end{cases}$$

Figure A.4 shows the contact angle, θ_c, for various pad locations when $r_p = 0.5\, r_w$.

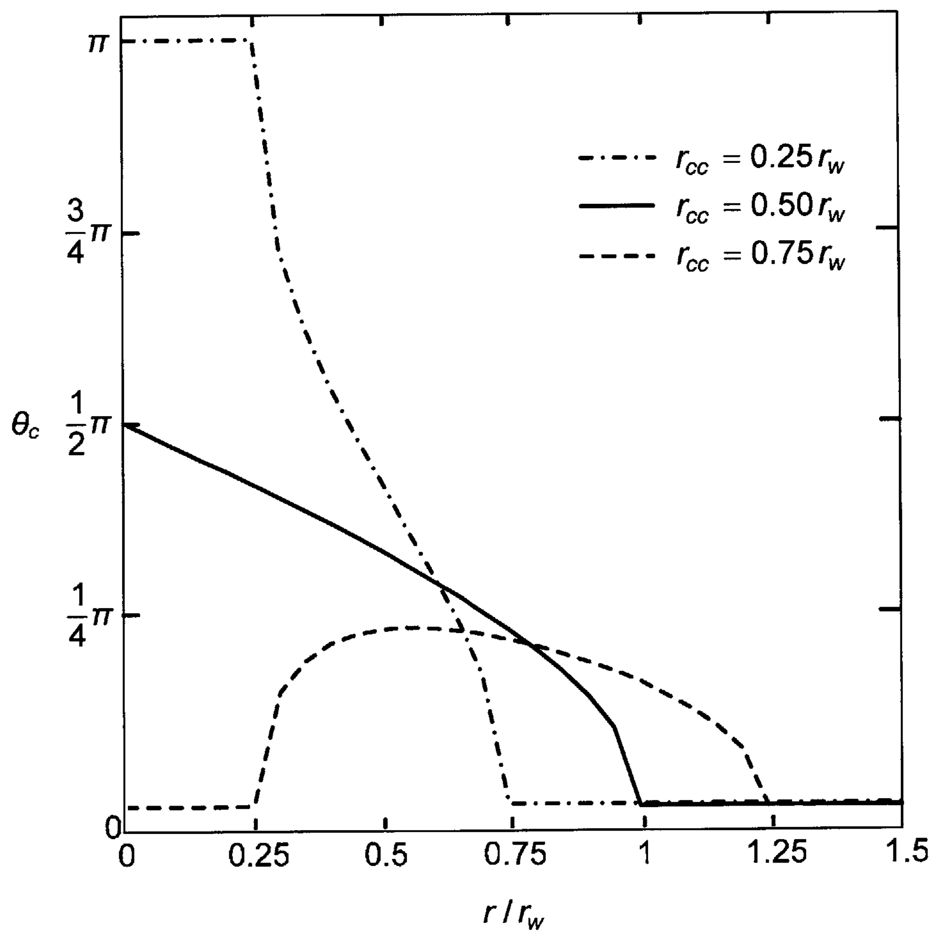


Figure A.4   Contact angle for three different pad positions.

A.2.2 Kinematics

Figure A.5 illustrates the coordinate systems for the conventional *face-down* CMP machine in Fig. A.1(b) and for the *face-up* polisher in Fig. A.1(e). The Cartesian and polar coordinate systems, fixed at the center of the wafer are represented by (x,y) and (r,θ), respectively. The polar coordinate systems, fixed to the pad and rotating with it, is represented by (r',θ'). Let the rotational centers of the wafer and the pad be O_w and O_p, and the angular velocities, respectively, be ω_w and ω_p. The two rotational axes are normal to the polishing plane with an offset r_{cc}. Additionally, the pad may translate horizontally with a velocity v_{cc} along the x-axis.

$$\boldsymbol{v}_{cc} = \dot{r}_{cc}\boldsymbol{e}_x = v_{cc}\boldsymbol{e}_x \tag{A.6}$$

The coordinates of a point P' on the pad (r',θ'), shown in Fig. A.5(b), can be readily converted to the wafer coordinates (r,θ), by the transformation relations:

$$\begin{aligned} r\cos\theta &= r_{cc} + r'\cos\theta' \\ r\sin\theta &= r'\sin\theta' \end{aligned} \tag{A.7}$$

The velocity at point $P'(r',\theta')$ on the pad can be expressed as:

$$\boldsymbol{v}_{P'} = \left(v_{cc} - \omega_p r'\sin\theta'\right)\boldsymbol{e}_x + \omega_p r'\cos\theta'\boldsymbol{e}_y \tag{A.8}$$

In the wafer coordinate (r,θ) system,

$$\boldsymbol{v}_{P'} = \left(v_{cc} - \omega_p r\sin\theta\right)\boldsymbol{e}_x + \left(\omega_p r\cos\theta - \omega_p r_{cc}\right)\boldsymbol{e}_y \tag{A.9}$$

The velocity at a point $P(r,\theta)$ on the wafer is expressed as:

$$\boldsymbol{v}_P = -\omega_w r\sin\theta\boldsymbol{e}_x + \omega_w r\cos\theta\boldsymbol{e}_y \tag{A.10}$$

Therefore, the velocity of the wafer relative to the pad, $\boldsymbol{v}_R$, at a point $P(r,\theta)$ is calculated in the wafer coordinate system as:

$$\boldsymbol{v}_R(r,\theta) = -\left[\left(\omega_w - \omega_p\right)r\sin\theta + v_{cc}\right]\boldsymbol{e}_x + \left[\left(\omega_w - \omega_p\right)r\cos\theta + \omega_p r_{cc}\right]\boldsymbol{e}_y \tag{A.11}$$

In the rotating pad coordinate system

$$\boldsymbol{v}_R(r',\theta') = -\left[\left(\omega_w - \omega_p\right)r'\sin\theta' + v_{cc}\right]\boldsymbol{e}_x + \left[\left(\omega_w - \omega_p\right)r'\cos\theta' + \omega_w r_{cc}\right]\boldsymbol{e}_y \tag{A.12}$$

The magnitude of the relative velocity, $v_R\ (= |\boldsymbol{v}_R|)$, is expressed as a function of r and θ in the wafer coordinate system as:

$$v_R(r,\theta) = \sqrt{\left[\left(\omega_w - \omega_p\right)r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right)r\cos\theta + \omega_p r_{cc}\right]^2} \tag{A.13}$$

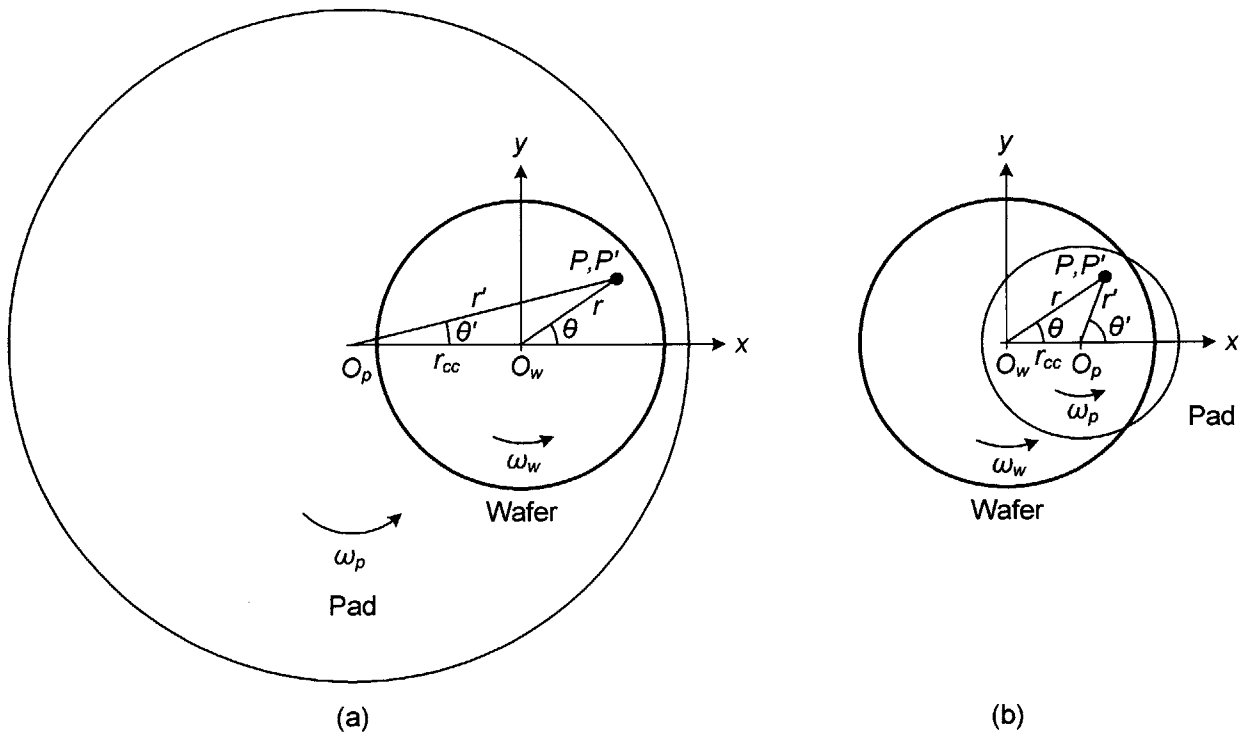


Figure A.5 Coordinate systems for (a) the conventional face-down and (b) the face-up CMP tools.

