MODELING PAD WEAR AND PAD CONDITIONING IN CHEMICAL-MECHANICAL POLISHING USING POPULATION BALANCE MODEL

by

Hong Shi

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemical Engineering

The University of Utah

December 2012

المنارة للاستشارات

www.manaraa.com

UMI Number: 3543361

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3543361

Published by ProQuest LLC (2012). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

المنارة للاستشارات

www.manaraa.com

Copyright © Hong Shi 2012

All Rights Reserved

المنارة للاستشارات

www.manaraa.com

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Hong Shi**

has been approved by the following supervisory committee members:

Name	Role	Date Approved
Terry A. Ring	, Chair	**May 10, 2012**
Kenneth L. Devries	, Member	**May 10, 2012**
Raj Rajamani	, Member	**May 10, 2012**
Ian Harvey	, Member	**May 10, 2012**
Edward Trujillo	, Member	**May 10, 2012**

and by **JoAnn S. Lighty** , Chair of

the Department of **Chemical Engineering**

and by Charles A. Wight, Dean of The Graduate School.

المنارة للاستشارات

www.manaraa.com

ABSTRACT

Chemical mechanical planarization (CMP) has become the leading planarization technology in integrated circuit manufacturing. The CMP process is highly dependent on the surface condition of the pad. The height distribution of the pad asperities (PDF) together with the process conditions determines the chemical-mechanical removal rate and within wafer nonuniformity (WIWNU) as well as defects in the form of scratches. Understanding the asperity height distribution during polishing and conditioning is important in controlling many aspects of CMP and forms the basis of this dissertation.

This dissertation attacks the problem of predicting the pad asperity PDF by considering the effects of pad asperity wear due to contact with the wafer and the generation of asperities during pad conditioning. In CMP, the aggressiveness (pad cut-rate) of the conditioner should overcome the CMP pad wear-rate during the *in situ* conditioning in order to sustain a stable polish-rate. In this work, a population balance model (PBE) for pad asperities is used to model the pad wear and pad conditioning processes. Fluid mechanics is built into the pad wear model. Pad plastic deformation and variable diamond height for the conditioner are built into the conditioner model that have not been done in previous works. The results with variable diamond height show that the conditioned pad asperity PDF is Gaussian. The variance of the conditioned pad asperity PDF is decided by the variance of the cutting diamonds (active diamonds). The model results also show that a different interface gap between the conditioner and pad will give

المنارة للاستشارات

www.manaraa.com

a different standard deviation of the resulting conditioned pad asperity PDF if the conditioner diamond height PDF is Gaussian. This is consistent with the experimental findings that with increasing load, the roughness of the conditioned pad asperity PDF increases. Model results in this work show that it is caused by the decreasing of the interface gap between the rough surface of the conditioner diamonds and the pad surface resulting in the change of the roughness of the active diamonds. Analytical solutions are also derived and found matching the Monte Carlo numerical results.

المنارة للاستشارات
www.manaraa.com

This thesis is dedicated to my parents, my husband and my daughters.

المنارة للاستشارات

www.manaraa.com

TABLE OF CONTENTS

ABSTRACT....iii

LIST OF FIGURES....viii

LIST OF TABLES....xii

ACKNOWLEDGMENTS....xiii

Chapters

1 INTRODUCTION....1

Semiconductor technology and device scaling....2
Interconnect delay with device scaling....4
Methods to reduce interconnect delay....6
The need for global planarization and the importance of CMP....8
The objective and structure of the thesis....12
References....20

2 BACKGROUND....31

Elastic contact model between wafer and pad....32
Mixed-lubrication approach....34
Wear of the pad....38
Wafer material removal rate....40
The asperity height population balance....41
References....45

3 ANALYTICAL SOLUTION FOR POLISH-RATE DECAY IN CHEMICAL-MECHANICAL POLISHING....49

Abstract....50
Introduction....50
The asperity height population balance....51
Wafer material removal rate (MRR)....52
Analytical solution of the asperity height population balance....52
Model results and comparison with the experiments....52
Conclusion....54

المنارة للاستشارات
www.manaraa.com

References……………………………………………………………………..54

4 CMP PAD WEAR AND POLISH-RATE DECAY MODELED BY ASPERITY POPULATION BALANCE WITH FLUID……………………………………..55

Abstract……………………………………………………………………..56
Introduction………………………………………………………..…………..56
Elastic contact model for CMP polish-rate decay……………..………...........58
Model result and comparison with the experiment……………………………....59
Conclusion…………………………………………………….…………...61
Appendix A………………………………………………...……………….62
References…………………………………………...………………...........63

5 PLASTIC DEFORMATION THEORY OF CHEMICAL-MECHANICAL POLISHING PAD CONDITIONING……………………………………….64

Abstract………………………………..……………………………………65
Introduction……………………………………………………………………..66
Model…………………………………………...……………………...69
Numerical model results and comparison with analytical solutions......................79
Conclusion…………………………………..…………………………...93
References…………………………..………………………………………94

6 A CONDITIONER MODEL WITH VARIABLE DIAMOND HEIGHTS……122

Abstract……………………………………………………………………....123
Introduction……………………………………………………………………..124
Model…………………………………………………………………..…….128
Results…………………………………………………………………………..134
Conclusion……………………………………..………………………….145
References…………………………………………………………………....147

7 CONCLUSION AND FUTURE WORK…………………………………….170

Conclusion……………………………………………..………………171
Future work…………………………………………………..……...........177
References………………………………………….….……………………..178

APPENDIX: SENSITIVITY TEST FOR FITTING PARAMETERS, C_a AND C_w…...179

www.manaraa.com

LIST OF FIGURES

Figure Page

1.1 A schematic of a modern IC……………………….…………………………………..25

1.2 Schematic of the CMP process…………………………………….……………….26

1.3 Schematics of the two options to fabricate interconnects…………………………27

1.4 Macro- and micrographs of a conventional diamond conditioner and the height distribution of the diamond grits on the conditioner…..........................………..……28

2.1 Control volume for population balance……………….………………………….48

3.1 PDF of asperity heights constructed from data in [5]……….…………….…………...53

3.2 Evolution of the tail of the initial PDF at 5 min intervals…………...……….……..53

3.3 MRR decay model in this work with our analytical solution versus Stein's experimental result……………………………...………………..……….........54

4.1 CMP process description……..........……………………………………………....57

4.2. PDF of asperity heights constructed from data in Stein….…....…..……………..60

4.3 Removed thickness results calculated with the present removal-rate model versus Stein's experimental results (Lot A) at successive 5-min intervals…..…………..60

4.4 Evolution of the tail of the PDF in Figure 4.2 at 5-min intervals using Equation (4.17) at Stein's experimental condition…………………………………………...….…...61

4.5 The decreasing of separation d with time due to wear at Stein's experimental condition…………….……………………………………………………………...62

4.6 The changing of pad asperity pressure and fluid pressure with time due to wear at Stein's experimental condition…………………………...………………………....62

4.7 The changing of the pad contact area fraction with time due to wear at Stein's

www.manaraa.com

experimental condition……………..……..……………...………………….…...62

5.1 Schematic illustration of CMP process………...…………………………………96

5.2 Schematic stress-strain curves of two categories of polymers…………………..….97

5.3 One-pass single-track scratch and two-pass intersecting scratch on two types of polymers……………………………………………………………………………98

5.4 Cross-section of a created groove………….………………………………………99

5.5 Illustration of a CMP tool with the circular conditioner and the equivalent bar conditioner ……………………………………………………………………..….100

5.6 The coordinate system and terminology used in [8] and in this work…………...102

5.7 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process by Equation (5.4)………………………………103

5.8 Development of the CCDF q(z,t) corresponding to the simulation in Figure 5.7…104

5.9 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process by Equation (5.7)…………………………………..105

5.10 Development of the CCDF q(z,t) corresponding to the simulation in Figure 5.9 …………………………………………………………………………………107

5.11 Process shows the oscillating tail for f_{add} = 1 case with constant height diamond conditioning corresponding to the simulation in Figure 5.9………………………108

5.12 Comparison of the CCDF produced by Monte Carlo simulation with the analytical solutions for constant cutting height……………………………….....110

5.13 Monte Carlo simulations for the changing of the average surface height (mean) $\bar{s}(t)$ and cutting depth h(t) for different chosen f_{add} values…………………….….111

5.14 The surface profile s(x,t) after 50 time steps for different f_{add} values…,…………113

5.15 The comparison of the Monte Carlo simulation results for the variation of the cut-rate c with f_{add} with the analytical solution, Equation (5.31)……………………...115

5.16 The surface profile s(x,t) for constant cut-rate case with h_0 = -5 μm and f_{add} = 0...116

5.17 Snap shots of the surface height CCDF q(z,t) with h_0 = -5 μm and f_{add} = 0………117

5.18 The comparison of analytical solution of Equation (5.34) with numerical solution

المنارة للاستشارات
www.manaraa.com

for CCDF and analytical solution of PDF with h_0 = -5 μm and f_{add} = 0..........118

5.19 The surface profile s(x,t) for constant cut-rate case with h_0 = - 2.5 μm, single diamond period ℓ = 80 μm and f_{add} = 0..119

5.20 Snapshots of the surface height CCDF with h_0 = -2.5 μm and f_{add} = 0..............120

5.21 The comparison of the analytical solution of Equation (5.34) with numerical solution for CCDF and analytical solution of PDF with h_0 = -5 μm and f_{add} = 0 121

6.1 Macro- and micrographs of a conventional diamond conditioner and the height distribution of the diamond grits on the conditioner..................................149

6.2 Illustration of a CMP tool with the circular conditioner and the bar conditioner..150

6.3 The coordinate system and terminology used in this work............................152

6.4 Steady-state CCDF with an exponential diamond cutting depth distribution and the corresponding PDF...153

6.5 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process for n = 1, 2, 3, 4, 250 time steps with initially flat surface and variable diamond cutting depth..155

6.6 Monte Carlo simulations for the changing of the average surface height (mean) $\bar{s}(t)$ and cutting depth h(t) for different chosen f_{add} values...............................157

6.7 The surface profile s(x,t) after 50 time steps for different f_{add} values................159

6.8 Analytical solutions for CCDF and PDF, Equation (6.12) and Equation (6.13) respectively, with single diamond cutting depth, h_0 = -5 μm 161

6.9 Monte Carlo simulations for the changing of the average surface height (mean), $\bar{s}(t)$ and cutting depth h(t) for the simulation in Figure 6.5..............................162

6.10 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process for n = 200, 400, and 600 time steps (pad rotations) with initially flat surface and variable diamond cutting dept........................163

6.11 Monte Carlo simulations for the changing of the average surface height (mean), $\bar{s}(t)$ and cutting depth h(t) for the simulation in Figure 6.10.........................164

6.12 Steady-state CCDF with an exponential diamond cutting depth distribution for the simulation in Figure 6.4...165

6.13 The changing of the standard deviation of the conditioned pad asperity PDF with d

المنارة للاستشارات
www.manaraa.com

(separation distance between conditioner diamond surface and pad) and σ (standard deviation of the diamond height PDF with a Gaussian distribution) by using Equation (6.28) and (6.29)… …………………………………………………..166

6.14 The measured conditioned pad asperity PDF at two different loads for 100 grit conditioner and the best fits of the data by Gaussian functions with different standard deviation…………………………………………………………...167

6.15 The measured standard deviations of the conditioned pad asperity PDF as a function of the load on the conditioner…………………………………………169

A.1 Removed thickness results calculated with different c_a values for the case without fluid……………………………………………………………………….181

A.2 Removed thickness results calculated with different c_a values for the case with fluid…………………………………………………………………………182

A.3 Removed thickness results calculated with different c_w values for the case without fluid…………………………………………………………………………183

A.4 Removed thickness results calculated with different c_w values for the case with fluid…..……………………………………………………………………..184

المنارة للاستشارات

www.manaraa.com

LIST OF TABLES

Table

1.1 Properties of interconnect metals…………….……………………………...……...29

1.2 Interconnect technology requirements according to ITRS roadmap…..………...….29

1.3 Dielectric constants of available low-k dielectric materials ……………………….30

للاستشارات المنارة

www.manaraa.com

ACKNOWLEDGEMENTS

First, I wish to express my deepest gratitude to my advisor, Dr. Terry A. Ring, for his tremendous help through the course of this study. His solid scientific background, rich personal experience both in academia and industry, unique teaching style, smart brain, professional altitude, and kind character have impressed me deeply. He always encourages me to move forward, stimulates my interests in this research and helps me when I need help. He teaches me the way to solve the problem and always tries to find the best way to help me to make sure I completely understand. Without his help, this dissertation would never have been produced.

I want to thank David Stein who provided me the linear speed for Cybeq 3900 wafer polisher used in his experiments. I used that speed in my calculation.

I also want to thank my committee members, Drs. K. Larry DeVries, Raj Rajamani, Edward Trujillo, Ian Harvey, for their advice and encouragement.

I thank the faculty and staff in the Chemical Engineering Department, especially Dr. Eric Eddings and Dr. Joann Lighty, for the encouragement during my graduate study here. They indeed care for students and help them succeed in their career.

I extend special thanks to my parents, my husband and my two lovely daughters, Anna and Angie, who bring joy to my life. Their love and support encouraged me to pass the most difficult times and gave me energy to finish my dissertation.

المنارة للاستشارات
www.manaraa.com

CHAPTER 1

INTRODUCTION

المنارة للاستشارات

www.manaraa.com

This thesis addresses the study of the mechanical interactions during chemical mechanical polishing (CMP) and develops theories that can guide CMP process control and tool design. Models of pad wear and pad conditioning for CMP processes are developed. This introduction chapter will give the reader a general overview of the thesis, the motivation and objectives of the thesis and the structure of the thesis.

Semiconductor technology and device scaling

The demand for increasing the computer speed, functionality and lowering the fabrication cost has driven the semiconductor industry to minimize device dimensions. Reducing device dimensions can allow more devices to be packed per unit area on a chip. The electronic industry has grown rapidly in the past three decades. Ultra-large-scale integrated circuits (ULSI) with 10^{11} or more devices on a chip, can now be fabricated on semiconductor substrates or wafers, to reduce cost and to increase device performance [1]. An integrated circuit consists of a large number of components (transistors, resistors, capacitors, inductors, etc.), fabricated side by side (known as front-end) and wired together (known as back-end) to perform a particular circuit function.

Figure 1.1 shows a schematic of a modern IC [2]. It shows the cross-section of an IC and the devices at the transistor level fabricated in the silicon substrate at the bottom of the IC. The source and drain regions of the transistor are the active conducting regions fabricated by doping the silicon substrate. When a voltage is applied on the gate, current will flow from the source to the drain completing the circuit. Hence the applied voltage will close or open the circuit corresponding to 0 or 1 state. It also shows a resistor fabricated side by side with the transistor. The resistor is a low-doped region and when a current is applied between its two contacts, it functions as a resistor. As shown in the

المنارة للاستشارات
www.manaraa.com

figure, the devices are built in the silicon substrate at the bottom of the IC and all of the above layers of metals and interlayer dielectrics (ILD) are connections (like wires). ILD are used to isolate tungsten plugs and copper interconnects. Tungsten plugs are used to connect the bottom level devices to the first level copper interconnects. One device is connected to the nearby devices by the first level copper and three devices are connected to other devices somewhere on the chip via the first level copper up through the eighth level copper. Figure 1.1 shows only one cross-section and there are many devices connected like this for the whole chip. Hence the chip is composed of many complex circuits.

An IC is manufactured by the patterned diffusion of trace elements into the surface of a thin substrate of semiconductor material. Additional materials are deposited and patterned to form interconnections between semiconductor devices. As shown in Figure 1.1, devices are made at the bottom layer inside silicon substrate first and then metal wires connecting these devices are made by multilayer copper interconnects at the top. Hence the IC chips are manufactured by the sequential layering process from "bottom up". Each layer depends on the previous layer. The copper interconnects are embedded in the insulating ILD at each layer. With successive layers of copper and insulating ILD, an eight-layer copper interconnect is formed. Copper interconnect lines connect different devices on a chip among themselves and to the outer world.

The semiconductor industry uses minimum feature size (f) or technology node (λ) to represent the standard of the transistor size. The minimum feature size is the distance between the source and drain of the transistor and the technology node $\lambda = f/2$. Since the beginning of the microelectronic era, the minimum feature size of an integrated circuit

المنارة للاستشارات
www.manaraa.com

has been reduced at a rate of about 13% per year [3]. Device miniaturization results in reduced unit cost per circuit function. For example, the cost per bit of memory chips has halved every two years for successive generations of DRAM circuits [4]. The timeline in the reduction of the minimum feature size is expressed as Moore's law, in which the number of transistors per square inch doubles approximately every two years at half the price. Computational power also increases a factor of two every two years. The current computation power is billions of instructions per second [4]. As device dimensions decrease, the intrinsic switching time decreases and device speed increases. Also as devices become smaller, they consume less power [4]. According to the Semiconductor Industry's Association's (SIA) ITRS, the present (2011) technology node of 40 nm will shrink to 36 nm and 32 nm by 2012 and 2013 respectively [5].

Interconnect delay with device scaling

The total delay of a device in processing a signal is composed of intrinsic gate delay (device delay) and interconnect delay. Interconnects are the metal wires that connect different devices on a chip among themselves and to the outer world. With the reduction of the device dimensions on a chip, the dimensions of the metal wires will also have to decrease. Device delay is the time to switch the transistor on or off when a voltage is applied on the gate [6]. It is a function of the feature size (f) and decreases with decrease in f. A signal propagating through the interconnection also experiences interconnect delay. Although decrease in λ can reduce gate delay and improve device performance, below 0.5 μm feature sizes, interconnect delay or RC (resistance-capacitance) delay becomes dominant and grows exponentially with a decrease in dimensions [7]. While continuous shrinking makes transistors faster, it makes interconnections between transistors work

المنارة للاستشارات

www.manaraa.com

slower and reducing interconnect delay is a key issue to allow the continuous shrinking of the IC [8].

The 'RC' product is used to measure the interconnect delay, where R is the resistance of the metal interconnects and C is the capacitance of the ILD. The "RC" product has the unit of time and it is the time to charge the capacitor through the resistor for the interconnect line. RC delay hinders the further increasing of speed in microelectronic integrated circuits. When the feature size becomes smaller and smaller to increase the clock speed, the RC delay plays an increasingly important role. The RC time delay associated with a metal line is given by Equation (1.1) [9] and the cut-off frequency of the line is related to RC delay by Equation (1.2) [10]. The clock frequency can not go above the cut-off frequency of the line.

$$RC = \frac{\rho k L^2}{TD} \tag{1.1}$$

$$f_c = \frac{1}{2\pi RC} \tag{1.2}$$

where ρ is the resistivity of the metal, k is the permittivity of the ILD, L is the length of the metal line, D is the thickness of the metal line and T is the thickness of the ILD. A signal propagating through the interconnection experiences RC delay. Shrinking the cross-section of a wire increases its resistance and bringing wires closer together increases capacitance between the wires. As a result, RC delay increases as device size decreases.

Apart from interconnect delay, device failure associated with electromigration

المنارة للاستشارات
www.manaraa.com

becomes significant as the current density increases with the minimization of the device dimensions. Electromigration is the transport of metal ions in the conductor because of the momentum transfer between electrons and metal ions [4]. It becomes significant when current density is large. Interconnects are the metal wires that connect different devices on a chip among themselves and to the outer world. With the reduction of the device dimensions on a chip, the dimensions of the metal wires will also have to decrease. This increases the current density and hence the susceptibility to electromigration.

Methods to reduce interconnect delay

Copper interconnect.

In order to reduce the RC delay and electromigration to meet the requirement of the ever-shrinking feature lengths, newer methods and materials have been explored. As circuits become more complex and device density increases, more levels of metal interconnects are needed on a chip. Also from Equation (1.1), RC delay is directly proportional to the square of the length of the metal line, the resistivity of the metal line and the dielectric constant of the ILD. Hence a decrease in the metal length can help reduce RC delay. This can be achieved by multilevel metal interconnects in which devices are connected together by metal interconnects at different levels through contact vias [11]. Additional improvement can be achieved by reducing the resistivity of the metal and dielectric constant of the ILD by using low-k dielectric material. Of all the metals that have lower resistivity than aluminum, only copper has good electromigration resistance. Gold has excellent electromigration resistance compared to aluminum alloy but it has only a marginal improvement in resistivity (2.35 μΩ-cm). The resistivity of copper is 1.67 μΩ-cm, while aluminum alloy has a resistivity of 3.5 μΩ-cm. Silver has

المنارة للاستشارات
www.manaraa.com

very good resistivity (1.59 μΩ-cm) but its electromigration resistance is very poor [11] - [13]. The summary is shown in Table 1.1.

Although aluminum has been used as interconnect material since 1960s, it suffers from electromigration at high electrical current. Copper was introduced in 1993 to replace aluminum for minimum feature size approaching 100 nm [14]. Copper has been the metal of choice since then.

Low-k dielectric materials

Device performance can be further improved by integrating copper with low-k dielectric materials for the ILD. Dielectric constant k (also called relative permittivity ε_r) is the ratio of the permittivity of a substance to that of free space. A material containing polar components, which are represented as electric dipoles (e.g., polar chemical bonds), has an relatively high dielectric constant. A capacitor with a dielectric medium of higher k will hold more electric charge at the same applied voltage and its capacitance will be higher [8]. Decreasing dipole strength and the number of dipoles can reduce k. These correspond to introducing nonpolar chemical bonds or porosity into the dielectric material respectively. The two methods can be combined to obtain even lower k values. Reducing k can increase the speed of the circuit and it can also reduce the cross-talk, parasitic capacitance and power consumption of the IC. This is because every time the voltage changes across the interconnect line, extra current has to charge or discharge the capacitor. When the voltage does not change very quickly, as in low frequency circuits, the extra current is usually negligible, but when the voltage is changing quickly the extra current is large and can dominate the operation of the circuit. Integrating low-k dielectric ILD with copper is critical for the continous minimization of the ICs. According to SIA's

المنارة للاستشارات
www.manaraa.com

ITRS (2006 Update), the 40 nm technology will require k between 2.1 and 2.4 for ILD with copper interconnect level of 12 as shown in Table 1.2 [5].

The first production example of the integration of copper with a true low-k dielectric material was in 2000, with the incorporation of the SiLK™ organic spin-on dielectric from Dow Chemical in an IBM 130 nm device. The 90 nm technology is the first node where true low-k material (k = 2.7) is required along with copper [15]. The choices for the low-k materials are between spin-on and CVD materials or between organic and inorganic materials. Spin-on materials can incorporate porosity easily and CVD materials can allow the usage of the current toolsets. Spin-on materials can be organic, inorganic or hybrid materials. Polymeric materials are generally deposited by spin-on approach and SiLK™ from Dow Chemical is an example in this category and used most widely in current ICs. CVD materials are usually carbon or fluorine – doped SiO_2 in which carbon or fluorine is introduced into SiO_2 matrix to lower the electric polarization and hence lower k [15]. Current low-k materials available are summarized in Table 1.3 [15] - [18]. Integrating copper with low-k dielectrics can effectively reduce RC delay and also reduce the number of the interconnect metal layers needed [19].

The need for global planarization and the importance of CMP

As the desired feature sizes of semiconductor devices continue to shrink to increase the computer speed, functionality, and lower the fabrication cost, the ability to globally planarize the wafer surface becomes increasingly important. Extremely high degree uniformity and repeatability are needed for high production yield. The requirement for the depth of focus becomes more stringent as the density of the devices on the wafer increases and more interconnect layers are stacked. Process quality becomes even stricter

www.manaraa.com

because of the accumulation of the nonuniformity with increasing interconnect levels. CMP which is a process of smoothing or leveling an uneven topographic surface using chemical and mechanical forces, has evolved as the only planarization technique in ICs and micro-electro-mechanical systems (MEMs) fabrication [20] - [23]. Figure 1.2 shows a schematic of a CMP process. In CMP, rotating wafers are immersed in an abrasive-containing slurry and pressed against a polymeric rotating CMP pad. Using gentle pressure and rotation, the pad/slurry system removes imperfections from the wafer and planarizes it before subsequent layers of circuitry are applied. The nano-sized abrasives in the slurry held by pad asperities perform the mechanical polishing work. The taller areas are subjected to higher mechanical forces and hence removed faster, aiding global planarization. As the name suggests, the process involves both chemical and mechanical effects, e.g., mechanical abrasion of the surface followed by chemical removal of the abraded material [24]. In general, the chemical effect contributes more to the selectivity and productivity and the mechanical effect contributes more to the surface planarity and surface roughness. The conditioner is a rotating disk whose surface contains embedded diamonds used to cut the pad surface. Because the pad asperities will be worn out as the polishing continues, the conditioner is used to regenerate the pad asperities to enable the continuous polishing with stable and controlled polishing rate and uniformity. CMP is used whenever global planarization is required.

There are two methods to fabricate interconnects as shown in Figure 1.3 [19] and consequently there are two applications of CMP in ULSI (Ultra-large-scale integrated circuits) fabrication: to smooth ILD (usually SO_2) or to remove excess material to produce inlaid metal structure or isolation trenches [25] - [27]. In 1989, CMP was

المنارة للاستشارات
www.manaraa.com

developed by Davari et al. for global planarization of ILD which is the key process for multilevel metallization [28]. The ILD CMP approach as shown in Figure 1.3 (a) is applied in aluminum interconnects fabrication in which aluminum is deposited on a oxide ILD layer then it is patterned and etched to form aluminum interconnects. Then another layer of ILD oxide is deposited on the aluminum interconnects and CMP is applied on the ILD layer to smooth it so that it is ready for the next photolithography step for the next layer of aluminum interconnects. The same procedure is repeated to form multilevel aluminum interconnects (three dimensional electrical wires) which connect all of the devices fabricated in the wafer substrate at the bottom to form ICs as shown in Figure 1.1.

Unlike ILD CMP, in copper interconnects, tungsten plug and shallow trench isolation (STI) processes, CMP is employed to remove the excess deposits covered on the trenches as shown in Figure 1.3 (b). An ILD layer is deposited and patterned to form trenches and then copper is deposited to fill the trenches. Then CMP is employed to remove the excess deposits and smooth the surface so that copper wires are formed. Then another layer of ILD is deposited and the process is repeated to form three dimensional copper interconnects. This additive (compared to the subtractive patterning process for traditional aluminum patterning) patterning process is called damascene process. Because copper is hard to pattern and etch, CMP is the only method for nanomachining the copper interconnects [29]. Copper interconnect was introduced in 1993 by Paraszczak et al. to replace aluminum for minimum feature size approaching 100 nm because aluminum suffers from electromigration and RC delay at submicron dimensions as discussed in the above sections [14]. Tungsten is used as an interconnect plug to the source, drain, and gates of transistors in Si chips. In chip processing, tungsten contact holes or vias are

www.manaraa.com

filled with CVD tungsten, which is then removed by CMP from planar areas leaving just the contact plugs filled with metal. The same structure can of course be produced by subtractive etch method as shown in Figure 1.3 (a) and the first implementation of tungsten plug indeed used this method. The drawback of this method is that at the end of the etch point, the etchant will attack the plug and results in significant plug loss. CMP is much better in this respect [30]. Trench isolation technology was introduced by Rung et al. in 1982 to isolate CMOS devices [31]. This approach eventually replaced all other isolation methods.

The advent of copper interconnects has indeed accelerated the need of CMP process. CMP has been the fastest growing process technology for IC manufacturing. Within three years (1995-1998), worldwide sales have more than tripled [32]. According to a report by Boston-based Linx Consulting, the global market for CMP slurries, pads, and pad-conditioning agents was worth almost $1.1 billion in 2005 [33]. The total market for CMP pads and slurry was $1.4 billion in 2008 [34]. Linx expects the market will show steady growth and expects to be worth more than $1.6 billion by 2014 [34]. Linx Consulting still sees Dow Advanced Materials and Cabot Microelectronics as the leading suppliers, with two-thirds market share for all slurry and pad consumables between them. The next tier of suppliers includes Asahi Glass, DA Nanomaterials, Fujimi, Hitachi Chemical, Planar Solutions and Samsung Cheil, who comprise an additional 25% share of the slurry and pad consumables market. Linx Consulting is the foremost electronic materials consultant to the semiconductor industry.

With the advent of integration of low-k dielectric constant material with copper, CMP becomes the leading planarization technology in the current and future IC

المنارة للاستشارات
www.manaraa.com

manufacturing [29]. Unfortunately, CMP is an experience-based process. The fundamental polishing mechanisms are still not well understood and it limits the advanced tool design and better process control. Some of the outstanding issues existing in the CMP process includes high cost, poor process controllability, limited local and global uniformity and defects [35].

Understanding the fundamentals of CMP can help solve the above mentioned issues which can lead to better tool design and more precise process control. Fundamentals include understanding the mechanical and chemical polishing mechanisms. These mechanisms are closely related to pad surface state, e.g., pad wear and pad conditioning. In this thesis, pad wear and pad conditioning processes are studied.

The objective and structure of the thesis

Chemical mechanical polishing (CMP) is a process that removes irregularity from silicon wafers in order to prepare them for the next photolithographic step, avoiding depth of focus problems during illumination of photosensitive layers. It is a process of planarizing surfaces with the combination of chemical and mechanical forces. Wet etching alone cannot attain good planarization because chemical reaction is conformal, while mechanical polishing alone causes too much surface damage. As Figure 1.2 shows, during polishing, a rotating wafer is pressed against a rotating polymeric pad that is flooded with reactive slurry. The slurry has nano particles in it which remove materials by combination of chemical and mechanical means [24]. The mechanical stress will change the solubility of the wafer surface and hence the material can be dissolved in the slurry and carried away by the flow or more easily removed by abrasive action [24], [36]. The abraded material is dissolved in the slurry and carried away by the flow. There is a

المنارة للاستشارات
www.manaraa.com

local height variation between chip areas and the taller areas are subjected to stronger forces and removed faster. Hence global planarization is achieved.

The CMP process is highly dependent on the surface condition of the pad. The height distribution of the pad asperities (PDF) together with the process conditions will determine the interface gap (d in Figure 1.2) and hence determine the chemical removal rate (the mass transport rate of the reactant and product to and from the wafer surface) [26]. The asperity height distribution also affects the applied load distribution on the wafer and hence influence the mechanical polish rate and within wafer non-uniformity (WIWNU). Understanding the asperity height distribution during polishing is important in controlling many aspects of CMP.

Many experimental observations have shown that the polish rate will drop dramatically if the pad is not conditioned [37] - [39]. In addition, some have shown a connection between pad surface asperity height distribution and polish-rate decay [38], [39]. Oliver [38] has shown that the average asperity height (roughness) is continuously reduced with polishing and so is the polish-rate. Lawing's experimental measurements [39] with interferometry of the pad surface topography have shown that the pad surface height distribution continuously changes with polishing time in that a second component peak at the high-end of the distribution continuously grows with polishing time without conditioning. These experimental observations indicate that the reduction of the polish-rate is caused by the wear of the pad asperities. For this reason, the pad needs to be regularly dressed by the conditioner. The conditioner is a diamond-embedded rotating disk dragged with down force across the pad surface in a similar manner to that of the wafer. The conditioner will create asperities on the pad surface to

المنارة للاستشارات
www.manaraa.com

enable the abrasive grit to perform their polishing action efficiently. The polishing operation will deform and flatten the pad surface and the conditioner will roughen the pad surface, so that these two operations work together to maintain a constant pad surface during polishing. A constant pad surface is important to maintain a stable polish performance.

Because the pad height distribution is dynamically changing during CMP, understanding how it changes and its effect on the polishing-rate are necessary to control CMP. To the best of our knowledge, Borucki [40] has made the first major step towards this understanding. In his model, he attributes the polishing-rate decay to the wear of the pad asperities and uses a numerical solution to a pad asperity population balance to calculate the changing of pad asperities with time. The wear rate of individual asperities is modeled by Archard's law [41]. Greenwood and Williamson's contact model with elastic Hertzian contact [42] is used to model the pad asperity and wafer contact. In this work, which is presented in Chapter 3 of this thesis, the asperity population balance equation has been solved analytically by the method of characteristics. This analytical solution is presented and used to calculate the polishing-rate decay that is presented as a numerical solution in [40]. The model results are compared with the numerical solution [40], [43] and with experimental data by Stein, et al. [37]. The model for the polishing-rate decay agrees with experimental results. The advantages of the analytical solution are that it is more accurate and numerically faster to compute. The results in Chapter 3 are published in *Journal of Engineering Mathematics* [44].

In CMP, the applied load on the wafer is supported by both pad asperities and fluid. The interface gap between the pad and wafer increases when applied load decreases

المنارة للاستشارات
www.manaraa.com

and/or the relative speed between the wafer and pad increases. The operating regime of CMP can be boundary-lubrication when asperities carry most of the load, mixed-lubrication when load partition between asperities and fluid are comparable, or hydrodynamic regime when load mostly carried by the fluid. The PDF of the pad, relative speed between the wafer and pad, and the applied load decide the interface gap and the operating regime of CMP. In boundary-lubrication regime, mechanical removal rate is important and on the contrary, in hydrodynamic regime, chemical removal rate (mass transport) dominates.

Although the model in [40] successfully captures the experimental data trend by Stein [37] by using two adjustable model parameters to match the experimental data, it has neglected the fluid effect in computing the load balance, e.g., the load is entirely carried by the pad asperities. This will overestimate the change of the interface gap, making the wear rate higher than the reality. The wear rate of the pad should be decided by the initial pad asperity PDF and particular process so the population balance equation needs to be solved simultaneously with fluid mechanics because it is done in the slurry. The applied load is balanced by the pad asperities and fluid. The fluid will reduce the load on the individual asperities resulting in less pad wear compared to the case without fluid. The wear rate of the pad is calculated from the rate of change of the pad-wafer separation distance because only asperities above this distance will be worn down and that portion will pile up at the pad-wafer separation distance on the PDF curve of the pad asperities. Accurately predicting the interface gap (wear rate of the pad) is important in determining both mechanical and chemical removal rate, especially in the case of metal polishing when chemical removal rate can not be neglected. In this work, we present a

المنارة للاستشارات
www.manaraa.com

model of the polish-rate decay by coupling population balance model with fluid mechanics which is presented in Chapter 4. Pad-wafer contact is modeled by Greenwood and Williamson's model with elastic Hertzian contact [42] and fluid is modeled by Reynold's equation [45]. Modeling results with and without fluid are compared. The pad asperity PDF evolution results with fluid show much less pad wear compared with the case without fluid. The results show that fluid is an important factor which needs to be put into the model in predicting the interface gap (pad wear) and surface profile evolution of the pad during polishing and conditioning. The results in Chapter 4 are published in *Microelectronic Engineering* [46].

Ref. [40], [44], [46] model pad wear only. For pad conditioning process, Borucki, et al. [47] have presented a model for conditioning and wear of solid pads and foamed pads. In his conditioner model, he made an assumption that the diamond grits create furrows on the pad surface by cleanly cutting the pad material, e.g., a polishing operation in which all the material cut by the diamonds is removed from the surface of the pad. That conditioner model is similar to the process of polishing brittle materials where materials are removed by brittle fracture. However, polyurethane which is commonly used as the CMP pad material is a polymer with plastic mechanical properties (in which material is simply pushed aside and not removed from the surface) that need to be considered in the conditioner model. In this work, which is presented in Chapter 5, we treat pad as a plastic material and have developed a conditioner model by combining the cutting and plastic deformation in which a groove with pile-up on both side-walls forms as the diamond moves over the pad surface. Plastic deformation is important in modeling the pad conditioning because the raised ridges along the sides of the grooves are the high points

المنارة للاستشارات
www.manaraa.com

(asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive particles used to polish the wafer. In [47], Borucki, et al. use a population balance equation with only a "death term" to model the cleanly cutting of the pad surface. In this work we add another term, a "birth term", by using the concept of a mass balance to represent the material piled-up on both side-walls of the groove. Consequently the pad material should be consumed much more slowly by adding this term than by solving the population balance equation with only the death term in which all of the materials cut by the diamond is removed from the pad surface. Pad cut-rate is important to estimate the aggressiveness of the pad conditioning, decide the time needed to bring the pad surface into steady-state in the pad "break-in" process and the time to change the pad. In this work, the analytical solutions for the cut-rate and steady-state pad conditioning PDF are derived and they match the numerical solutions obtained with Monte Carlo methods. These results are presented in Chapter 5 of this thesis.

Pad conditioning is an important step to restore the pad asperities to sustain a stable polish rate. Conditioning can be done either by *in situ* dressing, where the conditioning is performed while the wafers are polished, or by *ex situ* dressing, where the pad is dressed after the wafer has been polished. Current conditioner disks are made by attaching numerous similar diamond grits to the flat surface of a metal substrate using either nickel electroplating or brazing with alloys [48]. The size distribution of the diamond grits together with the deformation of the substrate cause the height variation of the diamond grits. In addition, it is difficult to control the height leveling of so many diamond grits. The embedded diamonds will have different shape, orientation and size. Diamond grit is typically sized using sieve techniques, resulting in a Gaussian distribution of sizes with a

المنارة للاستشارات
www.manaraa.com

standard deviation of 15-20% of the average size. Additional variations introduced by the shape of the particles can increase this deviation to as much as 30% of the average size. If the average penetration of grit into the pad at a given downforce is less than one or two standard deviations of the average particle size, then the actual fraction of grit in contact with the pad can be as low as a few percent of the total grit particles on the conditioner. This can produce extreme variations in the performance of the conditioner [49]. Diamond grit sizes can be found in the diamond grit manufacturer's website [50]. For example, a diamond disk with 100 grit size has an average diamond size of 137 μm [50]. An example of a conventional diamond disk is shown in Figure 1.4 [51].

To obtain good uniformity of the polished surface, the height distributions of pad asperities should be as uniform as possible. This requires the diamond heights of the conditioner should be as uniform as possible. For an ideal conditioner, each diamond particle would be precisely the same shape, size and orientation, and the spacing between particles would be identical. The penetration depth of the grit into the pad would increase monotonically with increasing applied downforce [52]. These can assure uniform force distribution over the whole disk. The force on each diamond grit is identical and the penetration depth of each diamond into the pad is identical. To obtain these goals, many novel conditioners have been designed [51], [53]. In [51], a novel conditioner is designed based on the forming of a pattern by wire electro-discharge micro-machining a strong polycrystalline diamond substrate that is sintered at high pressure and temperature. Each micro-machined diamond pyramid has the same shape and size and the spacing between them is identical. The roughness of the conditioned pad is about 3 μm by the novel conditioner compared to 5 μm by the conventional conditioner, which shows a

المنارة للاستشارات
www.manaraa.com

considerable improvement. In [53], the developed conditioner is composed of a lot of bundles of fine metal fibers to contact the pad surface independently. Constant conditioning force can be applied to the pad surface with fiber edges regardless of the large height variation of the pad surface. The uniformity of the conditioned pad is improved compared to the conventional conditioner.

Pad surface topography has a direct impact on the chemical and mechanical polishing rate and polishing uniformity and conditioner plays a key role in CMP. In Chapter 6, the model in Chapter 5 is extended to incorporate variable diamond heights for the conditioner. Our model results have shown that the PDF of the conditioned pad surface with a variable diamond height is Gaussian while it is not Gaussian (derivative of Gaussian) for constant diamond height case in Chapter 5 and in Ref. [47]. That the distribution of the pad asperity resulting from conditioning is Gaussian is consistent with the common experimental observations in the literature [54], [55]. The diamond height PDF only changes the standard deviation of the conditioned pad asperity PDF and it does not change its functional form which is Gaussian. The contact form of the pad asperities is important to determine the contact area and interface gap between conditioner diamonds (or the wafer) and the pad. The load on the conditioner, the rotation speed between the conditioner and the pad and the viscosity of the polishing slurry decides the interface gap between conditioner and the pad. Our results show that a different interface gap will give different standard deviation of the resulting conditioned pad PDF if the diamond height distribution is Gaussian. This is consistent to the literature that with increasing load, the roughness of the conditioned pad PDF increases [55]. According to our model results, it is the result of the decreasing of the interface gap between the mean

المنارة للاستشارات
www.manaraa.com

of the conditioner diamonds and the pad surface. For Gaussian diamond height distribution, the roughness of the active diamonds increases with decreasing conditioner-pad separation. The roughness of the conditioned pad asperity PDF is determined only by the roughness of the active diamonds, and it increases as the roughness of the active diamond increases. These results are presented in Chapter 6.

Chapter 2 of this thesis will give a general background introduction. It introduces the theories and models used in the simulation work of this thesis. Chapters 3, 4, 5, and 6 are independent chapters which have more detailed model derivation and results. Chapters 3 and 4 have been published. Chapter 5 has been submitted to *Journal of Engineering mathematics*. Chapter 7 will summarize the results of the whole dissertation and suggest future work.