In the rotating pad coordinate system,

$$v_R(r',\theta') = \sqrt{[(\omega_w - \omega_p)r'\sin\theta' + v_{cc}]^2 + [(\omega_w - \omega_p)r'\cos\theta' + \omega_w r_{cc}]^2} \qquad \text{(A.14)}$$

Equations (A.12)–(A.14) are applicable for the region in contact between a pad and a wafer at a given time. When the translational velocity is zero or negligible, i.e., $v_{cc} = 0$, the relative velocity can be rewritten as:

$$v_R(r,\theta) = \sqrt{[(\omega_w - \omega_p)r]^2 + (\omega_p r_{cc})^2 + 2(\omega_w - \omega_p)\omega_p r_{cc} r\cos\theta} \qquad \text{(A.15)}$$

If the wafer and the pad rotate at the same speed, $\omega_w = \omega_p = \omega$, the relative velocity from Eq. (A.13) is uniform over the wafer. Thus,

$$v_R = \sqrt{v_{cc}^{\,2} + (\omega r_{cc})^2} \qquad \text{(A.16)}$$

Additionally, if $v_{cc} = 0$ or $v_{cc} / \omega r_{cc} \ll 1$, then

$$v_R = \omega r_{cc} \qquad \text{(A.17)}$$

Suppose that the center of the wafer, O_w, is just covered by the pad, the relative velocity at the center of the wafer is:

$$v_R(0,\theta) = \omega_p r_{cc} \qquad \text{(A.18)}$$

In this case, $v_R(r,\theta)$ relative to $v_R(0,\theta)$ can be expressed as:

$$\frac{v_R(r,\theta)}{\omega_p r_{cc}} = \sqrt{\left[\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{r}{r_{cc}}\right)\sin\theta + \frac{v_{cc}}{\omega_p r_{cc}}\right]^2 + \left[\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{r}{r_{cc}}\right)\cos\theta + 1\right]^2} \qquad \text{(A.19)}$$

Again, if $v_{cc} = 0$

$$\frac{v_R(r,\theta)}{\omega_p r_{cc}} = \sqrt{\left[\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{r}{r_{cc}}\right)\right]^2 + 1 + 2\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{r}{r_{cc}}\right)\cos\theta} \qquad \text{(A.20)}$$

Equation (A.19) can be expressed in the wafer-centered Cartesian coordinate system as:

$$\frac{v_R(x,y)}{\omega_p r_{cc}} = \sqrt{\left[\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{y}{r_{cc}}\right) + \frac{v_{cc}}{\omega_p r_{cc}}\right]^2 + \left[\left(\frac{\omega_w}{\omega_p} - 1\right)\left(\frac{x}{r_{cc}}\right) + 1\right]^2} \qquad \text{(A.21)}$$

Figure A.6 shows examples of normalized relative velocity $(v_R / \omega_p r_{cc})$ profiles for various ω_w / ω_p ratios.

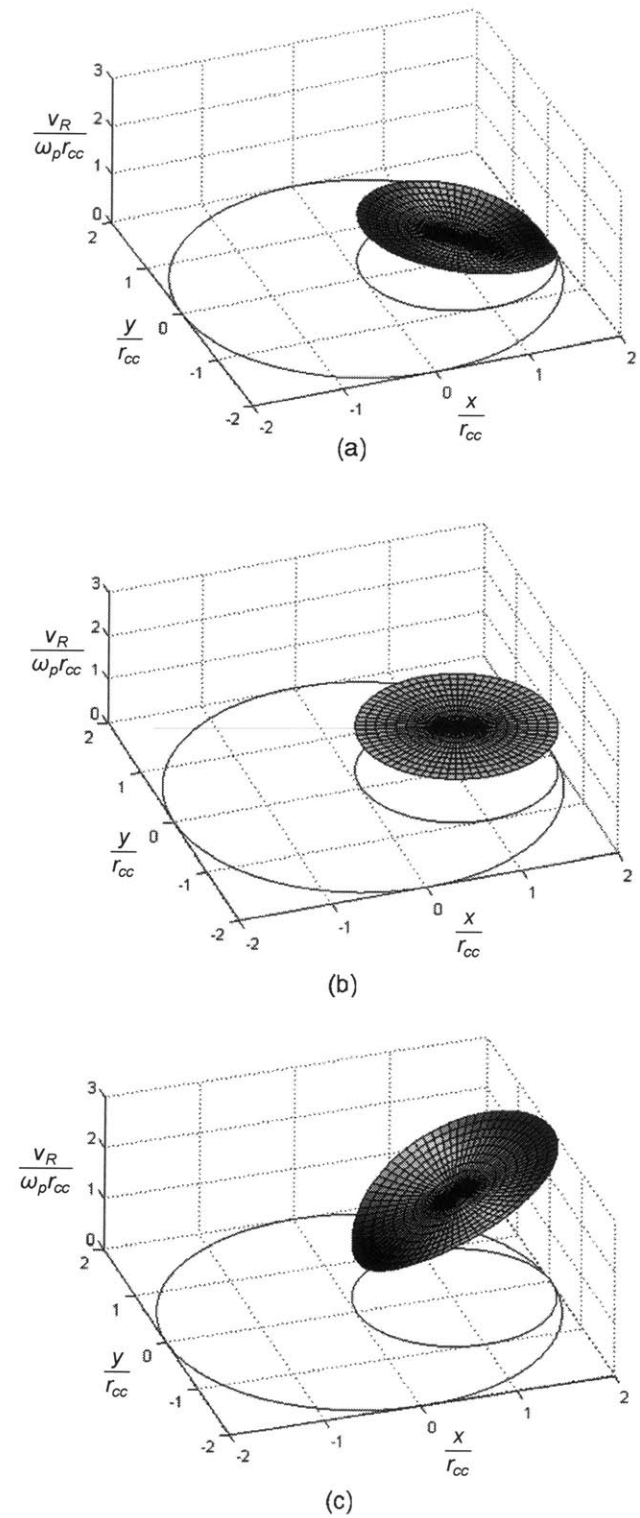


Figure A.6 Relative velocity profile for (a) $\omega_w / \omega_p = 0.5$, (b) $\omega_w / \omega_p = 1$, and (c) $\omega_w / \omega_p = 2$ when $r_w \geq r_{cc} + r_p$, $r_p / r_{cc} = 1$ and $v_{cc} = 0$.

A.2.3 Material Removal Rate

Material removal rate at $P(r,\theta)$ on the wafer at a given time t is expressed by the Preston equation [Preston, 1927]:

$$\left|\frac{dh}{dt}\right| = k_p \cdot p \cdot v_R \tag{A.22}$$

where h is the thickness of the layer removed, t the polishing time, p the nominal pressure, v_R the relative velocity, and k_p the Preston constant. Since Preston equation represents the local material removal rate, the amount of material removed Δh at $P(r,\theta)$, during the time for one wafer rotation, Δt, can be calculated as:

$$\Delta h = \int_0^{\Delta t} \left|\frac{dh}{dt}\right| dt = \int_0^{2\pi} \frac{1}{\omega_w} \left|\frac{dh}{dt}\right| d\theta \tag{A.23}$$

where $dt = d\theta / \omega_w$ and $\Delta t = 2\pi / \omega_w$.

In conventional *face-down* CMP tools, the pad is always in contact with the wafer during polishing. Therefore any point on the wafer is always under pressure. However, if there is a non-contact region in the wafer, as in the *face-up* CMP scheme shown in Fig. A.7, the material removal rate is zero. Thus, the pad size affects material removal rates and the Preston equation is valid only in the contact region. Thus,

$$\left|\frac{dh}{dt}\right| = \begin{cases} k_p p v_R & -\theta_{c_1} \le \theta \le \theta_{c_2} \\ 0 & \text{otherwise} \end{cases} \tag{A.24}$$

where θ_{c_1} and θ_{c_2} are pad contact angles at incoming point P_1 and outgoing point P_2 respectively. At any given time, the pad contact angle θ_c is determined by r, r_p and r_{cc} as described in the previous section. Therefore, if r_{cc} changes while the wafer rotates, θ_{c_1} and θ_{c_2} would be different.

$$\theta_{c_1} = \theta_{c_1}(r,t), \theta_{c_2} = \theta_{c_2}(r,t) \tag{A.25}$$

Combining Eq. (A.23) with Eq. (A.24), the amount of material removed during one wafer rotation at any radius r on the wafer can be expressed as:

$$\Delta h(r) = \int_{-\theta_{c_1}}^{\theta_{c_2}} \frac{1}{\omega_w} k_p p v_R \, d\theta \tag{A.26}$$

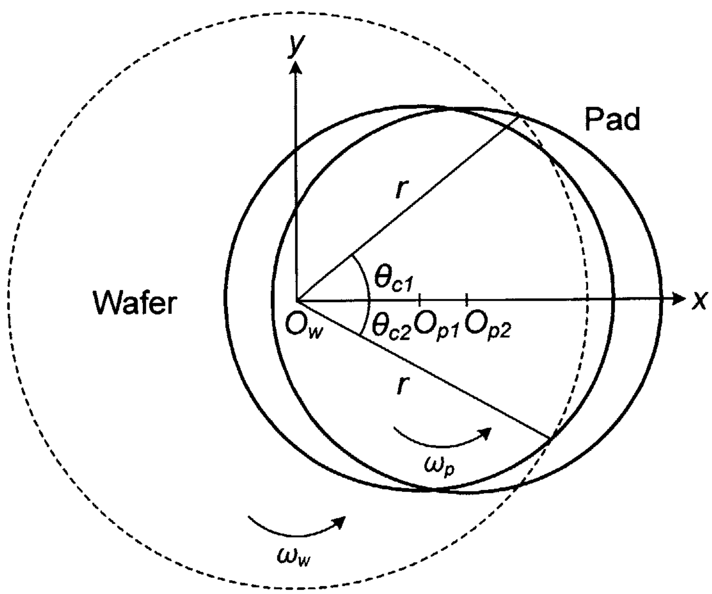


Figure A.7 Geometry for material removal rate in the face-up CMP configuration.

In general, the Preston constant, k_p, and the applied pressure, p, may not be constant across the wafer-scale for various reasons. In the current analysis, however, we assume that k_p and p are uniform in the contact area because the pad is much smaller than the wafer. Thus Eq. (A.26) can be rewritten as:

$$\Delta h(r) = k_p p \int_{-\theta_{c_1}}^{\theta_{c_2}} \frac{1}{\omega_w} v_R \, d\theta \tag{A.27}$$

Replacing v_R in (A.27) with (A.13),

$$\Delta h(r) = k_p p \int_{-\theta_{c_1}}^{\theta_{c_2}} \frac{1}{\omega_w} \sqrt{\left[(\omega_w - \omega_p) r \sin\theta + v_{cc}\right]^2 + \left[(\omega_w - \omega_p) r \cos\theta + \omega_p r_{cc}\right]^2} \, d\theta \tag{A.28}$$

The parameters, ω_w, ω_p and r_{cc} may be time-dependent. Similarly, θ_{c_1} and θ_{c_2} may be time-dependent but are determined by r_{cc}.

$$\begin{aligned} \omega_w &= \omega_w(t), \omega_p = \omega_p(t) \\ r_{cc} &= r_{cc}(t) \\ \theta_{c_1} &= \theta_{c_1}(r,t), \theta_{c_2} = \theta_{c_2}(r,t) \end{aligned} \tag{A.29}$$

Although these are time-dependent parameters, since $\Delta h(r)$ is defined as the thickness of material removed in each wafer rotation, it is a function of both radius, r, and the number of wafer rotations, n.

$$\Delta h = \Delta h(r,n) \tag{A.30}$$

Let the center of the wafer, O_w, be a reference point, when the periphery of the pad is at O_w. Thus,

$$r_{cc}(0) = r_p \tag{A.31}$$

In this initial position, the contact angle θ_c is:

$$\theta_{c_1}(0,0) = \theta_{c_2}(0,0) = \frac{\pi}{2} \tag{A.32}$$

The thickness of the material removed at the center of the wafer in the first revolution of the wafer can be expressed as:

$$\Delta h(0,1) = k_p p \frac{1}{\omega_w(0)} \int_{-\pi/2}^{\pi/2} \omega_p(0) r_{cc}(0) \, d\theta = k_p p \cdot \pi r_{cc}(0) \cdot \frac{\omega_p(0)}{\omega_w(0)} \tag{A.33}$$

Thus, a new dimensionless variable, Δh^*, is defined as the ratio of the material removed at radius r to the initial material removed at the center of the wafer during one wafer rotation as:

$$\Delta h^*(r,n) \equiv \frac{\Delta h(r,n)}{\Delta h(0,1)} \tag{A.34}$$

Therefore,

$$\Delta h^*(r,n) = \frac{1}{\pi}\int_{-\theta_{c1}}^{\theta_{c2}} \left[\frac{\omega_w(0)}{\omega_w}\right] \cdot \sqrt{\left\{\left[\frac{\omega_w-\omega_p}{\omega_p(0)}\right]\left[\frac{r}{r_{cc}(0)}\right]\sin\theta + \frac{v_{cc}}{\omega_p(0) r_{cc}(0)}\right\}^2 + \left\{\left[\frac{\omega_w-\omega_p}{\omega_p(0)}\right]\left[\frac{r}{r_{cc}(0)}\right]\cos\theta + \frac{\omega_p r_{cc}}{\omega_p(0) r_{cc}(0)}\right\}^2}\, d\theta \tag{A.35}$$

The amount of material removed ΔH, during the polishing time, t, $(t/\Delta t \gg 1)$, can be expressed as the sum of Δh for each wafer rotation.

$$\Delta H(r,N) = \sum_{n=1}^{N} \Delta h(r,n) = \sum_{n=1}^{N} \Delta h(0,1)\Delta h^*(r,n) \tag{A.36}$$

where $t = \sum_{n=1}^{N} \Delta t(n)$ and N the number of wafer rotations in time t.

A. Stationary pad: In this case, all variables are time-independent, i.e.,

$$\begin{aligned} \omega_w(t) &= \omega_w \\ \omega_p(t) &= \omega_p \\ r_{cc}(t) &= r_{cc} \end{aligned} \tag{A.37}$$

Thus,

$$\begin{aligned} &\theta_{c_1}(r,t) = \theta_{c_2}(r,t) = \theta_c(r) \\ &\Delta h(r,n) = \Delta h(r),\ \Delta h^*(r,n) = \Delta h^*(r) \end{aligned} \tag{A.38}$$

Additionally, if $v_{cc}(t) = 0$, Δh and Δh^* are even functions. Thus,

$$\Delta h(r) = k_p p \int_0^{\theta_c} \frac{2}{\omega_w} \sqrt{\left[(\omega_w - \omega_p) r \sin\theta\right]^2 + \left[(\omega_w - \omega_p) r\cos\theta + \omega_p r_{cc}\right]^2}\, d\theta \tag{A.39}$$

$$\Delta h^*(r) = \frac{2}{\pi}\int_0^{\theta_c} \sqrt{\left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\sin\theta\right]^2 + \left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\cos\theta + 1\right]^2}\, d\theta \tag{A.40}$$

Figure A.8 shows Δh^* plot versus radial location in wafer coordinates for various rotational speed ratios.

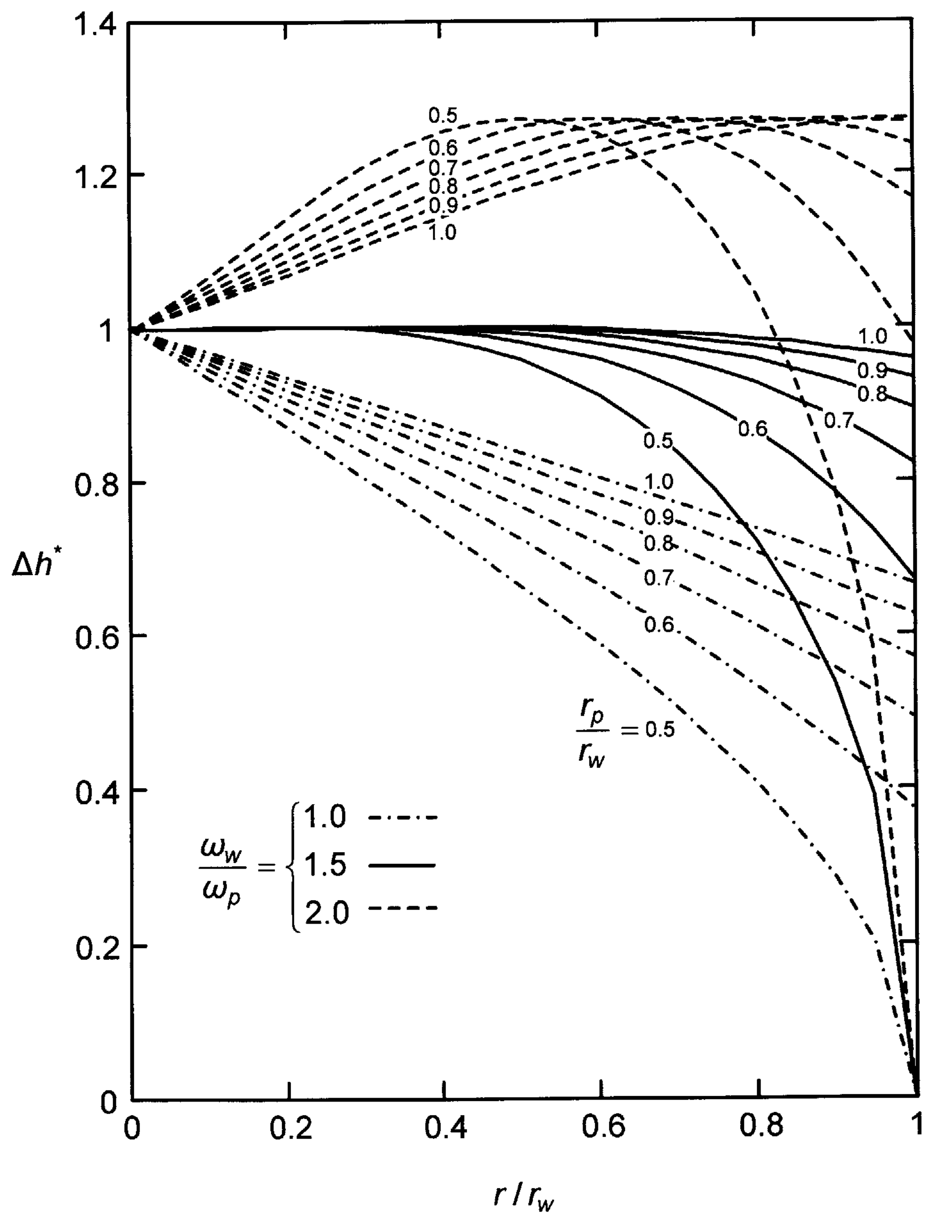


Figure A.8 Δh^* versus r / r_w plots for various ω_w / ω_p and r_p / r_w values.