References

[1] S. Wolf and R. N. Tauber, *Silicon Processing for the VLSI Era, Volume 1: Processing Technology*. Sunset Beach, CA: Lattice Press, 2001.

[2] A. Meled, "Optimization of polishing kinematics and consumables during chemical mechanical planarization processes," Ph.D. Dissertation, Dept. Chem. Eng., Univ. Arizona, Tucson, AZ, 2011.

[3] *International Technology Roadmap for Semiconductors, 2000 Update* [Online]. Available: http://www.itrs.net/Links/2000UpdateFinal/ORTC2000final.pdf

[4] G. S. May and S. M. Sze, *Fundamentals of Semiconductor Fabrication*. New York, NY: John Wiley & Sons, 2004.

[5] *International Technology Roadmap for Semiconductors, 2006 Update Interconnect* [Online]. Available: http://www.itrs.net/Links/2006Update/FinalToPost/09_Interconnect2006Update.pdf

[6] R. J. Gutmann, J. M. Stiegerwald, L. You, D. T. Price, J. Nierynck, D. J. Duquette, and S. P. Murarka, "Chemical-mechanical polishing of copper with oxide," *Thin Solid Films*, vol. 270, pp. 596-600, 1995.

المنارة للاستشارات
www.manaraa.com

[7] J. M. Stiegerwald, S. P. Murarka, J. Ho, R. J. Gutmann, and D. J. Duquette, "Mechanism of copper removal during chemical mechanical polishing," *J. Vac. Sci. Technol. B,* vol. 13, pp. 2215-2218, 1995.

[8] D. Shamiryan, T. Abell, F. Iacopi and K. Maex, "Low-*k* dielectric materials," *Mater. Today*, vol. **7**, no. 1, pp. 34-39, January 2004.

[9] N. H. Weste & D. Harris, *Principles of CMOS VLSI Design: A Circuit & Systems Perspective*. White Plains, NY: Addison-Wesley Publishers, 2005.

[10] *RC Time Constant* [Online]. Available: http://en.wikipedia.org/wiki/RC_time_constant

[11] S. P. Murarka, *Metallization: Theory and Practice for VLSI and ULSI*. Boston, MA: Butterworth Heinmann, 1993.

[12] S. P. Murarka and S. W. Hymes, "Copper metallization for ULSI and beyond," *Crit. Rev. Solid State Mat. Sci.*, vol. 20, no. 2, pp. 87-124, 1995.

[13] J. Li, T. E. Seidel, and J. W. Mayer, "Copper-based metallization in ULSI structures, part II: Is Cu ahead of its time as an on-chip interconnect material?," *MRS Bull.*, vol. 19, no. 8, pp. 15-21, Aug. 1994.

[14] J. Paraszczak, D. Edelstein, S. Cohen, E. Babich and J. Hummel, "High performance dielectrics and processes for ULSI interconnection technologies," in *Tech. Dig. Int. Electron Dev. Mtg.(IEDM)*, 1993, pp. 261-264.

[15] D. Merricks. (2007). *CMP in Copper/Low-k Dielectric Integration Schemes* [Online]. Available: http://www.ferro.com/NR/rdonlyres/261E44B2-C15D-4732-BCC6-B45B43231060/0/AVS2003.pdf

[16] S. V. Sathish Babu Janjam, "Chemical mechanical polishing of tantalum and tantalum nitride," Ph.D. Dissertation, Dept. Chem. Eng., Clarkson Univ., Potsdam, NY, Dec. 2009.

[17] *Low-k Dielectric* [Online]. Available: http://en.wikipedia.org/wiki/Low-k_dielectric

[18] M. Fayolle, G. Passemard, M. Assous, D. Louis, A. Beverina, Y. Gobil, J. Cluzel and L. Arnaud, "Integration of copper with an organic low-*k* dielectric in 0.12-µm node interconnect," *Microelectron. Eng.*, vol. 60, no. 1-2, pp. 119-124, 2002.

[19] K. C. Saraswat. *Interconnections: Copper and Low-k Dielectrics* [Online]. Available: http://www.stanford.edu/class/ee311/NOTES/Interconnect_Cu.pdf

[20] J. M. Steigerwald, S. P. Murarka, and R. J. Gutmann, *Chemical-Mechanical*

المنارة للاستشارات

www.manaraa.com

Planarization of Microelectronic Materials. New York: John Willey & Sons, 1997.

[21] M. R. Oliver, *Chemical-Mechanical Planarization of Semiconductor Materials*. Berlin, Germany: Springer-verlag Berlin Heidelberg, 2004.

[22] P. B. Zantye, A. Kumar, and A. K. Sikder, "Chemical mechanical planarization for microelectronics applications," *Mater Sci. Eng.*, R45, pp. 89-220, 2004.

[23] S. Sivaram, H. Bath, R. Legegett, A. Maury, K. Monning, and R. Tolles, "Planarizing interlevel dielectrics by chemical mechanical polishing," *Solid State Technol.*, pp. 87-91, May 1992.

[24] S. Sundararajan, D. G. Thakurta, D. W. Schwendeman, S. P. Murarka and W. N. Gill, "Two-dimensional wafer-scale chemical mechanical planarization models based on lubrication theory and mass transport," *J. Electrochem. Soc.*, vol. 146, no. 2, pp. 761-766, 1999.

[25] C. Yu, P. C. Fazan, V. K. Mathews, and T. T. Doan, "Dishing effects in a chemical mechanical polishing planarization process for advanced trench isolation," *Appl. Phys. Lett.*, vol. 61, no. 11, pp. 1344-1346, Sept. 1992.

[26] F. B. Kaufman, D. B. Thomson, R. E. Broadie, M. A. Jaso, W. L. Guthrie, D. J. Pearson and M. B. Small, "Chemical-mechanical polishing for fabricating patterned W metal features as chip interconnect," *J. Electrochem. Soc.*, vol. 138, no. 11, pp. 3460-3465, 1991.

[27] J. M. Steigerwald, R. Zirpoli, S. P. Muraka and R. J. Gutmann, "Pattern geometry effects in the chemical-mechanical polishing of inlaid copper," *J. Electrochem. Soc.*, vol. 142, no. 10, pp. 2841-2848, 1994.

[28] B. Davari, C. W. Koburger, R. Schulz, J. D. Warnock, T. Furukawa, M. Jost, Y. Taur, W. G. Schwittek, J. K. DeBrosse, M. L. Kerbaugh, and J. L. Mauer, "A new planarization technique using combination of RIE and chemical mechanical polishing (CMP)," in *Tech. Dig. Int. Electron Dev. Mtg. (IEDM)*, 1989, pp. 61-64.

[29] S. Jung, "Chemical mechanical polishing of copper using nanoparticle-based slurries," Ph.D. Dissertation, Dept. Mater. Sci. Eng., Univ. Florida, Gainesville, FL, 2005.

[30] S. Franssila, *Introduction to Microfabrication, 2nd edition*. London, UK: John Wiley & Sons, 2010.

[31] R. D. Rung, H. Momose, and Y. Nagakubo, "Deep trench isolation CMOS devices," in *Tech. Dig. Int. Electron Dev. Mtg. (IEDM)*, 1982, pp. 237-240.

[32] R. Dejule, "CMP grows in sophistication," *Semicond. Int.*, vol. 21, no. 13, pp. 56-

المنارة للاستشارات

www.manaraa.com

62, Nov., 1998.

[33] M. McCoy, "Many roads for CMP," *Chem. Eng. News*, vol. 84, no. 26, pp. 17-19, 2006.

[34] *CMP Markets & Technologies to the 22nd Node* [Online]. Available: http://www.linx-consulting.com/CMP-22node.html

[35] L. Shan, "Mechanical interactions at the interface of chemical mechanical polishing," Ph.D. dissertation, Dept. Mech. Eng., Georgia Tech., Ann Arbor, MI, 2000.

[36] D. Chen and B. Lee, "Parameter analysis of chemical mechanical polishing: an investigation based on pattern planarization model," *J. Electrochem. Soc.,* vol. 146, pp. 744-748, 1999.

[37] D. Stein, D. Hetherringgton, M. Dugger, and T. Stout, "Optical interferometry for surface measurements of CMP pads," *J. Electron. Mater.*, vol. 25, no. 10, pp. 1623-1627, 1996.

[38] M. R. Oliver, R. E. Schmidt and M. Robinson, "CMP pad surface roughness and CMP removal rate," in *Electrochem. Soc. Proc.,* vol. 2000-26, 2000, pp. 77-83.

[39] A. S. Lawing, "Polish rate, pad surface morphology and oxide chemical mechanical polishing," in *Electrochem. Soc. Proc*., vol. 2002-1, 2002, pp. 46-60.

[40] L. J. Borucki, "Mathematical modeling of polish-rate decay in chemical-mechanical polishing," *J. Eng. Math.*, vol. 43, pp. 105-114, 2002.

[41] J. Archard, "Contact and rubbing of flat surfaces," *J. Appl. Phys.*, vol. 24, pp. 981-988, 1953.

[42] J. A. Greenwood and J. B. P. Williamson, "Contact of nominally flat surfaces," *Proc. Roy. Soc. Lond. Math. Phys. Sci.*, vol. 295, no. 1442, pp. 300-319, 1966.

[43] C. Wang, P. Sherman, and A. Chandra, "A stochastic model for the effects of pad surface topography evolution on material removal rate decay in chemical-mechanical planarization," *IEEE. T. Semiconduct. M.*, vol. 18, no. 4, pp. 695-708, 2005.

[44] H. Shi and T. A. Ring, "Analytical solution for polish-rate decay in chemical-mechanical polishing," *J. Eng. Math.*, vol. 68, no. 2, pp. 207-211, 2010.

[45] W. A. Gross, L. A. Matsch, V. Castelli, A. Eshel, J. H. Vohr and M. Wildmann, *Fluid Film Lubrication*. New York: John Wiley & Sons, 1980.

المنارة للاستشارات
www.manaraa.com

[46] H. Shi and T. A. Ring, "CMP pad wear and polish-rate decay modeled by asperity population balance with fluid effect," *Microelectron. Eng.*, vol. 87, no. 11, pp. 2368-2375, 2010.

[47] L. J. Borucki, T. Witelski, C. Please, P. R. Kramer, and D. Schwendeman, "A theory of pad conditioning for chemical-mechanical polishing," *J. Eng. Math.*, vol. 50, pp.1-24, 2004.

[48] M. Y. Tsai and Y. S. Liao, "Dressing characteristics of oriented single diamond on CMP polyurethane pad," *Mach. Sci. Technol.,* vol. 13, pp. 92-105, 2009.

[49] E. Thear and F. Kimock. (2010, Jul. 13). *Improving Productivity Through Optimization of the CMP Conditioning Process* [Online]. Available: http://www. morgantechnicalceramics.com/ resources/technical_articles/improving-productivity-through-optimization-of-the-cmp-conditioning process/?page_index=1

[50] *Diamond Grit Size Designations, Approximate Particle Sizes, and Counts* [Online]. Available: http://www.inlandlapidary.com/user_area/gritchart.asp

[51] Y. Kim and S. L. Kang, "Novel CVD diamond-coated conditioner for improved performance in CMP processes," *Int. J. Mach. Tool Manu.,* vol. 51, pp. 565-568, 2011.

[52] J. Zimmer and A. Stubbmsnn, "Key factors influencing performance consistency of CMP pad conditioners," in *Proc. 2nd Int. Symp. on CMP,* 1998, pp. 87-92.

[53] H. Gao and J. Su, "Study on the surface characteristics of polishing pad used in chemical mechanical polishing," *Adv. Mate. Res.*, vol. 102-104, pp. 724-728, 2010.

[54] K. Park and H. Jeong, "Investigation of pad surface topography distribution for material removal uniformity in CMP process," *J. Electrochem. Soc.,* vol. 155, no. 8, pp. H595-H602, 2008.

[55] T. Sun, L. Borucki, Y. Zhuang and A. Philipossian, "Investigating the effect of the diamond size and conditioning force on chemical mechanical planarization pad topography," *Microelectron. Eng.*, vol. 87, pp. 553-559, 2010.

المنارة للاستشارات
www.manaraa.com

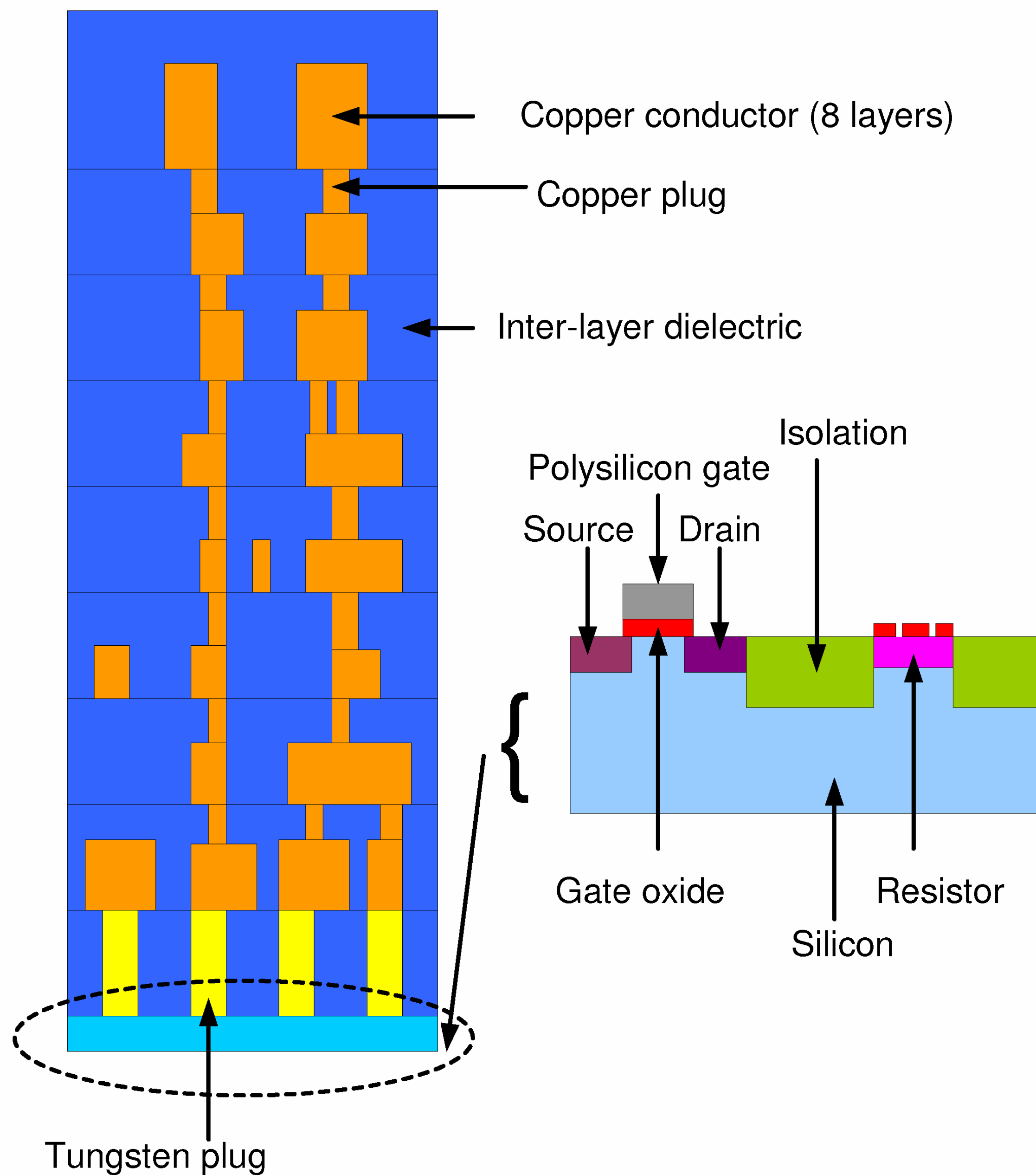


Figure 1.1 A schematic of a modern IC. Adapted from [2].

www.manaraa.com

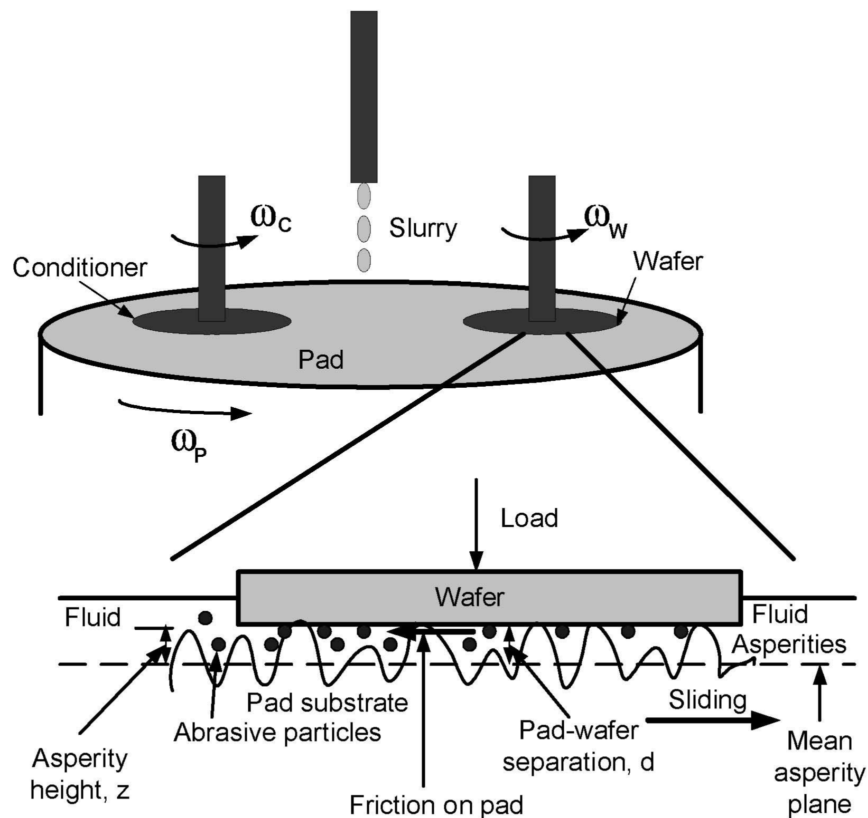
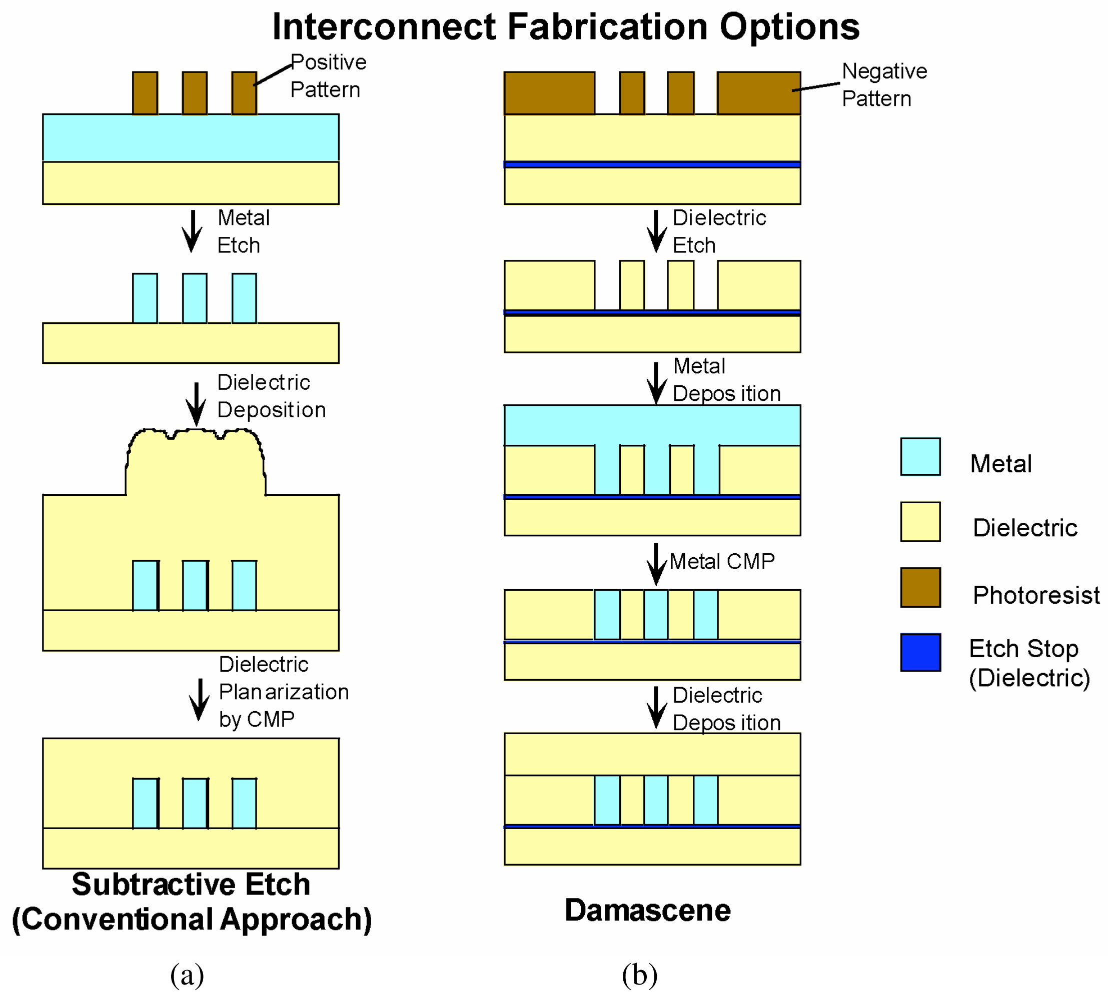


Figure 1.2 Schematic of the CMP process. The rotation rate of pad, conditioner and wafer are ω_P, ω_C, and ω_w respectively. Pad asperities drag abrasives to abrade the wafer surface and the abraded material is dissolved in the slurry by chemical reaction [24]. The separation d is measured from the mean plane of the pad and asperity heights are also measured from the mean plane of the pad.

المنارة للاستشارات
www.manaraa.com



(a) (b)

Figure 1.3 Schematics of the two options to fabricate interconnects. (a) Subtractive etch approach. (b) Damascene approach. Subtractive etch is used in aluminum interconnects fabrication and damascene is used in copper and tungsten interconnects fabrication. Reprinted with permission from Dr. Prof. Krishna C. Saraswat, Stanford University [19].

المنارة للاستشارات
www.manaraa.com

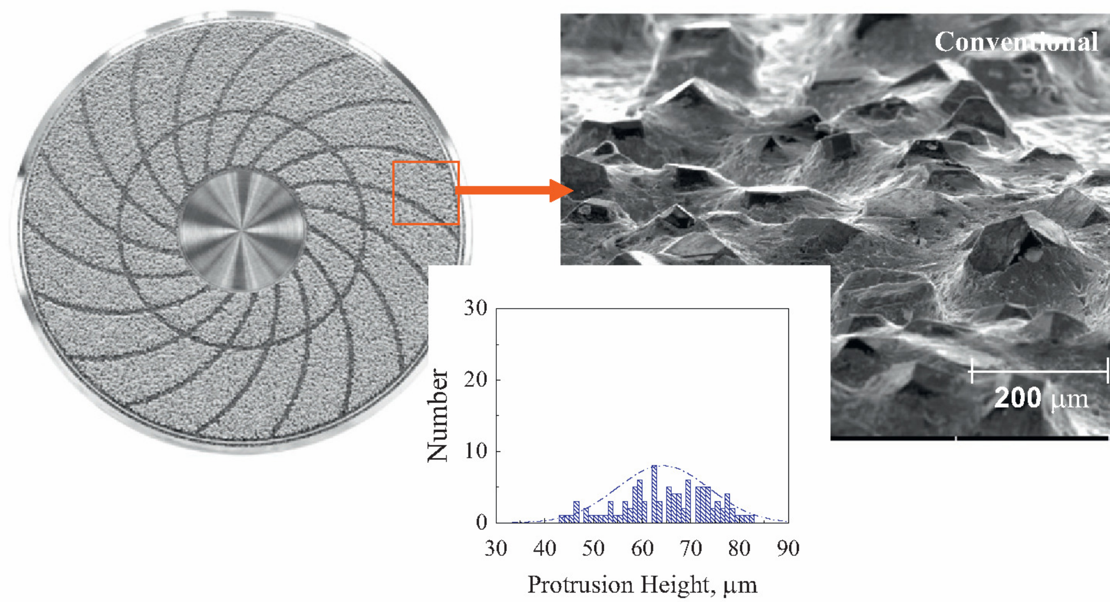


Figure 1.4 Macro- and micrographs of a conventional diamond conditioner and the height distribution of the diamond grits on the conditioner. Reprinted from [51], with permission from Elsevier.

www.manaraa.com

Table 1.1 Properties of interconnect metals [11] - [13].

Properties	Al alloy	Cu	Ag	Au
Resistivity (μΩ-cm)	~ 3.5	1.67	1.59	2.35
Electromigration resistance (at 0.5 μm)	moderate	good	poor	Very good

Table 1.2 Interconnect technology requirements according to ITRS roadmap [5].

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013
Technology node, nm	80	70	65	57	50	45	40	36	32
Number of Cu levels	11	11	11	12	12	12	12	12	13
ILD dielectric constant (k)	2.6-3.0	2.6-3.0	2.3-2.7	2.3-2.7	2.1-2.4	2.1-2.4	2.1-2.4	1.8-2.1	1.8-2.1

المنارة للاستشارات
www.manaraa.com

Table 1.3 Dielectric constants of available low-k dielectric materials [15] - [18].

<table>
<tr><th>Deposition method</th><th colspan="2">Material</th><th>Dielectric constant, k</th></tr>
<tr><td rowspan="4">CVD</td><td rowspan="2">Inorganic</td><td>Porous SiO_2</td><td>1.8 – 2.4</td></tr>
<tr><td>Fluorine-doped SiO_2</td><td>3.4 – 3.6</td></tr>
<tr><td rowspan="2">Organic</td><td>Carbon-doped SiO_2</td><td>2.7 – 2.9</td></tr>
<tr><td>Porous carbon-doped SiO_2</td><td>2.2 – 2.5</td></tr>
<tr><td rowspan="6">Spin-on</td><td>Inorganic</td><td>Hydrogen silsesquioxane (HSQ)</td><td>2.8 – 3.0</td></tr>
<tr><td rowspan="5">Organic</td><td>Polyimide</td><td>3.0 – 3.5</td></tr>
<tr><td>SiLK^{TM}</td><td>2.7</td></tr>
<tr><td>Porous SiLK^{TM}</td><td>1.9</td></tr>
<tr><td>Methylsilsesquioxane (MSQ)</td><td>2.7 – 2.9</td></tr>
<tr><td>Porous MSQ</td><td>1.8 – 2.5</td></tr>
</table>

المنارة للاستشارات
www.manaraa.com

CHAPTER 2

BACKGROUND

المنارة للاستشارات
www.manaraa.com

This chapter gives a general introduction of the theory and model used in this thesis. The first section introduces the elastic contact model between wafer and pad which is the foundation for the pad wear model. The second section introduces the mixed-lubrication approach in modeling pad wear which considers fluid effect. The third section introduces wear of the pad between wafer and pad. The fourth section introduces the model for wafer material removal rate (MRR). The last section introduces the population balance model. It starts with the general population balance model used in particulate systems and then shows how it is applied to the pad wear and pad conditioning problems in this thesis. Chapters 3, 4 and 5 are independent chapters that are focused on pad wear and pad conditioning problems with more detailed model derivation and model results.

Elastic contact model between wafer and pad

Greenwood and Williamson developed a theory for the contact between a flat surface and a rough surface with Hertzian contact theory [1]. As Figure 1.2 shows, the wafer surface is assumed flat and the pad surface is assumed rough. The mean plane of the pad surface is taken as the reference plane. The height z of any asperity relative to the mean plane of the pad surface is assigned a PDF (population density function), denoted by $\phi(z)$. All asperities are assumed to have spherical summits with identical curvatures k_s [=1/m]. The separation distance between the pad mean plane and the wafer surface is denoted by d. An asperity with height z is in contact with the wafer surface when $z > d$. The load on an individual asperity

المنارة للاستشارات

www.manaraa.com

$$L = \frac{4E^*}{3k_s^{1/2}}(z-d)^{3/2} \tag{2.1}$$

is given by its compliance, (z-d). This load is distributed over a circular contact area at the tip of the asperity

$$A = \frac{\pi(z-d)}{k_s} \tag{2.2}$$

where $E^* = E/(1-\nu^2)$, E is the pad Young's modulus and ν is the Poisson ratio.

Now let us consider the total load carried by all of the pad asperities, N, in contact with the wafer of nominal area A_0 when the asperities are assumed to be uniformly spread over the pad with an area density of $\eta_s = N/A_0$. The nominal asperity pressure, the total load carried by all of the asperities divided by the nominal area A_0 corresponding to the separation d for this elastic model of the rough pad surface in contact with a flat wafer surface is given by

$$P_a = \frac{4}{3}\eta_s E^* / k_s^{1/2} \int_d^\infty (z-d)^{3/2} \phi(z) dz \tag{2.3}$$

and the ratio of the actual or total asperity contact area to the nominal area (contact area fraction) is given by

$$A_c = \pi\eta_s / k_s \int_d^\infty (z-d)\phi(z) dz \tag{2.4}$$

Equations (2.3) and (2.4) form the basis of the Greenwood and Williamson model of rough surface contact. The original derivation of the Greenwood and Williamson

المنارة للاستشارات
www.manaraa.com

model used elastic contact, but any type of contact can be used for the individual asperity to arrive at Equations (2.3) and (2.4).

Mixed-lubrication approach

CMP is a complex process. The interactions between the wafer, pad and slurry make the CMP process difficult to predict. The hydrodynamics of slurry flow at the wafer-pad interface has an important effect on the wafer material removal rate and pad wear. The CMP lubrication regime depends largely on how the applied load is distributed between the fluid film layer and the pad asperities. In hydrodynamic lubrication, solid contact between pad and wafer is negligible and the fluid pressure balances the entire applied load. Material removed in this regime is by chemical reaction and slurry particle abrasion. At the other extreme, boundary lubrication, the fluid film is thin and there is significant solid-solid contact between the wafer and the pad. In this regime, material removal is mainly by mechanical abrasion of the asperities against the wafer surface. The abrasion is enhanced by chemical activation of the wafer surface and silica slurry particles entrapped between the pad asperities and wafer surface. The consensus hypothesis is that CMP operates somewhere between these two extremes in what is known as the partial lubrication (or mixed-lubrication) regime [2].

In the literature there are many publications on hydrodynamics of slurry flow in CMP, including theoretical and numerical models as well as experimental measurements. For the experimental measurements of slurry film thickness and hydrodynamic pressure distribution, the work done by a group of people represented by A. M. Scarfo, J. Lu and Bullen, et al. are considered as a more complete set of measurements [2] - [4]. Their measurements for pressure, friction and slurry film thickness are performed while

www.manaraa.com

both wafer and pad are rotating, which are closer to real CMP condition. While others are performed by assuming static wafer and rotating pad for ease of modeling and experimentation [5] - [7]. Scarfo's experimental facility consists of a tabletop polisher which is pressed upon by a drill press which serves as a wafer carrier. The wafer is attached to the drill press spindle via a spherical gimbal that allows it to pivot freely while rotating. The wafer is outfitted with a series of pressure taps at various radii that are connected to pressure transducers, which rotate with the wafer. They use dual emission laser induced fluorescence (DELIF) to measure fluid film thickness in real-time CMP process. BK7 glass wafers are used because DELIF technique requires transparent wafers. They measured fluid film thickness, slurry pressure at various applied loads, wafer/pad relative velocities, and wafer curvatures. These measurements are done under process conditions scaled directly from those used in industry. As wafer shape affects the lubrication regime during CMP, three types of wafers are used. Concave wafers (type A) with a center to edge bow ranging from 5 -15 μm, convex wafers (type B) with 5 - 15 μm center to edge convexity, and wafers close to flat (type C). At an applied down load of 27.6 kPa, for wafer type B, only positive pressure is observed, increasing parabolically toward the wafer center. For wafer type A, subambient pressure is obtained near the middle portion of the wafer. Pressure seems to start out slightly positive near the wafer edge, but suction is created upon moving closer to the wafer center. The behavior of wafer type C falls between the behaviors of the other two wafer types. The pressure is all positive over all speeds for wafer type C and increases at the wafer center. Subambient pressure is only observed for the concave wafers over all speeds. When applied load is increased to 41.4 kPa, it induces a transition from subambient to positive pressure for

المنارة للاستشارات
www.manaraa.com

wafer type A over all speeds. Wafer type A is most sensitive to down force. One explanation is that as the down force is increased, the outer edge of the concave wafer digs into the pad causing the pad to deform. The pad conforms to the bowed area at the wafer center, creating a converging gap. According to lubrication theory, a converging gap will produce a positive pressure and a diverging gap will produce a negative pressure. As the wafer shape transitions from concave to convex (i.e., type A to type B to type C), the fluid film supports a greater percentage of the applied load. Increasing applied load results in more classic positive parabolic pressure profile. However, fully hydrodynamic lubrication is not observed under any conditions tested. Hence CMP operates at partial hydrodynamic lubrication. The works done by this group of people have shown that in studying the slurry hydrodynamics, the wafer shape and pad deformation are very important. A small amount of wafer concavity on the order of 3-5 μm from edge to center can cause subambient pressure.

The work done by Bullen, et al. [3] has also shown that the interfacial pressure distribution under a static wafer is radically different from a rotating wafer. Their experimental pressure measurements have been collected from both stationary and rotating wafer trials. It is found that the pressure distribution differs greatly between static and dynamic wafers. The static wafer pressure distribution is characterized by a wide range of pressures including some subatmospheric values. The dynamic wafer pressure distribution is characterized by a smaller range and no subatmospheric values. Overall the data show that modeling CMP with a nonrotating wafer is not an adequate approximation.

In the literature, there are many publications on the analytical and numerical

المنارة للاستشارات

www.manaraa.com

modeling of the slurry flow for CMP. Some assume smooth pad surface and neglect solid contact between wafer and pad. In this way, the applied load is entirely supported by the slurry [8] - [10]. Some consider both the hydrodynamics of the slurry and contact mechanics between the rough pad surface and the wafer [6], [11], which is called the mixed-lubrication approach. Tsai et al. has presented a model which applies the average lubrication equation with partial hydrodynamic lubrication theory and considers the elastic-plastic microcontact theory [11]. Their average lubrication equation with partial hydrodynamic lubrication theory has considered the slurry particle and pad roughness effects to the slurry flow. For the microcontact model, the contact conditions for wafer and polishing pad are determined by the interface gap. In CMP, the elastic, elastoplastic, and plastic deformations of the polishing pad may be produced. In this work, they have considered elastic deformation [1], plastic deformation [12], and elastoplastic deformation [13]. The total contact pressure is supported by these three types of contacts. They assumed convex wafer curvature. The applied pressure is the sum of the partial hydrodynamic pressure and the wafer-pad asperity contact pressure. Momentum balance is also performed around the wafer pivot point to solve the slurry film thickness and slurry fluid pressure distribution. They use experimental parameters in Scarfo, Lu and Bullen's experiments. Their modeling results for slurry pressure and slurry film thickness have matched experimental results by A. M. Scarfo, J. Lu and Bullen, et al. The film thickness increases with the speed and decreases with the applied load. The attack angle (i.e., wafer tilt angle) creates converging flow and provides hydrodynamic load (positive fluid pressure).

In this dissertation, we use the mixed-lubrication approach to model pad wear by

المنارة للاستشارات
www.manaraa.com

combining Reynold's equation for hydrodynamic lubrication theory and microcontact theory by Greenwood and Williamson [1]. Fluid pressure is assumed positive and we use the averaged pressure obtained by dimensional analysis of Reynold's equation [14], [15].

The pad-wafer separation d is found by the load balancing. In the mixed-lubrication approach, the total applied normal pressure P_{app} between pad and wafer is carried by both asperity pressure P_a and fluid pressure P_f as shown by Equation (2.5). We assume that fluid pressure and asperity pressure distribute over the same area without influencing each other. Hence the separation d is found by finding the root of the load balancing equation, Equation (2.5). The asperity pressure is defined by Equation (2.3). The fluid pressure is defined by Equation (2.6) which is obtained by dimensional analysis of Reynold's equation [14], [15].

$$P_{app} = P_a + P_f \tag{2.5}$$

$$P_f = \frac{D\mu U}{4d^2} \tag{2.6}$$

where D is the wafer diameter, μ is the viscosity of the slurry, U is the linear velocity between the pad and wafer and d is the pad-wafer separation distance.

<u>Wear of the pad</u>

During wafer polishing, the pad asperities are abraded by abrasive particles. This wear can be determined from Archard's law [16], [17].

$$dV = -k_w L dx \tag{2.7}$$

المنارة للاستشارات

www.manaraa.com

where dV is the volume of material removed in a sliding distance dx, k_w is a wear coefficient which is a function of material hardness, microstructure, and lubrication, and L is the load on the indenter which is the asperity in this case. The wear depth can be estimated from the wear volume by dividing Equation (2.7) by the contact area of the asperity

$$dz = \frac{dV}{A} = -k_w \frac{L}{A} dx \tag{2.8}$$

If we divide Equation (2.8) by the increment time dt, then we obtain the wear rate defined in wear depth per time

$$\frac{dz}{dt} = -k_w \frac{L}{A} U \tag{2.9}$$

where U is the velocity between the wafer and pad.

Next if we define $c_a = k_w U$ and substitute the local contact pressure L/A from Equation (2.1) and (2.2), we arrive at the wear rate for a single asperity

$$\frac{dz}{dt} = -c_a \frac{L}{A} = -c_a \frac{4}{3\pi} E^* k_s^{1/2} (z-d)^{1/2} \tag{2.10}$$

where z is the height of the asperity.

An important concept is that L/A in Equation (2.10) can be obtained from elastic or inelastic contact models whose formulations depend upon material properties, such as elastic deformation, viscoelastic deformation, viscoplastic deformation and strain hardening/softening effects, etc. In this work, however, only elastic contact is used.

المنارة للاستشارات

www.manaraa.com

Wafer material removal rate

The average material removal rate (MRR) can be derived from Equation (2.7) analogously. Equation (2.7) is the wear volume abraded from the wafer surface by a single asperity. If we add up the wear volume removed by all of the asperities, the left side of Equation (2.7) is the total wear volume and the right side is the total load carried by asperities as Equation (2.11) shows. Here we assume every point on the wafer has the same velocity relative to the pad.

$$dV_t = k_w^{'} L_a dx \tag{2.11}$$

where dV_t is the total differential volume removed in a differential time dt and $k_w^{'}$ is the wear coefficient for the abrasive wear of the wafer. Dividing Equation (2.11) by dt, the wear rate in terms of total differential volume removed is obtained

$$\frac{dV_t}{dt} = k_w^{'} L_a U \tag{2.12}$$

Then dividing Equation (2.12) by the total contact area $A_c A_0$, the wear rate in terms of thickness removed is obtained

$$MRR(t) = c_w \frac{P_a}{A_c(t)} \tag{2.13}$$

where P_a is the asperity pressure, A_c is the asperity contact area fraction, A_0 is the nominal contact area and $c_w = k_w^{'}U$. Wear will change asperity PDF, and hence will make d, MRR, P_a, and A_c time-dependent assuming P_{app} and U are constants. In [18], Borucki uses the normal applied pressure P_{app} for P_a in Equation (2.13) because the fluid is

المنارة للاستشارات
www.manaraa.com

neglected. In that case, P_{app} equals to P_a and P_f equals to zero everywhere. Motivated by this approach and considering the fluid effect, we use P_a instead of P_{app}. Fundamentally, Equation (2.13) is for mechanical removal rate. Hence, it is more accurate for the case where chemical removal can be neglected.

The asperity height population balance

In this section, the general population balance equation which is used in particulate systems is introduced first, and then how it is applied to pad wear and pad conditioning problems is explained.

General population balance equation

The population balance is particle continuity equation. It is used as a method of accounting for particles as they go through a process, such as grinding, classification, crystallization, aggregation, or grain growth. The number density of particles N(L) (with units of number of particles per unit volume) is equal to the integral of the population $\eta_0(L)$ from size L to size L+ΔL and is defined as [19]

$$N(L) = \int_{L}^{L+\Delta L} \eta_0(L)dL \tag{2.14}$$

The units on the population $\eta_0(L)$ are number of particles per unit volume per particle length, L.