Expressing Eq. (A.39) in a dimensionless form

$$\Delta h(r) = \left[k_p p\left(\pi r_{cc}\right)\frac{\omega_p}{\omega_w}\right]\left\{\frac{2}{\pi}\int_0^{\theta_c}\sqrt{\left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\sin\theta\right]^2+\left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\cos\theta+1\right]^2}\,d\theta\right\}$$

(A.41)

The thickness of material removed, ΔH, in time t can be calculated as:

$$\Delta H(r,N) = \left[\frac{1}{2}k_p p\left(\omega_p r_{cc}\right)\right]\cdot\Delta h^*(r)\cdot t \qquad \text{(A.42)}$$

where $t = 2\pi N / \omega_w$.

Assuming that the initial thickness of Cu, h_{Cu}, is uniform across the wafer, the required polishing time or end-point, $t_{ep}(r)$, to remove excess copper at a radius r is expressed as:

$$t_{ep}(r) = \frac{h_{Cu}}{\left[\frac{1}{2}k_p p\left(\omega_p r_{cc}\right)\right]\Delta h^*(r)} \qquad \text{(A.43)}$$

Thus at the center of a wafer, $r = 0$,

$$t_{ep}(0) = \frac{h_{Cu}}{\left[\frac{1}{2}k_p p\left(\omega_p r_{cc}\right)\right]\Delta h^*(0)} = \frac{h_{Cu}}{\left[\frac{1}{2}k_p p\left(\omega_p r_{cc}\right)\right]} \qquad \text{(A.44)}$$

A new dimensionless variable, t_{ep}^*, which represents the ratio of the required polishing time at radius r to that at the center of a wafer to remove any given Cu thickness, is represented as:

$$t_{ep}^*(r) \equiv \frac{t_{ep}(r)}{t_{ep}(0)} = \frac{1}{\Delta h^*(r)} \qquad \text{(A.45)}$$

The dimensionless polishing time t_{ep}^* versus the radial location of the wafer for various wafer/pad rotational speeds are shown in Fig. A.9.

Incidentally, in the conventional face-down CMP tool, Δh^* is represented by the same expression as in face-up configuration.

$$\Delta h^*(r) = \frac{1}{\pi}\int_0^{\pi}\sqrt{\left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\sin\theta+\left(\frac{v_{cc}}{\omega_p r_{cc}}\right)\right]^2+\left[\left(\frac{\omega_w}{\omega_p}-1\right)\left(\frac{r}{r_{cc}}\right)\cos\theta+1\right]^2}\,d\theta \qquad \text{(A.46)}$$

Figure A.10 shows Δh^* across the wafer for various wafer/pad speed ratios in the face-up CMP. Since the wafer is always in contact with the pad now, $\theta_c = \pi$, there is no geometric effect of the pad and thus Δh^* is affected by relative velocity profile only.

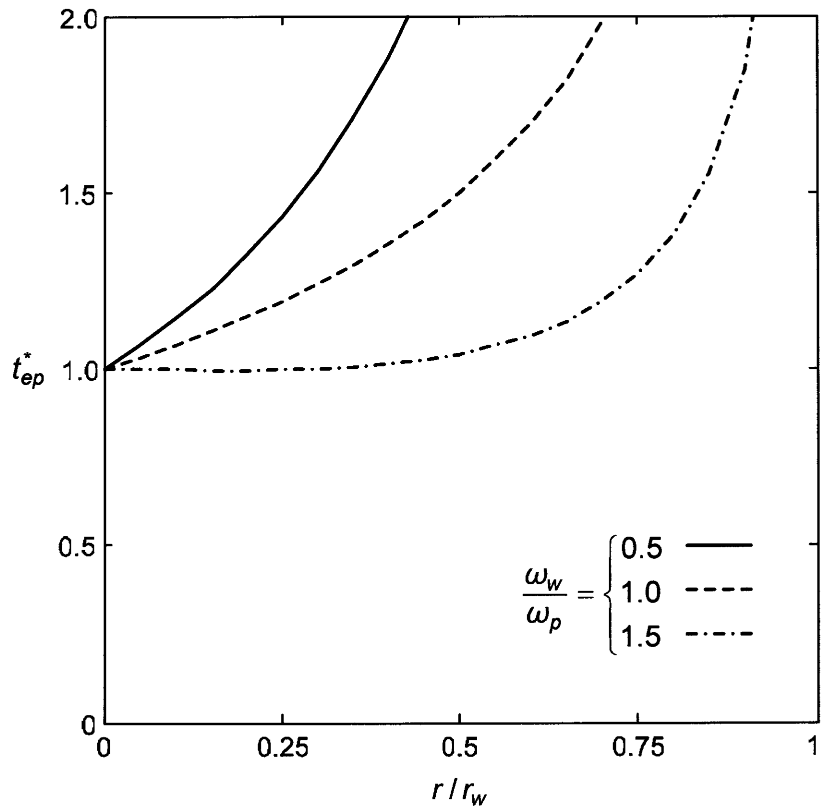
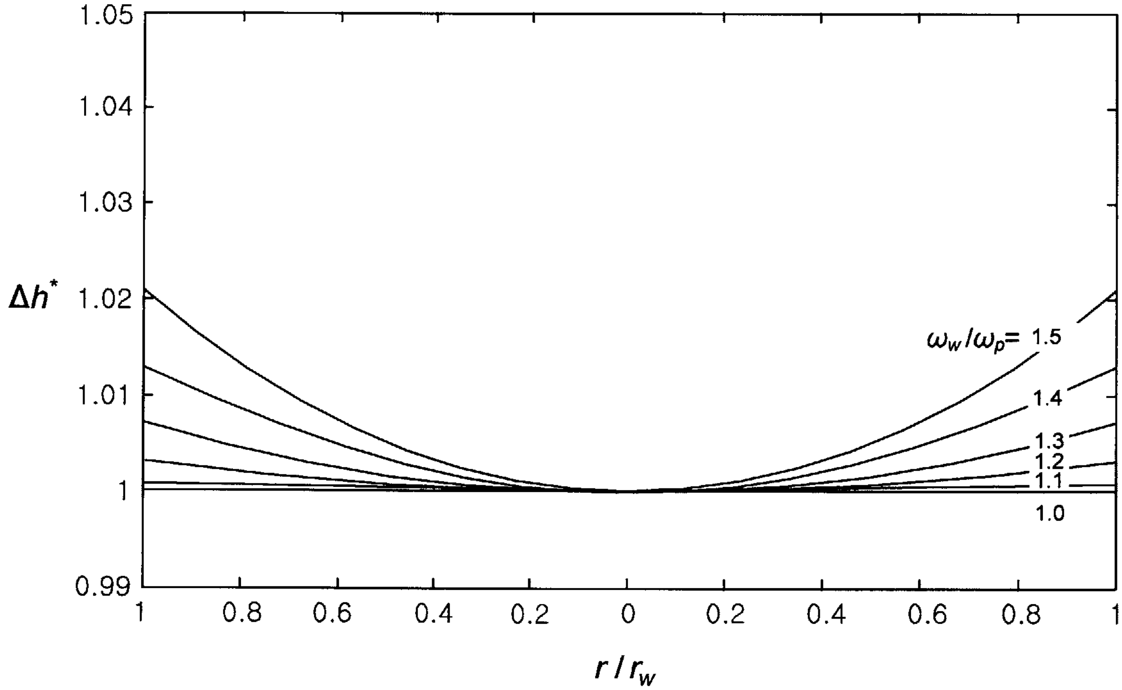
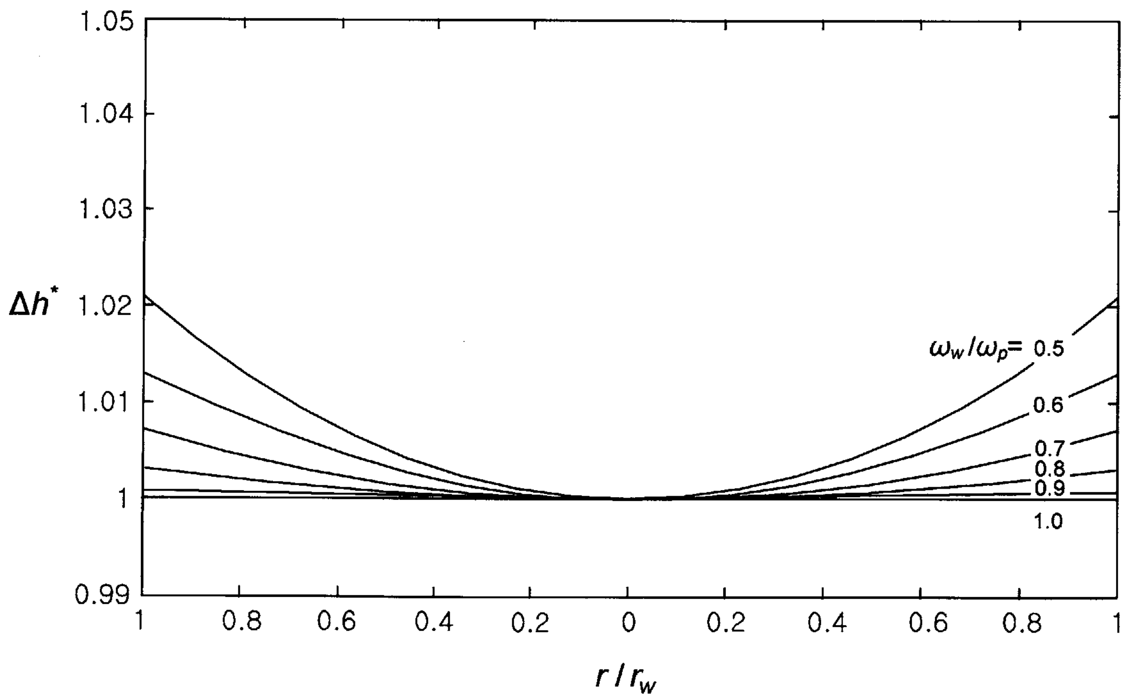


Figure A.9 t_{cp}^* versus $r\,/\,r_{cc}$ plots for various $\omega_w\,/\,\omega_p$ when $r_p\,/\,r_{cc}\,=1$ and $v_{cc}\,=0$.



(a)



(b)

Figure A.10 Plot for material removal rate ratio Δh^* for various wafer/pad rotational speed ratios in the face-down CMP tool: (a) $\omega_w / \omega_p > 1$ and (b) $\omega_w / \omega_p < 1$.

B. Moving pad: In this case, it will be assumed that there is only one time-dependent variable, $r_{cc} = r_{cc}(t)$. Thus

$$\begin{aligned} \omega_w(t) &= \omega_w \\ \omega_p(t) &= \omega_p \end{aligned} \tag{A.47}$$

Since r_{cc} is not the same during one turn of the wafer, $\theta_{c_1} \neq \theta_{c_2}$,

$$\Delta h(r,n) = k_p p \int_{-\theta_{c1}}^{\theta_{c2}} \frac{1}{\omega_w} \sqrt{\left[(\omega_w - \omega_p) r \sin\theta + v_{cc}\right]^2 + \left[(\omega_w - \omega_p) r \cos\theta + \omega_p r_{cc}\right]^2}\, d\theta \tag{A.48}$$

$$\Delta h^*(r,n) = \frac{1}{\pi} \int_{-\theta_{c1}}^{\theta_{c2}} \sqrt{\left\{\left(\frac{\omega_w}{\omega_p} - 1\right)\left[\frac{r}{r_{cc}(0)}\right]\sin\theta + \frac{v_{cc}}{\omega_p r_{cc_o}}\right\}^2 + \left\{\left(\frac{\omega_w}{\omega_p} - 1\right)\left[\frac{r}{r_{cc}(0)}\right]\cos\theta + \frac{r_{cc}}{r_{cc}(0)}\right\}^2}\, d\theta \tag{A.49}$$

In many practical cases, however, the pad may translate in a quasi-static manner, i.e., $r_{cc}(t)$ is a step-function. In these cases, $v_{cc} = 0$ during most of the polishing period, and thus the following conditions can be considered:

$$\begin{aligned} &\theta_{c_1}(r,t) = \theta_{c_2}(r,t) = \theta_c(r,t) \\ &v_{cc}(t) = 0 \end{aligned} \tag{A.50}$$

Therefore, Eq. (A.49) is reduced to

$$\Delta h^*(r,n) = \frac{2}{\pi} \int_0^{\theta_c} \sqrt{\left\{\left(\frac{\omega_w}{\omega_p} - 1\right)\left[\frac{r}{r_{cc}(0)}\right]\sin\theta\right\}^2 + \left\{\left(\frac{\omega_w}{\omega_p} - 1\right)\left[\frac{r}{r_{cc}(0)}\right]\cos\theta + \frac{r_{cc}}{r_{cc}(0)}\right\}^2}\, d\theta \tag{A.51}$$

Figure A.11 shows Δh^* plot versus radial position in the wafer coordinate for various r_{cc} and rotational speed ratios between the wafer and the pad. The total material removed after polishing time, t, is

$$\Delta H(r,N) = \left[k_p p \cdot \pi r_{cc}(0) \cdot \frac{\omega_p}{\omega_w}\right] \sum_{n=1}^{N} \Delta h^*(r,n) \tag{A.52}$$

To remove excess Cu layer uniformly over the wafer, the following condition should be satisfied by adjusting $r_{cc}(t)$ based on the speed ratio, ω_w / ω_p.

$$h_{Cu} = \Delta h(0,1) \cdot \sum_{n=1}^{N} \Delta h^*(r,n) \tag{A.53}$$

There may be many scenarios for choosing $r_{cc}(t)$ under the condition in (A.53). Since the center of the wafer has the highest value of Δh when $\omega_w / \omega_p < 1.5$ for the initial condition

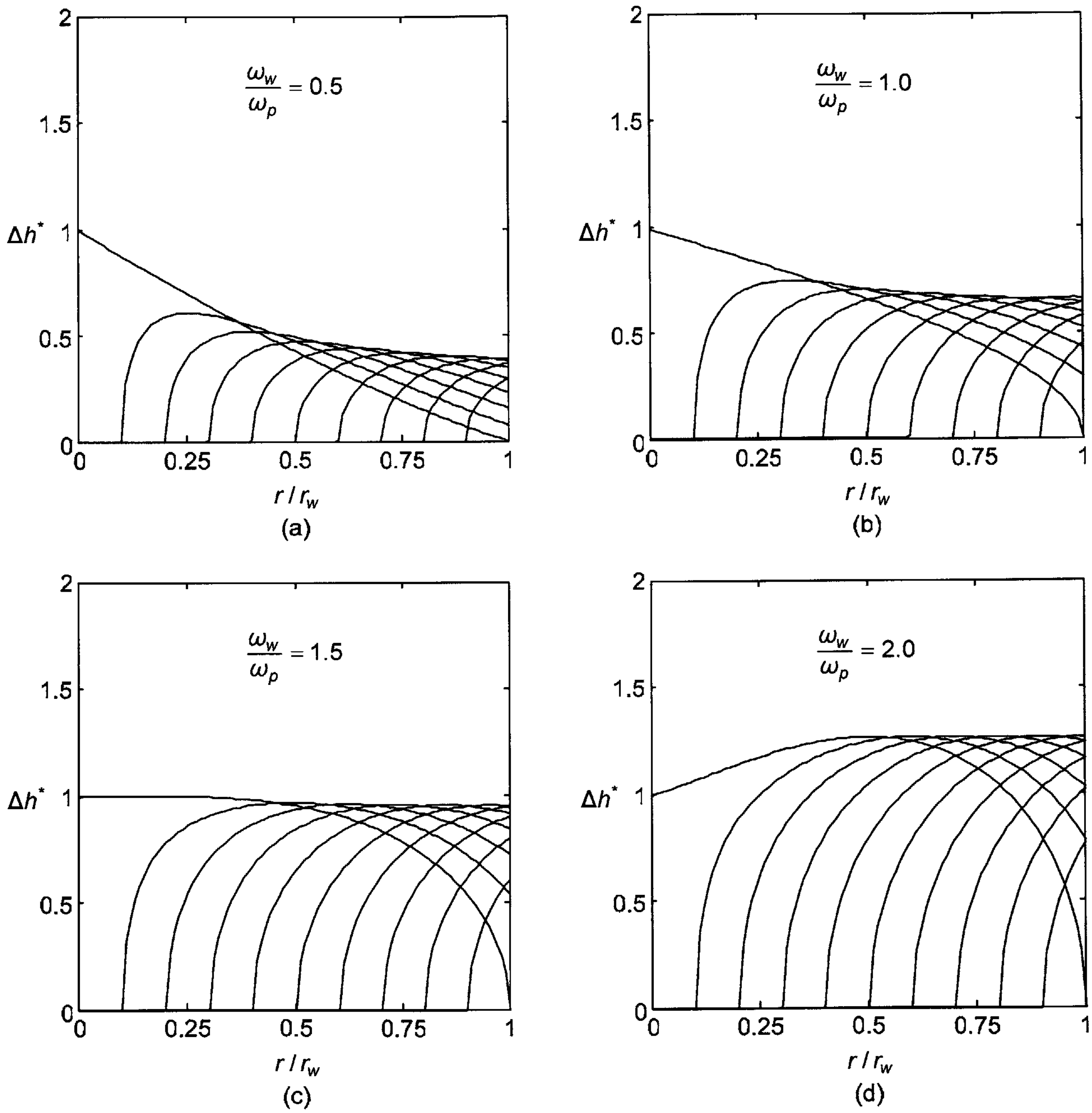


Figure A.11 Δh^* versus r / r_{cc} plots for various ω_w / ω_p values when the pad is moving and $r_p / r_{cc} = 1$ initially.