The microscopic population balance is obtained by accounting for all the particles in a differential volume dV = dx*dy*dz as shown in Figure 2.1. The population for particles in a differential volume is

المنارة للاستشارات
www.manaraa.com

$$\text{Accumulation} = \text{Input} - \text{Output} + \text{Net generation} \tag{2.15}$$

The accumulation term is the change in the population with time, $d\eta_0/dt$. The input and output terms are considered together. For the input-output terms, we have one term due to flow and another term due to internal variables like particle growth. For output-input due to flow we have

$$\nabla \cdot v\eta_0 = \frac{d(v_x\eta_0)}{dx} + \frac{d(v_y\eta_0)}{dy} + \frac{d(v_z\eta_0)}{dz} \tag{2.16}$$

where v_x, v_y, v_z are the velocity components in the x, y, and z directions.

For output-input due to growth or other internal variable x_i we must add

$$\nabla \cdot v_i\eta_0 = \frac{d(v_i\eta_0)}{dx_i} \tag{2.17}$$

where v_i is the velocity of an internal variable, for example, the grow rate, G, and x_i is an internal variable, for example, the size of the particle.

For the net generation terms we have birth, minus death. Combining all these terms we have the microscopic population balance [19]:

$$\frac{\partial\eta_0}{\partial t} + \frac{\partial}{\partial x}(v_x\eta_0) + \frac{\partial}{\partial y}(v_y\eta_0) + \frac{\partial}{\partial z}(v_z\eta_0) + \sum_{i=1}^{m}\frac{\partial}{\partial x_i}(v_i\eta_0) - birth + death = 0 \tag{2.18}$$

This differential equation is the fundamental population balance. This equation together with mass and energy balances for a system form a dynamic multidimensional accounting of a process where there is a change in the particle size distribution. This

المنارة للاستشارات
www.manaraa.com

equation is completely general and is used when the particles are distributed along both external and internal coordinate space. External coordinate space is simply the position x, y, z in Cartesian coordinates. Internal coordinate space x_i is, for example, the shape, chemical composition, and the size of the particles [19].

The asperity population balance for pad wear

The population $\eta_0(L)$ in equation (2.18) can be normalized so that

$$\int_0^\infty \phi(L)dL = 1 \tag{2.19}$$

where $\phi(L)$ is normalized $\eta_0(L)$ which is called population density function (PDF) and it has the units $(\text{size})^{-1}$.

The PDF can be normalized to the original population because it is usually known. For pad asperity problem, our model is one-dimensional, which is z, asperity height. The term due to flow is zero. The birth and death terms are caused by the conditioner because the conditioner will generate some new asperities and also kill some that the diamonds run over. Since for pad wear problem the conditioner is turned off, the birth and death terms are zero. The velocity term in Equation (2.17) should be the wear rate of the asperities defined by Equation (2.10). From Equation (2.10), we can see that taller asperities will be subjected to higher wear rate. The wear rate on asperities will change the asperity PDF, $\phi(z,t)$, with time. CMP is usually operated at constant applied normal pressure, so the load balance causes separation d(t) to vary with time too. Wear of the pad can be formulated by applying the wear rate on the PDF and calculate the time evolution of the PDF. So the asperity population balance to calculate wear of the asperity is

المنارة للاستشارات
www.manaraa.com

$$\frac{\partial \phi(z,t)}{\partial t} = -\frac{\partial}{\partial z}\left(\frac{dz}{dt}\phi(z,t)\right) \tag{2.20}$$

Plugging Equation (2.10) into Equation (2.20), the final form for the PDF evolution with time is obtained

$$\frac{\partial \phi(z,t)}{\partial t} = \frac{c_a 4E^* k_s^{1/2}}{3\pi}\frac{\partial}{\partial z}\left((z-d(t))^{1/2}\phi(z,t)\right) \tag{2.21}$$

Equation (2.3), (2.5), (2.6), (2.13) and (2.21) form the model to predict MRR decay caused by the pad asperities as they wear down. In the next two chapters, an analytical solution to Equation (2.21) is used, instead of numerical solution. In Chapter 3, the accuracy of the derived analytical solution is verified by comparing it with the numerical solution in [18] and experimental MRR data. Chapter 4 uses the mixed-lubrication approach defined in this chapter to show the influence of the fluid to pad wear. Chapter 3 and 4 are independent chapters and have more detailed model derivation and results.

Asperity population balance for pad conditioning

Equation (2.18) is also valid for q(z,t), which is complementary cumulative density function (CCDF) with its corresponding birth and death terms

$$\frac{\partial q(z,t)}{\partial t} + \frac{\partial}{\partial z}(v_w q(z,t)) + \frac{\partial}{\partial z}(cq(z,t)) - B + D = 0 \tag{2.22}$$

where q(z,t) is similar to N(L) in Equation (2.14) which accounts for the total particles (asperities) in an interval

$$q(z,t) = \int_z^\infty \phi(z,t)dz \tag{2.23}$$

www.manaraa.com

and

$$\phi(z,t) = -\frac{\partial q(z,t)}{\partial z} \tag{2.24}$$

The pad conditioning problem is easier to be formulated with CCDF than PDF. In Equation (2.22), v_w is the wear rate because of the polishing operation between pad and wafer and c is the cut-rate because of the conditioner. The birth and death terms, B and D are the asperity generation and killing rate because of the conditioner. The conditioner will generate some new asperities and kill some that the diamond run over. The wear rate v_w term can be derived from Equation (2.10) because of the relationship between CCDF and PDF, Equation (2.23) and (2.24). B and D terms for CCDF are derived in Chapter 5. When CCDF is known, PDF can be derived from Equation (2.24). Chapter 5 will give more detailed model derivation and results. Chapter 6 extends the results in Chapter 4 by including variable diamond heights. In this thesis, only pad conditioning without pad wear is modeled so the second term in Equation (2.22) is set to zero in the pad conditioning modeling. But the derived result for the cut-rate can be used in the modeling for simultaneous pad wear and pad conditioning.

References

[1] J. A. Greenwood and J. B. P. Williamson, "Contact of nominally flat surfaces," *Proc. Roy. Soc. Lond. Math. Phys. Sci.*, vol. 295, no. 1442, pp. 300-319, 1966.

[2] A. M. Scarfo, V. P. Manno, C. B. Rogers, S. P. Anjur, and M. Moinpour, "*In situ* measurement of pressure and friction during CMP of contoured wafers," *J. Electrochem. Soc.*, vol. 152, no. 6, pp. G477-G481, 2005.

[3] J. Lu, C. Rogers, V. P. Manno, A. Philipossian, S. Anjur, and M. Moinpour, "Measurements of slurry film thickness and wafer drag during CMP," *J. Electrochem. Soc*, vol. 151, no. 4, pp G241-G247, 2004.

المنارة للاستشارات

www.manaraa.com

[4] D. Bullen, A. Scarfo, A. Koch, D. P. Y. Bramono, J. Coppeta, and L. Racz, "*In situ* technique for dynamic fluid film pressure measurement during chemical mechanical polishing," *J. Electrochem. Soc.*, vol. 147, no.2, pp 2741-2743, 2000.

[5] L. Shan, J. Levert, L. Meade, J. Tichy and S. Danyluk, "Interfacial fluid mechanics and pressure prediction in chemical mechanical polishing," *J. Tribol.*, vol. 122, pp. 539-543, 2000.

[6] C. F. Higgs III, S. H. Ng, L. Borucki, I. Yoon, and S. Danyluk, "A mixed-lubrication approach to predicting CMP fluid pressure modeling and experiments," *J. Electrochem. Soc.*, vol. 152, no. 3, pp. G193-G198, 2005.

[7] S. H. Ng, L. Borucki, C. F. Higgs III, I. Yoon, A. Osorno and S. Danyluk, "Tilt and interfacial fluid pressure measurements of a disk sliding on a polymeric pad," *Trans. ASME*, vol. 127, pp. 198-205, 2005.

[8] S. Sundararajan, D. G. Thakurta, D. W. Schwendeman, S. P. Murarka, and W. N. Gill, "Two-dimensional wafer-scale chemical mechanical planarization models based on lubrication theory and mass transport," *J. Electrochem. Soc.*, vol. 146, no. 2, pp.761-766, 1999.

[9] J. M. Chen and Y. C. Fang, "Hydrodynamic characteristics of the thin fluid film in chemical-mechanical polishing," *IEEE T Semiconduct. M.,* vol. 15, no. 1, pp. 39-44, 2002.

[10] C. H. Cho, S. S. Park, and Y. Ahn, "Three-dimensional wafer scale hydrodynamic modeling for chemical mechanical polishing," *Thin Solid Film*, vol. 389, pp. 254-260, 2001.

[11] H. J Tsai, Y. R Jeng, and P. Y Huang, "An improved model considering elastic-plastic contact and partial hydrodynamic lubrication for chemical mechanical polishing," *Proc. IMechE, Part J: J. Eng. Tribol.*, vol. 222, pp. 761-770, 2008.

[12] E. J. Abbott and F. A. Firestone, "Specifying surface quality – a method based on accurate measurement and comparison," *Proc. Instn. Mech. Engrs.*, vol. 55, pp. 569-572, 1933.

[13] Y. R. Jeng, and P. Y. Wang, "An elliptical microcontact model considering elastic, elastoplastic and plastic deformation," *Trans. ASME, J. Tribol.*, vol. 125, pp. 232-240, 2003.

[14] W. A. Gross, L. A. Matsch, V. Castelli, A. Eshel, J. H. Vohr and M. Wildmann, *Fluid Film Lubrication*. New York: John Wiley & Sons, 1980.

[15] D. G. Thakurta, C. L. Borst, D. W. Schwendeman, R. J. Gutmann, and W. N. Gill, "Pad porosity, compressibility and slurry delivery effects in chemical-mechanical

المنارة للاستشارات
www.manaraa.com

planarization: modeling and experiments," *Thin Solid Films*, vol. 366 , pp. 181-190, 2000.

[16] J. F. Archard, "Contact and rubbing of flat surfaces," *J. Appl. Phys.*, vol. 24, pp. 981-988, 1953.

[17] W. D. Miner, "*A tool wear comparative study in turning versus computer simulation in 1018 steel*," M.S. thesis, School Tech., Brigham Young Univ., Provo, UT, 2005.

[18] L. J. Borucki, "Mathematical modeling of polish-rate decay in chemical-mechanical polishing," *J. Engr. Math.*, vol. 43, pp. 105-114, 2002.

[19] T. A. Ring, *Fundamentals of Ceramic Powder Processing and Synthesis*. San Diego, CA: Academic Press, 1996.

المنارة للاستشارات

www.manaraa.com

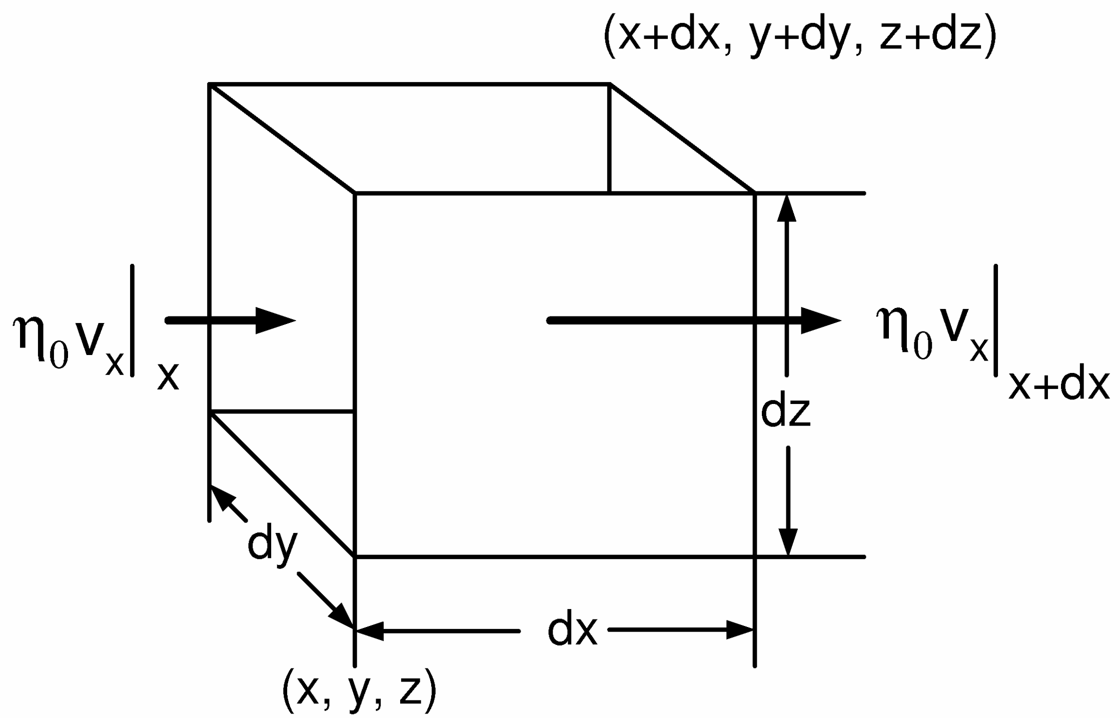


Figure 2.1 Control volume for population balance. Adapted from [19].

المنارة للاستشارات
www.manaraa.com

CHAPTER 3

ANALYTICAL SOLUTION FOR POLISH-RATE DECAY IN CHEMICAL-MECHANICAL POLISHING

Reprinted from J. Eng Math (2010) 68: 207-211, Hong Shi and Terry A. Ring, Analytical solution for polish-rate decay in chemical-mechanical polishing, with permission from Springer

المنارة للاستشارات

www.manaraa.com

J Eng Math (2010) 68:207–211
DOI 10.1007/s10665-010-9369-9

LETTER TO THE EDITOR

Analytical solution for polish-rate decay in chemical–mechanical polishing

Hong Shi · Terry A. Ring

Received: 17 August 2009 / Accepted: 15 March 2010 / Published online: 30 March 2010
© Springer Science+Business Media B.V. 2010

Abstract An analytical solution to the population balance for the heights of pad asperities worn down during chemical–mechanical polishing of silicon wafers is used to calculate the chemical–mechanical removal-rate decay. This analytical solution is compared with experimental data from Journal of Electronic Materials, 25, No. 10 (1996) 1623–1627 by Stein et al. The model results agree with the experimental removal-rate decay results.

Keywords Chemical–mechanical polishing · Wear of pad asperities · Rate decay · Polymer wear

1 Introduction

In chemical–mechanical polishing (CMP), circuit-line-covered silicon wafers are immersed in chemically reactive slurry containing a small weight percent of nano-sized particles and pressed face-down against a polymeric pad. The pad and wafer are rotated in the same direction around their respective and differently located centers. The surface of the pad contains small asperities which hold the abrasive particles and drag them across the pad surface during the polishing operation. Materials are removed by a combination of chemical and mechanical means (i.e., by the action of the reactive slurry that dissolves the wafer surface and the nano-particle abrasive grit) [1,2]. Any local height variations in the wafer surface after layer deposition are subjected to stronger mechanical abrasive forces and are removed faster thus aiding global planarization. The CMP process is highly dependent on the surface condition of the pad. The height distribution of the pad asperities (PDF) together with the processing conditions determine the interface gap and hence determine the chemical reaction rate (the mass transport rate of the reactant and product to and from the wafer surface) [1]. The asperity height distribution also affects the applied load distribution on the wafer and hence influence the mechanical polish rate and within-wafer non-uniformity (WIWNU). Understanding the pad asperity height distribution during polishing is important in controlling many aspects of CMP.

Many experimental observations have shown that the polish rate will drop dramatically if the pad is not conditioned [3–5]. Some have shown a connection between pad surface asperity height distribution and polishing-rate

H. Shi (✉) · T. A. Ring
Department of Chemical Engineering, University of Utah, 50 S. Central Campus Drive, Room 3290, Salt Lake City, UT 84112-9203, USA
e-mail: hongshi888@gmail.com

T. A. Ring
e-mail: ring@eng.utah.edu

المنارة للاستشارات
www.manaraa.com

decay [3,4]. Oliver et al. [3] has shown that the average asperity height (roughness) is continuously reduced with polishing and so is the polishing-rate. Lawing's experimental measurements [4] with interferometry of the pad surface topography have shown that the pad surface-height distribution continuously changes with polishing time in that a second peak at the high-end of the distribution continuously grows with polishing time without conditioning. The conditioner is a diamond-embedded rotating disk dragged with down force across the pad surface in a similar manner to that of the wafer. The conditioner creates asperities on the pad surface. The polishing operation deforms and flattens the pad surface and the conditioner roughens the pad surface, so that these two operations work together to maintain a constantly renewed pad surface during polishing. Because the pad-height distribution is dynamically changing during CMP, understanding how it changes and its effect on the polishing-rate are necessary to control CMP. To the best of our knowledge, Borucki [6] has made the first major step towards this understanding. In his model, he attributes the polishing-rate decay to the wear of the pad asperities and uses a numerical solution to a pad-asperity population balance to calculate the changing of pad asperities with time. The wear rate of individual asperities is modeled by Archard's law [7]. Greenwood and Williamson's contact model with elastic Hertzian contact [8] is used to model the pad asperity and wafer contact. In this work, the asperity population balance equation has been solved analytically by the method of characteristics. This analytical solution is presented and used to calculate the polishing-rate decay that is presented as a numerical solution in [6]. The model results are compared with the numerical solution [6,9] and with experimental data by Stein et al. [5]. The model for the polishing-rate decay agrees with the experimental results.

2 The asperity height population balance

Wear of the pad asperities can be formulated by applying the wear rate, dz/dt, to the PDF and calculate the time evolution of PDF from the population balance model [10].

$$\frac{\partial \phi(z,t)}{\partial t} = -\frac{\partial}{\partial z}\left(\frac{dz}{dt}\phi(z,t)\right) + B - D, \tag{1}$$

where B and D are the birth and death terms for asperities caused by the conditioner. The conditioner will generate some new asperities and also kill some that the diamonds run over. These two terms will be set to zero since the conditioner was turned off in the experimental cases used for comparison.

Using Archard's law for the wear of the asperities for a given load and Greenwood and Williamson for the elastic compliance of the pad for a given load, the individual asperity wear rate is given by:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -k_w U \frac{L}{A} = -c_a \frac{4}{3\pi} E^* k_s^{1/2} (z - d(t))^{1/2}, \tag{2}$$

where $c_a = k_w U$, k_w is the wear rate constant, U is the relative velocity between pad and wafer, L is the load and A is the area of contact between the pad asperity and wafer, $E^* = E/(1-\nu^2)$, E is the pad Young's modulus and ν is the Poisson ratio, k_s is the radius of curvature of the asperities and $(z - d(t))$ is the pad compliance[1] under load which is time dependent due to asperity wear. The pad compliance is determined by balancing the applied pressure with the elastic compliance of all the pad asperities that are in contact with the wafer. Plugging this into (1), the final form for the PDF evolution with time is obtained

$$\frac{\partial \phi(z,t)}{\partial t} = \frac{c_a 4 E^* k_s^{1/2}}{3\pi} \frac{\partial}{\partial z}\left((z - d(t))^{1/2}\phi(z,t)\right), \tag{3}$$

which is the governing equation for asperity wear as a result of the wafer-pad contact during CMP.

[1] $d(t)$ is calculated from $P = \frac{4}{3}\eta_s E^*/k_s^{1/2} \int_{d(t)}^{\infty} (z - d(t))^{3/2}\phi(z)\mathrm{d}z$, where P is the applied pressure and η_s is the asperity area density.

المنارة للاستشارات
www.manaraa.com

3 Wafer material removal rate (MRR)

The average wafer material removal rate (MRR) can also be derived from Archard's Law giving rise to:

$$MRR(t) = c_w \frac{P}{A_f(t)}, \quad (4)$$

where P is the total applied normal pressure, $A_f(t)$ is the contact-area fraction[2] calculated from the pad compliance using Greenwood and Williamson [8] and $c_w = k'_w U$. It is important to note that wear will make *MRR* and A_f time-dependent assuming P and U are constants.

4 Analytical solution of the asperity height population balance

The dynamic analytical solution to the asperity height distribution is obtained by solving (3) using the method of characteristics by assuming that *d*, the pad compliance, is a constant and the initial condition is $\phi(z, t = 0) = \phi_o(z)$. The solution is as follows:

$$\phi(z,t) = \begin{cases} \phi_0(z), & z < d \\ \dfrac{w}{2\sqrt{z-d}}\left(t + \dfrac{2}{w}\sqrt{z-d}\right)\phi_o\left(d + \dfrac{1}{4}w^2\left(t + \dfrac{2}{w}\sqrt{z-d}\right)^2\right), & z > d, \end{cases} \quad (5)$$

where $w = c_a 4E^*\kappa_s^{1/2}/3\pi$. In a later paper [11], Borucki solved the population balance equation (3) analytically. Unfortunately, the analytical solution [11, Eq. 1.2] has a typographical error: the square on w of the term $(w^2/4)$ is missing.

5 Model results and comparison with the experiments

Equations 4 and 5 form a model to predict wafer material removal rate (*MRR*) decay caused by the pad asperities as they wear down. The two coefficients, c_a and c_w, are for the wear of pad asperities and wafer, respectively. They are treated as adjustable parameters and obtained by fitting with experimental results for *MRR*(*t*). The procedure for predicting the removal rate decay, *MRR*(*t*), based on Eqs. 4 and 5, begins by specifying the initial pad-asperity-height distribution $\phi_0(z)$. Next, the load-balancing separation distance between wafer and pad $d(t)$ is calculated, the contact area fraction $A_f(t)$ is then calculated, and finally the *MRR*(*t*) is obtained from (4). The evolution of pad-asperity-height distribution, $\phi(z, t)$ with time due to asperity wear is calculated from (5) at any point in time. The separation $d(t)$, contact area fraction $A_f(t)$ and *MRR*(*t*) are recalculated after each time step. Wear-rate coefficients, c_a and c_w, are treated as adjustable parameters by fitting model *MRR* results with experimental data. For the initial condition, $\phi_o(z)$, a Pearson type IV taken from Stein's experimental measurements on Rodel IC1000 pads [5,9] was used with mean of 0, standard deviation of 15.625, skewness of −1.25, and kurtosis of 6.875 (taken from Stein et al. [5]) as shown in Fig. 1. It should be noted that these values are not those quoted by Borucki et al. in the caption of their Fig. 5 in [6], e.g. skewness of −0.98 and kurtosis of 4.1 because these skewness and kurtosis do not satisfy the criteria for Pearson type IV distribution[3] [12]. In addition, values of the asperity density $\eta_s = 2 \times 10^8/\text{m}^2$, summit curvature $k_s = 2 \times 10^4/\text{m}$, pad modulus $E^* = 119$ MPa, the nominal contact pressure $P = 50$ kPa, wear coefficients $c_a = 2.4 \times 10^{-16}$ m/s/Pa and $c_w = 1.85 \times 10^{-16}$ m/s/Pa, similar to those used in [6]

[2] $A_f(t) = \pi\eta_s/k_s \int_d^{\infty} (z - d(t))\,\phi(z)\,dz$.

[3] From [12], the following condition needs to be met in order to satisfy Pearson-Type-IV distribution:

$$0 < \frac{\text{Sk}^2(\text{Ku}+3)^2}{4(4\text{Ku} - 3\text{Sk}^2)(2\text{Ku} - 3\text{Sk}^2 - 6)} < 1,$$

in which Sk is the skewness and Ku is the kurtosis.

المنارة للاستشارات
www.manaraa.com

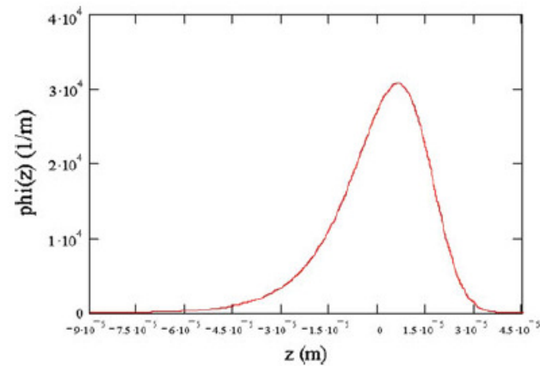


Fig. 1 PDF of asperity heights constructed from data in [5]. The mean of the PDF is 0, and the standard deviation, skewness and kurtosis are 15.6 μm, −1.25, and 6.875

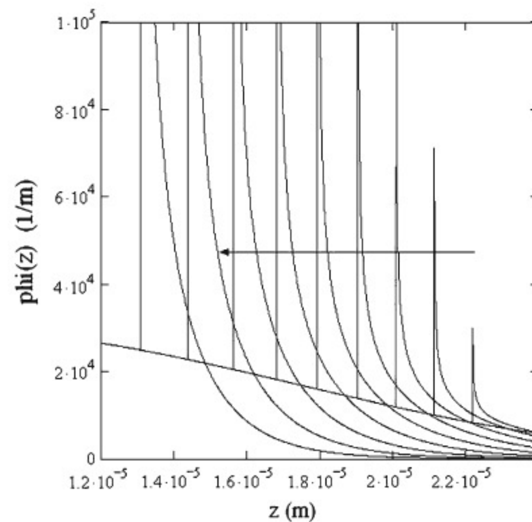


Fig. 2 Evolution of the tail of the initial PDF at 5 min intervals. The *solid lines* are the results obtained by using (5). The first peak is $\phi(z,t)$ after 5 min, and the last is that after 45 min. The *vertical line* is the separation d at each time. Below d, $\phi(z,t)$ keeps the initial population-density function, $\phi_0(z)$. All populations are normalized to the number of asperities in the initial population

and [9] were used in this work. Figure 2 with the analytical solution shows similar result to the numerical solution in [6], except that the sharp peak is broadened by numerical inaccuracy in the numerical solution obtained by a 4th-order Runga-Kutta. In Fig. 2, a piece-wise integrable singularity is present at $z = d(t).z > d(t)$ represents the portion of the pad asperities that are in contact with the wafer and have been worn down by the wafer. With time, this singularity broadens with its amplitude approaching a limiting value which is the fraction of the asperities with height above $d(t)$ at the load-balancing separation. This peak travels to smaller z and hence the load-balancing separation $d(t)$ decreases with time due to asperity wear (Fig. 2).

Figure 3 shows the predicted and measured MRR decay with polishing time. From (4), the polishing-rate decay is attributed to the increase of the total contact area fraction, $A_f(t)$. In our results, $A_f(t)$ increases from 0.31% at 5 min to 0.49% at 45 min polishing time, which are similar to those in [6]. The decrease in $d(t)$ after 8 min of polishing is 1.7 μm versus Borucki's [6] numerical result of 1.8 μm. The decrease in $d(t)$ after 45 min is 10.1 μm versus Borucki's [6] numerical result of 7.9 μm. These differences are due to the sharp peak in the asperity-height distribution above which the asperities are in contact with the wafer. The numerical solution by Borucki does not capture the shape of this sharp peak accurately.

المنارة للاستشارات
www.manaraa.com

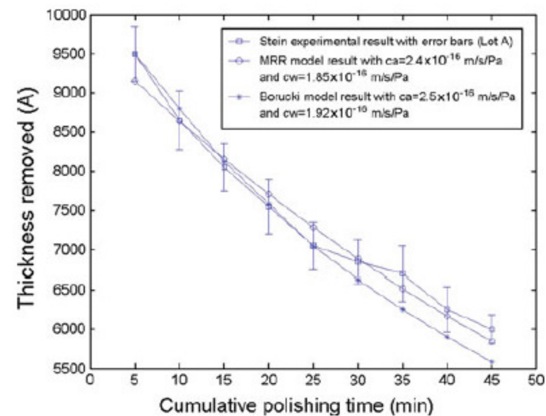


Fig. 3 MRR decay model in this work with our analytic solution versus Stein's experimental result (Lot A) for $c_a = 2.4 \times 10^{-16}$ m/s/Pa and $c_w = 1.85 \times 10^{-16}$ m/s/Pa. Model result predicts experimental trend well. The RMS value (root mean square) of the fit for Stein's experimental data is 17.4 nm compared to 25.6 nm for Borucki's numerical solution [6]

6 Conclusion

The polishing-rate decay predicted by the analytical solution to asperity population balance is similar to the numerical prediction. The analytical solution matches well with the experimental data by Stein and the correlation coefficient for the data fit is better than that for the numerical method. The advantages of the analytical solution are that it is more accurate and numerically faster to compute.

References

1. Sundararajan S, Thakurta DG, Schwendeman DW, Murarka SP, Gill WN (1999) Two-dimensional wafer-scale chemical mechanical planarization models based on lubrication theory and mass transport. J Electrochem Soc 146(2):761–766
2. Chen D, Lee B (1999) Pattern planarization model of chemical–mechanical polishing. J Electrochem Soc 146:744–748
3. Oliver MR, Schmidt RE, Robinson M (2000) CMP pad surface roughness and CMP removal rate. Electrochem Soc Proc 26:77–83
4. Lawing AS (2002) Polish rate, pad surface morphology and oxide chemical mechanical polishing. Electrochem Soc Proc 1:46–60
5. Stein D, Hetherringgton D, Dugger M, Stout T (1996) Optical interferometry for surface measurements of CMP pads. J Electr Mater 25(10):1623–1627
6. Borucki LJ (2002) Mathematical modeling of polish-rate decay in chemical–mechanical polishing. J Eng Math 43:105–114
7. Archard J (1953) Contact and rubbing of flat surfaces. J Appl Phys 24:981–988
8. Greenwood JA, Williamson JBP (1966) Contact of norminally flat surfaces. Proc R Soc Lond A 295(1442):300–319
9. Wang C, Sherman P, Chandra A (2005) A stochastic model for the effects of pad surface topography evolution on material removal rate decay in chemical–mechanical planarization. IEEE Trans Semicond Manuf 18(4):695–708
10. Ring TA (1996) Fundamentals of ceramic powder processing and synthesis. Academic Press, San Diego
11. Borucki LJ, Witelski T, Please C, Kramer PR, Schwendeman D (2004) A theory of pad conditioning for chemical–mechanical polishing. J Eng Math 50:1–24
12. Nie H, Chen S (2007) Lognormal sum approximation with type IV Pearson distribution. IEEE Commun Lett 11(10):790–792

المنارة للاستشارات
www.manaraa.com

CHAPTER 4

CMP PAD WEAR AND POLISH-RATE DECAY MODELED BY ASPERITY POPULATION BALANCE WITH FLUID EFFECT

Reprinted from Microelectronic Engineering 87 (2010) 2368-2375,
Hong Shi and Terry A Ring, CMP pad wear and polish-rate decay
modeled by asperity population balance with fluid effect,
with permission from Elsevier

المنارة للاستشارات

www.manaraa.com

Microelectronic Engineering 87 (2010) 2368–2375

Contents lists available at ScienceDirect

Microelectronic Engineering

journal homepage: www.elsevier.com/locate/mee

CMP pad wear and polish-rate decay modeled by asperity population balance with fluid effect

Hong Shi *, Terry A. Ring

Chemical Engineering Department, University of Utah, 50 S. Central Campus Drive, Room 3290, Salt Lake City, UT 84112-9203, USA

ARTICLE INFO

Article history:
Received 26 November 2009
Received in revised form 29 March 2010
Accepted 7 April 2010
Available online 13 April 2010

Keywords:
Chemical–mechanical polishing
Wear of pad asperities
Rate decay
Pad wear
Polymer wear
Surface statistics
Probability density function
Population balance
Method of characteristics
Reynold's equation
Mixed-lubrication

ABSTRACT

An important task in chemical–mechanical polishing (CMP) is to determine when the pad should be changed or reconditioned. A model which can predict the pad asperity probability distribution function (PDF) during polishing and conditioning is valuable for this purpose. Previous work has been done without incorporating fluid mechanics into the model in L.J. Borucki, J. Eng. Math. 43 (2002) 105–114, but that will overestimate the pad wear because the fluid reduces the load applied on the individual asperities. This work models the wear of pad asperities and polish-rate decay in CMP by coupling the population balance model with fluid mechanics. Modeling results with and without fluid effect are compared. Polish-rate model results are compared with experimental data in D. Stein et al., J. Electron. Mater. 25 (10) (1996) 1623–1627, and the results agree with experimental results for both cases by using different wear rate coefficients to fit experimental data. A lesser wear rate coefficient must be used to fit Stein's data for the fluid case compared to the case without fluid. The wear rate of the pad is calculated from the rate of change of the pad–wafer separation distance during polishing because only asperities above this distance will be in contact and worn down and that portion will be piled up at the pad–wafer separation distance on the PDF curve of the pad asperities. The PDF evolution model results with fluid show much less pad wear compared to the case without fluid. Also if fluid is neglected, pad–wafer separation distance will continue to drop until it is zero, while for the fluid case, it will eventually reach a dynamic steady-state with no pad wear indicating that the load is entirely supported by the fluid. Accurate predicting the interface gap (pad wear) is critical for controlling both mechanical and chemical removal rate and within wafer non-uniformity (WIWNU). These results show that fluid is an important factor which needs to be included into the model in predicting the interface gap and surface profile evolution of the pad during polishing and conditioning.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Chemical–mechanical polishing (CMP) has become the most widely used planarization technology in the semiconductor manufacturing [1] because of its ability to meet the stringent depth-of-focus limitations in advanced integrated-circuit manufacturing. But the mechanism of CMP has not been understood completely and it limits the advanced tool design and better process control. Some of the outstanding issues existing in the CMP process includes high cost, poor process controllability, limited local and global uniformity and defects [2]. Studying the fundamentals of chemical-mechanical polishing can help resolve these issues which can improve production throughput and lower chip fabrication cost.

A schematic CMP process is shown in Fig. 1a. In CMP, circuit-line-covered silicon wafers are immersed in a chemically-reactive slurry containing a small weight percent of nano-sized particles and pressed face-down against a polymeric pad. The pad and wafer are rotating in the same direction around their respective centers. The slurry is delivered to the pad. This slurry contains both reactive chemicals and nano-sized abrasives. The surface of the pad contains small asperities which trap and hold the abrasive particles and drag them across the pad surface during polishing operation. Materials are removed by the combination of chemical and mechanical means (i.e. by the action of the nano particle abrasive grit). The chemical reaction will change the solubility or mechanical properties of the wafer surface and hence the material can be dissolved in the slurry and carried away by the flow or more easily removed by abrasive action [3,4]. There is a local height variation after layer deposition where the taller areas are subjected to stronger mechanical abrasive forces hence are removed faster aiding global planarization.

Cybeq 3900 is a wafer polishing and planarization system used by Stein [5] whose experimental data are used for comparison with the model results in this paper. It is shown in Fig. 1b. This tool is a

* Corresponding author. Tel.: +1 8018280899; fax: +1 8015859291.
E-mail address: hongshi888@gmail.com (H. Shi).

0167-9317/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.mee.2010.04.010

المنارة للاستشارات

www.manaraa.com

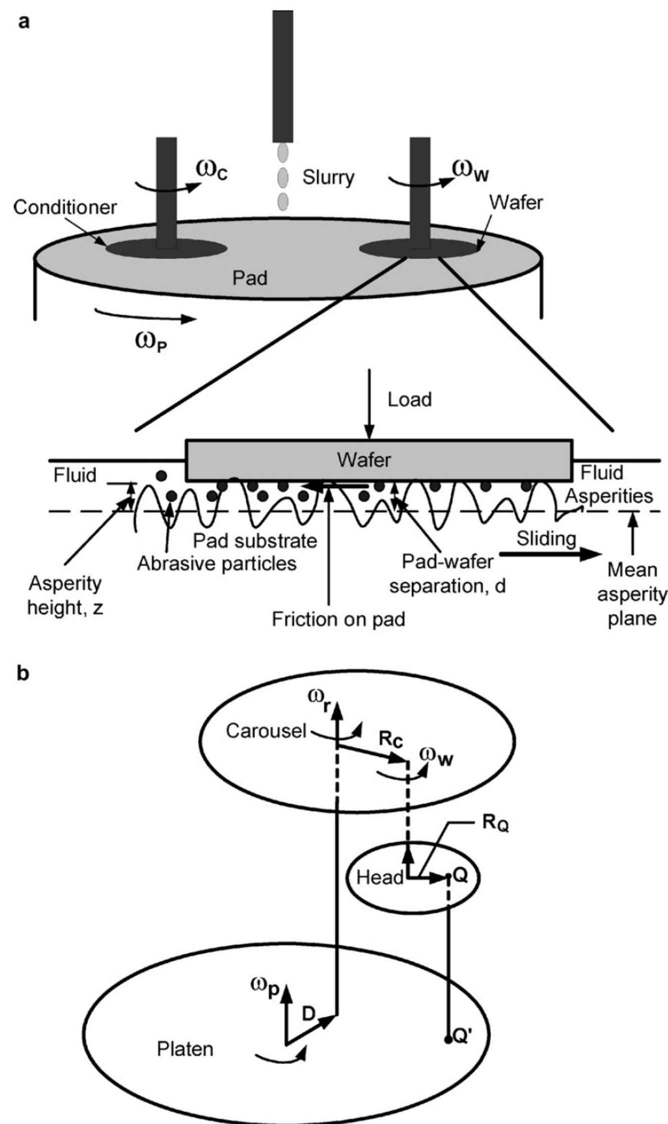


Fig. 1. CMP process description. (a) A schematic of CMP process. The separation d is measured from the mean plane of the pad and asperity heights are also measured from the mean plane of the pad. (b) Geometry of a three-axis polisher. The rotation speed of the carousel, platen and wafer are ω_r, ω_P, and ω_w, respectively. Q is a point fixed to the wafer, and Q' is the projection of Q onto the platen. The relative speed of Q with respect to the platen is $V_Q - V_{Q'} = \omega_r \times R_c - \omega_P \times (R_c + D) + (\omega_r + \omega_w - \omega_P) \times R_Q$. If $\omega_r + \omega_w - \omega_P$ equals zero, then the velocity is the same for all points on the wafer at any given time [7]. For Cybeq 3900, the carousel and platen axis have an offset of 1 in. and the stainless steel platen has a diameter of 32 in. [6].

three-axis wafer polisher consisting of a head (wafer carrier), a carousel, and a platen rotating at the same direction. The rotation rates of the carousel, platen, and head are all independently adjustable. If the platen rotation rate is set equal to the sum of the wafer carrier and carousel rotation rates, all points on a wafer will have the same instantaneous speed relative to the platen during CMP [6,7], and kinematic uniformity is achieved.

The CMP process is highly dependent on the surface condition of the pad. The pad asperity probability distribution function (PDF) together with the process conditions will determine the interface gap and hence determine the chemical removal rate (the mass transport rate of the reactant and product to and from the wafer surface) [3]. The asperity height distribution also affects the applied load distribution on the wafer and hence influences the mechanical polish-rate and within wafer non-uniformity (WIW-NU). Understanding the asperity height distribution during polishing is important in controlling many aspects of CMP.

In CMP, the applied load on the wafer is supported by both pad asperities and fluid. The interface gap between the pad and wafer increases when applied load decreases and/or the relative speed

المنارة للاستشارات
www.manaraa.com

between the wafer and pad increases. The operating regime of CMP can be boundary-lubrication when asperities carry most of the load, mixed-lubrication when load partition between asperities and fluid is comparable, or hydrodynamic regime when load is mostly carried by the fluid. The PDF of the pad, relative speed between the wafer and the pad and the applied load decide the interface gap and the operating regime of CMP. In boundary-lubrication regime, mechanical removal rate is important and on the contrary, in hydrodynamic regime, chemical removal rate (mass transport) dominates. In mixed-lubrication, both chemical and mechanical removal rate play a role.

Many experimental observations have shown that the polish-rate will drop dramatically if the pad is not conditioned [5,8,9]. In addition, some have shown a connection between pad surface asperity height distribution and polish-rate decay [8,9]. Oliver et al. [8] has shown that the average asperity height (roughness) is continuously reduced with polishing and so is the polish-rate. Lawing's experimental measurements [9] with interferometry of the pad surface topography have shown that the pad PDF continuously changes with polishing time in that a second component peak at the high-end of the distribution continuously grows with polishing time without conditioning. These experimental observations indicate that the reduction of the polish-rate is caused by the wear of the pad asperities. For this reason, the pad needs to be regularly dressed with a conditioner. The conditioner is a diamond-embedded rotating disk dragged with down force across the pad surface in a similar manner to that of the wafer. The conditioner will create asperities on the pad surface to enable the abrasive grit to perform their polishing action efficiently. The polishing operation will deform and flatten the pad surface and the conditioner will roughen the pad surface, so that these two operations work together to maintain a constant pad surface during polishing.