$r_{cc} = r_p$ as shown in Fig. A.8, the simplest way to satisfy Eq. (A.53) is to move the pad in one direction, from the center to the edge of the wafer.

In this case, the pad maintains the initial condition $r_{cc} = r_p$ until the total material removed at the center of the wafer reaches h_{Cu} as:

$$h_{Cu} = \Delta h(0,1) \cdot N = \Delta h(0,1) \cdot \frac{t_{ep}(0)}{\left(\dfrac{2\pi}{\omega_w}\right)} \tag{A.54}$$

After the polishing time t reaches $t_{ep}(0)$, the pad starts to move outward in a quasi-static way. After the pad is moved to the next position by the amount of Δr_{cc}, it stays until the closest point to the center of a wafer, which is $r = r_{cc} - r_p$ in this case, is completely polished. Thus,

$$h_{Cu} = \Delta h(0,1) \cdot \sum_{n=1}^{N} \Delta h^*(r_{cc} - r_p, n) \tag{A.55}$$

Since the pad is moving from the center to the edge of a wafer, the end-point of polishing, t_{ep}, is when the edge of a wafer is completely polished. That is,

$$h_{Cu} = \Delta h(0,1) \cdot \sum_{n=1}^{N} \Delta h^*(r_w, n) \tag{A.56}$$

Since the pad moves in a quasi-static way, the value of Δr_{cc} must be limited by the required wafer-scale uniformity specification. Figure A.12 shows the location of the center of the pad, r_{cc}, during polishing to satisfy 1% of wafer-scale uniformity, $\beta = 0.99$, for various wafer/pad rotational speed ratios. The final wafer surface profile is shown in Fig. A.13.

A.2.4 Determination of the Preston Constant

From the Preston equation, the decrease of film thickness can be expressed by

$$dh = -k_p \cdot p \cdot ds \tag{A.57}$$

For one wafer revolution, the decrease of film thickness, Δh, and the average volume, ΔV, at any radius r can be expressed as:

$$\Delta h(r) = k_p \cdot p \cdot \Delta s(r) \tag{A.58}$$

$$dV(r) = k_p \cdot p \cdot \Delta s(r) \cdot 2\pi r dr \cdot N \tag{A.59}$$

where $N = \Delta t \cdot \omega_w / 2\pi$.

Therefore, the total material loss during polishing can be expressed as:

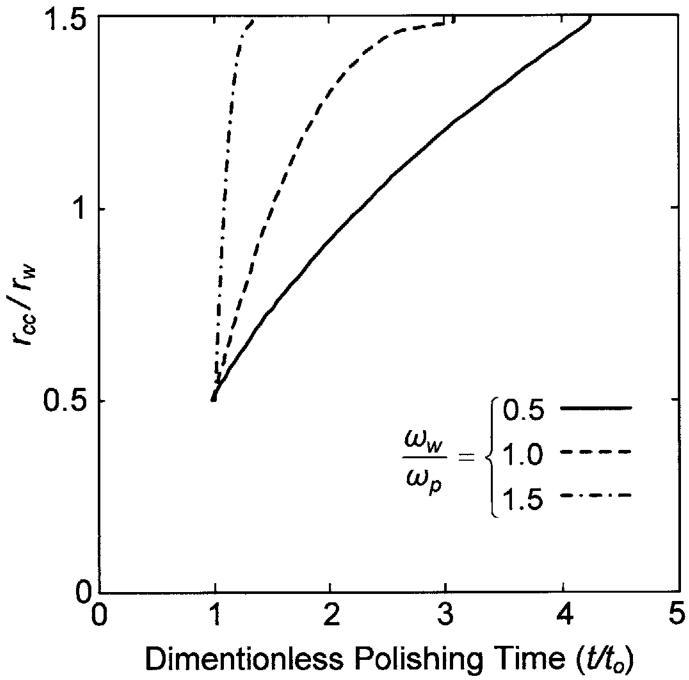


Figure A.12 Position of the pad, r_{cc}, profile versus dimensionless polishing time for a moving pad.

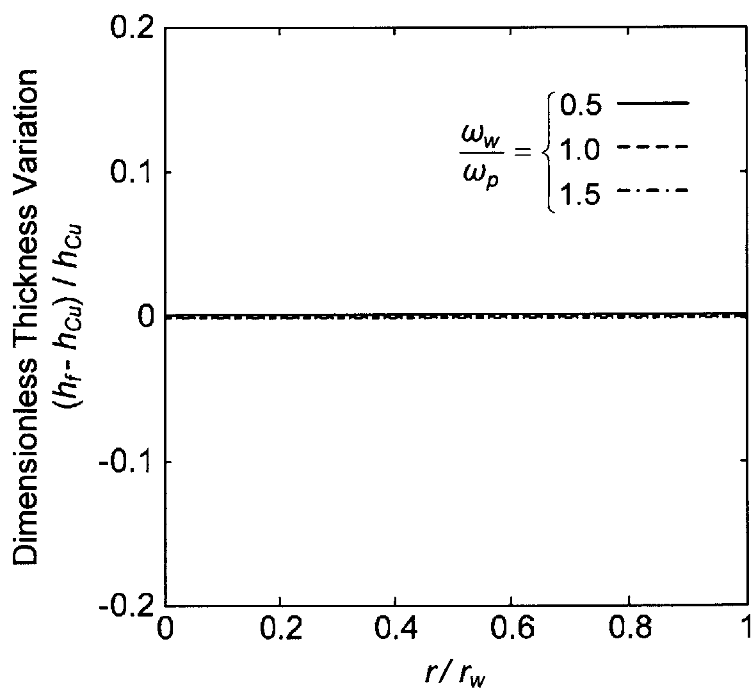


Figure A.13 Final Cu profile for the moving pad scheme.

$$\Delta V = \int_{r_1}^{r_2} dV(r) = \int_{r_1}^{r_2} k_p p s(r) \cdot 2\pi r dr \cdot N = \int_{r_1}^{r_2} k_p p s(r) \omega_w r dr \cdot \Delta t \qquad \text{(A.60)}$$

However, it is more convenient to use the pad coordinate to evaluate the total loss of material as:

$$\Delta V = \int_0^{\Delta t} \int_0^{r_p} \int_{-\pi}^{\pi} k_p p v_R r' d\theta' dr' dt \qquad \text{(A.61)}$$

Assuming constant angular velocities and substituting the relative velocity in pad coordinates,

$$\Delta V = k_p p \Delta t \int_0^{r_p} \int_{-\pi}^{\pi} \sqrt{\left[(\omega_w - \omega_p) r' \sin\theta' + v_{cc}\right]^2 + \left[(\omega_w - \omega_p) r' \cos\theta' + \omega_w r_{cc}\right]^2} r' d\theta' dr' \qquad \text{(A.62)}$$

If $\omega_w = \omega_p = \omega$, the relative velocity becomes constant, thus,

$$\Delta V = k_p p v_R \Delta t \int_0^{r_p} \int_{-\pi}^{\pi} r' d\theta' dr' = k_p p v_R \Delta t A_p \qquad \text{(A.63)}$$

where A_p is the area of the pad. Therefore, the Preston constant can be expressed as

$$k_p = \frac{\Delta V}{p v_R \Delta t A_p} = \frac{\Delta m}{\rho p v_R \Delta t A_p} \qquad \text{(A.64)}$$

where Δm is the mass loss of the wafer during the polishing time Δt and ρ is the density of the material being polished.

A.2.5 Friction Force and Torque

The frictional force on the wafer at a point P, $d\boldsymbol{F}_w$, is oppositely directed to the relative velocity, $\boldsymbol{v}_R$, as shown in Fig. A.14. The direction of $\boldsymbol{v}_R$ can be expressed by

$$\frac{\boldsymbol{v}_R}{v_R} \equiv \cos\theta_R \boldsymbol{e}_x + \sin\theta_R \boldsymbol{e}_y \qquad \text{(A.65)}$$

where

$$\begin{aligned} \cos\theta_R &= \frac{-\left[(\omega_w - \omega_p) r \sin\theta + v_{cc}\right]}{\sqrt{\left[(\omega_w - \omega_p) r \sin\theta + v_{cc}\right]^2 + \left[(\omega_w - \omega_p) r \cos\theta + \omega_p r_{cc}\right]^2}} \\ \sin\theta_R &= \frac{(\omega_w - \omega_p) r \cos\theta + \omega_p r_{cc}}{\sqrt{\left[(\omega_w - \omega_p) r \sin\theta + v_{cc}\right]^2 + \left[(\omega_w - \omega_p) r \cos\theta + \omega_p r_{cc}\right]^2}} \end{aligned} \qquad \text{(A.66)}$$

If a friction coefficient is μ and a nominal pressure p, the amount of frictional force on the differential area $dA = rd\theta dr$ on the wafer is expressed as:

$$dF_w \equiv |d\boldsymbol{F}_w| = \mu p dA = \mu p r d\theta dr \qquad \text{(A.67)}$$

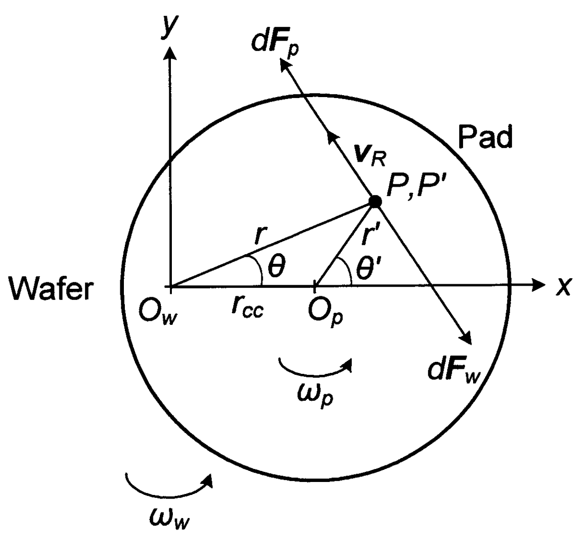


Figure A.14 Coordinate systems for force and torque calculation in the face-up CMP configuration.

$$\begin{aligned} d\boldsymbol{F}_w &= -\mu p r d\theta dr\left(\cos\theta_R \boldsymbol{e}_x + \sin\theta_R \boldsymbol{e}_y\right) \\ &= dF_{w_x}\boldsymbol{e}_x + dF_{w_y}\boldsymbol{e}_y \end{aligned} \tag{A.68}$$

where,

$$\begin{aligned} dF_{w_x} &= \frac{\mu p\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right] r d\theta dr}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} \\ dF_{w_y} &= \frac{-\mu p\left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right] r d\theta dr}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} \end{aligned} \tag{A.69}$$

At a given time t, the pad and the wafer contacts in the region $-\theta_c \le \theta \le \theta_c$. Here θ_c is dependent on the radius r and the center to center distance r_{cc}, which may be time-dependent.

$$\boldsymbol{F}_w = F_{w_x}\boldsymbol{e}_x + F_{w_y}\boldsymbol{e}_y \tag{A.70}$$

$$\begin{aligned} F_{w_x} &= \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{\mu p\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right] r}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} d\theta dr \\ F_{w_y} &= \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{-\mu p\left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right] r}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} d\theta dr \end{aligned} \tag{A.71}$$

Since $\boldsymbol{F}_p = -\boldsymbol{F}_w$,

$$\boldsymbol{F}_p = F_{p_x}\boldsymbol{e}_x + F_{p_y}\boldsymbol{e}_y \tag{A.72}$$

$$\begin{aligned} F_{p_x} &= \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{-\mu p\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right] r}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} d\theta dr \\ F_{p_y} &= \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{\mu p\left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right] r}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} d\theta dr \end{aligned} \tag{A.73}$$

The frictional torque about the center of the wafer O_w, $\boldsymbol{T}_w$, is defined as:

$$d\boldsymbol{T}_w = \boldsymbol{r} \times d\boldsymbol{F}_w \tag{A.74}$$

$$d\boldsymbol{T}_w = \left(r\cos\theta dF_{w_y} - r\sin\theta dF_{w_x}\right)\boldsymbol{e}_z \tag{A.75}$$

$$dT_w = \frac{-\mu p\left[\left(\omega_w - \omega_p\right) r^3 + \omega_p r_{cc} r^2\cos\theta + v_{cc} r^2 \sin\theta\right]}{\sqrt{\left[\left(\omega_w - \omega_p\right) r\sin\theta + v_{cc}\right]^2 + \left[\left(\omega_w - \omega_p\right) r\cos\theta + \omega_p r_{cc}\right]^2}} d\theta dr \tag{A.76}$$

Thus,

$$T_w = \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{-\mu p\left[\left(\omega_w-\omega_p\right)r^3+\omega_p r_{cc} r^2\cos\theta+v_{cc}r^2\sin\theta\right]}{\sqrt{\left[\left(\omega_w-\omega_p\right)r\sin\theta+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r\cos\theta+\omega_p r_{cc}\right]^2}}\,d\theta dr \tag{A.77}$$

The frictional torque about the center of the pad O_p is defined as $\boldsymbol{T}_p$.

$$d\boldsymbol{T}_p = \boldsymbol{r}'\times d\boldsymbol{F}_p \tag{A.78}$$

$$d\boldsymbol{T}_p = \left(r'\cos\theta' dF_{p_y} - r'\sin\theta' dF_{p_x}\right)\boldsymbol{e}_z \tag{A.79}$$

By using relationship between $\boldsymbol{T}_w$ and $\boldsymbol{T}_p$,

$$dT_p = \frac{\mu p\left[\left(\omega_w-\omega_p\right)r\cos\theta+\omega_p r_{cc}\right]r r_{cc}}{\sqrt{\left[\left(\omega_w-\omega_p\right)r\sin\theta+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r\cos\theta+\omega_p r_{cc}\right]^2}}\,d\theta dr - dT_w \tag{A.80}$$

Thus,

$$T_p = \int_0^{r_w}\int_{-\theta_c}^{\theta_c} \frac{\mu p\left[\left(\omega_w-\omega_p\right)r\cos\theta+\omega_p r_{cc}\right]r r_{cc}}{\sqrt{\left[\left(\omega_w-\omega_p\right)r\sin\theta+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r\cos\theta+\omega_p r_{cc}\right]^2}}\,d\theta dr - T_w \tag{A.81}$$

If the wafer fully covers the pad, the range of integration will be simplified to $-\pi \le \theta' \le \pi$ and $0 \le r' \le r_p$. Therefore, the frictional force is

$$\begin{aligned}
F_{w_x} &= \int_0^{r_p}\int_{-\pi}^{\pi} \frac{\mu p\left[\left(\omega_w-\omega_p\right)r'\sin\theta'+v_{cc}\right]r'}{\sqrt{\left[\left(\omega_w-\omega_p\right)r'\sin\theta'+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]^2}}\,d\theta' dr' \\
F_{w_y} &= \int_0^{r_p}\int_{-\pi}^{\pi} \frac{-\mu p\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]r'}{\sqrt{\left[\left(\omega_w-\omega_p\right)r'\sin\theta'+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]^2}}\,d\theta' dr'
\end{aligned} \tag{A.82}$$

and,

$$T_w = \int_0^{r_p}\int_{-\pi}^{\pi} \frac{-\mu p\left[\left(\omega_w-\omega_p\right)r'^2+\omega_p r_{cc}\left(r_{cc}+r'\cos\theta'\right)+v_{cc}r'\sin\theta'\right]r'}{\sqrt{\left[\left(\omega_w-\omega_p\right)r'\sin\theta'+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]^2}}\,d\theta' dr' \tag{A.83}$$

$$T_p = \int_0^{r_p}\int_{-\pi}^{\pi} \frac{\mu p\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]r' r_{cc}}{\sqrt{\left[\left(\omega_w-\omega_p\right)r'\sin\theta'+v_{cc}\right]^2+\left[\left(\omega_w-\omega_p\right)r'\cos\theta'+\omega_w r_{cc}\right]^2}}\,d\theta' dr' - T_w \tag{A.84}$$

When the pad and the wafer rotates at the same speed, $\omega_w = \omega_p = \omega$:

$$\begin{aligned}
F_{w_x} &= \mu p\left(\pi r_p^2\right)\frac{v_{cc}}{\sqrt{v_{cc}^2+\left(\omega r_{cc}\right)^2}} \\
F_{w_y} &= -\mu p\left(\pi r_p^2\right)\frac{\omega r_{cc}}{\sqrt{v_{cc}^2+\left(\omega r_{cc}\right)^2}}
\end{aligned} \tag{A.85}$$

$$
\begin{aligned}
T_w &= \mu p\left(\pi r_p^2\right) r_{cc} \frac{\omega r_{cc}}{\sqrt{{v_{cc}}^2 + (\omega r_{cc})^2}} \\
T_p &= 0
\end{aligned}
\tag{A.86}
$$

Additionally, if $v_{cc} = 0$,

$$
\begin{aligned}
F_{w_x} &= 0 \\
F_{w_y} &= \mu p\left(\pi r_p^2\right)
\end{aligned}
\tag{A.87}
$$

$$
\begin{aligned}
T_w &= \mu p\left(\pi r_p^2\right) r_{cc} \\
T_p &= 0
\end{aligned}
\tag{A.88}
$$

Figures A.15 and A.16 show the frictional force and torques, respectively, in the face-up configuration for various wafer/pad rotational speed ratios.