Because the pad height distribution is dynamically changing during CMP, understanding and modeling how it changes and its effect on the polish-rate are necessary to control CMP. To the best of our knowledge, Borucki [10] has made the first major step towards this understanding. In his model, he attributes the polishing-rate decay to the wear of the pad asperities and uses a population balance equation to calculate the changing of pad asperities with time. In that paper, Borucki presents a numerical solution to the asperity population balance. The wear rate of individual asperities is modeled by Archard's law [11]. Greenwood and Williamson's contact model with elastic Hertzian contact [12] is used to model the pad and wafer contact. His model successfully captured the experimental data trend by Stein et al. [5] by using two adjustable model parameters to match the experimental data. However, he has neglected the fluid effect in computing the load balance, e.g., the applied load is entirely carried by the pad asperities. This will overestimate the change of the interface gap, making the wear rate higher than the reality. The asperity population balance equation needs to be solved simultaneously with fluid mechanics because CMP is done in the slurry and the wear rate depends upon the asperity PDF. The applied load is balanced by the pad asperities and the fluid. The fluid will reduce the load on the individual asperities resulting in less pad wear compared to the case without fluid. The wear rate of the pad is calculated from the rate of change of the pad–wafer separation distance because only asperities above this distance will be in contact and worn down to the pad–wafer separation distance. Accurately predicting the interface gap (wear rate of the pad) is important in determining both mechanical and chemical removal rate, especially in the case of metal polishing when chemical removal rate (mass transfer rate) can not be neglected. In this work, we present a model of the polish-rate decay by coupling population balance model with fluid mechanics. Pad–wafer contact is modeled by Greenwood and Williamson's model with elastic Hertzian contact [12] and fluid is modeled by Reynold's equation [13]. The results show that fluid is an important factor which needs to be put into the model in predicting the interface gap (pad wear) and surface profile evolution of the pad during polishing and conditioning.

2. Elastic contact model for CMP polish-rate decay

2.1. Contact between wafer and pad

Greenwood and Williamson developed a theory for the contact between a flat surface and a rough surface with Hertzian contact theory [12]. As Fig. 1a shows, the wafer surface is assumed flat and the pad surface is assumed rough. The mean plane of the pad surface is taken as the reference plane. The height z of any asperity relative to the mean plane of the pad surface is assigned a PDF (probability density function), denoted by $\phi(z)$. All asperities are assumed to have spherical summits with identical curvatures $k_s[=1/m]$. The separation distance between the pad mean plane and the wafer surface is denoted by d. An asperity with height z is in contact with the wafer surface when $z > d$. The load on an individual asperity

$$L = \frac{4E^*}{3k_s^{1/2}}(z-d)^{3/2} \tag{1}$$

is given by its compliance, $(z - d)$. This load is distributed over a circular contact area at the tip of the asperity

$$A = \frac{\pi(z-d)}{k_s} \tag{2}$$

where $E^* = E/(1 - \nu^2)$, E is the pad Young's modulus and ν is the Poisson ratio.

Now let us consider the total load carried by all of the pad asperities, N, in contact with the wafer of nominal area A_0 when the asperities are assumed to be uniformly spread over the pad with an area density of $\eta_s = N/A_0$. The nominal asperity pressure, the total load carried by all of the asperities divided by the nominal area A_0 corresponding to the separation d for this elastic model of the rough pad surface in contact with a flat wafer surface is given by

$$P_a = \frac{4}{3}\eta_s E^*/k_s^{1/2}\int_d^\infty (z-d)^{3/2}\phi(z)dz \tag{3}$$

and the ratio of the actual or total asperity contact area to the nominal area (contact area fraction) is given by

$$A_c = \pi\eta_s/k_s\int_d^\infty (z-d)\phi(z)dz \tag{4}$$

Eqs. (3) and (4) form the basis of the Greenwood and Williamson model of rough surface contact. The original derivation of the Greenwood and Williamson model used elastic contact, but any type of contact can be used for the individual asperity and arrive at Eqs. (3) and (4).

2.2. Mixed-lubrication approach

The pad–wafer separation d is found by the load-balancing. In the mixed-lubrication approach, the total applied normal pressure P_{app} between pad and wafer is carried by both asperity pressure P_a and fluid pressure P_f as shown by Eq. (5). We assume that fluid pressure and asperity pressure distribute over the same area without influencing each other. Hence the separation d is found by finding the root of the load-balancing equation, Eq. (5). The asperity pressure is defined by Eq. (3). The fluid pressure is defined by Eq. (6) which is obtained by dimensional analysis of Reynold's equation [13,14]

المنارة للاستشارات

www.manaraa.com

$$P_{app} = P_a + P_f \quad (5)$$

$$P_f = \frac{D\mu U}{4d^2} \quad (6)$$

where D is the wafer diameter, μ is the viscosity of the slurry, U is the linear velocity between the pad and wafer and d is the pad–wafer separation distance.

2.3. Wear of the pad

Wear of the pad asperities can be determined from Archard's law [11,15]

$$dV = -k_w L dx \quad (7)$$

where dV is the volume of material removed in a sliding distance dx, k_w is a wear coefficient which is a function of material hardness, microstructure, and lubrication, and L is the load on the indenter which is the asperity in this case. The wear depth can be estimated from the wear volume by dividing Eq. (7) by the contact area of the asperity

$$dz = \frac{dV}{A} = -k_w \frac{L}{A} dx \quad (8)$$

If we divide Eq. (8) by the increment time dt, then we obtain the wear rate defined in wear depth per time

$$\frac{dz}{dt} = -k_w \frac{L}{A} U \quad (9)$$

where U is the velocity between the wafer and pad.

Next if we define $c_a = k_w U$ and substitute the local contact pressure L/A from Eqs. (1) and (2), we arrive at the wear rate for a single asperity

$$\frac{dz}{dt} = -c_a \frac{L}{A} = -c_a \frac{4}{3\pi} E^* k_s^{1/2} (z-d)^{1/2} \quad (10)$$

where z is the height of the asperity.

An important concept is that L/A in Eq. (10) can be obtained from elastic or inelastic contact models whose formulations depend upon material properties, such as elastic deformation, viscoelastic deformation, viscoplastic deformation and strain hardening/softening effects, etc. In this work, however, only elastic contact is used.

2.4. The asperity height population balance

From Eq. (10), we can see that taller asperities will be subjected to higher wear rate. The wear rate on asperities will change the asperity PDF, $\phi(z, t)$, with time. CMP is usually operated at constant applied normal pressure, so the load balance causes separation $d(t)$ to vary with time too. Wear of the pad can be formulated by applying the wear rate on the PDF and calculate the time evolution of PDF with time from the population balance model [16]

$$\frac{\partial \phi(z,t)}{\partial t} = -\frac{\partial}{\partial z}\left(\frac{dz}{dt}\phi(z,t)\right) + B - D \quad (11)$$

where B and D are the birth and death terms for the asperities which are caused by the conditioner because the conditioner will generate some new asperities and also kill some that the diamonds run over. In comparison with the experimental data available these two terms will be set to zero since the conditioner is assumed to be turned off in these experimental cases.

Plugging Eq. (10) into Eq. (11), the final form for the PDF evolution with time is obtained

$$\frac{\partial \phi(z,t)}{\partial t} = \frac{c_a 4 E^* k_s^{1/2}}{3\pi} \frac{\partial}{\partial z}\left((z-d(t))^{1/2}\phi(z,t)\right) \quad (12)$$

While Borucki derived Eq. (12) by using a probability analysis [8], Eq. (12) can also be directly obtained from a general population balance equation which is readily formulated [16].

2.5. Wafer material removal rate (MRR)

The average material removal rate (MRR) can be derived from Eq. (7) analogously. Eq. (7) is the wear volume abraded from the wafer surface by a single asperity. If we add up the wear volume removed by all of the asperities, the left side of Eq. (7) is the total wear volume and the right side is the total load carried by asperities as Eq. (13) shows. Here we assume every point on the wafer has the same velocity relative to the pad

$$dV_t = k'_w L_a dx \quad (13)$$

where dV_t is the total differential volume removed in a differential time dt and k'_w is the wear coefficient for the abrasive wear of the wafer. Dividing Eq. (13) by dt, the wear rate in terms of total differential volume removed is obtained

$$\frac{dV_t}{dt} = k'_w L_a U \quad (14)$$

Then dividing Eq. (14) by the total contact area $A_c A_0$, the wear rate in terms of thickness removed is obtained

$$\mathrm{MRR}(t) = c_w \frac{P_a}{A_c(t)} \quad (15)$$

where P_a is the asperity pressure, A_c is the asperity contact area fraction and $c_w = k'_w U$. Wear will change asperity PDF, and hence will make d, MRR, P_a, and A_c time-dependent assuming P_{app} and U are constants. In [10], Borucki uses the normal applied pressure P_{app} for P_a in Eq. (15) because the fluid is neglected. In that case, P_{app} equals to P_a and P_f equals to zero everywhere. Motivated by this approach and considering the fluid effect, we use P_a instead of P_{app}. Fundamentally, Eq. (15) is for mechanical removal rate. Hence, using it for the case where chemical removal can be neglected is more accurate. In Stein's experiment, silicon dioxide wafer is polished in which chemical removal is less important, hence Eq. (15) is accurate model.

The data that we use for comparison with Stein's experimental data are the total thicknesses removed over successive time intervals. The total amount removed $TR(t_0, t_1)$ between time interval t_0 and t_1 can be obtained by integration of the instantaneous removal rate MRR(t) over this time interval as shown in Eq. (16)

$$TR(t_0, t_1) = \int_{t_0}^{t_1} \mathrm{MRR}(t) dt \quad (16)$$

2.6. Analytical solution of the asperity height population balance

The dynamic analytical solution to the asperity height population balance equation is obtained using the method of characteristics (see Appendix) by assuming a given d and an initial condition $\phi(z, t = 0) = \phi_0(z)$. The solution is shown in Eq. (17)

$$\phi(z,t) = \begin{cases} \phi_0(z) & z < d \\ \frac{w}{2\sqrt{z-d}}\left(t + \frac{2}{w}\sqrt{z-d}\right)\phi_0\left(d + \frac{1}{4}w^2\left(t + \frac{2}{w}\sqrt{z-d}\right)^2\right) & z > d \end{cases} \quad (17)$$

where $w = 4c_a E^* \kappa_s^{1/2}/(3\pi)$.

3. Model result and comparison with the experiment

Eqs. (3)–(6), (15)–(17) form a model to predict material removal rate (MRR) decay caused by the pad asperities as they wear down.

المنارة للاستشارات

www.manaraa.com

The two wear rate coefficients, c_a and c_w, are for the wear of pad asperities and wafers, respectively. They are treated as adjustable parameters and are obtained by matching modeled removed thickness results with experimental results. The procedure for predicting the removal rate decay, MRR(t), begins by specifying the initial pad asperity height distribution $\phi_0(z)$. Next, the load-balancing separation distance between wafer and pad $d(t)$ is calculated from load-balancing equation, Eq. (5), the contact area fraction $A_c(t)$ is calculated from Eq. (4), and then the MRR(t) is obtained from Eq. (15). The evolution of the pad asperity height distribution, $\phi(z, t)$ with time due to asperity wear is calculated from Eq. (17) at any point in time. The separation $d(t)$, contact area fraction $A_c(t)$ and MRR(t) are calculated after each time step of interest.

Wear rate coefficients, c_a and c_w, are treated as adjustable parameters by matching modeled removed thickness results with experimental results. For the initial condition, $\phi_0(z)$, a Pearson type IV distribution taken from Stein's experimental measurements on Rodel IC1000 pads [5,17] was used with mean of 0, standard deviation of 15.625, skewness of −1.25, and kurtosis of 6.875 as shown in Fig. 2. The skewness and kurtosis of −0.98 and 4.1, respectively, shown in the caption of Fig. 5 in [10] are very likely a typographical error because these skewness and kurtosis do not satisfy the criteria for Pearson type IV distribution [17,18]. In addition, we used values of the asperity density $\eta_s = 2\text{e}8/\text{m}^2$, summit curvature $k_s = 2\text{e}4/\text{m}$, pad modulus $E^* = 119$ MPa, the nominal applied pressure $P_{app} = 50$ kPa, similar to those used in [5,10,17].

Stein's experiment was done with Cybeq 3900 wafer polishing and planarization system as shown in Fig. 1b. In [5], the platen, carousel, and head speeds were 20, 15, and 5 rpm, respectively. Then the relative linear speed between wafer and pad is $U = 0.153$ m/s for Stein's experiment [6,7,19]. We used this same speed in the modeling of the fluid case. For the viscosity of the slurry, we used $\mu = 0.0016$ Pa S which is the viscosity for conventional silica slurry at 25 °C [20]. The wafers polished were unpatterned 150 mm diameter silicon oxide wafers as in [5].

Wear rate coefficients $c_a = 2.4\text{e}(-16)$ m/s/Pa and $c_w = 1.85\text{e}(-16)$ m/s/Pa are used for no fluid case and $c_a = 2.1\text{e}(-16)$ m/s/Pa and $c_w = 1.88\text{e}(-16)$ m/s/Pa are used for fluid case for the best fits of the modeled removed thickness results to Stein's experimental data for Lot A [5] as shown in Fig. 3. The RMS values (root mean square) for the best fits to the experimental data for both cases are very close which are 19.3 nm and 21.1 nm for no fluid case and fluid case, respectively. Since the error bars on this experimental data are ±400 Å, the RMS values for with fluid and without fluid fits are essentially the same. So our results match Stein's experimental data very well for both cases by using different wear rate coefficient c_a. A smaller wear rate coefficient c_a must be used

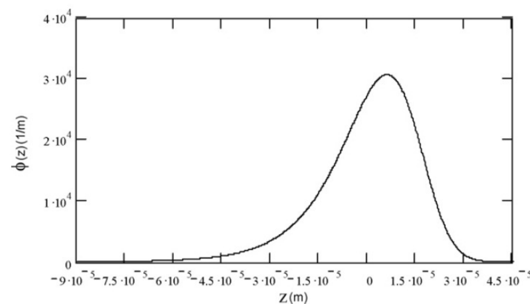


Fig. 2. PDF of asperity heights constructed from data in Stein [5,17]. The mean of the PDF is 0, and the standard deviation, skewness and kurtosis are 15.6 μm, −1.25, and 6.875.

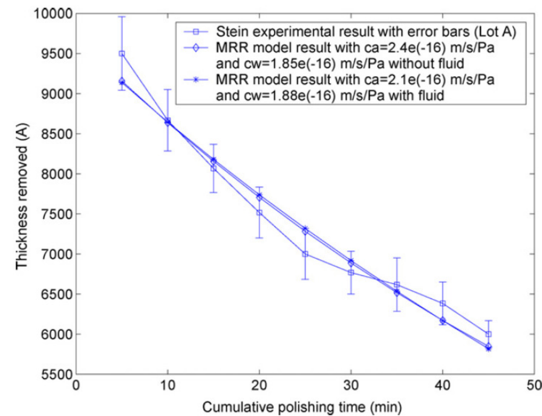


Fig. 3. Removed thickness results calculated with the present removal-rate model versus Stein's experimental results (Lot A) at successive 5-min intervals. Model results predict experimental trend well for both cases. A smaller wear coefficient C_a needs to be used to fit the data for the fluid case compared to the case without fluid. The RMS values (root mean square) of the best fits of our prediction to Stein's experimental data are very close for both cases with 19.3 nm and 21.1 nm for no fluid and fluid case, respectively.

in order to fit the experimental data for fluid case compared to the case without fluid. Fig. 4 shows the time evolution of PDF due to pad wear for both cases with and without fluid. In Fig. 4, a peak is present at $z = d$, and $z > d$ represents the portion of the pad asperities that are in contact with the wafer and have been worn down by the wafer. The location of the peak is at the load-balancing separation d at each time. This peak moves to smaller z with time and hence the load-balancing separation d decreases with time due to asperity wear. The change of d with time is plotted in Fig. 5. A significant feature to mention in Figs. 4 and 5 is that the pad–wafer separation d decreases much slower for the case with fluid compared to the case without fluid indicating less pad wear. This is because fluid reduces wear by taking some load off the asperities. After 45 min, the decrease of d is 6.5 μm which is only 36% of that for the case without fluid which is 10.14 μm as shown in Figs. 4 and 5. Without fluid, d continues to decrease until it is zero and the model will stop working eventually. This indicates the pad is worn down to a flat surface. According to the simulation, this happens at about 78 min of cumulative wear time. This is not the real physical situation because as the surface becomes flat, the CMP process moves to hydrodynamic lubrication and the gap should be determined by the fluid mechanics. Our model with fluid shows significant improvement over the non-fluid case because it clearly shows this trend as shown in Fig. 5. After about 100 min of wear time, d reaches a dynamic steady-state which is 13.6 μm at the applied pressure and velocity for Stein's experiment. At that point, the fluid pressure $P_f = P_{app} = 50$ kPa and the asperity pressure is trivial. This is shown in Fig. 6. As wear progresses, the role of the fluid becomes more and more important until P_f equals to P_{app}, while solid contact becomes less and less important until the asperity pressure becomes trivial. Wear has changed the operating regime of CMP from boundary-lubrication to hydrodynamic lubrication. This transition can not be seen if fluid is neglected as in the wear model in [10]. So our model has shown that in order to accurately predict the interface gap when wear is present, fluid must be taken into account. The closer the process condition is to hydrodynamic regime, the faster the transition to steady-state d occurs with wear. This is especially important for copper polishing where small P_{app} and/or faster velocity is generally used and wear quickly pushes the operating regime from mixed-lubrication

المنارة للاستشارات
www.manaraa.com

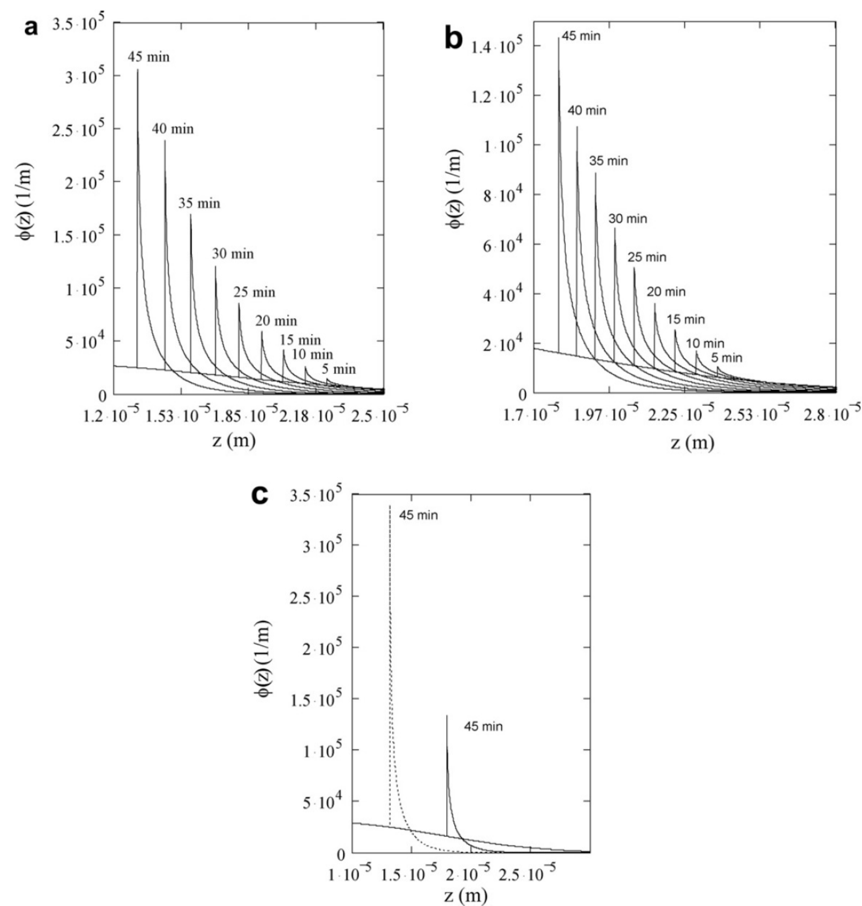


Fig. 4. Evolution of the tail of the PDF in Fig. 2 at 5-min intervals using Eq. (17) at Stein's experimental condition. (a) PDF without fluid. (b) PDF with fluid. (c) PDF with fluid and without fluid together after 45 min of cumulative wear time. The dot line is the case without fluid.

to hydrodynamic lubrication in which only the chemical removal mechanism (mass transfer) is left.

Another interesting result is the changing of the contact area fraction $A_c(t)$ which is shown in Fig. 7 for both with fluid and without fluid cases. Polish-rate decay is attributed to the reduction of the average contact pressure according to Eq. (15). This is true for both cases according to the model results. If fluid is neglected, $P_a = P_{app}$. Wear changes PDF and makes d smaller which results in increasing $A_c(t)$ with time. According to the results, $A_c(t)$ increases from 0.31% at 5 min to 0.49% at 45 min of polishing time similar to Borucki's in [10]. If fluid is considered, pad wear continuously reduces P_a and increases P_f. The load balance between P_a and P_f will make $A_c(t)$ either increase or decrease depending on the magnitude of wear coefficient c_a and the operating regime of CMP determined by the process conditions. As Fig. 7 shows, at Stein's experimental conditions, the contact area will increase slightly in the beginning and then level off and finally decrease. The reason that $A_c(t)$ increases a little in the beginning is because the starting point is boundary-lubrication when P_a is much larger than P_f for Stein's experiment. The increase of $A_c(t)$ in the beginning is similar to the case without fluid. The decrease of $A_c(t)$ later is because wear has changed the operating regime of CMP by shifting the load to P_f. If the starting point is in mixed-lubrication, then $A_c(t)$ will start to decrease from the beginning. Also at larger c_a, $A_c(t)$ will decrease faster. No matter if $A_c(t)$ increases or decreases, the average contact pressure P_a/A_c always decreases with wear, so does the removal rate as shown in Fig. 3.

4. Conclusion

This paper models the wear of pad asperities and polish-rate decay during CMP of silicon oxide wafers by coupling the population balance model with fluid mechanics. Modeling results with and without fluid effect are compared. Polish-rate model results are compared with experimental data by Stein et al., and the results agree with experimental results for both cases by using different wear rate coefficients to fit the experimental data. A smaller wear rate coefficient must be used to fit Stein's data for fluid case compared to the case without fluid. The wear rate of the pad is calculated from the rate of change of the pad–wafer separation distance during polishing because only asperities above this distance will be

المنارة للاستشارات
www.manaraa.com

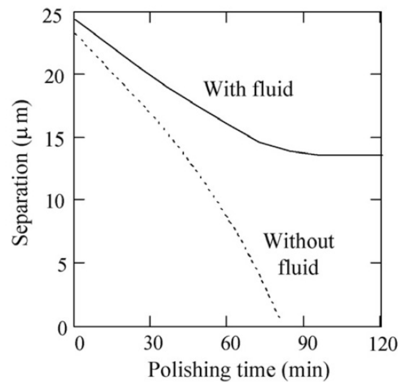


Fig. 5. The decreasing of separation d with time due to wear at Stein's experimental condition. The solid line is the case with fluid and the dot line is the case without fluid. The difference for the separation between the two cases increases with time due to wear. After about 78 min, the separation becomes about zero for the case without fluid and the model stops working eventually. While for the fluid case, a dynamic steady-state d has reached after about 100 min of cumulative wear time indicating that wear has stopped and entirely hydrodynamic regime has been reached, e.g., the load is entirely supported by the fluid.

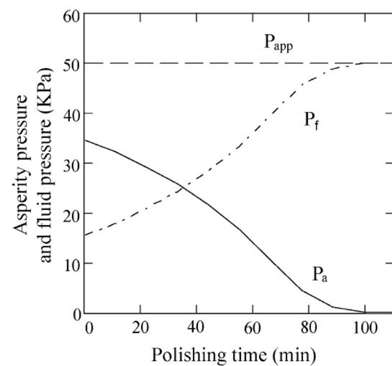


Fig. 6. The changing of pad asperity pressure and fluid pressure with time due to wear at Stein's experimental condition. The dash-dot line is the fluid pressure, the solid line is the asperity pressure, and the dash line is the normal applied pressure.

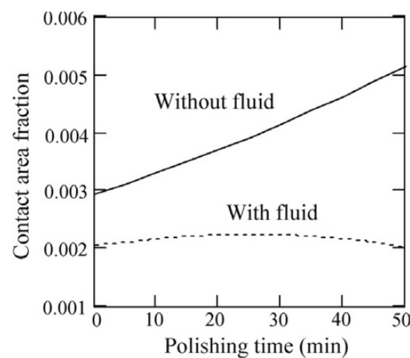


Fig. 7. The changing of the pad contact area fraction with time due to wear at Stein's experimental condition. The dot line is the case with fluid. The solid line is the case without fluid.

in contact and worn down and that portion will pile up at the pad–wafer separation distance on the PDF curve of the pad asperities. The results with fluid show much less pad wear because the decrease of pad–wafer separation distance is only 36% of that without fluid after 45 min of cumulative wear time at Stein's experimental conditions. During this time, wear has driven the operating regime of CMP from boundary-lubrication to hydrodynamic lubrication by shifting the load to P_f. This reflects in the pad–wafer separation distance and contact area fraction, both of which have shown increasing difference comparing the cases without fluid and with fluid during 45 min of wear time at Stein's experimental condition. Another important result from our model is that if fluid is neglected, pad–wafer separation distance will continue to drop until it is zero. While for the fluid case, it will reach a dynamic steady-state with no pad wear eventually, indicating that entirely hydrodynamic regime has reached where only chemical removal mechanism is left which is the real physical situation. The closer the process condition is to hydrodynamic regime, the faster the transition to dynamic steady-state d occurs with wear. This is especially important for copper polishing where small P_{app} or faster velocity is generally used and wear quickly pushes the operating regime from mixed-lubrication to hydrodynamic lubrication in which only chemical removal mechanism (mass transfer) is left.

Polish-rate is a sensitive function of pad surface PDF. Accurately estimating the interface gap (pad wear) is critical for controlling both mechanical and chemical removal rate. As we have shown in this work, population balance equation needs to be solved with fluid mechanics to accurately predict the interface gap and pad wear. For stable removal rate (both chemical and mechanical), constant interface gap resulting from constant pad surface needs to be maintained. Accurately estimating pad wear is also important to decide when the pad needs to be changed. Pad wear and pad conditioning work together to maintain a constant pad surface. Model without fluid effect overestimates the pad wear. It will result in over-conditioning of the pad, and making the task of maintaining a constant pad surface difficult.

Appendix A

The procedure to solve Eq. (12) is briefly explained in the following.

Rearrange Eq. (12), Eq. (A1) is obtained.

$$\frac{\partial \phi(z,t)}{\partial t} - w\sqrt{z-d}\frac{\partial \phi(z,t)}{\partial z} = \frac{w}{2\sqrt{z-d}}\phi(z,t) \qquad \text{(A1)}$$

Eq. (A1) can be broken up to three first-order, ordinary differential equations by using a new variable, s. All of the original variables are transformed by this new variable, e.g., $t(s)$, $z(s)$ and $\phi(s)$, and the original PDE becomes ODE as shown in Eqs. (A2)–(A4)

$$\frac{dt}{ds} = 1 \qquad \text{(A2)}$$

$$\frac{dz}{ds} = -w\sqrt{z-d} \qquad \text{(A3)}$$

$$\frac{d\phi}{ds} = \frac{w}{2\sqrt{z-d}}\phi(z,t) \qquad \text{(A4)}$$

The above three equations can be solved with initial conditions, e.g., at $s = 0$, $t(s) = 0$, $z(s) = z_0$, and $\phi(s) = \phi_0(z_0)$. Solving Eq. (A2), $s = t$ is obtained. Solving Eq. (A3), the characteristic curve, Eq. (A5), is obtained. Finally, by solving Eq. (A4) and substitute z_0 using Eq. (A5), the final solution is obtained which is Eq. (17)

$$z_0 = \left(\frac{w}{2}t + \sqrt{z-d}\right)^2 + d \qquad \text{(A5)}$$

المنارة للاستشارات
www.manaraa.com

References

[1] J. Su, X. Chen, X. Zhang, J. Du, D. Guo, J. Comput. Theor. Nanos 5 (8) (2008) 1656–1660.
[2] L. Shan, Mechanical interactions at the interface of chemical mechanical polishing, Ph.D. Thesis, Georgia Institute of Technology, 2000, p. 13.
[3] S. Sundararajan, D.G. Thakurta, D.W. Schwendeman, S.P. Murarka, W.N. Gill, J. Electrochem. Soc. 146 (2) (1999) 761–766.
[4] D. Chen, B. Lee, J. Electrochem. Soc. 146 (1999) 744–748.
[5] D. Stein, D. Hetherringgton, M. Dugger, T. Stout, J. Electron. Mater. 25 (10) (1996) 1623–1627.
[6] P. Wrschka, J. Hernandez, Y. Hsu, T.S. Kuan, G.S. Oehrlein, H.J. Sun, D.A. Hansen, J. King, J. Electrochem. Soc. 146 (7) (1999) 2689–2696.
[7] R. Kolenkow, R. Nagahara, Solid State Technol. (June) (1992) 112–114.
[8] M.R. Oliver, R.E. Schmidt, M. Robinson, Proc. Electrochem. Soc. 26 (2000) 77–83.
[9] A.S. Lawing, Proc. Electrochem. Soc. 1 (2002) 46–60.
[10] L.J. Borucki, J. Eng. Math. 43 (2002) 105–114.
[11] J. Archard, J. Appl. Phys. 24 (1953) 981–988.
[12] J.A. Greenwood, J.B.P. Williamson, Proc. R. Soc. Lond., Ser. A 295 (1442) (1966) 300–319.
[13] W.A. Gross, L.A. Matsch, V. Castelli, A. Eshel, J.H. Vohr, M. Wildmann, Fluid Film Lubrication, John Wiley & Sons, New York, 1980.
[14] D.G. Thakurta, C.L. Borst, D.W. Schwendeman, R.J. Gutmann, W.N. Gill, Thin Solid Films 366 (2000) 181–190.
[15] W.D. Miner, A tool wear comparative study in turning versus computer simulation in 1018 steel, M.S. thesis, 2005, Eq. (3.10), p. 39.
[16] T.A. Ring, Fundamentals of Ceramic Powder Processing and Synthesis, Academic Press, Inc., 1996. Eq. (3.7), p. 87.
[17] C. Wang, P. Sherman, A. Chandra, IEEE Trans. Semiconduct. 18 (4) (2005) 695–708.
[18] H. Nie, S. Chen, IEEE Commun. Lett. 11 (10) (2007) 790–792.
[19] Personal Communication with David Stein.
[20] T. Nishioka, K. Sekine, Y. Tateyama, Interconnect technology, in: IEEE International Conference, 24–26 May 1999, pp. 89–91.

المنارة للاستشارات

www.manaraa.com

CHAPTER 5

PLASTIC DEFORMATION THEORY OF CHEMICAL-MECHANICAL POLISHING PAD CONDITIONING

Submitted to *Journal of Engineering Mathematics*

المنارة للاستشارات

www.manaraa.com

Abstract

In chemical-mechanical polishing, the pad asperities hold abrasive grits in the polishing slurry to polish wafers. Without conditioning, pad asperities will be continuously worn down by polishing of the wafers and it results in reduced wafer polish-rate. Conditioning is a very important process to create pad asperities on the pad surface to enable abrasive grits to perform polishing efficiently. The polishing operation will flatten the pad surface because of the wear of the pad asperities and the conditioner will roughen the pad surface so that these two operations work together to maintain a constant pad surface. A constant pad surface is important to maintain a stable polish-rate. A model that can predict the pad asperity probability distribution function (PDF) during polishing and conditioning is valuable for this purpose. In this work, a population balance model (PBE) for pad asperities is used to model the pad conditioning process. Previous work has been done without incorporating pad plastic deformation into the model in Borucki L, et al. (2004), J Eng Math 50:1-24, but that will overestimate pad cut-rate during conditioning because the model assumes that all of the material being cut are removed from the pad surface. In their PBE conditioner model, only a "death" term is used. Pad cut-rate is important to estimate the aggressiveness of the pad conditioning, decide the time needed to bring the pad surface into steady-state in the pad "break-in" process and the time to change the pad. Plastic deformation is important in modeling the pad conditioning because the raised ridges along the sides of the grooves are the high points (asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive particles used to polish the wafer. In this work, plastic deformation is modeled by adding a "birth" term by mass balance to the population

www.manaraa.com

balance model to take account of the material pileups (plastic deformation) during pad conditioning. This material pileup is modeled by a f_{add} parameter to represent the fraction of the material removed that is added back to the pad surface at the edge of the groove cut by the diamond. The derived steady-state pad surface conditioning profile only depends on the diamond cutting depth into the pad which depends on the load applied to the conditioner. The cut-rate derived from the model depends on both diamond cutting depth and f_{add}. Larger load will give more pad roughness (larger polish-rate) and larger cut-rate (shorter pad life). Larger f_{add} will give smaller cut-rate. The model can be used to model a series of pad materials with plastic properties because f_{add} is a variable parameter. The analytical solutions derived for the pad surface conditioning profile and cut-rate agree with the Monte Carlo simulations.

Introduction

Chemical mechanical polishing (CMP) has become the most widely used planarization technology in the semiconductor manufacturing [1] because of its ability to meet the stringent depth-of-focus limitations in advanced integrated-circuit manufacturing. But the mechanism of CMP has not been well understood and it limits the advanced tool design and better process control. Some of the outstanding issues existing in the CMP process includes high cost, poor process controllability, limited local and global uniformity and defects [2]. Studying the fundamentals of chemical mechanical polishing can help resolve these issues which can improve production throughput and lower chip fabrication cost.

In CMP, circuit-line-covered silicon wafers are immersed in a chemically-reactive slurry containing a small weight percent of nano-sized particles and pressed face-down

المنارة للاستشارات
www.manaraa.com

against a polymeric pad as shown in Figure 5.1. The pad and wafer both rotate in the same direction around their respective centers. The slurry is delivered to the pad. This slurry contains both chemicals and nano-sized abrasives. The surface of the pad contains small asperities which hold the abrasive particles and drag them across the pad surface during polishing operation. Materials are removed by the combination of chemical and mechanical means (i.e., by the action of the nanoparticle abrasive grit). The chemical reaction will change the solubility or mechanical properties of the wafer surface and hence the material can be dissolved in the slurry and carried away by the flow or more easily removed by abrasive action [3, 4]. There is a local height variation after layer deposition where the taller areas are subjected to stronger mechanical abrasive forces hence are removed faster aiding global planarization.

The CMP process is highly dependent on the surface condition of the pad. The height distribution of the pad asperities described by a probability distribution function (PDF) together with the process conditions will determine the interface gap and hence determine the chemical reaction rate (the mass transport rate of the reactant and product to and from the wafer surface) [3]. The asperity height distribution also affects the applied load distribution on the wafer and hence influence the mechanical polish rate and within wafer non-uniformity (WIWNU). Understanding the asperity height distribution during polishing is important in controlling many aspects of CMP including removal rate and scratching rates.

Many experimental observations have shown that the polish rate will drop dramatically if the pad is not constantly conditioned [5 - 7]. In addition, some have shown a connection between pad surface asperity height distribution and polishing-rate

المنارة للاستشارات
www.manaraa.com

decay [5, 6]. Oliver [5] has shown that the average asperity height (roughness) is continuously reduced with polishing and so is the polishing-rate. Lawing's experimental measurements [6] with interferometry of the pad surface topography have shown that the pad surface height distribution continuously changes with polishing time in that a second component peak at the high-end of the distribution continuously grows and shifts to the low-end of the distribution with polishing time without conditioning. This is because taller asperities that perform most of the polishing have been worn down during the polishing. Hence another process step is introduced in CMP - pad conditioning. The conditioner is a diamond-embedded rotating disk dragged with down force across the pad surface in a similar manner to that of the wafer. The conditioner creates asperities on the pad surface to enable the abrasive grit to perform their polishing action efficiently. The polishing operation deforms and flattens the pad surface and the conditioner roughens the pad surface, so that these two operations work together to maintain a constant pad surface during polishing. Because the pad height distribution is dynamically changing during CMP, understanding and modeling how asperity height changes and its effect on the polishing-rate are necessary to control CMP. Borucki [8] has presented a model for conditioning and wear of solid pads and foamed pads. In his conditioner model, he made an assumption that the diamond grits create furrows on the pad surface by cleanly cutting the pad material, e.g., a polishing operation in which all the material cut by the diamonds is removed from the surface of the pad. That conditioner model is similar to the process of polishing brittle materials where materials are removed by brittle fracture. However, polyurethane which is commonly used as the CMP pad material is a polymer with plastic mechanical properties (in which material is simply pushed aside and not removed from

المنارة للاستشارات
www.manaraa.com

the surface) that need to be considered in the conditioner model. In this work, we treat pad as a plastic material and have developed a conditioner model by combining the cutting and plastic deformation in which a groove with pile-up on both side-walls forms as the diamond moves over the pad surface. Plastic deformation is important in modeling the pad conditioning because the raised ridges along the sides of the grooves are the high points (asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive particles used to polish the wafer. In [8], Borucki uses a population balance equation with only a "death term" to model the cleanly cutting of the pad surface. In this work we add another term, a "birth term", by using the concept of a mass balance to represent the material piled-up on both side-walls of the groove. Consequently the pad material should be consumed much more slowly by adding this term than by solving the population balance equation with only the death term in which all of the materials cut by the diamond is removed from the pad surface.

Model

Abrasive wear model for the single diamond dressing of CMP polyurethane pad

There are two general types of polymers based on their mechanical properties: ductile polymers and brittle polymers. Epoxy is a good example of brittle ones and polyethylene and polyurethane are good examples of ductile ones [9, 10]. Ductile materials can be deformed without failure. These properties are contrasted in the stress-strain curves shown in Figure 5.2. In the beginning of the curve, there is a linear (elastic) region with a slope given by the Young's modulus (tensile modulus) for that material. Above the linear region is the plastic region and ductile material has significant

المنارة للاستشارات
www.manaraa.com

elongation before the point of the rupture [9, 11]. The first point on the plot of Figure 5.2 where the slope is zero is called the "yield point". Strain at the yield point is called "elongation at yield" (ε_Y). Stress at the yield point is called "yield strength" (σ_Y). Elongation continues until the sample breaks. Strain at the break point is called "elongation at break" (ε_B). Stress at the break point is called "tensile strength" (σ_B) [12]. The brittle material has little elongation beyond the elastic zone before the rupture and hence the total elongation for brittle materials is very small before they break [9 - 12]. Thus from Figures 5.2 (a) and 5.2 (b) we can see the difference between ductile materials and brittle materials. Polyurethanes are ester-amide derivatives of carboxylic acids. They are also called polycarbamates, from carbamic acid – R_2NHCO_2H. They are used in a wide variety of applications including elastic fibers, adhesives, coatings, elastomers, both flexible and rigid foams and CMP pads [13]. Because of the inherent flexibility of the urethane chain, the urethane polymer has a wide window of formulation flexibility. Based on the formulation and manufacture method, the mechanical properties of polyurethane can vary from 300% elongation on adhesives used in applications where ductility is desired, to only 50% elongation in structural applications where higher strength and stiffness are desired [14]. Epoxy shows brittle behavior (no yield point) and percent elongation at failure is about 3% [14]. The percent elongation at failure for glass is 0.12%. Copper is a ductile metal and the percent elongation at failure is 53% [11].

Scratch testing is a very popular tool for understanding material deformation and removal mechanisms in abrasive wear. Abrasive wear can be defined as wear due to the penetration of hard asperities into the softer solid surface in dynamic contact [10]. The scratch test is defined as a mechanical deformation process where a controlled force or

المنارة للاستشارات
www.manaraa.com

displacement is exerted on a hard spherical tip to indent to a polymer substrate and move across its surface at a certain speed [9]. There are two main types of damage found in polymers: ductile damage (e.g., shear yielding and ironing) and brittle damage (e.g., crazing and cracking). Their occurrences depend on the material characteristics (ductile/brittle) and the applied stress. Examples of the two types of damage from scratch test results [10] are shown in the SEM images in Figure 5.3. The cross-section of the created groove in Figure 5.3 (a) is drawn schematically in Figure 5.4. In Figure 5.4, the areas A_1 and A_2 are the two side-ridge areas, and A_V is the groove area. A parameter we have defined in the modeling of the abrasive wear is f_{add} [10, 15], which is given by

$$f_{add} = \frac{A_1 + A_2}{A_V} \tag{5.1}$$

and consists of a ratio of the material added back above the original surface due to a scratch to the material removed below the original surface.

In practice, only part of the volume of the wear groove is plastically deformed making edges of the groove, the material pileups. It should be noted that f_{add} is equal to zero for ideal micro-cutting (brittle fracture), and equal to unity for ideal micro-plowing (plastic deformation). For polymers with high ductility (percent elongation at failure), f_{add} is close to one [10]. For brittle polymers, f_{add} is close to zero.

In this work, we will use f_{add} to model the plastic deformation. Plastic deformation is important to incorporate into the conditioner model because the raised ridges are the high points (asperities) on the pad that pressurize the wafer and polish away the contacting region of the wafer. Another feature that needs to be mentioned in Figure 5.3

المنارة للاستشارات

www.manaraa.com

is that for the intersecting scratch test, at the four corners of the intersection, the intersections are higher because the ridge created in the first scratch is raised further in the second scratch. In other words, these ridges created at intersections become the highest points on the pad surface. As a consequence, the ridge intersections will make contacts with the wafer and pressurize the rotating wafer causing the polishing to take place at these contacting regions.

Pad conditioning problem

We start with the basic assumptions and model in [8] and then we will derive our plastic deformation model and compare it with the brittle fracture model in [8].

The top view of a CMP tool geometry is shown in Figure 5.5 (a). Like the wafer, the conditioner is a rotating disk pressed onto a rotating pad-surface. The conditioning disk has a much smaller radius r_c and is periodically swept over a fixed radial track with a constant translational speed U between points P_1 and P_2 and constant rotational speed ω_C as shown in Figure 5.5 (a). The length between P_1 and P_2 is P_1P_2. Conditioning disks are covered with diamond cutting tips distributed randomly over the disk with controlled heights. As the conditioner traverses the track, it will make random cuts on the pad surface. The pad will rotate several times during one conditioner trip between P_1 and P_2. To model the conditioning process, we need to find out how many cuts the conditioner will create at the end of one pad rotation. These cuts are distributed between points P_1 and P_2. Hence we can calculate the long time averaged effect for the total cuts between P_1 and P_2 which is equal to the number of pad rotation times the cuts made after one pad rotation. This concept can be illustrated by replacing the circular conditioner by an equivalent one-dimensional bar conditioner [8]. The one-dimensional bar conditioner is

المنارة للاستشارات
www.manaraa.com

shown in Figure 5.5 (b). The N diamond cutting elements of the bar conditioner are assumed to be arranged on a line between points P_1 and P_2 with mean spacing ℓ, $\ell < r_{pad}$. N is the equivalent number of cuts that the pad will be obtained from the circular conditioner after one pad rotation. During each pad rotation, these N cutting elements will displace randomly and independently from each other in their intervals of length ℓ. So the long time averaged effect is that the bar conditioner will create the same number of random cuts per pad rotation between points P_1 and P_2 as the circular conditioner does. A special case for this is when the conditioner and pad rotate at the same direction with the same rotation rate. For this situation, $\ell = P_1P_2/N_D$, in which N_D is the total number of the cutting elements (total number of active diamonds) on the circular conditioner. If the rotation speed of the pad and conditioner is different, $\ell = \omega_P * P_1P_2/(N_D*\omega_c)$, making ℓ scaled by the ratio of the speed between pad and conditioner. Usually the rotation speed of pad and conditioner is very close. The parameter ℓ is just used to calculate how many cuts between P_1 and P_2 after one pad rotation (density of cuts per pad rotation).

As in [8], we assume the diamond tip is triangular with tip angle α. In our model, the conditioning process is statistically homogeneous along the radial direction. The coordinate system is based on the initial vertical position of the pad surface which is at z = 0. Vertical positions will be measured relative to the initial vertical position of the pad surface. For solid pads, the initial pad surface is assumed smooth. The coordinate system and the terminology are shown in Figure 5.6. Figure 5.6 (a) is for the model in [8] and Figure 5.6 (b) is for our model.

The profile of each of the diamond cutting tips will be represented by the function

$$V(x) = \nu * |x - X| + h(t) \tag{5.2}$$

المنارة للاستشارات
www.manaraa.com

in which X is a uniformly-distributed random variable on the simulation domain $[-\ell/2, \ell/2]$ that gives the horizontal position of the tip. We choose ℓ large enough to contain enough independent cuts so that the spatial statistics are preserved. $\nu = \cot(\alpha/2)$ and $h(t) = h_0 + c^*t$ gives the horizontal position of the tip which is either a constant ($c = 0$, $h_0 < 0$) for constant-height conditioning or time-dependent ($h_0 + c^*t$, $h_0 < 0$ and $c < 0$) for constant cut-rate conditioning as in the following sections. The width of the cutting tip is a function of the height z, defined as

$$w(z - h(t)) = 2(z - h(t))/\nu, \qquad z \leq 0 \tag{5.3}$$

The vertical position of a solid pad surface can be expressed as a single-valued function $z = s(x,t)$. If the rotation rate of the platen is ω_p radians/sec, then conditioning happens at discrete times $t_n = 2\pi n/\omega_p$. We model the conditioning process as a discrete random process in which the surface of the pad is modified at each occurrence with the conditioner within any interval of one diamond cutter $[-\ell/2, \ell/2]$, resulting in transformation $s(x,t_n) \rightarrow s(x,t_n+1)$ for each diamond cut. For the model in [8], ideal micro-cutting is assumed, e.g., $f_{add} = 1$, as shown in Figure 5.6 (a). However, as discussed in the previous section, polyurethane which is commonly used as the CMP pad material is a polymer with plastic mechanical properties (in which material is simply pushed aside and not removed) that need to be considered in the conditioner model. We have introduced a new parameter f_{add} to model the material pile-up along the side-walls of the groove. As shown in Figure 5.4 and 5.6 (b), these two "bumps" are represented by two half-circles at the immediate vicinity of the cut. In our model, the area of these two half-circles is calculated by value of f_{add} and it can range from 0 to 1. Hence in this work, we

المنارة للاستشارات
www.manaraa.com

have introduced another new parameter to allow the model to incorporate a range of brittle and plastic deformation. For Monte Carlo simulation, the following represents the model in [8]

$$s(x,t_{n+1}) = \min(s(x,t_n), V(x - X_n) + h(t_n)) \quad (5.4)$$

while the following represents our new model which incorporates plastic deformation,

$$s(x,t_{n+1}) = \min(s(x,t_n), V(x - X_n) + h(t_n)) + left\quad bump + right\quad bump \quad (5.5)$$

It means that each time when the diamond makes the cut, the material being cut is piled up onto the pad surface at the edge of the cut as two bumps and the new surface is formed by the new cut and the surfaces of these two bumps. We use two half-circles with equal area to represent the two bumps as shown in Figure 5.4. However, the simulation does not depend on the shape of the bumps used as long as the mass (area) is preserved by f_{add} parameter. Circular-shaped bumps make physical sense because when the new surface is formed, the material will minimize the surface energy by forming a nearly circle shape. In our model, the two half-circular bumps are mathematically defined by the function

$$Circ\,(x,R,x_c) = \begin{Bmatrix} \left[R^2 - (x-x_c)^2\right]^{0.5}, & x_c - R \le x \le x_c + R \\ 0 & ,\ otherwise \end{Bmatrix} \quad (5.6)$$

in which R is the radius of the bump and x_c is the center of the bump, and Equation (5.5) becomes

المنارة للاستشارات
www.manaraa.com

$$s(x,t_{n+1}) = \min(s(x,t_n), V(x - X_n) + h(t_n)) + Circ(x, RL_n, xL_n - RL_n) + Circ(x, RR_n, xR_n + RR_n) \tag{5.7}$$

In Equation (5.7), RL_n is the calculated radius for the left bump and xL_n is the calculated left root between the surface, $s(x,t_n)$ and diamond cutting tip, $V(x-X_n)$. RR_n is the calculated radius of the right bump and xR_n is the calculated right root between the surface $s(x,t_n)$ and diamond cutting tip $V(x-X_n)$. During each cut, the program will calculate the first term in Equation (5.7) to obtain the new-cut surface without bump as in [8]. Then the program will calculate the roots, xL_n and xR_n, and then calculate the areas cut by the left half and right half of the cutting tip. Then these areas are multiplied by the f_{add} parameter and the radii, RL_n and RR_n are calculated from these scaled areas by the f_{add} parameter. The program then adds these two half-circles on top of the new-cut surface at the locations of the roots, xL_n and xR_n respectively as shown in Figure 5.4 and 5.6 (b). RL_n and RR_n may be different and depend on the areas cut by the left-half and right-half of the tip respectively at each time step, n. We will show this combined cutting and plastic deformation process (micro-plowing) in the next section and compare our results with the results in [8].

Periodic boundary conditions are used at the boundaries of the computation domain with one diamond cutter $[-\ell/2, \ell/2]$ in the computing procedure of Equation (5.4) and (5.7) so that the statistical homogeneity of the conditioning process is preserved and it does not behave differently at the edges of the computational interval. It ensures the conservation of mass for the pad material at the interval $[-\ell/2, \ell/2]$.

As in [8], we will model two cases according to the position of the conditioner as measured by the position of the cutting tips, $z = h(t)$: (1) constant height, $h(t) = h_0$, and (2)

المنارة للاستشارات
www.manaraa.com

when the tip decreases at the constant cut rate c, h(t) = h_0 + c*t. Such a situation will occur if a constant force is applied to the conditioner and the conditioner will be taken into the pad with a constant speed c. After the initial transients have subsided, the density distribution will obtain a steady shape and also move down with a constant speed c. A special case for this case is that when f_{add} = 1, no material loss from the surface and the conditioner always stays at z = 0. This is equivalent to the constant cutting height case because the cut-rate, c = 0 and cutting height, h(t) = h_0. For other f_{add} values, c is not zero, and the conditioner will move down toward negative z axis with speed of the cut-rate, c. We will seek the structure of this steady distribution in section 3.

The structure of the surface is described by the probability function q(z,t), defined as the fraction of the surface which remains above height z after time t. To calculate q at height z, draw a line at height z within the sampling interval $[-\ell/2, \ell/2]$, then add up all of the line sections which is above z and then divide it by the total length ℓ, repeating the procedure for the whole z domain. Hence the integration of q(z,t) with z will give the area of the whole surface which is the conservation of mass. The definition of q(z,t) can be written

$$q(z,t) = \Pr ob(z < s(x,t)) \tag{5.8}$$

for any choice of horizontal position x since the conditioning process is statistically homogeneous along the radial direction. The function q(z,t) is the complementary cumulative density function (CCDF). The cumulative density function (CDF) p(z,t) describes probabilities for a random variable z to fall in the interval $(-\infty, z)$. Hence CCDF describes probabilities for a random variable z to fall in the interval (z, ∞). The

المنارة للاستشارات
www.manaraa.com

probability density function (PDF) of a random variable describes the relative frequencies of different values for that random variable. The PDF for the pad surface height is defined by

$$\phi(z,t) = -\frac{\partial q(z,t)}{\partial z} = \frac{\partial p(z,t)}{\partial z} \tag{5.9}$$

To calculate q(z,t) and p(z,t) from ϕ(z,t), Equation (5.10) and (5.11) are used

$$q(z,t) = \int_{z}^{\infty} \phi(z,t)dz \tag{5.10}$$

$$p(z,t) = 1 - q(z,t) = \int_{-\infty}^{z} \phi(z,t)dz \tag{5.11}$$

The average surface height over a sample interval with one diamond cutter $[-\ell/2, \ell/2]$ is defined by

$$\bar{s}(t_n) = \frac{1}{\ell}\int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} s(x,t_n)dx = \int_{h(t_n)}^{\infty} z\phi(z,t_n)dz \tag{5.12}$$

The average surface height shows if there is a material lost from the surface and hence it is a mass balance. It defines the position of the surface. By assuming that the initial position of the surface is at z = 0, if $\bar{s}(t_n)$ is a constant, e.g., it stays at z = 0, then there is no material loss and this corresponds to the case of ideal micro-plowing ($f_{add} = 1$) which means that the material just moves around without leaving the surface. If f_{add} is smaller than 1, then the average surface height (the position of the surface) will move towards the negative z axis, e.g., the surface is abraded. In the right-hand side of the

المنارة للاستشارات
www.manaraa.com

second formula of Equation (5.12), for the model in [8], the upper integration limit is 0 because of the assumption of clean-cut, while for our model, because of material pileups, the upper integration limit is infinity.

Numerical model results and comparison with analytical solutions

As in [8], we start with the simplest case for a solid pad where the conditioner is maintained at a constant height, $h(t) = h_0$ ($h_0 < 0$, $c = 0$). The sequences of the pad surfaces generated from Equation (5.4) by Monte Carlo simulation are shown in Figure 5.7. The initial pad surface is flat and is at $z = 0$. The surface roughness of the pad increases as the diamond cuts are made. Then the surface roughness decreases for much longer times as a large amount of the pad material is removed. The average surface height continuously moves to the negative z axis and eventually equals to h_0. This is because the diamond height is a constant at $h_0 = -5$ μm and $f_{add} = 0$. The development for the CCDF q(z,t) corresponding to the simulation in Figure 5.7 is shown in Figure 5.8. This shows how an initially smooth surface is transformed to another worn-smooth surface at h_0. CCDF can be interpreted as the fraction of the remaining pad material at a given height z at time t.

In contrast, for our model (micro-plowing, $f_{add} = 1$, $h(t) = h_0$, $h_0 < 0$, $c = 0$), the sequence of the pad surface generated from Equation (5.7) by Monte Carlo simulation is shown in Figure 5.9. During the first cut, two circular-shaped bumps are formed at the edge of the groove representing the materials plowed onto the pad surface by the diamond. The graph also shows the implementation of the periodic boundary condition – part of the right-bump is moved to the left side of the period because it is out of the right boundary, ensuring that the mass is conserved in the period $[-\ell/2, \ell/2]$. During

www.manaraa.com

simulation, Equation (5.12) (integration of s(x,t)) is used to calculate the average surface height for the period to ensure that the mass is conserved. For example, for $f_{add} = 1$, the average surface height after each cut should always stay at the original plane (z = 0) because there is no material loss. The third and fourth cuts show that the material fills into an old groove adjacent to the cut. After a long time a rough surface is formed as shown in the graphs after 50 cuts. The process clearly shows the concept of material is moved around without leaving the surface. Compared to Figure 5.7, the surface is extended above the original surface (z = 0). The development for the CCDF q(z,t) corresponding to the simulation in Figure 5.9 is shown in Figure 5.10.

In Figure 5.10, in the first several cuts, most of the surface is uncut so that the q plot is divided into two regions – those less than zero and those greater than zero (the original plane). The curve below zero represents the groove cuts, the curve above zero represents the bumps. As more cuts happen, more material is plowed onto the pad surface so the area under the curve below zero becomes smaller and the area under the curve above zero becomes larger. Eventually the two curves combine into one because all of the surface has been cut. Finally 80% of the curve with the value of q(z,t) from 1.0 to 0.2 stabilizes with a half-Gaussian shape with an "oscillating" tail. The tail is oscillating is because the high peaks of the surface are constantly killed by the diamond running over them and then they grow again. This oscillating tail behavior is shown in Figure 5.11. In Figure 5.11 (a), a high peak is killed by a diamond running over it and the mass is redistributed. Figure 5.11 (b) shows the q(z,t) plot associated with it. Figure 5.11 (c) shows the oscillating tail with time. To get the good statistical data, we averaged q(z,t) plot for a total of 858 time steps and the result of the averaged q is shown in Figure 5.12 (b)

المنارة للاستشارات
www.manaraa.com

showing a smooth curve. The averaged q is the steady-state solution of the CCDF for f_{add} = 1 case. Because the material is just moved around, q should reach steady-state and this steady-state can be observed by averaging over many oscillations. This averaged q has a fixed shape without oscillating tail which is a half-Gaussian shape. Comparing Figure 5.12 (a) with Figure 5.12 (b), we can see that plastic deformation has changed q(z,t) from exponential shape to a half-Gaussian shape with constant diamond cutting height h_0. Figure 5.12 (a) also shows the analytical solution for this $f_{add} = 0$ case which is Equation 3.2 or Equation 3.7 in [8]. Figure 5.12 (b) shows our derived analytical solution for f_{add} = 1 case which is a half-Gaussian function with standard deviation of $\sqrt{2/\pi}|h_0|$. Good agreements between analytical and Monte Carlo numerical simulations are obtained for both $f_{add} = 0$ and $f_{add} = 1$ cases with constant diamond height conditioning. The derivations of these analytical solutions are demonstrated next.

We have tailored the population balance equation, Equation 3.7 in [16] specifically for the CMP pad asperity problem as shown in Equation (5.13). This is a general equation and applies to both fixed-height and constant cut-rate conditioning.

$$\frac{\partial q(z,t)}{\partial t}+\frac{\partial}{\partial z}(v_w q(z,t))+\frac{\partial}{\partial z}(cq(z,t))-B+D=0 \qquad h(t)\le z\le\infty \qquad (5.13)$$

Equation (5.13) is obtained by accounting for all of the asperities in a differential distance along the z direction, dz. The first term is the accumulation term. The second term is due to the wear of the pad asperities because of wafer polishing. The velocity term, v_w is the wear rate of the individual pad asperity because of wafer polishing. The velocity term c is the speed the cutting tip moves into the pad. For constant height conditioning, c = 0. Hence the third term is due to pad conditioning because the cutting

www.manaraa.com

tip moves into the pad with cut-rate c. From Equation (5.12), c should be proportional to 1-f_{add}. Cut-rate is maximum when $f_{add} = 0$ and when $f_{add} = 1$, cut-rate is zero. B and D are the asperity birth and death terms because of pad conditioning. The cut-rate c depends on the B and D terms. If there is no conditioning, only the first two terms are nonzero. This problem has been modeled in [17 - 19]. Ref. [18] and [19] use an analytical solution, while Ref. [17] uses numerical solution. Ref. [18] has considered fluid effects in the wear of the pad asperities. The wear rate v_w is originally derived in Ref. [17]. It is a z dependent variable and derived from Archard's law [20] and Hertz contact theory [21] for individual asperities as shown in Equation (5.14). However, for conditioning only (without polishing of wafers), the second term is zero. Ref. [8] and this work focus on this case with the difference being that there is only a death term in the conditioner model in [8] for the clean-cut assumption. In this work we add the birth term to model the bumps created by plastic deformation.

$$v_w = \frac{dz}{dt} = -c_a \frac{4}{3\pi} E^* k_s^{1/2} (z - d(t))^{1/2} \tag{5.14}$$

In Equation (5.14), c_a is a parameter that is proportional to the sliding velocity of the asperity of the asperity tip relative to the wafer and it is treated as a constant in Equation (5.13). $E^* = E/(1-\upsilon^2)$, E is the pad Young's modulus and υ is the Poisson ratio. All asperities are assumed to have spherical tips with identical curvature k_s and behave elastically. The separation distance d(t) between the mean of the pad surface and the wafer depends on the load on the wafer and is calculated by Greenwood and Williamson contact mechanics [21]. It should be mentioned that Ref. [8], Ref.[17] and this work have all neglected the wear between the pad asperities and the conditioner. This wear is similar

المنارة للاستشارات
www.manaraa.com

to Equation (5.14) and happens when the pad under the conditioner is not cut by the diamonds. However this wear should be small compared to the cut-rate of the pad. The focus of this work is the plastic deformation of the pad conditioning and hence we have neglected both this wear and the second term in Equation (5.13).

The coordinate system of Equation (5.13) is based on the initial vertical position of the pad surface which is at $z = 0$. Equation (5.13) can be understood that the mean of the pad asperities moves down towards negative z axis because the diamonds are cutting the pad with speed c and at the same time the pad asperities are worn down by the pad because of wafer polishing. The cut-rate c caused by the cutting of the diamonds depends on the load on the conditioner, active diamond density and f_{add}. By neglecting the wear of the pad asperities because of wafer polishing and considering only the cutting by the diamonds, the density distribution should obtain a steady shape and moves down with constant speed c. Such a situation will happen when a constant force is applied on the conditioner and initial transient has subsided. The conditioner then is taken into the pad with a constant speed c. We will seek this steady-state density distribution and we will also derive the cut-rate c from this steady-state, which are the focus of this work. Although we did not perform the simultaneous wear and conditioning of the pad asperities, the derived cut-rate c can be used in Equation (5.13) for that purpose. A term similar to the second term in Equation (5.13) can be added to Equation (5.13) to include wear between pad and conditioner. The definition for that wear is similar to Equation (5.14) with d(t) replaced by the distance between the mean of the conditioner surface and the pad. The parameter c_a will be decided by the characteristics between the pad and the conditioner.

المنارة للاستشارات
www.manaraa.com

Because of the mass balance, the birth term has the same functional form as the death term, e.g., they are proportional to each other, and when f_{add} =1, the summation of these two terms is zero, e.g., the mass loss is zero. The following population balance equation for the pad conditioning is obtained

$$\begin{aligned}\frac{\partial q(z,t)}{\partial t}+\frac{\partial}{\partial z}\left(c\,q(z,t)\right)&=-\frac{\omega_p}{\pi\nu\,\ell}(z-h(t))q(z,t)\\&\quad+f_{add}\frac{\omega_p}{\pi\nu\,\ell}(z-h(t))q(z,t)\end{aligned}\qquad h(t)\le z\le\infty \quad (5.15)$$

in which at the right-hand side, the first term represents the death term, e.g., the material removed as in [8], and the second term represents the birth term, e.g., the fraction of the material removed that is added back to the pad surface. In Equation (5.15), ω_p is the pad rotation rate, $\nu = \cot(\alpha)$ and l is the distance between diamond cuts per pad rotation as explained before. Because the death term and birth term have the same functional form, they can be combined into one term for $h(t)\le z\le\infty$ as shown in Equation (5.16)

$$\frac{\partial q(z,t)}{\partial t}+\frac{\partial}{\partial z}\left(cq(z,t)\right)=\left(f_{add}-1\right)\frac{\omega_p}{\pi\nu\,\ell}(z-h(t))q(z,t)\quad h(t)\le z\le\infty \qquad (5.16)$$

For constant diamond height conditioning, h(t) = h_0, c = 0 and the second term in Equation (5.16) disappears. So Equation (5.16) becomes

$$\frac{\partial q(z,t)}{\partial t}=\left(f_{add}-1\right)\frac{\omega_p}{\pi\nu\,\ell}(z-h_0)q(z,t)\qquad h_0\le z\le\infty \qquad (5.17)$$

The analytical solution for PDF is derived from Equation (5.17) and the result is shown in Equation (5.18). The result for f_{add} = 0 case is shown in Equation (3.7) in [8]

المنارة للاستشارات
www.manaraa.com

and repeated here for q(z,t) in Equation (5.19). Our analytical solution matches that in Ref. [8]. CCDF q(z,t) for this $f_{add} = 0$ case can also be derived directly from the definition of q as done in [8] which is shown in Equation (5.20) here. Both of the equations will give the same answer at longer time. The comparison of Equation (5.19) with Monte Carlo numerical solution is shown in Figure 5.12 (a).

$$q(z,t) = \exp\left(\frac{\omega_p (f_{add} - 1)t}{2\pi\ell} w(z - h_0)\right) \qquad h_0 \le z \le 0 \qquad (5.18)$$

$$q(z,t) = \exp\left(-\frac{\omega_p t}{2\pi\ell} w(z - h_0)\right) \qquad h_0 \le z \le 0 \qquad (5.19)$$

$$q(z,t_n) = \left(1 - \frac{1}{\ell} w(z - h_0)\right)^n \qquad h_0 \le z \le 0 \qquad (5.20)$$

In the above equations, the width of the diamond tip is given by Equation (5.3) for triangular diamond cut tips. As f_{add} becomes larger and larger, the first term in Equation (5.17) becomes smaller and smaller and eventually the CCDF reaches steady-state at $f_{add} = 1$, e.g., the first term in Equation (5.16) becomes zero. So for the case of $f_{add} = 1$, the solution can not be obtained from Equation (5.18) since it is a singular limit. Equation (5.18) becomes more accurate as f_{add} becomes smaller. However, the clean-cut case ($f_{add} = 0$) with constant diamond cutting height is not plausible because the conditioner should be taken into the pad with nonzero c because of the load balance. More plausible case is the solution for constant cut-rate case. Because for polyurethane, f_{add} is close to 1, we can clearly see the difference between the cases for $f_{add} = 0$ and $f_{add} = 1$ with the constant diamond cutting height as demonstrated in the numerical solutions in Figure 5.12. The

المنارة للاستشارات
www.manaraa.com

analytical solution we have derived for $f_{add} = 0$ also matches the numerical solution well as shown in Figure 5.12 (a). It is also the same as that in Ref. [8]. Next we will seek the constant cut-rate case (traveling wave). The analytical solution for $f_{add} = 1$ should be obtained from this constant cut-rate case because it is also a steady-state solution.

When the conditioner is balanced with a constant load applied to it, the conditioner will move into the pad at a constant speed c called the pad cut-rate. After the initial transients had subsided, the density distribution would then obtain a steady shape that would also move down at a constant speed c. We now derive the structure of this traveling-wave distribution. At steady-state, the change of q with time is zero so the PDE in Equation (5.16) becomes an ODE as shown in Equation (5.21). Then by solving Equation (5.21), a half-Gaussian function for q(z,t) is obtained as shown in Equation (5.22) for diamond height $h(t) = h_0 + c*t$. The corresponding PDF is obtained in Equation (5.23) by using Equation (5.9).

$$\frac{d}{dz}\left(cq(z,t)\right) = -\frac{(1-f_{add})\omega_p}{\pi\nu\ell}\left(z-h(t)\right)q(z,t) \tag{5.21}$$

$$q(z,t) = \exp\left(-\frac{(1-f_{add})\omega_p}{2\pi c\nu\ell}\left(z-h_0-ct\right)^2\right) \quad h_0+ct \le z \le \infty \tag{5.22}$$

$$\phi(z,t) = \frac{(1-f_{add})\omega_p}{\pi c\nu\ell}\left(z-h_0-ct\right)\exp\left(-\frac{(1-f_{add})\omega_p}{2\pi c\nu\ell}\left(z-h_0-ct\right)^2\right) \quad h_0+ct \le z \le \infty \tag{5.23}$$

At steady-state conditioning, the penetration depth of the diamond tip into the pad h_0 is a constant and hence the cut rate c is a constant. When $f_{add} < 1$, there is steady-state loss from the surface, the mean surface height will move down with speed c and because

المنارة للاستشارات
www.manaraa.com

of the load balance, the cutting tip will move down with speed c, $h(t) = h_0 + c*t$, $c < 0$. According to the mass balance, the cut-rate c should be linearly proportional to $1-f_{add}$ so the ratio of these two should be a constant. To determine the cut-rate at the particular load with corresponding penetration depth h_0, we use the first moment calculation, Equation (5.12). To simplify the calculation, as in [8], adopting the moving coordinate Z with $Z = z - h(t)$ and $q(Z,t) = q(z,t)$, Equation (5.22) and (5.23) become

$$q(Z,t) = \exp\left(-CZ^2\right) \qquad 0 \le Z \le \infty \qquad (5.24)$$

$$\phi(Z,t) = 2CZ\exp(-CZ^2) \qquad 0 \le Z \le \infty \qquad (5.25)$$

in which

$$C = \frac{(1-f_{add})\omega_p}{2\pi c \nu \ell} \qquad (5.26)$$

and

$$Z = z - h_0 - ct \qquad (5.27)$$

By using Equation (5.12), the first moment of the PDF in Equation (5.25) is calculated and shown in Equation (5.28).

$$\int_0^\infty 2CZ^2\exp(-CZ^2)dZ = \frac{\sqrt{\pi}}{2\sqrt{C}} \qquad 0 \le Z \le \infty \qquad (5.28)$$

Because of the mass balance, the ratio between cut-rate c and $1 - f_{add}$ is always a constant. At the steady-state conditioning, C is a constant and the traveling q(Z,t) profile

المنارة للاستشارات

www.manaraa.com

has the same standard deviation for all f_{add} values under this diamond penetration depth h_0. Because we know that at $f_{add} = 1$, the mean always stays at $-h_0$ ($h_0 < 0$) in the new coordinate ($0 \le Z \le \infty$), we can calculate the cut-rate for other f_{add} cases with this h_0. Hence from the first moment result of the PDF in Equation (5.28), the following result is obtained,

$$\frac{\sqrt{\pi}}{2\sqrt{C}} = -h_0 \tag{5.29}$$

Plugging Equation (5.26) into Equation (5.29), we obtain the cut-rate c which is shown in Equation (5.30) with unit of μm per minute or Equation (5.31) with unit of μm per pad rotation or per cut in the period $[-\ell/2, \ell/2]$ for a single diamond cutter.

$$c = \frac{2h_0^2 \omega_p}{\pi^2 \nu \ell}\left(f_{add} - 1\right) \tag{5.30}$$

$$c = \frac{4h_0^2}{\pi \nu \ell}\left(f_{add} - 1\right) \tag{5.31}$$

From Equation (5.31),

$$C = \frac{\pi}{4h_0^2} \tag{5.32}$$

From the Gaussian distribution, we obtain the standard deviation for the traveling wave profile of Equation (5.24) and (5.25),

المنارة للاستشارات
www.manaraa.com

$$\sigma_t = \sqrt{\frac{2}{\pi}}|h_0| \tag{5.33}$$

Hence we rewrite Equation (5.22) and (5.23) in Equation (5.34) and (5.35) respectively for the group of the traveling wave profiles with the same h_0 and standard deviation defined by Equation (5.33) but with different f_{add} and cut-rates defined by Equation (5.30). If c = 0, Equation (5.34) and (5.35) become Equation (5.36) and (5.37).

$$q(z,t) = \exp\left(-\frac{\pi}{4h_0^2}(z-h_0-ct)^2\right) \qquad h_0 + ct \le z \le \infty \tag{5.34}$$

$$\phi(z,t) = \frac{\pi}{2h_0^2}(z-h_0-ct)\exp\left(-\frac{\pi}{4h_0^2}(z-h_0-ct)^2\right) \quad h_0 + ct \le z \le \infty \tag{5.35}$$

$$q(z,t) = \exp\left(-\frac{\pi}{4h_0^2}(z-h_0)^2\right) \qquad h_0 \le z \le 0 \tag{5.36}$$

$$\phi(z,t) = \frac{\pi}{2h_0^2}(z-h_0)\exp\left(-\frac{\pi}{4h_0^2}(z-h_0)^2\right) \quad h_0 \le z \le 0 \tag{5.37}$$

From Equation (5.34) and (5.35), we can see that, the standard deviation of the steady-state profile depends on h_0 which depends on the load applied to the conditioner, and the cut-rate depends on the plastic property of the pad material f_{add} as well as the load. When f_{add} = 1 (pure plastic deformation), the cut-rate is zero and when f_{add} = 0 (pure brittle fracture), the cut-rate is maximum. This makes physical sense because under high load, the diamonds penetrate the pad deeper resulting in a larger h_0 and also more active diamonds in contact, and this increases the cut-rate. With the same h_0, larger f_{add} means

المنارة للاستشارات
www.manaraa.com

that less material is lost from the surface resulting in a lower cut-rate. With the same h_0, f_{add} only influences the cut-rate and has no effect on the pad surface PDF as shown in Equation (5.35). The pad surface PDF depends entirely on h_0. The case with $f_{add} = 1$ is a special case of the traveling wave because the cut-rate is zero, the mean of the asperity heights stays at zero and the cut height stays at h_0 at all the times during conditioning. This corresponds to the case of the constant diamond cutting height conditioning with $f_{add} = 1$. The CCDF and PDF for $f_{add} = 1$ are shown in Equation (5.36) and (5.37) by making c = 0. From Equation (5.32), for h_0 = -5 μm, the standard deviation of the Gaussian q(z) function is about 4 μm. The Monte Carlo numerical simulation for this case and the analytical solution matches numerical solution very well as shown in Figure 5.12 (b).

Now we will verify the traveling wave analytical solution using Monte Carlo simulation for another case, $f_{add} = 0$. Firstly, we use Monte Carlo simulation to verify Equation (5.30) or (5.31). After each diamond cut in the interval $[-\ell/2, \ell/2]$ for single diamond cutter, the mean is recalculated by the first equation in Equation (5.12). Then the tip position will be moved down according to the new mean. The depth of the tip is always h_0 relative to the new mean at each cut and this ensures the constant cutting depth h_0 which is imposed by the load balance. At each cut, the circular bumps are added onto the pad surface at the edge of the groove cut by f_{add} value to ensure mass balance as explained before. The f_{add} value is varied from 0 to 1 at 0.1 intervals at a constant h_0. For each f_{add} value, the program tracks the mean $\bar{s}(t)$ and the cutting tip depth h(t). 50 time steps (cuts) are used to calculate mean and h(t) for each f_{add} case. Then the slope of $\bar{s}(t)$ or h(t) is used to calculate the cut-rate c. Figure 5.13 shows the examples of $\bar{s}(t)$ and h(t) for several f_{add} values. From these plots, we can see that $\bar{s}(t)$ and h(t) are moving down

المنارة للاستشارات
www.manaraa.com

with the same speed which is the cut-rate c and $\bar{s}(t)$ lags h(t) by an amount of h_0 which is the cutting tip depth imposed by the load balance. The mean of the surface height $\bar{s}(t)$ intercepts the z axis at 0 which is what it starts with (z = 0) and h(t) intercepts the z axis at h_0 which means that at the steady-state conditioning, the cutting depth is always h_0 for each cut. Figure 5.14 shows the surface profile s(x,t) after 50 time steps for different f_{add} values. We can see that the average surface height moves down slower for larger f_{add} values as expected. The cut-rates for f_{add} values between 0 to 1 are calculated by the program this way and the results are shown in Figure 5.15. In these calculations we have used ℓ = 80 µm and h_0 = -5 µm with triangle diamond cutter. We can see that at constant h_0, the cut-rate is linearly varying with f_{add}. Larger f_{add} values give smaller cut-rates. Linear regression for these data in Figure 5.15, c (µm/pad rotation) = 0.225*f_{add} - 0.226 is obtained with R^2 = 0.9981. We also plotted analytical solution for c (µm/pad rotation) which is Equation (31) in Figure 5.15 for ℓ = 80 µm and h_0 = -5 µm. The analytical solution for the cut-rate gives -0.23 µm/pad rotation. We can see that the analytical solution fits numerical solution very well.

Next, to confirm the analytical solution of Equation (5.34), we perform the Monte Carlo simulation for f_{add} = 0 case using ℓ = 80 µm and h_0 = -5 µm. The result is shown in Figure 5.16 for the surface height s(x,t) and Figure 5.17 for the traveling wave q(z,t). Because c is -0.23 µm/pad rotation, the mean should travel to about -230 µm after 1001 time-steps which is shown in Figure 5.16. Figure 5.17 (a) shows the developing of q(z,t) in the first 50 time-steps and it stabilizes at about n = 50, then the stable q(z,t) travels down into the pad as shown in Figure 5.17 (b) for n = 200, 600, 1001 time-steps. To get data with good statistics, we again average 951 time-steps as we have done for the

المنارة للاستشارات
www.manaraa.com

constant diamond height case to get the averaged q(z,t) as shown in Figure 5.18 (a). The averaged q(z,t) is shown at the position of q(z,t) at n = 1001 time-step. This half-Gaussian curve has a standard deviation of 4 μm which is consistent with the analytical solution of Equation (5.33) and (5.34). The comparison of the analytical solution of Equation (5.34) with the numerical solution is also shown in Figure 5.18 (a). We can see that the analytical solution and numerical solution matches very well. The analytical curve sits on top of the numerical curve. In Figure 5.18 (b), the analytical solution for PDF $\phi(z,t)$ given by Equation (5.35) is shown.

Next we repeat the numerical simulation with the same parameters but with $h_0 = -2.5$ μm. For this case, from Equation (5.33) and (5.35), the standard deviation should be half of that for the case with $h_0 = -5$ μm. And from Equation (5.30) or (5.31), for the specific f_{add}, the cut-rate should be 0.25 times of that for the case with $h_0 = -5$ μm. We choose $f_{add} = 0$, which is again perfect brittle case as done in [8]. The surface height results are shown in Figure 5.19. The average surface height has traveled to about z = -58 μm which is consistent with the analytical cut-rate. The developing of q(z,t) for the first 50 time-steps is shown in Figure 5.20 (a). Again, q(z,t) stabilizes at about n = 50 time-step and then the stable q(z,t) travels down the z axis. Figure 5.20 (b) shows q(z,t) at time-steps of n = 200, 600 and 1001 respectively. To get data with good statistics, we again average q(z,t) for 951 time-steps and the result is shown in Figure 5.21 (a). Figure 5.21 (a) also shows the analytical solution of Equation (5.34) with standard deviation of 4 μm and we can see that they match very well. The numerical solution overlaps the analytical solution. Figure 5.21 (b) shows the analytical solution for PDF $\phi(z,t)$.

In this section, we have derived the analytical solutions for the constant cutting

المنارة للاستشارات
www.manaraa.com

height ($h(t) = h_0$, $c = 0$) and constant cut-rate cases (load-balanced case with a known constant h_0, $h(t) = h_0 + c*t$, $c \leq 0$) and the analytical solutions compare very well to the numerical solutions for $q(z,t)$ and cut-rate c with varying values of f_{add} or h_0. The results show that the standard deviation of the surface $q(z,t)$ only depends on the diamond cutting depth h_0. With the same h_0, the group of the materials with different plasticity characterized by f_{add} will have the same variance. While the cut-rate depends on both h_0 and f_{add}. The cut-rate varies with f_{add} linearly and with the same h_0, larger f_{add} shows less abrasive wear.

The derived cut-rate is a function of the plastic property of the pad expressed by f_{add} parameter, the load and the RPM between the pad and the conditioner. The dependence of the load is reflected in the dependence of the diamond penetration depth h_0 and the density of the active diamonds reflected in the parameter ℓ. Plastic deformation can explain the low wear rate of the pad during conditioning [22]. For example, the cut-rate is 0.465 μm/min for IC1000 pad for the conditioner type of Saesol at 7 lbf load with 4 in. diameter in SS-12 slurry. At the same process condition, the cut-rate for D100 pad is 0.309 μm/min [22]. For material with high plasticity like polyurethane, the pad is mostly "worn down" by the adhesive wear instead of "cut down" by the abrasive wear from the diamond cuts.

Conclusion

In this work, we have demonstrated a method to simulate pad plastic deformation by population balance model during the pad conditioning process. Plastic deformation is important in modeling the pad conditioning because the raised ridges along the sides of the grooves are the high points (asperities) that make contacts with the wafer. These

المنارة للاستشارات
www.manaraa.com

asperities act as pressure enhancers for the abrasive particles used to polish the wafer. The steady-state conditioned PDF depends on the load only. The derived cut-rate equation depends on the load, pad plasticity, and the RPM between the pad and the conditioner. Plastic deformation can explain the observed low cut-rate during the pad conditioning process and it suggests that the polyurethane pad is mostly “worn down” by the adhesive wear instead of “cut down” by the abrasive wear from the diamonds on the conditioner.

References

1. Su J, Chen X, Zhang X, Du J, Guo D (2008) Study on characteristic of material removal rate in chemical mechanical polishing of silicon wafer. J Comput Theor Nanosci 5(8):1656-1660.

2. Shan L (2000) Mechanical Interactions at the Interface of Chemical Mechanical polishing. Dissertation, Georgia Institute of Technology.

3. Sundararajan S, Thakurta DG, Schwendeman DW, Murarka SP, Gill WN (1999) Two-dimensional wafer-scale chemical mechanical planarization models based on lubrication theory and mass transport. J Electrochem Soc 146(2):761-766.

4. Chen D, Lee B (1999) Pattern planarization model of chemical-mechanical polishing. J Electrochem Soc. 146:744-748.

5. Oliver MR, Schmidt RE, Robinson M (2000) CMP pad surface roughness and CMP removal rate. Electrochem Soc Proc 2000-26:77-83.

6. Lawing AS (2002) Polish rate, pad surface morphology and oxide chemical mechanical polishing. Electrochem Soc Proc 2002-1:46-60.

7. Stein D, Hetherringgton D, Dugger M, Stout T (1996) Optical interferometry for surface measurements of CMP pads. J Electron Mater 25(10):1623-1627.

8. Borucki LJ, Witelski T, Please C, Kramer PR, Schwendeman D, A theory of pad conditioning for chemical-mechanical polishing. J Eng Math 50:1-24.

9. Jiang H, Browning R, Sue SJ (2009), Understanding of scratch-induced damage mechanisms in polymers. Polymer 50:4056-4065.

المنارة للاستشارات
www.manaraa.com

10. Sinha SK, Chong WLM, Lim SC (2007) Scratching of polymers - modeling abrasive wear. Wear 262:1038-1047.

11. Ring TA, Feeney P, Boldridge D, Kasthurirangan J, Li S, Dirksen JA (2007) Brittle and ductile fracture mechanics analysis of surface damage caused during CMP. J Electrochem Soc 154(3): H239-H248.