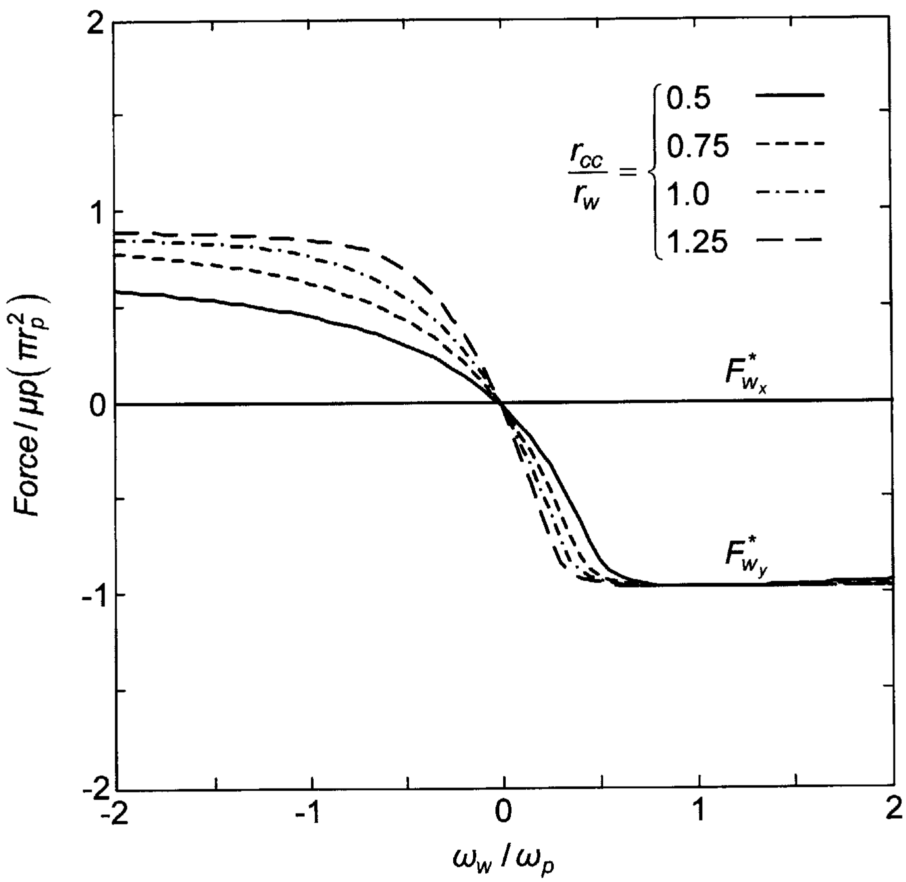


Figure A.15 Normalized friction force versus wafer/pad rotational speed ratio for various pad locations in face-up CMP.

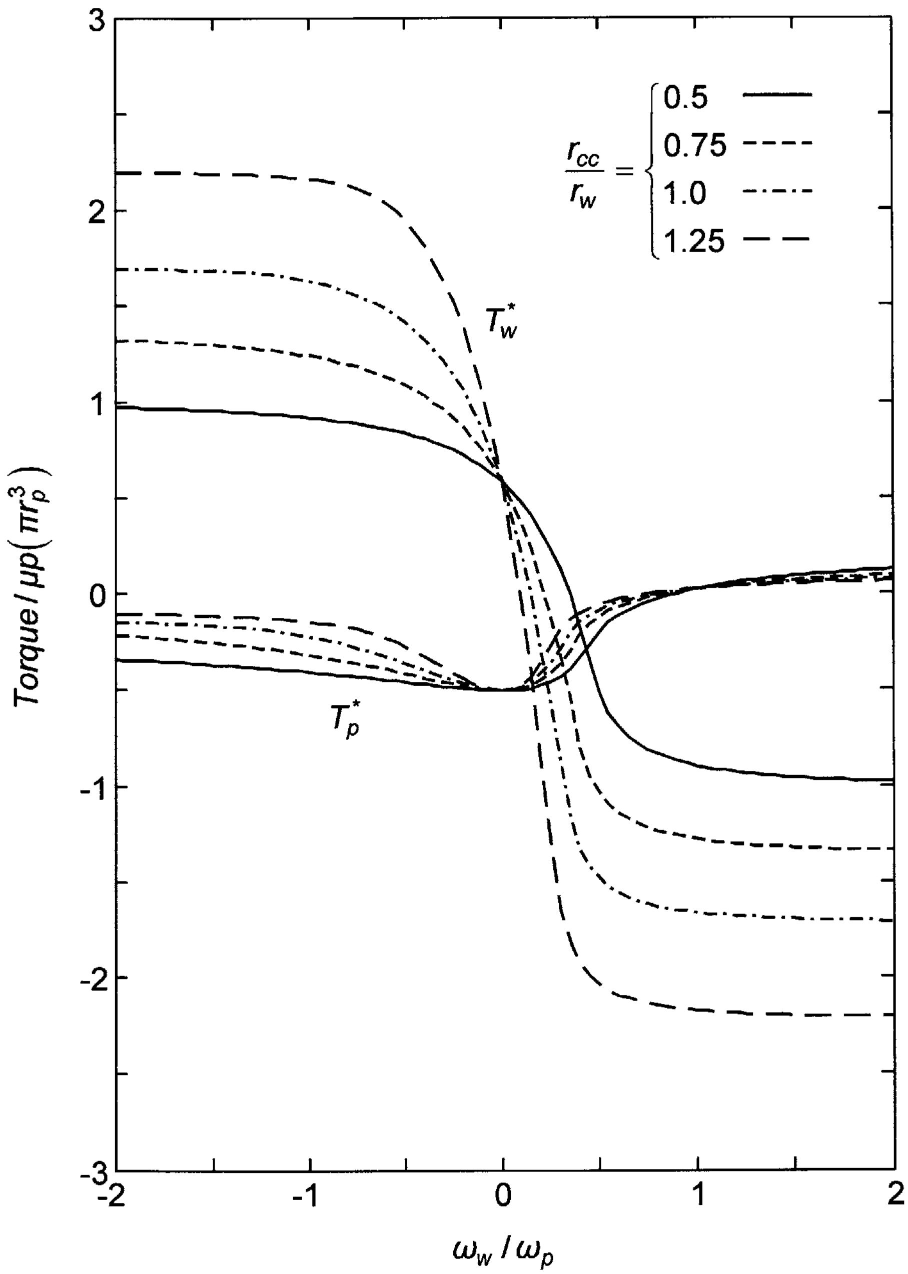


Figure A.16 Normalized wafer and pad torques versus the wafer/pad rotational speed ratio for various pad locations in face-up CMP.

Nomenclature

A_w, A_p = areas of the wafer and the pad (m^2)
e_x, e_y, e_z = unit vector in the x-, y- and z-direction
F = applied normal force (N)
$\boldsymbol{F}_w, \boldsymbol{F}_p$ = frictional forces applied on the wafer and the pad (N)
F_{p_x}, F_{p_y} = the x- and y-direction frictional force components on the pad (N)
F_{w_x}, F_{w_y} = the x- and y-direction frictional force components on the wafer (N)
h = thickness of the film (m)
k_p = Preston constant (m^2/N)
N = number of wafer rotations during polishing time t_p
n = index of wafer rotations
O_w, O_p = center of the wafer and the pad
P, P' = point in the wafer and the pad
p = nominal pressure (N/m^2)
r, θ = stationary polar coordinates of the wafer
r', θ' = rotating polar coordinates of the pad
r_{cc} = distance between the centers of the wafer and the pad (m)
r_w, r_p = radii of the wafer and the pad (m)
T_w, T_p = torques on the wafer and the pad (Nm)
t = polishing time (s)
t_{ep}, t_{ep}^* = polishing end-point (s) and dimensionless end-point
v_{cc} = rate of change of r_{cc} (m/s)
$\boldsymbol{v}_R$ = relative velocity of the wafer with respect to the pad (m/s)
x, y, z = Cartesian coordinates
ΔH = material removed during polishing time t_{ep} (m)
Δh, Δh^* = material removed during one wafer rotation (m) and normalized value
μ = friction coefficient
θ_c = pad contact angle (rad)
θ_R = angle of $\boldsymbol{v}_R$ with respect to the x-axis (rad)
ω_w, ω_p = angular velocities of the wafer and the pad (rad/s)

References

Hoshino, S., Uda, Y. and EYamamoto, E.,"Characteristics of Copper CMP Process for Ultra-Low Pressure", *CMP-MIC 2004*, pp. 17-22, February 24-26, 2004, Marina Del Rey, CA.

Lai, J.-Y., 2001, "Mechanics, Mechanisms, and Modeling of the Chemical Mechanical Polishing Process," Ph.D. thesis, Department of Mechanical Engineering, MIT, Cambridge, MA.

Oliver, M. R. (editor), 2004, **Chemical-Mechanical Planarization of Semiconductor Materials,** Springer Series in Materials Science, Berlin.

Preston, F. W., 1927, "The Theory and Design of Plate Glass Polishing Machines," *J. Soc Glass Technology*, vol. 11, pp. 214-256.

Senga, T., Shinjo, K., Ueda, Y., Nakahiro, H., Ushio, Y. and Une, A., "Profile Control by the Smart Simulator in Partial Polishing CMP", *CMP-MIC 2003*, pp. 200-205, February 19-21, 2003, Marina Del Rey, CA.

Yashio, A., 2000, "CMP Technology for ULSI Interconnection," *SEMICON West*, pp. K1-K10, San Franciso, CA.