12. Koleske JV (1995) Paint and Coating Testing Manual: Fourteenth Edition of the Gardner-Sward Handbook. ASTM International, West Conshohocken, PA.

13. Stevens MP (1990) Polymer Chemistry – An Introduction, Second edition. Oxford University Press, New York.

14. Daggett S (2004) A guide to selection of methacrylate, urethane and epoxy adhesives. http://www.compositesworld.com/articles/a-guide-to-selection-of-methacrylate-urethane-and-epoxy-adhesives. Cited 19 Sept 2012.

15. Zum Gahr KH (1988) Modeling of two-body abrasive wear. Wear 124:87-103.

16. Ring TA (1996) Fundamentals of Ceramic Powder Processing and Synthesis. Academic Press, San Diego.

17. Borucki LJ (2002) Mathematical modeling of polish-rate decay in chemical-mechanical polishing. J Eng Math 43:105-114.

18. Shi H, Ring TA (2010) CMP pad wear and polish-rate decay modeled by asperity population balance with fluid effect. Microelectron Eng 87:2368-2375.

19. Shi H, Ring TA (2010) Analytical solution for polish-rate decay in chemical-mechanical polishing. J Eng Math 68:207-211.

20. Archard JF (1953) Contact and rubbing of flat surfaces. J Appl Phys 24:981-988.

21. Greenwood JA, Williamson JBP (1966) Contact of nominally flat surfaces. Proc Roy Soc Lond A Mat 295(1442):300-319.

22. Prasad A (2011) Private communication. Cabot Microelectronic Corporation.

المنارة للاستشارات
www.manaraa.com

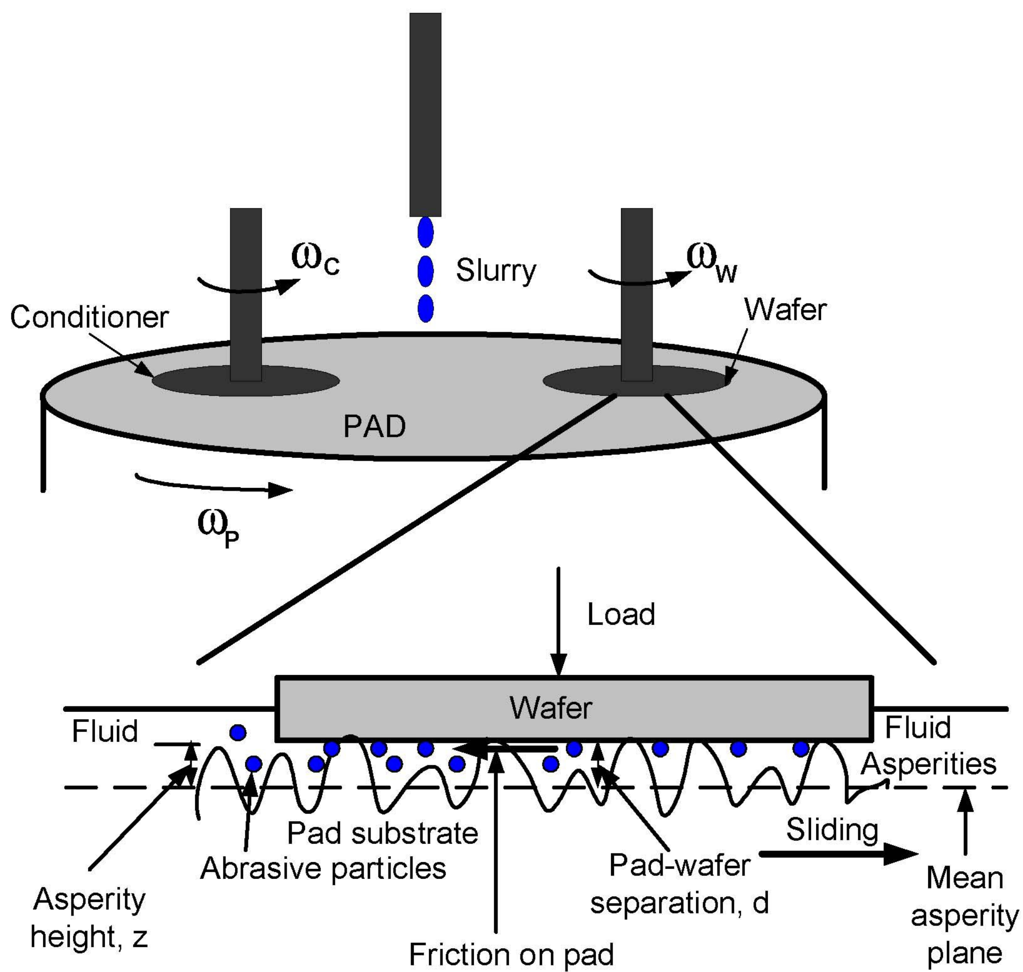
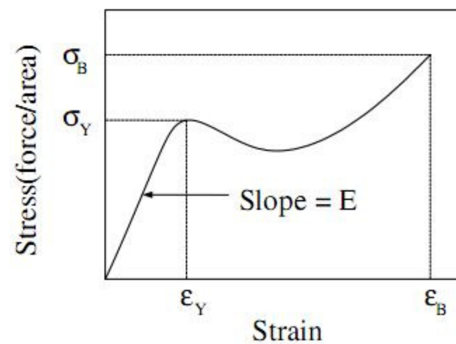
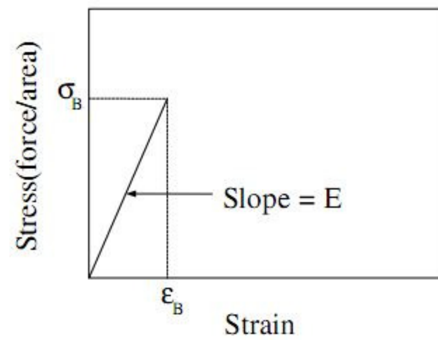


Figure 5.1 Schematic illustration of CMP process. The separation d is measured from the mean plane of the pad and asperity heights are also measured from the mean plane of the pad.

www.manaraa.com



(a)

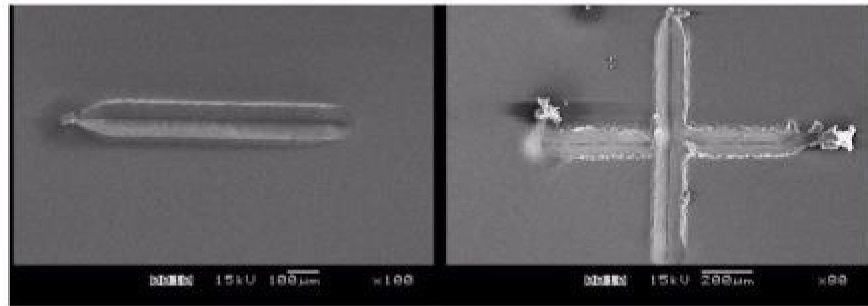


(b)

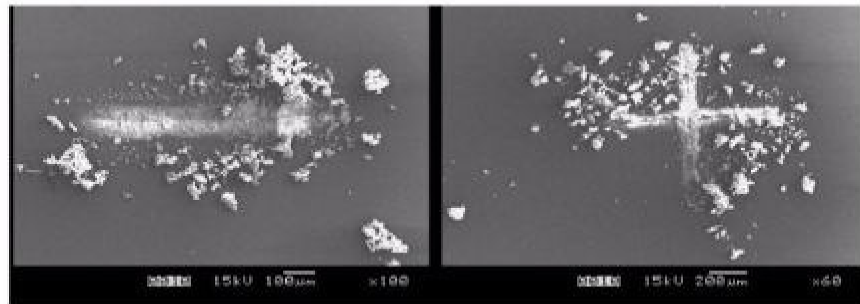
Figure 5.2 Schematic stress-strain curves of two categories of polymers. (a) Ductile polymer. (b) Brittle polymer. Tensile properties are defined: tensile modulus or Young's modulus (E), elongation at yield (ε_Y), elongation at break (ε_B), yield stress (σ_Y), and tensile strength (σ_B). Ductile material has significant elongation before break, while brittle material has no yield point and little elongation before break.

المنارة للاستشارات

www.manaraa.com



(a)



(b)

Figure 5.3 One-pass single-track scratch and two-pass intersecting scratch on two types of polymers. (a) Ductile polymer like polyurethane. (b) Brittle polymer like epoxy. Ductile polymer generates little wear debris and most of the materials are piled-up on both side walls of the cutting groove to form raised ridges. Brittle material generates significant amount of wear debris and most of the materials from the groove are removed from the surface. Reprinted from [10], with permission from Elsevier.

المنارة للاستشارات
www.manaraa.com

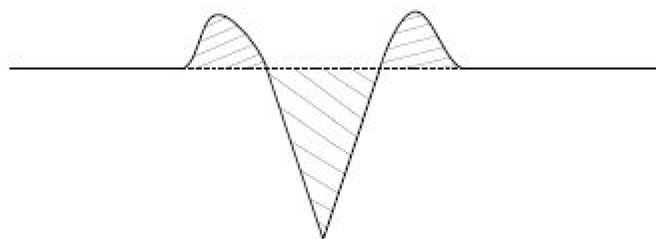


Figure 5.4 Cross-section of a created groove.

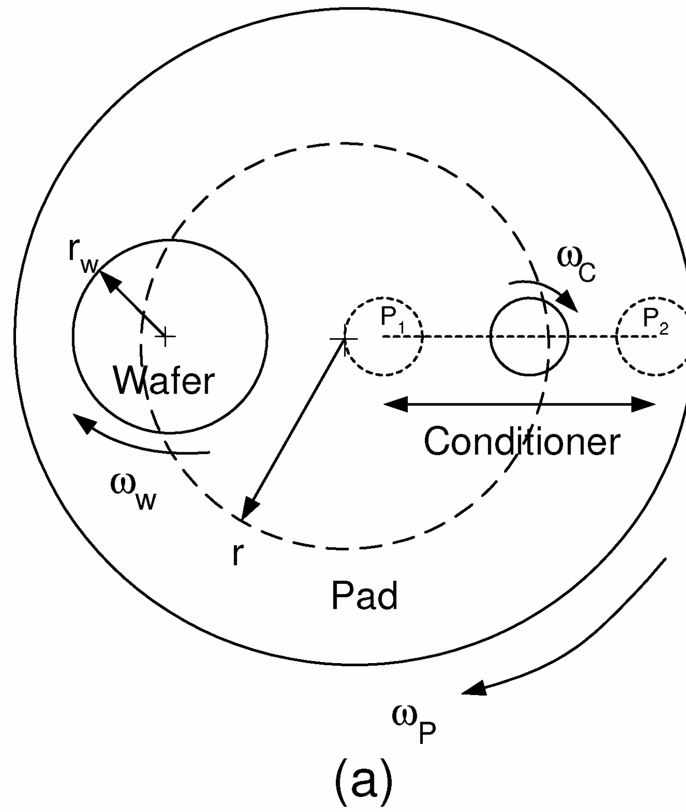
المنارة للاستشارات

www.manaraa.com

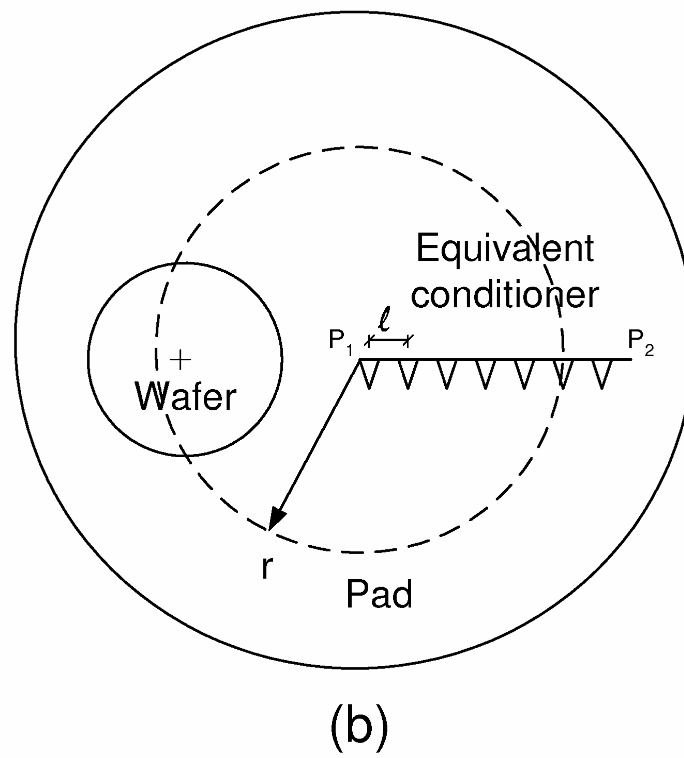
Figure 5.5 Illustration of a CMP tool with the circular conditioner and the equivalent bar conditioner. (a) Top view of a CMP tool geometry. The pad section between P_1 and P_2 travels along a circular path as shown in the figure with a small radial cross-section of the pad at the end of a vector r as an example and encounters the wafer once per rotation. It will also encounter the circular conditioner during a few of the periods in each conditioning sweep between P_1 and P_2. (b) Illustration of the bar conditioner. The circular conditioner is replaced by an equivalent bar conditioner. A small portion at r of a radial cross-section encounters both the wafer and the conditioner once per period.

المنارة للاستشارات

www.manaraa.com

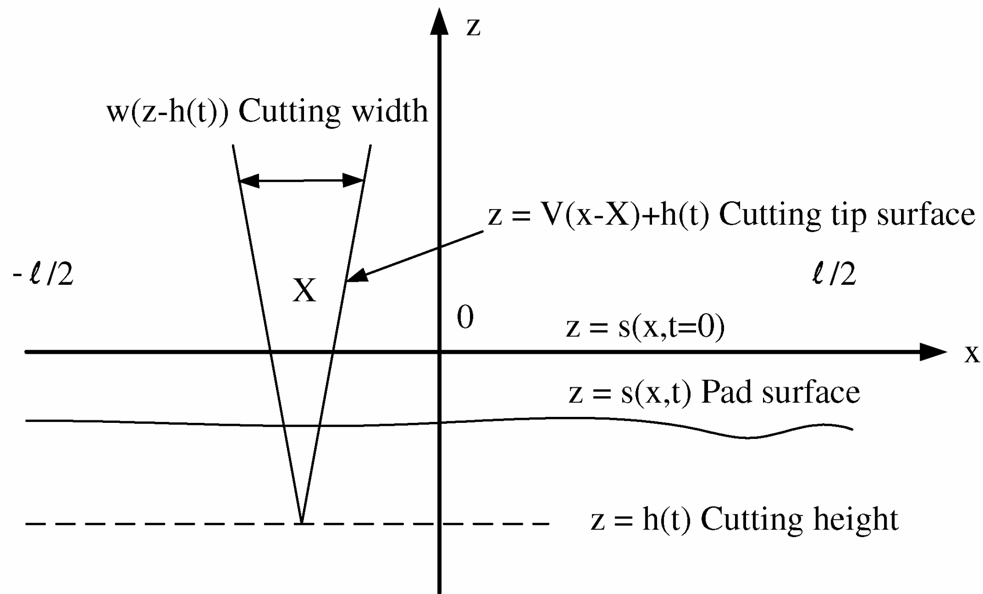


(a)

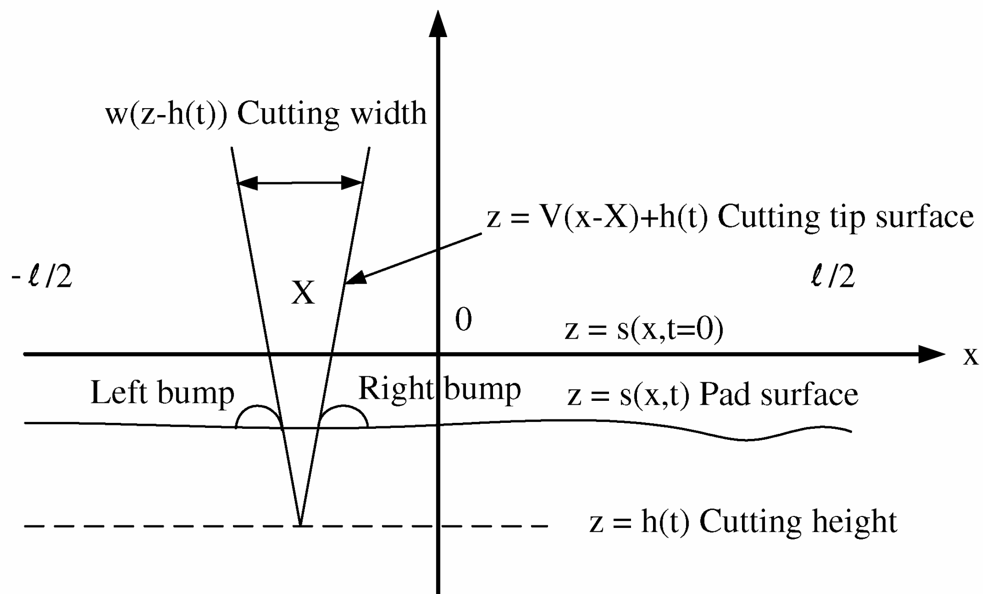


(b)

للاستشارات المنارة
www.manaraa.com



(a)



(b)

Figure 5.6 The coordinate system and terminology used in [8] and in this work. (a) The coordinate system and terminology used in [8]. (b) The coordinate system and terminology used in this work.

www.manaraa.com

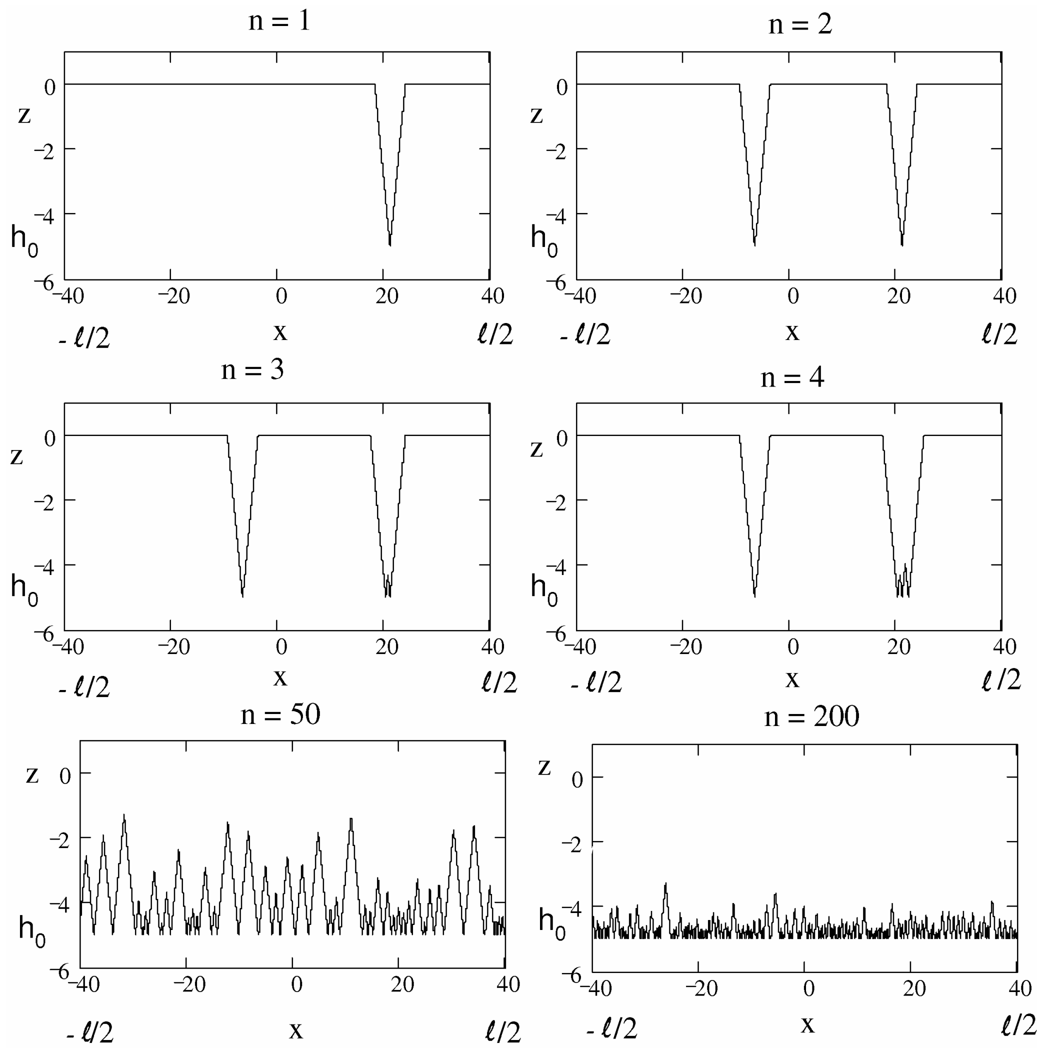


Figure 5.7 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process by Equation (5.4), for n = 1, 2, 3, 4, 50, and n = 200 time steps with initially flat surface. The parameters are $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm and $f_{add} = 0$. All the units are in μm.

www.manaraa.com

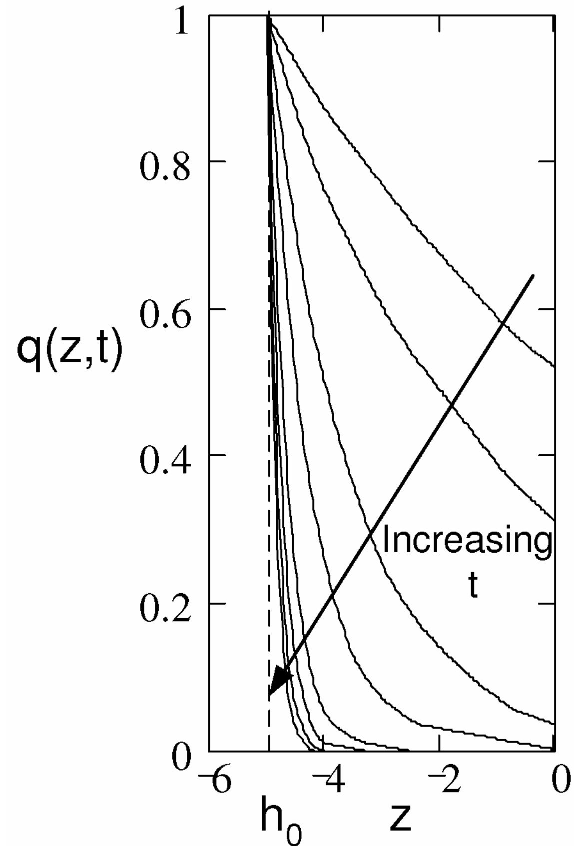
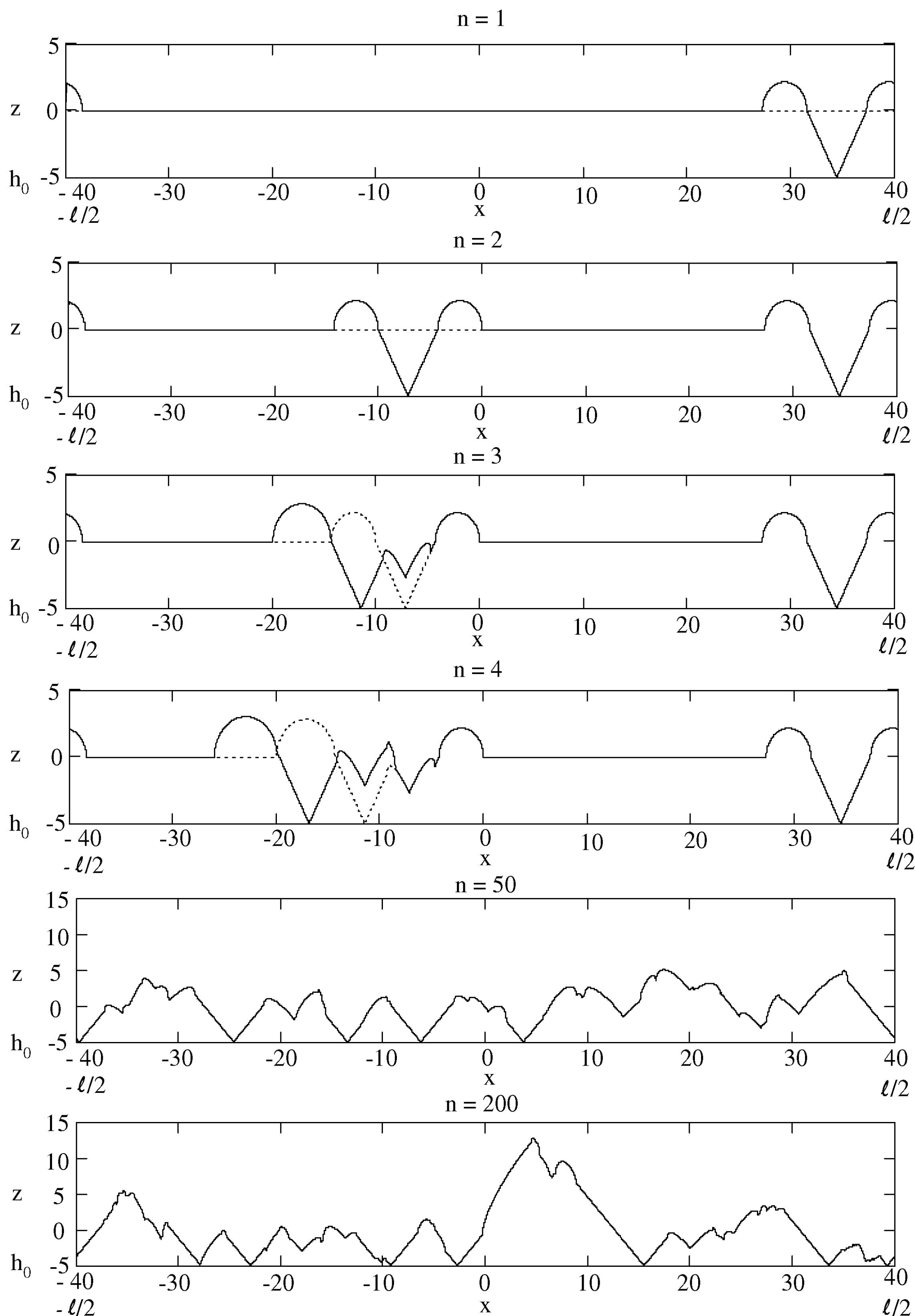


Figure 5.8 Development of the CCDF q(z,t) corresponding to the simulation in Figure 5.7 for n = 10, 20, 50, 100, 200, 300, 400, 500 time steps. The parameters are $\alpha = \pi/3$, $\ell =$ 80 μm, $h_0 = -5$ μm, and $f_{add} = 0$.

www.manaraa.com

Figure 5.9 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process by Equation (5.7), for n = 1, 2, 3, 4, 50, and n = 200 time steps with initially flat surface. The dotted line is the surface in the previous step. The parameters are $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm, and $f_{add} = 1$. The process shows that the material just moves around on the surface by micro-plowing. All the units are in μm.

المنارة للاستشارات
www.manaraa.com



المنارة للاستشارات
www.manaraa.com

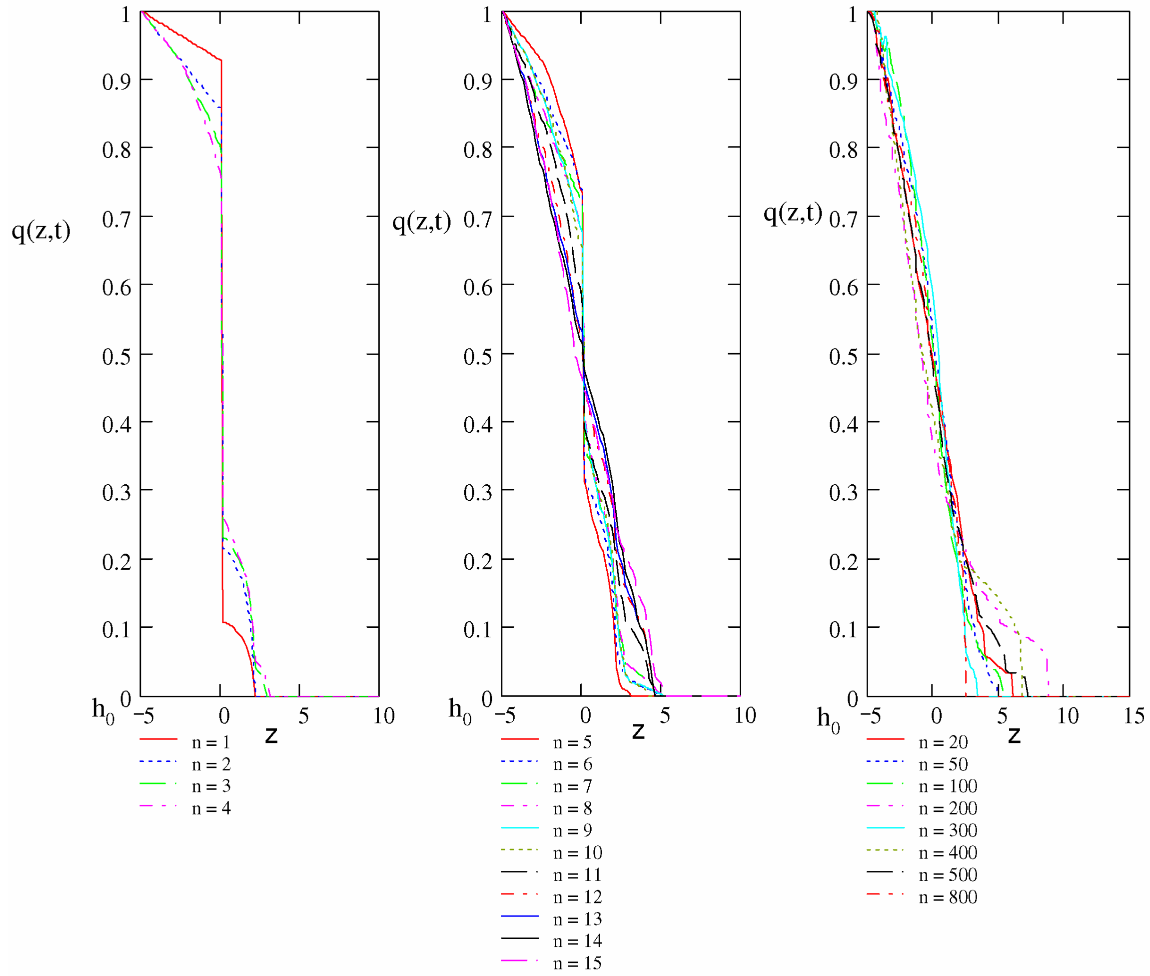
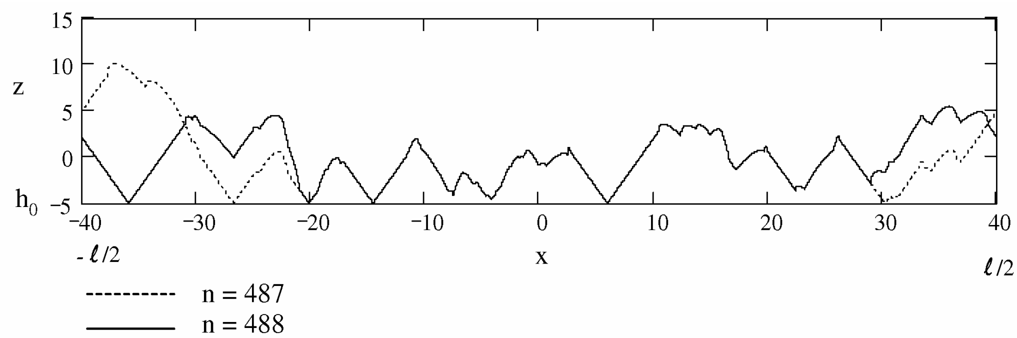


Figure 5.10 Development of the CCDF q(z,t) corresponding to the simulation in Figure 5.9 for n = 1, 2, …15, 20, 50, 100, 200, 300, 400, 500, 800 time steps. The parameters are $\alpha = \pi/3$, $\ell = 80$ µm, $h_0 = -5$ µm, and $f_{add} = 1$.

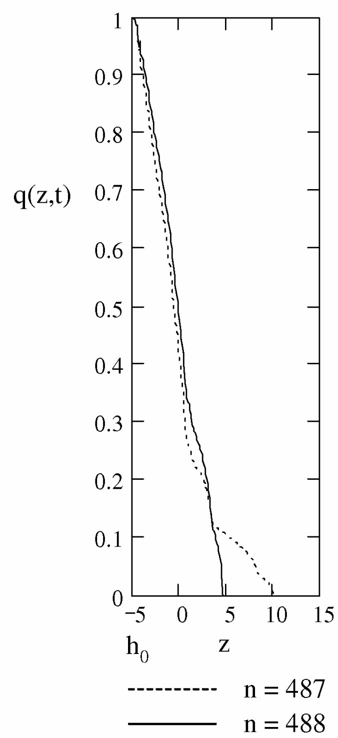
www.manaraa.com

Figure 5.11 Process shows the oscillating tail for $f_{add} = 1$ case with constant height diamond conditioning at h_0 = -5 μm, $\alpha = \pi/3$, and ℓ = 80 μm, corresponding to the simulation in Figure 5.9. (a) At time step n = 488, a high peak at time step n = 477 is killed by the diamond cutter running over it and the mass of the peak is redistributed to the adjacent surface. (b) The corresponding q(z,t) plot for these two time steps corresponding to Figure 5.11(a), showing the oscillating tail. (c) The tail of q(z,t) for all of the time steps corresponding to the simulation in Figure 5.9. The graph shows the value of q(z,t) at z = 12 μm, which corresponds to the location of the tail. It shows about five clusters showing that the tail grows, is killed and grows and is killed again.

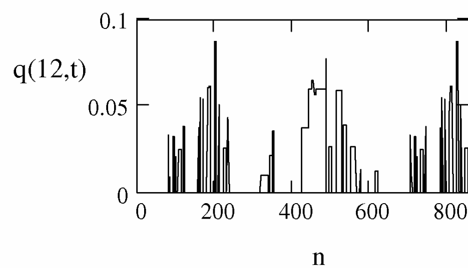
المنارة للاستشارات
www.manaraa.com



(a)



(b)



(c)

المنارة للاستشارات

www.manaraa.com

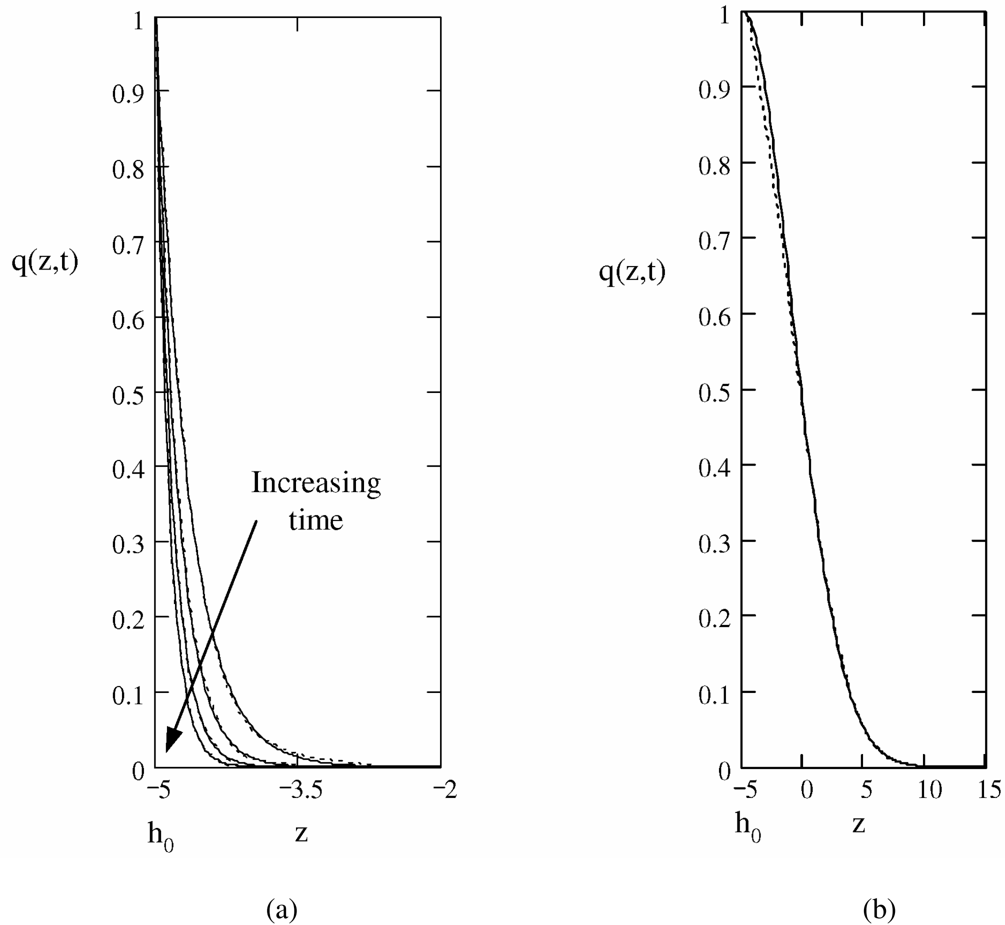


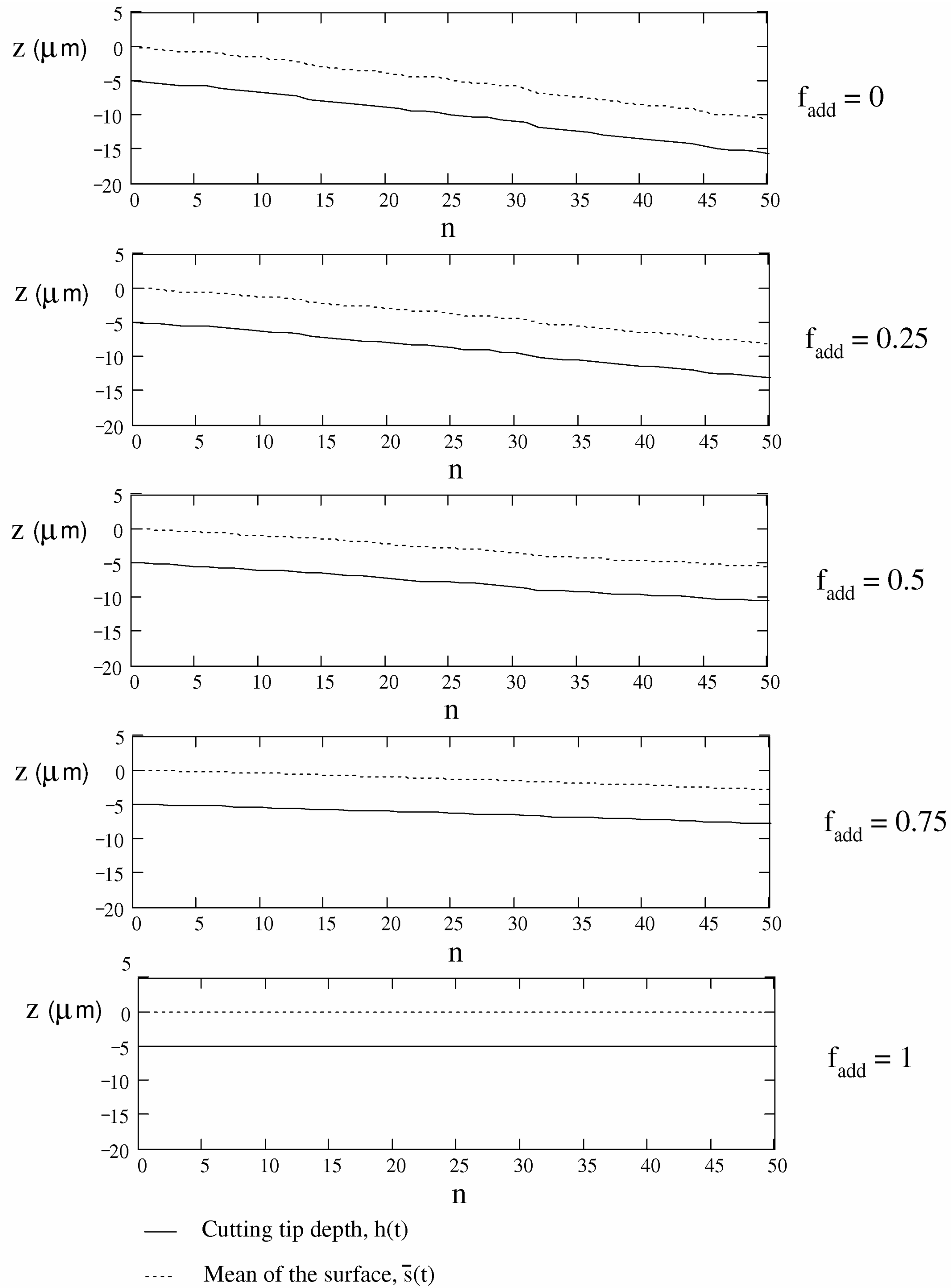
Figure 5.12 Comparison of the CCDF produced by Monte Carlo simulation with the analytical solutions for constant cutting height. Solid lines are analytical solutions and dotted lines are numerical solutions. (a) Comparison of the CCDF produced by Monte Carlo simulation in Figure 5.7 at time steps, n = 200, 300, 400, 500, with the analytical solution Equation (5.19). The parameters are $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm, and $f_{add} = 0$. (b) Comparison of the CCDF produced by Monte Carlo simulation in Figure 5.9 with the analytical solution, Equation (5.36). The parameters are $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm, and $f_{add} = 1$.

المنارة للاستشارات
www.manaraa.com

Figure 5.13 Monte Carlo simulations for the changing of the average surface height (mean) $\bar{s}(t)$ and cutting depth h(t) for different chosen f_{add} values. The program calculates the new mean after each cut and moves down the cutting tip relative to the new mean to ensure that the cutting depth is always h_0 relative to the mean at each cut. The constant cutting depth h_0 is imposed by the load balance for the conditioner under a constant load. The mean $\bar{s}(t)$ lags the cutting depth h(t) by h_0 for all the cases. The slope of $\bar{s}(t)$ or h(t) is the cut-rate. Lager f_{add} value gives smaller cut-rate. At $f_{add} = 1$, the average surface height remains at z = 0 (pure plastic deformation). At $f_{add} = 0$, the average surface height moves down fastest (pure brittle fracture) which is the case in [8]. In these simulations, $\ell = 80$ μm, $\alpha = \pi/3$, and $h_0 = -5$ μm.

المنارة للاستشارات

www.manaraa.com

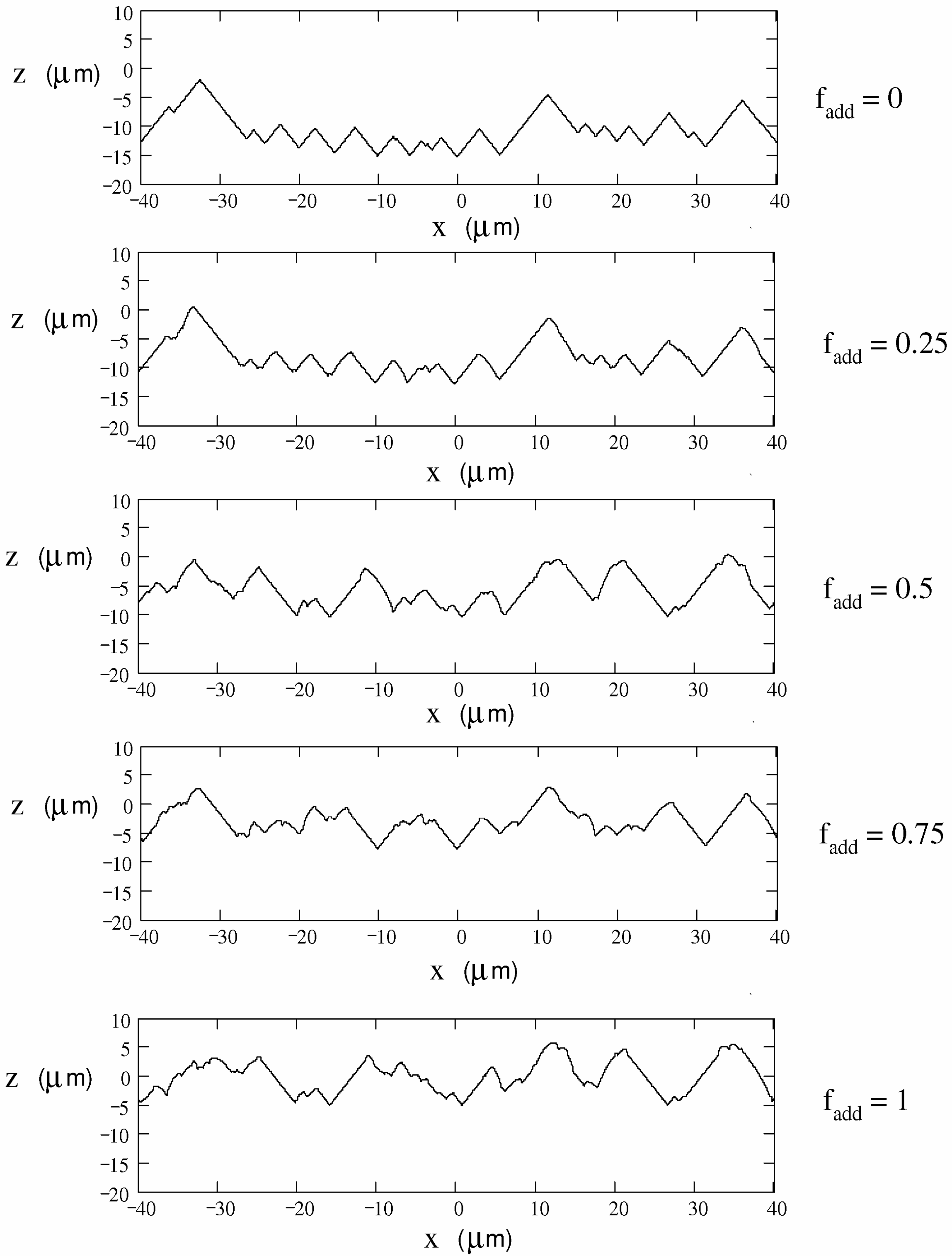


www.manaraa.com

Figure 5.14 The surface profile s(x,t) after 50 time steps for different f_{add} values. The profile moves down slower with increasing f_{add} values. At $f_{add} = 1$, the average surface height remains at z = 0 (pure plastic deformation). At $f_{add} = 0$, the average surface height moves down fastest (pure brittle fracture) which is the case in [8]. In these simulations, $\ell = 80\ \mu m$, $\alpha = \pi/3$, and $h_0 = -5\ \mu m$.

المنارة للاستشارات

www.manaraa.com



المنارة للاستشارات
www.manaraa.com

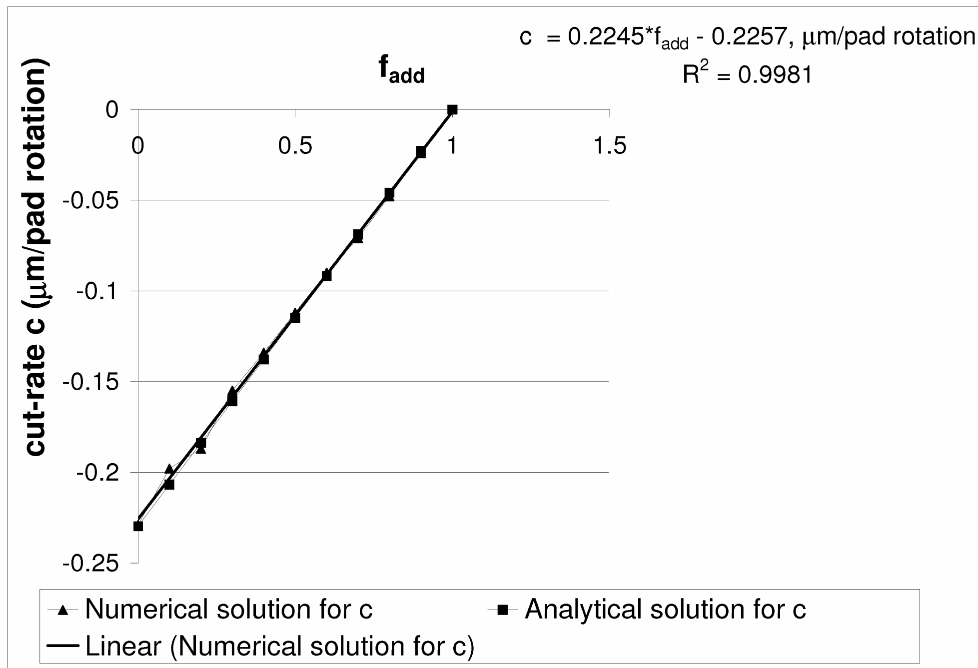


Figure 5.15 The comparison of the Monte Carlo simulation results for the variation of the cut-rate c with f_{add} with the analytical solution, Equation (5.31). The cutting depth h_0 = -5 µm and single diamond period ℓ = 80 µm. Linear fit of the numerical data gives the slope of 0.22 µm/pad rotation and Equation (5.31) gives the slope of 0.23 µm/pad rotation. Numerical solution fits analytical solution very well.

المنارة للاستشارات
www.manaraa.com

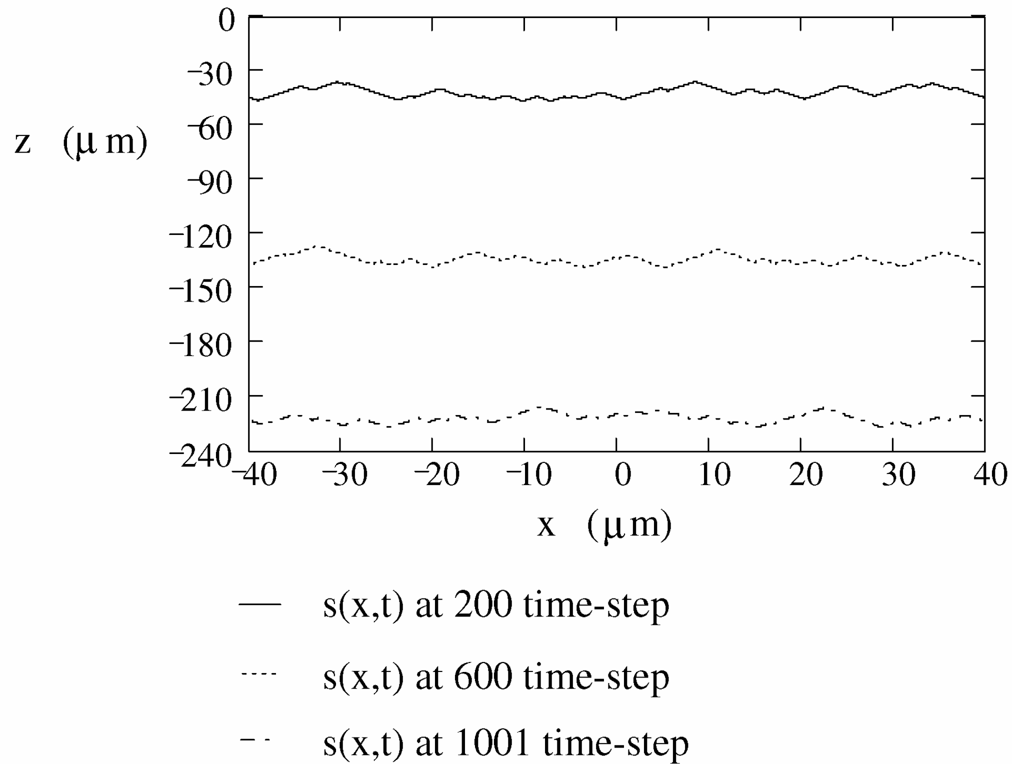
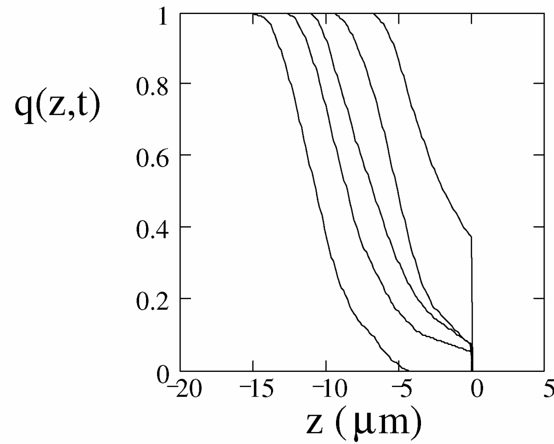
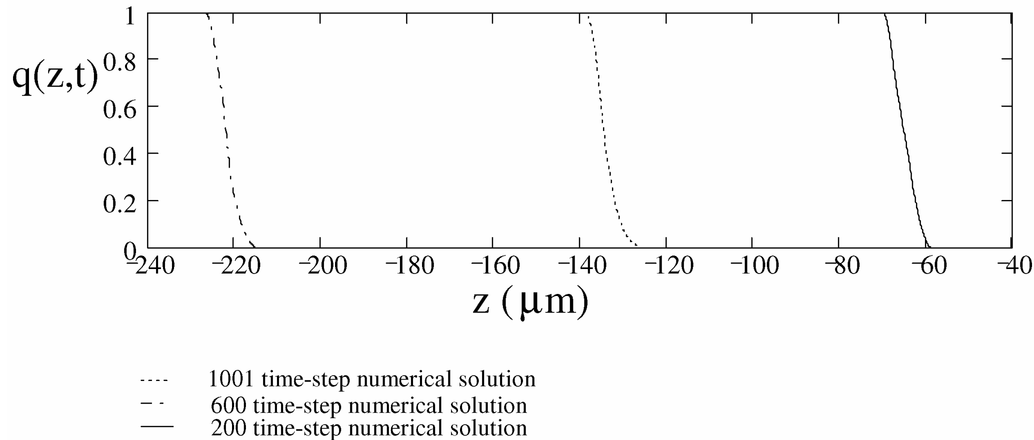


Figure 5.16 The surface profile s(x,t) for constant cut-rate case with $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm, and $f_{add} = 0$. With cut-rate c = -0.23 μm, the average surface height has traveled to about -230 μm at n = 1001 time-step. The original pad surface is smooth and at z = 0.

المنارة للاستشارات
www.manaraa.com



(a)



(b)

Figure 5.17 Snap shots of the surface height CCDF q(z,t) with time. (a) The developing of q(z,t) for the first 50 time-steps is shown and q(z,t) stabilizes at about n = 50 time-step. In the figure, n = 10, 20, 30, 40, 50 respectively from right to left. (b) The stable q(z,t) travels down the z axis with cut-rate c = -0.23 μm for $\alpha = \pi/3$, $\ell = 80$ μm, $h_0 = -5$ μm, and $f_{add} = 0$.

المنارة للاستشارات

www.manaraa.com

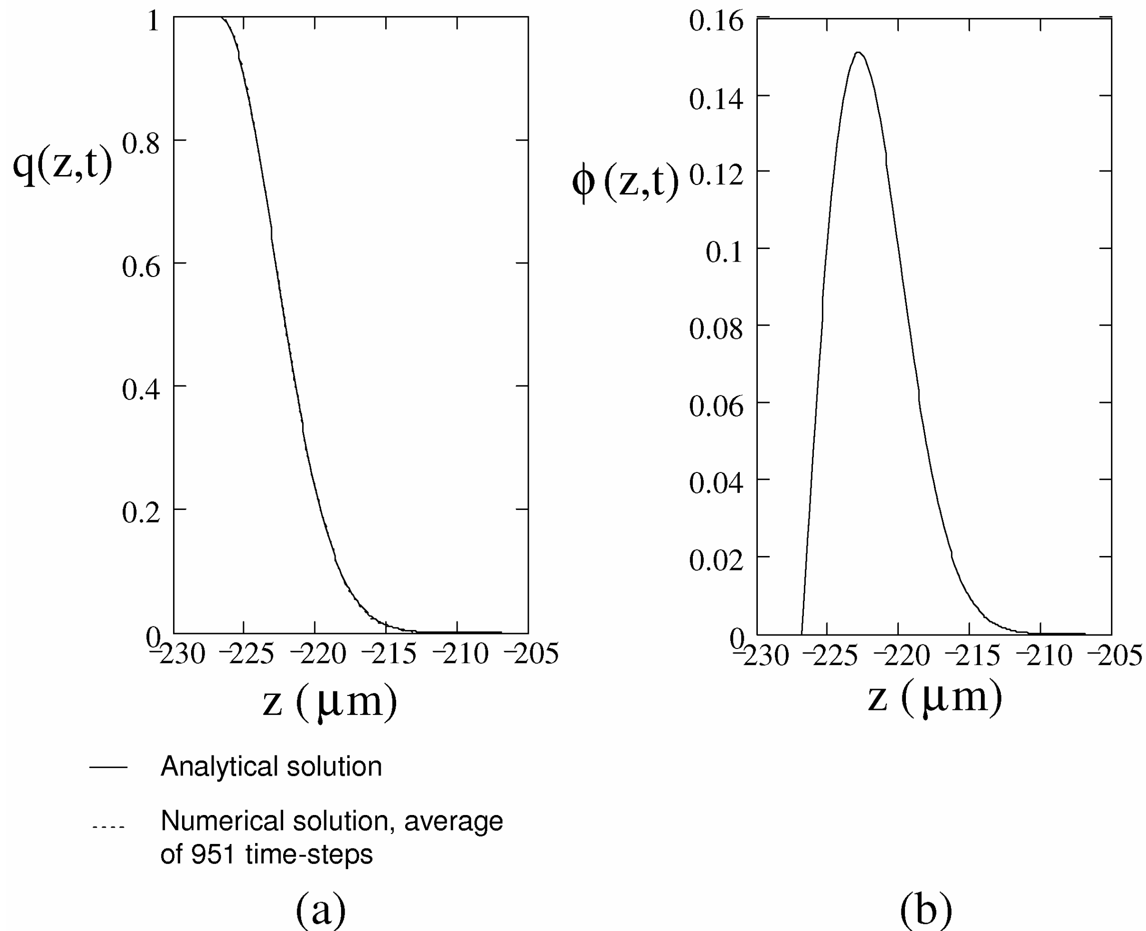


Figure 5.18 The comparison of analytical solution of Equation (5.34) with numerical solution for CCDF and analytical solution of PDF with h_0 = -5 μm and f_{add} = 0. (a) Averaged CCDF q(z,t) from Figure 5.17 for a total of 951 time-steps. Analytical solution of Equation (5.34) with h_0 = -5 μm, single diamond period ℓ = 80 μm and f_{add} = 0 is also shown. The numerical solution matches the analytical solution very well. (b) Analytical solution of Equation (5.35) for PDF φ(z,t) with h_0 = -5 μm, single diamond period ℓ = 80 μm and f_{add} = 0 is shown.

المنارة للاستشارات
www.manaraa.com

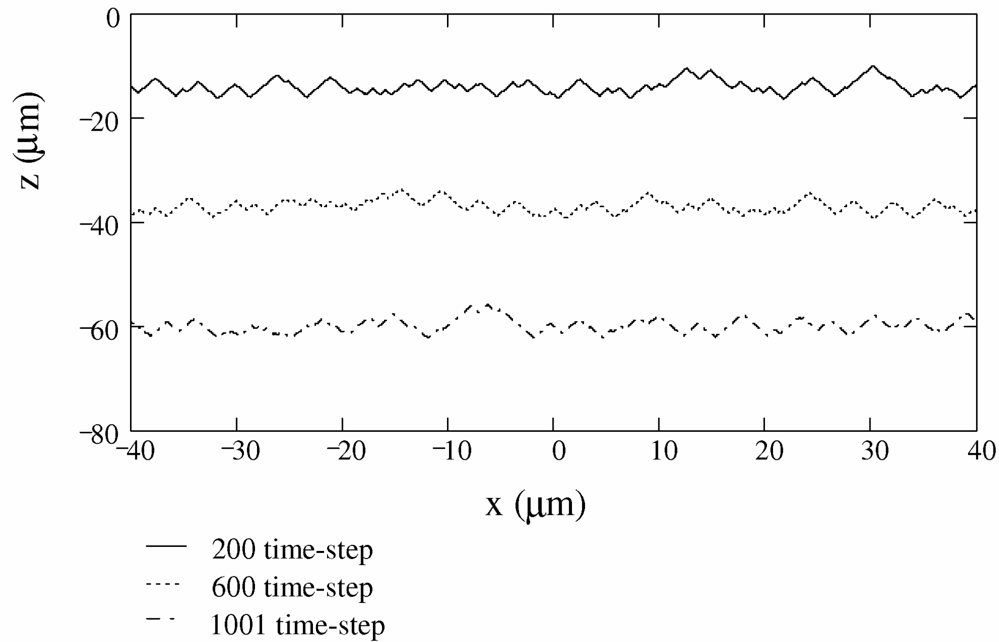
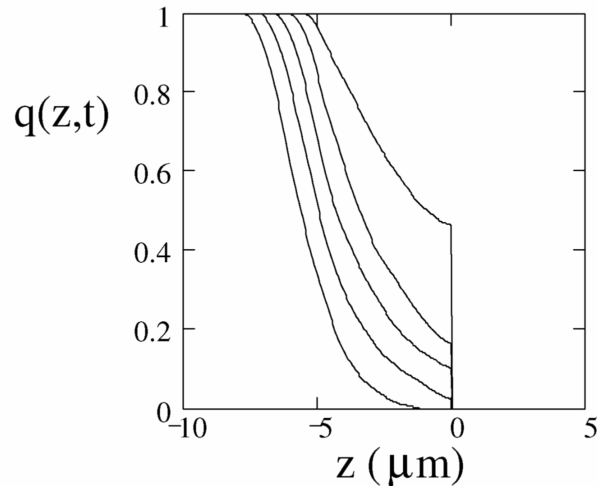


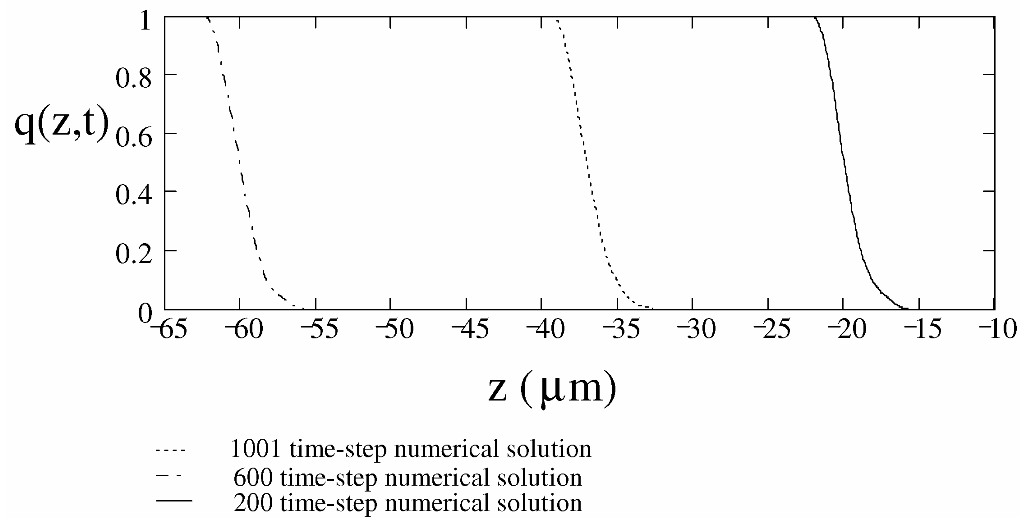
Figure 5.19 The surface profile s(x,t) for constant cut-rate case with h_0 = - 2.5 μm, single diamond period ℓ = 80 μm and f_{add} = 0. With cut-rate c = -0.057 μm, the average surface height has traveled to about z = - 58 μm at n = 1001 time-step. The original pad surface is smooth and at z = 0.

المنارة للاستشارات

www.manaraa.com



(a)



(b)

Figure 5.20 Snapshots of the surface height CCDF with time. (a) The developing of q(z,t) for the first 50 time-steps is shown and q(z,t) stabilizes at about n = 50 time-step. In the figure, n = 10, 20, 30, 40, 50 respectively from right to left. (b) The stable q(z,t) travels down the z axis with cut-rate c = -0.057 μm for h_0 = -2.5 μm, single diamond period ℓ = 80 μm and f_{add} = 0. The profile q(z,t) has traveled to about z = - 58 μm after 1001 time steps (1001 random cuts). The original pad surface is smooth and at z = 0.

المنارة للاستشارات
www.manaraa.com

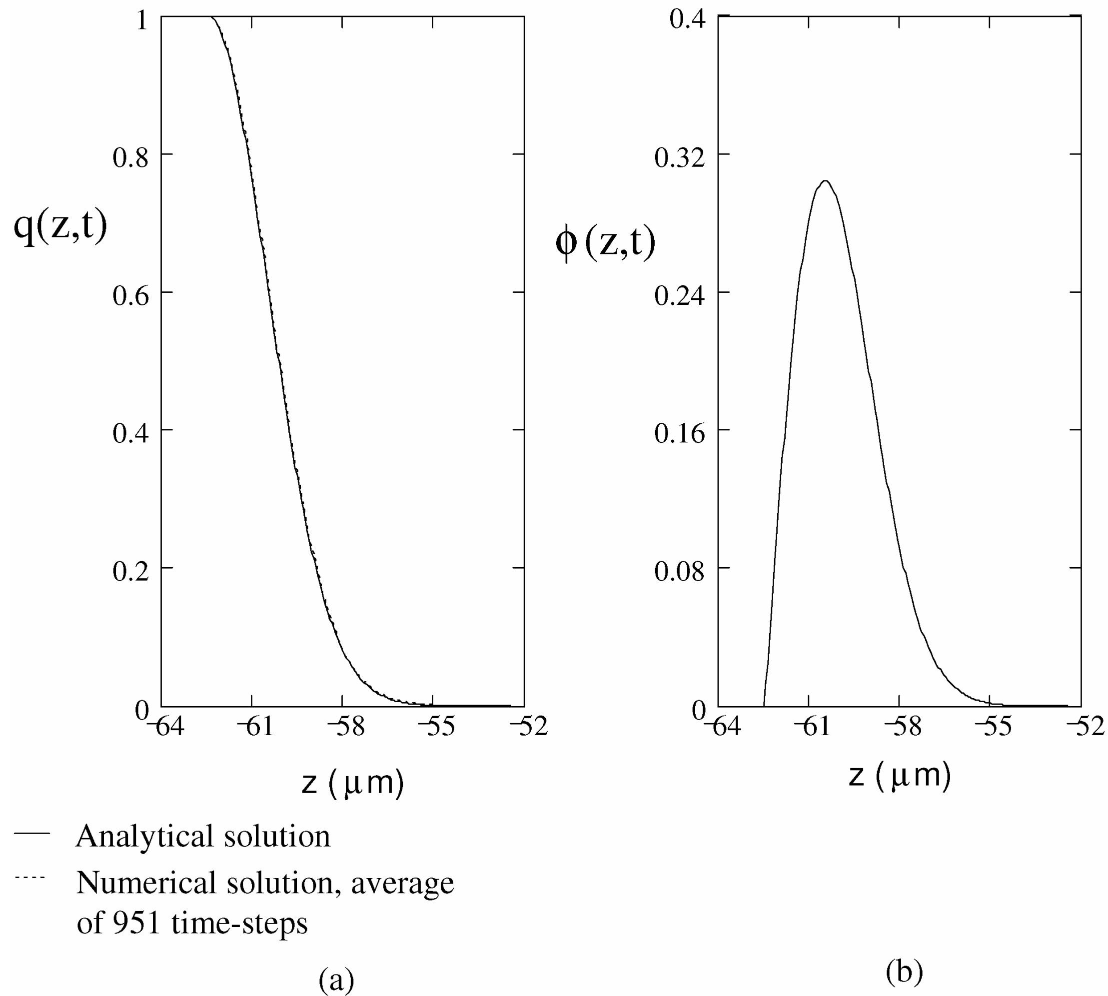


Figure 5.21 The comparison of analytical solution of Equation (5.34) with numerical solution for CCDF and analytical solution of PDF with h_0 = -5 μm and f_{add} = 0. (a) Averaged CCDF q(z,t) from Figure 5.20 for a total of 951 time-steps. Analytical solution of Equation (5.34) with h_0 = -2.5 μm, single diamond period ℓ = 80 μm and f_{add} = 0 is also shown. The numerical solution matches the analytical solution very well and both have standard deviations of about 2 μm which is half of that for h_0 = -5 μm. (b) Analytical solution of Equation (5.35) for PDF φ(z,t) with h_0 = - 2.5 μm, single diamond period ℓ = 80 μm and f_{add} = 0 is shown.

المنارة للاستشارات
www.manaraa.com

CHAPTER 6

A CONDITIONER MODEL WITH VARIABLE DIAMOND HEIGHTS

المنارة للاستشارات

www.manaraa.com

Abstract

Pad conditioning is an important step to restore the pad asperities to sustain a stable polish rate in chemical-mechanical polishing. In this work, a population balance model (PBE) for pad asperities is used to model the pad conditioning. Pad plastic deformation and variable diamond heights for the conditioner are built into the model that are not done in the previous works. Plastic deformation is important in modeling the pad conditioning because the raised ridges along the sides of the cutting grooves are the high points (asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive particles used to polish the wafer. Current conditioner disks are made by attaching numerous similar diamond grits to a metal substrate. The tips of the diamonds are not leveled to the same height resulting in a Gaussian height distribution. This height distribution can induce variance into the conditioned pad asperity probability distribution function (PDF). The model results show that the conditioned pad asperity PDF is Gaussian and is obtained by integrating the average of the conditioned pad asperity PDF from single diamond height. The variance of the conditioned pad asperity PDF is decided solely by the variance of the height of cutting diamonds (active diamonds). If the conditioner diamond height PDF is Gaussian, the model results show that a different interface gap between the conditioner and pad will give a different standard deviation of the resulting conditioned pad asperity PDF. This is consistent with the experimental findings that with increasing load, the roughness of the conditioned pad asperity PDF increases. Model results in this work show that this is caused by the decreasing of the interface gap between the rough surface of the conditioner and the pad surface, which results in the change of the roughness of the active

المنارة للاستشارات

www.manaraa.com

diamonds. When the roughness of the active diamonds is independent of the conditioner-pad separation (as in the case for exponential diamond height distribution), the roughness of the conditioned pad asperity PDF does not depend on the conditioner-pad separation. Hence the diamond height PDF and how its roughness changes with the conditioner-pad separation are very important in pad conditioning. It accounts for polish rate, within wafer nonuniformity and defects. Analytical solutions are derived and found matching the Monte Carlo numerical results.

Introduction

The CMP is carried out as the wafer surface moves across a polishing pad with a slurry. Mechanical motion and down force are applied to the polishing head. The polishing pad surface is very important because it has asperities which make contacts with the wafer [1]. The most important process in CMP is pad conditioning which generates asperities on the pad surface. These asperities force abrasive particles in the slurry against the wafer enabling the polishing operation. Pad asperities must be constantly regenerated by the pad conditioning because they are worn down during polishing. Accordingly, the pad conditioning is essential to maintaining the asperities on the surface of the pad [2].

Pad conditioning is an important step to restore the pad asperities to sustain a stable polish rate. Conditioning can be done either by *in situ* dressing, where the conditioning is performed while the wafers are polished, or by *ex situ* dressing, where the pad is dressed after the wafer has been polished. Current conditioner disks are made by attaching numerous similar diamond grits to the flat surface of a metal substrate using either nickel electroplating or brazing with alloys [3]. The size distribution of the diamond grit

المنارة للاستشارات
www.manaraa.com

together with the deformation of the substrate cause the height variation of the diamond grits. In addition, it is difficult to control the height of so many diamonds. The embedded diamonds will have different shape, orientation and size. Diamond grit is typically sized using sieve techniques, resulting in a Gaussian distribution of sizes with a standard deviation of 15-20% of the average size. Additional variations introduced by the shape of the particles can increase this deviation to as much as 30% of the average size. If the average penetration of diamond grit into the pad at a given downforce is less than one or two standard deviations of the average diamond size, then the actual fraction of grit in contact with the pad can be as low as a few percent of the total grit particles on the conditioner This can produce extreme variations in the performance of the conditioner [4]. Diamond grit sizes can be found in the diamond grit manufacturer's website [5]. For example, a diamond disk with 100 grit size has an average diamond size of 137 μm [5]. An example of a conventional diamond disk is shown in Figure 6.1 [6].

To obtain good uniformity of the polished surface, the height distributions of pad asperities should be as uniform as possible. This requires the diamond heights of the conditioner to be as uniform as possible. For an ideal conditioner, each diamond particle should be precisely the same shape, size and orientation, and the spacing between particles should be uniform. The penetration depth of the grit into the pad would increase monotonically with increasing applied down-force [7]. These can assure uniform force distribution over the whole disk. The force on each diamond is identical and the penetration depth of each diamond into the pad is identical. To obtain these goals, many novel conditioners have been designed [6], [8]. In [6], a novel conditioner is designed based on the forming of a pattern by wire electro-discharge micro-machining of a

المنارة للاستشارات
www.manaraa.com

polycrystalline diamond substrate that is sintered at high pressure and temperature. Each micro-machined diamond pyramid has the same shape and size and the spacing between them is identical. The roughness of the conditioned pad is about 3 μm by the novel conditioner compared to 5 μm by the conventional conditioner, which shows a considerable improvement. In [8], the conditioner is composed of bundles of fine metal fibers to contact pad surface independently. Constant conditioner down-force can be applied to pad surface with fiber edges regardless of long range of height variation of pad surface. In fact, this conditioner made pad surface to be roughened with conforming to a protruding area of 50 μm on pad surface. The uniformity of the conditioned pad is improved as a result.

Pad surface topography has direct impact on the chemical and mechanical polishing rate and polishing uniformity. The conditioner plays a key role in CMP. We have presented a conditioner model with plastic deformation for a constant diamond penetrating depth in Chapter 5 of this thesis. In this work, the model in Chapter 5 is extended to incorporate a variable diamond penetrating depth for the conditioner. Chapter 5 has been submitted to *Journal of Engineering mathematics* and it is referenced as Ref. [9] in this chapter. This chapter is also an independent chapter like other chapters of this thesis. The model in this chapter is based on the model in Chapter 5, so there is some repetition of the material from Chapter 5 in this Chapter to better explain the variable diamond height model which is the focus of this chapter.

Plastic deformation is important in modeling of the pad conditioning because the raised ridges along the sides of the cutting grooves are the high points (asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive

المنارة للاستشارات
www.manaraa.com

particles used to polish the wafer. Current conditioner disks are made by attaching numerous similar size diamond grits to a metal substrate resulting in a Gaussian distribution of diamond height distribution. The tips of the diamonds are not leveled to the same height. This height distribution can induce variance into the conditioned pad asperity PDF. The model results in this work show that the conditioned pad asperity PDF is Gaussian which is different from the previous work that assumes a constant diamond cutting depth [9], [10]. The Gaussian asperity PDF is obtained by integrating the average of the conditioned pad asperity PDF for single diamond height. The variance of the conditioned pad asperity PDF is decided solely by the variance of the cutting diamonds (active diamonds). The distribution of the pad asperities resulting from conditioning being Gaussian is consistent with the common experimental observations in the literature [11], [12]. The model results also show that a different interface gap between the conditioner and pad will give a different standard deviation of the resulting conditioned pad asperity PDF if the conditioner diamond PDF is Gaussian. This is consistent with the experimental findings that with increasing load, the roughness of the conditioned pad asperity PDF increases. Model results in this work show that it is caused by the decreasing of the interface gap between the rough surface of the conditioner and the pad surface resulting in the change of the roughness of the active diamonds. If the roughness of the active diamonds is independent of the conditioner-pad separation (as in the case for exponential diamond height distribution), the roughness of the conditioned pad asperity PDF is independent of the conditioner-pad separation. Hence the diamond height PDF and how its roughness changes with the conditioner-pad separation are very important in pad conditioning. The contact form of the pad asperities is important to determine the

المنارة للاستشارات
www.manaraa.com

contact area and interface gap between the rough pad surface and the wafer. The load on the conditioner, the rotation speed between the conditioner and the pad and the viscosity of the polishing slurry decides the interface gap between the conditioner and the pad. In this work, analytical solutions are derived and found to match Monte Carlo numerical results.

Model

We start with the basic assumptions and model in [9] and then we will derive the model with variable diamond heights and compare it with the model with a single diamond height condition which neglects the variance of the conditioner diamond heights in [9] and [10]. Monte Carlo numerical simulations are performed first and then the corresponding analytical solutions to the population balance for pad asperities are derived.

The top view of a CMP tool geometry is shown in Figure 6.2 (a). Like the wafer, the conditioner is a rotating disk pressed onto a rotating pad-surface. The conditioning disk has a much smaller radius r_c and is periodically swept over a fixed radial track with a constant translational speed U between points P_1 and P_2 and constant rotational speed ω_C as shown in Figure 6.2 (a). The length between P_1 and P_2 is P_1P_2. Conditioning disks are covered with diamond cutting tips distributed randomly over the disk with controlled heights. As the conditioner traverses the track, it makes random cuts on the pad surface. The pad rotates several times during one conditioner trip between P_1 and P_2. To model the conditioning process, we need to find out how many cuts the conditioner will create in one pad rotation. These cuts are distributed between points P_1 and P_2. Hence we can calculate the long time averaged effect for the total cuts between P_1 and P_2 which is equal to the number of pad rotation times the cuts made after one pad rotation. This concept can

المنارة للاستشارات
www.manaraa.com

be illustrated by replacing the circular conditioner by an equivalent one-dimensional bar conditioner [10]. The one-dimensional bar conditioner is shown in Figure 6.2 (b). The N diamond cutting elements of the bar conditioner are assumed to be arranged on a line between points P_1 and P_2 with mean spacing ℓ, $\ell < r_{pad}$. N is the equivalent number of cuts that the pad will be obtained from the circular conditioner after one pad rotation. During each pad rotation, these N cutting elements will displace randomly and independently from each other in their intervals of length ℓ. So the long time averaged effect is that the bar conditioner will create the same number of random cuts per pad rotation between points P_1 and P_2 as the circular conditioner does. A special case for this is when the conditioner and pad rotate at the same direction with the same rotation rate. For this situation, $\ell = P_1P_2/N_D$, in which N_D is the total number of the cutting elements (total number of active diamonds) on the circular conditioner. If the rotation speed of the pad and conditioner is different, $\ell = \omega_P * P_1P_2/(N_D*\omega_c)$, making ℓ scaled by the ratio of the speed between pad and conditioner. Usually the rotation speed of the pad and conditioner is very close. The parameter ℓ is just used to calculate how many cuts between P_1 and P_2 after one pad rotation (density of cuts per pad rotation).

As in [9] and [10], we assume the diamond tip is triangular with tip angle α. In our model, the conditioning process is statistically homogeneous along the radial direction. The coordinate system is based on the initial vertical position of the pad surface which is at $z = 0$. Vertical positions will be measured relative to the initial vertical position of the pad surface. For solid pads, the initial pad surface is assumed smooth. The coordinate system and the terminology are shown in Figure 6.3.

The profile of each of the diamond cutting tips will be represented by the function

www.manaraa.com

$$V(x) = \nu * |x - X| + h(t) \quad (6.1)$$

in which X is a uniformly-distributed random variable on the simulation domain [-ℓ/2, ℓ/2] that gives the horizontal position of the tip. We choose ℓ large enough to contain enough independent cuts so that the spatial statistics are preserved. $\nu = \cot(\alpha/2)$ and $h(t) = h_0 + c_v{*}t$ gives the vertical position of the tip which displaces vertically with time, t. The diamond cutting depth h_0 can be a constant as in previous works [9], [10] or a variable with a PDF for a rough conditioner surface which is the focus of this chapter. The cutting depth is the depth the diamond cut relative to the mean of the surface. The cut-rate c_v here represents the cut-rate for the variable diamond height. Cut-rate represents the speed of the decrease of the mean of the pad surface. The width of the cutting tip is a function of the height z, defined as

$$w(z - h(t)) = 2(z - h(t))/\nu, \qquad z \leq 0 \quad (6.2)$$

The vertical position of a solid pad surface can be expressed as a single-valued function z = s(x,t). If the rotation rate of the platen is Ω radians/sec, then conditioning happens at discrete times $t_n = 2\pi n/\Omega$. We model the conditioning process as a discrete random process in which the surface of the pad is modified at each occurrence with the conditioner within any interval of one diamond cutter [-ℓ/2, ℓ/2], resulting in the transformation $s(x,t_n) \rightarrow s(x,t_n+1)$ for each diamond cut. Polyurethane, which is commonly used as the CMP pad material, is a polymer with plastic mechanical properties (in which material is simply pushed aside and not removed). In [9], we have introduced a new parameter f_{add} to model the material pile-up along the side-walls of the groove cut by

المنارة للاستشارات
www.manaraa.com

the diamond. As shown in Figure 6.3, these two "bumps" are represented by two half-circles at the immediate vicinity of the cut. In our model, the area of these two half-circles is calculated by value of f_{add} and it can range from 0 to 1. The definition of f_{add} in [9] is repeated here in Equation (6.3). The areas A_1 and A_2 are the two side-ridge areas, and A_V is the groove area. It consists of a ratio of the material added back above the original surface due to a scratch to the material removed below the original surface. In practice, only part of the volume of the wear groove is plastically deformed making edges of the groove, the material pileups. It should be noted that f_{add} is equal to zero for ideal micro-cutting (brittle fracture, e.g., all material is removed from the pad), and equal to unity for ideal micro-plowing (plastic deformation, e.g., all materials remains attached to the pad). For polymers with high ductility (percent elongation at failure), f_{add} is close to one. For brittle polymers, f_{add} is close to zero.

$$f_{add} = \frac{A_1 + A_2}{A_V} \tag{6.3}$$

For Monte Carlo simulation, Equation (6.4) represents our model.

$$s(x, t_{n+1}) = \min\left(s(x, t_n), V(x - X_n) + h(t_n)\right) + left \quad bump + right \quad bump \tag{6.4}$$

It means that each time when the diamond makes the cut, the material being cut is piled up onto the pad surface at the edge of the cut as two bumps and the new surface is formed by the new cut and the surfaces of these two bumps. We use two half-circles with equal area to represent the two bumps. However, the simulation does not depend on the shape of the bumps used as long as the mass (area) is preserved by f_{add} parameter.

المنارة للاستشارات
www.manaraa.com

Circular-shaped bumps make physical sense because when the new surface is formed, the material will minimize the surface energy by forming a nearly circular shape. We will show this combined cutting and plastic deformation process (micro-plowing) with variable diamond height in the next section and compare our results with the results done previously with the assumption of a single diamond height in [9] and [10].

Periodic boundary conditions (PBC) are used at the boundaries of the computation domain with one diamond cutter $[-\ell/2, \ell/2]$ in the computing procedure of Equation (6.4) so that the statistical homogeneity of the conditioning process is preserved within the interval and it does not behave differently at the edges of the computational interval. It ensures the conservation of mass for the pad material at the interval $[-\ell/2, \ell/2]$.

For $f_{add} < 1$, we will model the case when the tip decreases at the constant cut rate c_v, and $h(t) = h_0 + c_v*t$. It corresponds to the case when the conditioner eats into the pad by abrasion with time. Constant cut-rate happens at steady-state conditioning.

The structure of the surface is described by the probability function q(z,t), defined as the fraction of the surface which remains above height z after time t. Hence the integration of q(z,t) with z will give the area of the whole surface which conserves mass. The definition of q(z,t) can be written

$$q(z,t) = \Pr ob(z < s(x,t)) \tag{6.5}$$

for any choice of horizontal position x since the conditioning process is statistically homogeneous along the radial direction. The function q(z,t) is the complementary cumulative density function (CCDF). The cumulative density function (CDF) p(z,t) describes probabilities for a random variable z to fall in the interval $(-\infty, z)$. Hence

www.manaraa.com

CCDF describes probabilities for a random variable z to fall in the interval (z,∞). The probability density function (PDF) of a random variable describes the relative frequencies of different values for that random variable. The PDF for the pad surface height is defined by

$$\phi(z,t) = -\frac{\partial q(z,t)}{\partial z} = \frac{\partial p(z,t)}{\partial z} \tag{6.6}$$

To calculate q(z,t) and p(z,t) from ϕ(z,t), Equation (6.7) and (6.8) are used

$$q(z,t) = \int_{z}^{\infty} \phi(z,t)dz \tag{6.7}$$

$$p(z,t) = 1 - q(z,t) = \int_{-\infty}^{z} \phi(z,t)dz \tag{6.8}$$

The average surface height over a sample interval with one diamond cutter [-ℓ/2, ℓ/2] is defined by

$$\bar{s}(t_n) = \frac{1}{\ell}\int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} s(x,t_n)dx = \int_{h(t_n)}^{\infty} z\phi(z,t_n)dz \tag{6.9}$$

The average surface height shows the material lost from the surface and defines the position of the pad surface. By assuming that the initial position of the surface is at z = 0, if $\bar{s}(t_n)$ is a constant, e.g., it stays at z = 0, then there is no material loss and this corresponds to the case of ideal micro-plowing ($f_{add} = 1$) which means that the material just moves around without leaving the surface. If f_{add} is smaller than 1, then the average surface height (the position of the surface) will move towards the negative z axis, e.g., the

المنارة للاستشارات
www.manaraa.com

surface is abraded.

Results

In Monte Carlo simulation, we use a random number generator (RNG) with an exponential distribution and a standard deviation of 5 μm for diamond height PDF. During each cut, the RNG generates a random number greater than the assumed separation depth, d, between the mean of the rough surface of the conditioner diamonds and the pad surface. We assume that the pad surface is smooth initially. The program calculates the new mean after each cut and moves down the cutting tip relative to the new mean. The cutting depth is h_0 which is the diamond height randomly taken from exponential PDF subtracting conditioner-pad separation d at each cut. The results for steady-state complementary cumulative density function (CCDF), q(z,t), with ℓ = 80 μm, $\alpha = \pi/3$, and $f_{add} = 1$ (pure plastic deformation) are shown in Figure 6.4 (a). The parameter ℓ is the average spacing of new cuts per pad rotation which is related to the number of active diamonds on the conditioner, α is the angle of the diamond tip and f_{add} represents the fraction of the material removed from the cutting groove and plowed back onto the pad surface at the edges of the groove during each cut. $1- f_{add}$ represents the fraction of the material that is lost from the pad during each cut. Figure 6.4 (a) shows the averaged q(z,t) from 483 cuts by assuming pad-conditioner separation d = 28.9 μm and the averaged q(z,t) from 448 cuts by assuming d = 9.5 μm. The penetration depth for each cut is the diamond protrusion length as determined from the exponential distribution subtracting the separation d. Figure 6.4 (a) also shows that the two cases are identical and can be fit well by a complementary error function (erfc) with a standard deviation of 5.8 μm [12]. The cumulative distribution function (CDF) describes probabilities for a random

المنارة للاستشارات
www.manaraa.com

variable z to fall in the interval (-∞, z). The CDF of a generic normal random variable z with standard deviation σ is defined in Equation (6.10). The CCDF of a generic normal random variable z with standard deviation σ is defined in Equation (6.11). It describes probabilities for a random normal variable z to fall in the interval (z,∞). Equation (6.11) is used in fitting q(z,t) in Figure 6.4 (a). Hence from the CCDF, the PDF which is the derivative of CCDF is Gaussian with a standard deviation of 5.8 μm. This PDF is shown in Figure 6.4 (b). The results also indicate that for an exponential distribution of cutting diamond heights, the conditioned pad PDF has no d dependence. The derivation of the analytical solutions will be shown later in this work and the derived analytical solutions are found matching the numerical solutions.

$$CDF(z) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sigma\sqrt{2}}\right)\right] \tag{6.10}$$

$$CCDF(z) = \frac{1}{2}\left[1 - erf\left(\frac{z}{\sigma\sqrt{2}}\right)\right] = \frac{1}{2} erfc\left(\frac{z}{\sigma\sqrt{2}}\right) \tag{6.11}$$

The sequence of solid pad surface profiles after successive steps of discrete conditioning process for n = 1, 2, 3, 4, 250 time steps with initially flat surface and variable diamond cutting depth are shown in Figure 6.5. The parameters are $\alpha = \pi/3$, $\ell =$ 80 μm, $f_{add} = 1$ and $\sigma = 5$ μm (diamond height distribution) with exponential distribution. From mass balance, no material is leaving the surface with $f_{add} = 1$, so the mean of the surface stays at the original position (z = 0) during conditioning. By integrating the surface profile at each step and the mean of the surface, $\bar{s}(t)$, is found to remain at zero for all the time steps. In [9], the analytical solutions for q(z,t) (Equation (5.22) and (5.36))

المنارة للاستشارات
www.manaraa.com

and Figure 5.13 show that the mean lags the cutting depth h(t) by h_0 for all the cases of f_{add}. Figure 5.13 in [9] is repeated here in Figure 6.6. Figure 5.14 in [9] which shows the surface profile s(x,t) after 50 time steps for different f_{add} values is repeated here in Figure 6.7. During the simulation, the program calculates the new mean after each cut and moves down the cutting tip relative to the new mean to ensure that the cutting depth is always h_0 relative to the mean at each cut. Analytical solutions for CCDF and PDF for single diamond height in [9] are repeated here in Equation (6.12) and (6.13). Figure 6.6 shows that the mean is behind the cutting tip by h_0 at all times. For $f_{add} = 1$, the mean stays at $z = 0$ and for other $f_{add} < 1$ cases, the mean travels down and lags the cutting tip by h_0. Also the profile is the same for all f_{add} cases as long as h_0 is the same. The surface profiles of CCDF and PDF for single diamond height in [9] are repeated here in Figure 6.8 for comparison purpose with the solutions for variable diamond height in this work. These results for single diamond cutting depth indicate that for variable diamond cutting depth, the mean will lag the cutting depth by the expected value (average) of h_0 for all f_{add} cases. It also indicates that for variable diamond cutting depth, the profile is the same for all f_{add} cases as long as the expected value of h_0 is the same. It is confirmed by the results from Monte Carlo simulation for the changing of the average surface height (mean), $\bar{s}(t)$, and cutting depth, h(t) which are plotted in Figure 6.9 for $f_{add} = 1$ case. In Figure 6.9, the average of h_0, $\bar{h}_0$ is 5.130 μm, which is very close to the standard deviation of the cutting diamond height distribution with exponential PDF, σ, which is 5 μm. For other f_{add} cases, the difference between the mean and the cutting depth should be the same which is the expected value (average) of the cutting diamond penetrating depth, h_0, which is the conclusion from the results of constant diamond height case in [9]. In the constant

www.manaraa.com

diamond height case, the mean lags h(t) by a constant value h_0 (constant cutting diamond penetrating depth) for all the f_{add} cases. The variance of the conditioned pad asperity PDF is determined by the diamond cutting depth h_0 for single diamond cutting depth. Similarly, the variance of the resulting conditioned pad asperity PDF is determined by the average of the cutting diamond penetrating depth h_0 as shown later in the derivation of the analytical solutions and the derived analytical solutions match Monte Carlo results.

$$q(z,t) = \exp\left(-\frac{\pi}{4h_0^2}(z-h_0-ct)^2\right) \qquad h_0 + ct \le z \le \infty \tag{6.12}$$

$$\phi(z,t) = \frac{\pi}{2h_0^2}(z-h_0-ct)\exp\left(-\frac{\pi}{4h_0^2}(z-h_0-ct)^2\right) \quad h_0 + ct \le z \le \infty \tag{6.13}$$

The same simulation but with $f_{add} = 0$ (pure brittle fracture) is also performed in order to determine the cut-rate and the conditioned pad asperity PDF. The results are shown in Figure 6.10, 6.11 and 6.12. Similarly to the results for single diamond height shown in Figure 6.6 and 6.7, the mean $\bar{s}(t)$ and cutting depth h(t) move down with an effectively constant cut-rate at steady-state conditioning. Linearly fitting for the mean or the cutting depth gives the cut-rate. The mean $\bar{s}(t)$ lags the cutting depth h(t) by the expected value of h_0. At each cut, the cutting depth h_0 is determined by a diamond height randomly chosen from exponential distribution subtracting the conditioner-pad separation d. Comparing Figure 6.6 with Figure 6.9 and Figure 6.11, we can see that the only difference between the results of variable diamond height and those of single diamond height is that the cutting depth h(t) is not a straight line any more, instead it has a variance introduced by the variable diamond height. For non-zero cut-rate, the mean is

www.manaraa.com

also not a straight line any more, it has a variance introduced by the variable diamond cutting depth. It should be mentioned that the more the random numbers are used to do the simulation, the more accurate and reproducible the results are. But more random numbers increase the computation time. Comparing Figure 6.8 with Figure 6.4 and Figure 6.12, we can also see that the CCDF is half-Gaussian with a sharp turn-off at the cutting depth h_0 for single diamond height. While the CCDF for variable diamond height is a full-Gaussian with a left-hand tail. Consequently, the PDF for variable diamond height is Gaussian and the PDF for single diamond height is not Gaussian for all f_{add} values. Hence the variance from the diamond height distribution has changed the shape of the conditioned pad asperity PDF and the diamond height distribution for the conditioner is very important for CMP. Next we will show the derivation of the analytical solutions for the cut-rate and conditioned pad asperity PDF for the case of variable diamond height.

In [9], the analytical solutions for the cut-rate and steady-state conditioned pad asperity PDF for the case of constant diamond height have been derived and verified by the numerical solutions. These analytical solutions are repeated here in Equations (6.12) to Equation (6.15). Equation (6.12) is the CCDF and Equation (6.13) is the PDF. Equation (6.14) is the cut-rate with unit of µm per minute and Equation (6.15) is the cut-rate with unit of µmper pad rotation.

$$c = \frac{2h_0^2 \omega_p}{\pi^2 \nu \ell}\left(f_{add} - 1\right) \tag{6.14}$$

$$c = \frac{4h_0^2}{\pi \nu \ell}\left(f_{add} - 1\right) \tag{6.15}$$

www.manaraa.com

The cut-rate can be derived by replacing the variable, h_0^2, with its expected value to get the average cut-rate as shown in Equation (6.16). For an exponential distribution with a standard deviation of σ shown in Equation (6.17), the expected value of h_0^2 is $2*\sigma^2$.

$$\bar{h}_0^2 = \frac{\int_d^\infty (h-d)^2 \phi_d \, dh}{\int_d^\infty \phi_d \, dh} \tag{6.16}$$

$$\phi_d = \frac{1}{\sigma} e^{-\frac{h}{\sigma}} \tag{6.17}$$

So the analytical solutions for the cut-rate are shown in Equation (6.18) and (6.19). These are the averaged cut-rates for all the active (cutting) diamonds. Equation (6.18) is the cut-rate in units of μm per minute and Equation (6.19) is the cut-rate in units of μm per pad rotation. So the d dependence of the cut-rate only shows up in ℓ which decreases as d decreases because more diamonds are cutting and that will increase the cut-rate. As explained previously, ℓ is the length between the cuts per pad rotation. Reducing the load will increase the number of the active diamonds and consequently ℓ is decreased.

$$c_v = \frac{4\sigma^2 \omega_p}{\pi^2 v \ell} \left(f_{add} - 1 \right) \tag{6.18}$$

$$c_v = \frac{8\sigma^2}{\pi v \ell} \left(f_{add} - 1 \right) \tag{6.19}$$

The expected value of h_0 can be calculated using Equation (6.20). For exponential diamond height distribution, the expected value of h_0 is calculated as σ.

المنارة للاستشارات

www.manaraa.com

$$\bar{h}_0 = \frac{\int_d^\infty (h-d)\phi_d \, dh}{\int_d^\infty \phi_d \, dh} \tag{6.20}$$

To derive the analytical solution for the conditioned pad asperity PDF, if we replace the cut-rate c and cutting depth h(t) by their corresponding expected values (average values) in the population balance equation, Equation (5.13) in [9] and repeated as Equation (6.21) here, the averaged population q(z,t) and ϕ(z,t) are calculated as shown in Equation (6.22) and (6.23) for exponential distribution.

$$\frac{\partial q(z,t)}{\partial t} + \frac{\partial}{\partial z}\left(cq(z,t)\right) = \left(f_{add} - 1\right)\frac{\omega_p}{\pi \nu \, \ell}\left(z - h(t)\right)q(z,t) \qquad h_0 + ct \le z \le \infty \tag{6.21}$$

$$\bar{q}(z,t) = \exp\left(-\frac{(1-f_{add})\omega_p}{2\pi c_v \nu \ell}\left(z - \sigma - c_v t\right)^2\right) \qquad \sigma + c_v t \le z \le \infty \tag{6.22}$$

$$\bar{\phi}(z,t) = \frac{(1-f_{add})\omega_p}{\pi c_v \nu \ell}\left(z - \sigma - c_v t\right)\exp\left(-\frac{(1-f_{add})\omega_p}{2\pi c_v \nu \ell}\left(z - \sigma - c_v t\right)^2\right) \tag{6.23}$$

$$\sigma + c_v t \le z \le \infty$$

in which the cut-rate c_v is defined in Equation (6.18). Equation (6.22) and (6.23) can also be obtained by replacing h_0^2 and h_0 by their expected values $2*\sigma^2$ and σ respectively in Equation (6.12) and Equation (6.13).

Then the CCDF and PDF of the conditioned pad surface for variable diamond cutting depth can be obtained by integrating Equation (6.22) and (6.23) in the range (-∞, z) as shown in Equation (6.24) and (6.25) in the following.

المنارة للاستشارات

www.manaraa.com

$$q_v(Z) = \frac{1}{2}\left[1 - erf\left(\frac{Z}{\sigma_v\sqrt{2}}\right)\right] = \frac{1}{2}erfc\left(\frac{Z}{\sigma_v\sqrt{2}}\right) \qquad -\infty < Z < \infty \tag{6.24}$$

$$\phi_v(Z) = \sqrt{\frac{(1-f_{add})\omega_p}{2\pi^2 c_v \nu \ell}}\exp\left(-\frac{(1-f_{add})\omega_p}{2\pi c_v \nu \ell}Z^2\right) \qquad -\infty < Z < \infty \tag{6.25}$$

in which, $Z = z - \sigma - c_v{*}t$ and σ_v is the standard deviation of the conditioned pad PDF with variable diamond heights of exponential distribution. By replacing c_v with Equation (6.18) for exponential distribution of diamond height PDF, σ_v and $\phi_v(Z)$ are

$$\sigma_v = \frac{2}{\sqrt{\pi}}\sigma \tag{6.26}$$

$$\phi_v(Z) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left(-\frac{1}{2\sigma_v^2}Z^2\right) \qquad -\infty < Z < \infty \tag{6.27}$$

So from the analytical solutions in Equation (6.26) and (6.27), the standard deviation of the conditioned pad asperity PDF, σ_v is 5.642 μm if σ is 5 μm for the diamond height PDF. The results in Figure 6.4 and Figure 6.12 for the average of 556 cuts which are 5.8 μm ($f_{add} = 1$) and 5.6 μm ($f_{add} = 0$) respectively match the analytical solution well. From Equation (6.19), the cut-rate for $\ell = 80$ μm and $f_{add} = 0$ is 0.46 μm per pad rotation. The numerical result from Figure 6.11 is 0.49 μm, which matches the analytical solution well. It should be mentioned that when more random numbers are used to do the simulation, more accurate results are achieved. The numerical results show the trends of the data very well with reasonable accuracy.

It should be mentioned that the numerical results are done by using a large number

المنارة للاستشارات
www.manaraa.com

of active diamonds. Because when the number of the active diamonds increases, the expected values of h_0^2 and h_0 approach to $2\sigma^2$ and σ respectively. In the real situation, if the number of diamonds on the conditioner is small, the result will be different and it depends on the population of the active diamonds. For statistical purpose, to get accurate results, a large number of active diamonds need to be used. When we do the Monte Carlo simulation, we only use the diamond heights greater than d corresponding to the heights of the active diamonds. The diamond cutting depth is the active diamond height randomly taken from exponential distribution after subtracting d. For example, in Figure 6.4 (a), 483 active diamonds are used to do the cuts. This number is good enough to get the results with reasonable accuracy.

Both analytical solution and numerical solution show Gaussian distribution of the conditioned pad asperity PDF with exponential diamond height PDF and this indicates other diamond height distributions will all give Gaussian distribution and the standard deviation of the conditioned pad asperity PDF will depend on the expected value of the cutting depths which will depend on the diamond height PDF. In general, from Equation (6.14) and (6.25), the standard deviation of the conditioned pad asperity PDF is

$$\sigma_v = \sqrt{\frac{2\bar{h_0^2}}{\pi}} \tag{6.28}$$

in which $\bar{h_0^2}$ is the expected value of h_0^2 and it depends on the diamond height PDF calculated by using Equation (6.16).

For normal distribution, $\bar{h_0^2}$ is calculated by replacing ϕ_d in Equation (6.16) with the normal distribution PDF as shown in Equation (6.29).

المنارة للاستشارات

www.manaraa.com

$$\bar{h_0^2} = \frac{\int_d^\infty (h-d)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{h^2}{2\sigma^2}} dh}{\int_d^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{h^2}{2\sigma^2}} dh} \quad (6.29)$$

From Equation (6.16), (6.17) and (6.28), the standard deviation of the conditioned pad asperity PDF does not depend on the gap, d, if the diamond height PDF is exponential. And also from Equation (6.18) and (6.19), the cut-rate depends on d only through ℓ which depends on d. Increasing the load on the conditioner will reduce d and increase the number of the active diamonds. This will reduce ℓ and increase the cut-rate as shown by Equation (6.18). That cut-rate is increased by increasing load has been verified by experiments [14]. Increasing cut-rate means the aggressiveness of the conditioning is increased and the pad surface can reach steady-state faster. With exponential diamond height PDF, the surface profile of the conditioned pad is determined by the conditioner diamond height PDF, not by process parameters.

Not all of the conditioned pad asperity PDF is independent of d. It depends on how the roughness of the active diamonds changes with d. By using Equation (6.28) and (6.29), we have calculated σ_v for the conditioned pad asperity PDF for a Gaussian distribution of diamond heights. The results show that the standard deviation of the conditioned pad asperity PDF strongly depends on the separation d for a Gaussian distribution of the diamond height PDF. The results are shown in Figure 6.13. It shows that σ_v increases as the standard deviation of diamond height PDF, σ, increases as expected. While σ_v decreases fast as d increases. This is consistent with the experimental findings that with increasing load, the roughness of the conditioned pad asperity PDF

المنارة للاستشارات
www.manaraa.com

increases [12]. Model results in this work suggests that it is caused by the decreasing of the interface gap between the rough surface of the conditioner diamonds and the pad surface resulting in the change of the roughness of the active diamonds. If the roughness of the active diamonds is independent of the conditioner-pad separation (as in the case for exponential diamond height distribution), the roughness of the conditioned pad asperity PDF does not depend on the conditioner-pad separation. Hence the diamond height PDF and how its roughness changes with the conditioner-pad separation are very important in pad conditioning. In CMP, when the load increases, d decreases and this results in increasing in the roughness of the active diamonds, which results in increasing in the roughness of the conditioned pad asperity PDF. Increasing velocity and/or increasing the viscosity of the polishing slurry will increase the fluid pressure, which results in increasing in d and decreasing in the standard deviation of the conditioned pad asperity PDF. For the conventional diamond disks made by embedded diamond grits into the substrate, as shown in Figure 6.1, the diamond height PDF is a normal distribution and this is consistent with our results because normal diamond PDF is used in the simulation.

Figure 6.14 and Figure 6.15 show the measured newly-conditioned pad [12]. The conditioners are made by diamonds with 100 and 200 grit sizes which correspond to standard deviations of about 16 μm and 8 μm respectively [5]. The type of pad is Rohm and Haas IC 1000 plain pad. The conditioners are provided by Mitsubishi Materials Corporation. Figure 6.14 shows the changes of the conditioned pad PDF with two different conditioner loads which are 3.6 lb and 8 lb respectively for 100 grit conditioner. The standard deviation of the conditioned pad PDF is increased as load is increased which shows the same trend as predicted by our model. Both PDFs can be fit well by

www.manaraa.com

Gaussian distribution which is consistent with our model results. Figure 6.15 shows the variation of σ_v with the load on the conditioner for 100 grit and 200 grit conditioners. As σ increases, σ_v increases. However because when σ increases (the grit size increases), d is increased at the same load, σ_v only increases a small amount as shown in Figure 6.15. For example, at 9 lb of load, σ_v increases only 1 μm from 2.7 μm to 3.7 μm when the grit size changes from 200 grit to 100 grit. Also as load increases, σ_v increases, which is consistent with the trend predicted by our model. When the load increases the cut-rate increases because σ increases and ℓ decreases (or the contact area between the conditioner and pad increases), which shows the same trend with the experimental results [14].

As shown by Equation (6.27) and (6.28), the profile of the conditioned pad surface only depends on the expected value of h_0, which depends on the diamond height PDF and conditioner-pad separation. The plastic deformation parameter f_{add} only changes the cut-rate as shown by Equation (6.18). This is confirmed by the results in Figure 6.4 and Figure 6.12 which show the same standard deviation but different cut-rate by using f_{add} = 1 and f_{add} = 0. For constant diamond height, the profile of the conditioned pad surface only depends on the diamond penetration depth h_0 and f_{add} only changes cut-rate [9].

Conclusion

In this work, we have presented a conditioning model with plastic deformation and variable diamond heights. Our model results have shown that the PDF of the conditioned pad surface with a variable diamond height is Gaussian, which is consistent with the common experimental observations in the literature. According to our model results, the diamond height PDF only changes the standard deviation of the conditioned pad asperity PDF. The contact form of the pad asperities is important to determine the contact area

المنارة للاستشارات

www.manaraa.com

and interface gap between the rough surface of the pad and the wafer. The load on the conditioner, the rotation speed between the conditioner and the pad and the viscosity of the polishing slurry decide the interface gap between the conditioner and the pad. Our results show that a different interface gap will give different standard deviation of the resulting conditioned pad PDF if the diamond height PDF is Gaussian. This is consistent with the experimental results in the literature that with increasing load, the roughness of the conditioned pad asperity PDF increases. According to our model results, it is the result of the decreasing of the interface gap between the rough surface of the conditioner diamonds and the pad surface, which results in increasing of the standard deviation of the active diamonds.

For exponential distribution of the diamond height PDF, the reason that the standard deviation of the conditioned pad asperity PDF does not depend on conditioner-pad separation is that the standard deviation of the PDF in contact with the pad does not vary with d. For normal distribution of diamond height PDF, because the standard deviation of the PDF in contact with the pad varies with d, this results in a strong dependence of the standard deviation of the conditioned pad asperity PDF on d. This is more pronounced when the standard deviation of the diamond height PDF is increased as was observed in experiments. These results have shown that the diamond height PDF is very important in chemical-mechanical polishing because pad surface topography has direct impact on the chemical and mechanical polishing rate, polishing uniformity and defects. The diamond height PDF is also important in deciding the cut-rate, an important parameter to determine the time to bring the pad surface to steady-state and the time to change the pad. Our model shows that cut-rate increases as load increases which is

المنارة للاستشارات
www.manaraa.com

consistent with experimental results in the literature. Our model has incorporated pad plastic deformation into the model which shows improvement in calculating the cut-rate than the model with only brittle fracture done previously.

References

[1] M. R. Oliver, *Chemical Mechanical Planarization of Semiconductor Materials*. Berlin, Germany: Springer, 2004.

[2] A. S. Lawing, "Polish rate, pad surface morphology and oxide chemical-mechanical polishing," in *Electrochem. Soc. Proc.*, vol. 2002-1, 2002, pp. 46-60.

[3] M. Y. Tsai and Y. S. Liao, "Dressing characteristics of oriented single diamond on CMP polyurethane pad," *Mach. Sci. Technol.*, vol.13, pp. 92-105, 2009.

[4] E. Thear and F. Kimock, *Improving Productivity Through Optimization of the CMP Conditioning Process* [Online]. Available: http://www.morgantechnicalceramics.com /resources/technical_articles/improving-productivity-through-optimization-of-the-cmp-conditioning-process/?page_index=1

[5] *Diamond Grit Size Designations, Approximate Particle Sizes, and Counts* [Online]. Available: http://www.inlandlapidary.com/user_area/gritchart.asp

[6] Y. Kim and S. L. Kang, "Novel CVD diamond-coated conditioner for improved performance in CMP processes," *Int. J. Mach. Tool Manu.*, vol. 51, pp. 565-568, 2011.

[7] J. Zimmer and A. Stubbmsnn, "Key factors influencing performance consistency of CMP pad conditioners," in *Proc. 2nd Int. Symp. on CMP*, 1998, pp. 87-92.

[8] H. Gao and J. Su, "Study on the surface characteristics of polishing pad used in chemical mechanical polishing," *Adv. Mater. Res.*, vol. 102-104, pp. 724-728, 2010.

[9] H. Shi and T. A. Ring, "Plastic deformation theory of chemical-mechanical polishing pad conditioning," submitted to *J. Eng. Math.*, 2011.

[10] L. J. Borucki, T. Witelski, C. Please, P.R. Kramer, and D. Schwendeman, "A theory of pad conditioning for chemical-mechanical polishing," *J. Eng. Math.*, vol. 50, pp. 1-24, 2004.

[11] K. Park and H. Jeong, "Investigation of pad surface topography distribution for material removal uniformity in CMP process," *J. Electrochem. Soc.*, vol. 155, no. 8, pp. H595-H602, 2008.

المنارة للاستشارات

www.manaraa.com

[12] T. Sun, L. Borucki, Y. Zhuang and A. Philipossian, “Investigating the effect of the diamond size and conditioning force on chemical mechanical planarization pad topography,” *Microelectron. Eng.*, vol. 87, pp. 553-559, 2010.

[13] L. Borucki, H. Lee, Y. Zhuang, N. Nikita, R. Kikuma, and A. Philipossian, “Theoretical and experimental investigation of conditioner design factors on tribology and removal rate in copper chemical mechanical planarization,” *Jpn. J. Appl. Phys.*, vol. 48, pp. 115502 (1-6), 2009.

[14] A. R. Baker, “Conditioning for removal rate stability,” in *CMP-MIC Conference*, 1997, pp. 339-342.

المنارة للاستشارات

www.manaraa.com

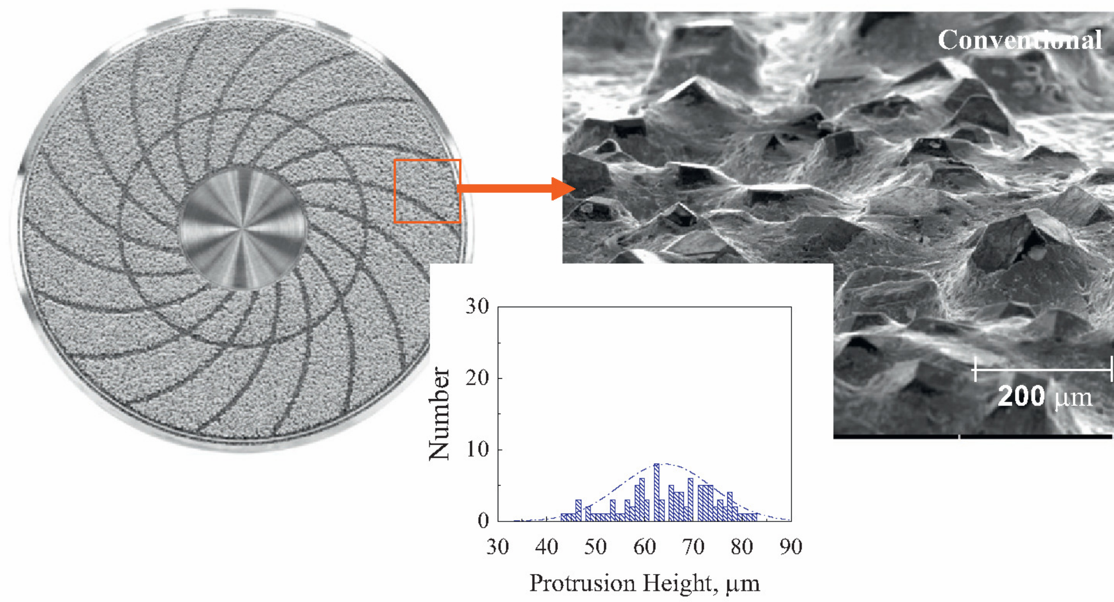


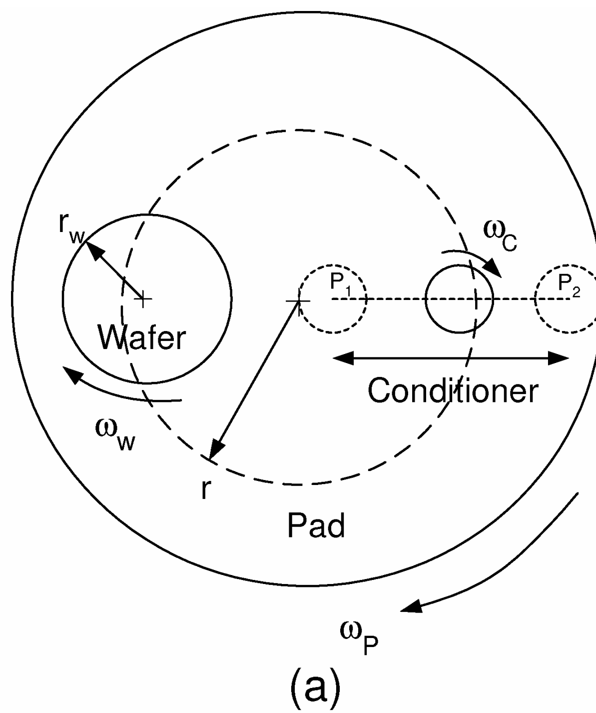
Figure 6.1 Macro- and micrographs of a conventional diamond conditioner and the height distribution of the diamond grits on the conditioner. Reprinted from [6], with permission from Elsevier.

المنارة للاستشارات
www.manaraa.com

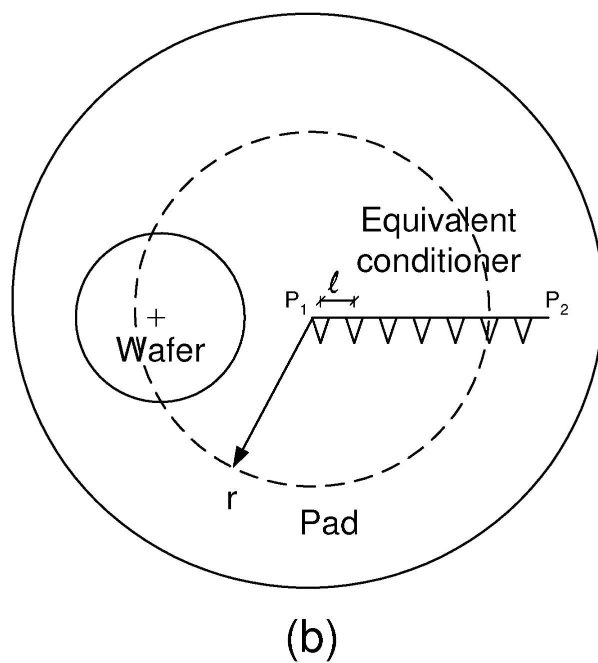
Figure 6.2 Illustration of a CMP tool with the circular conditioner and the bar conditioner. (a) Top view of a CMP tool geometry. The pad section between P_1 and P_2 travels along a circular path as shown in the figure with a small radial cross-section of the pad at the end of a vector r as an example and encounters the wafer once per rotation. It will also encounter the circular conditioner during a few of the periods in each conditioning sweep between P_1 and P_2. (b) The circular conditioner is replaced by an equivalent bar conditioner. A small portion at r of a radial cross-section encounters both the wafer and the conditioner once per period.

المنارة للاستشارات

www.manaraa.com



(a)



(b)

المنارة للاستشارات
www.manaraa.com

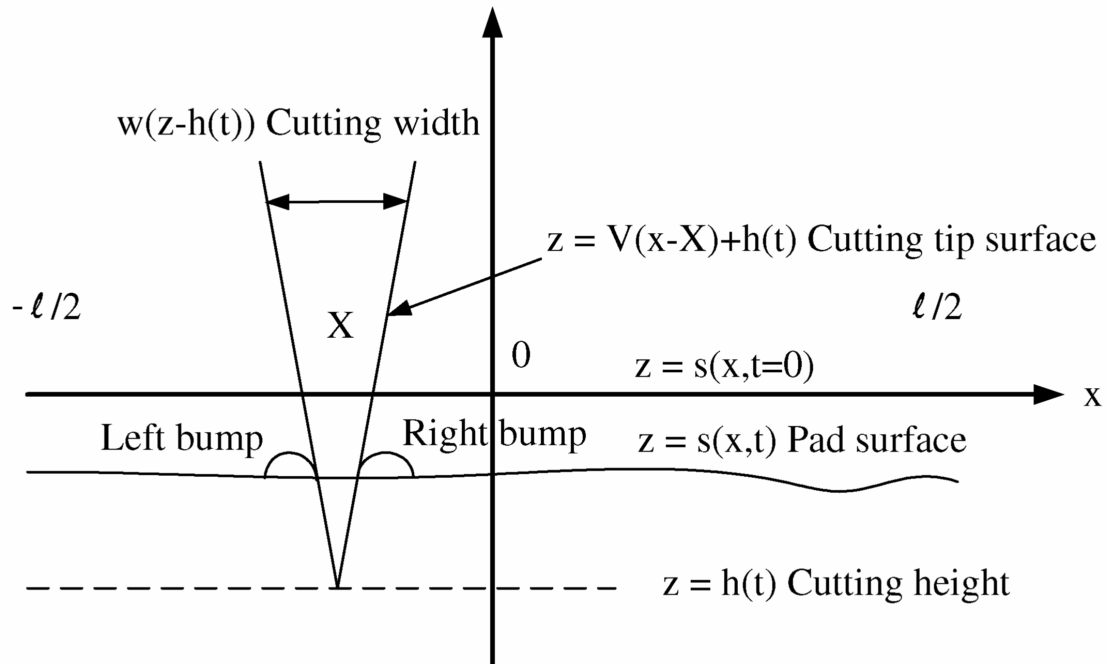


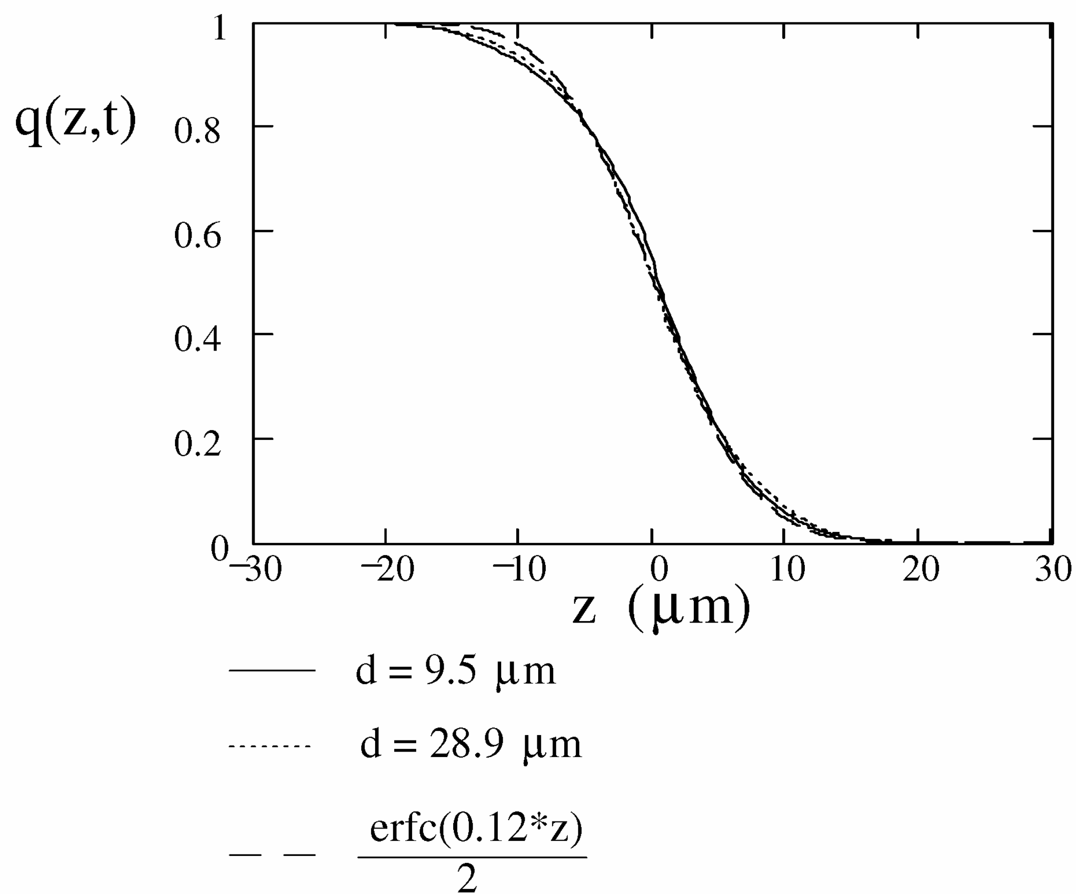
Figure 6.3 The coordinate system and terminology used in this work.

المنارة للاستشارات
www.manaraa.com

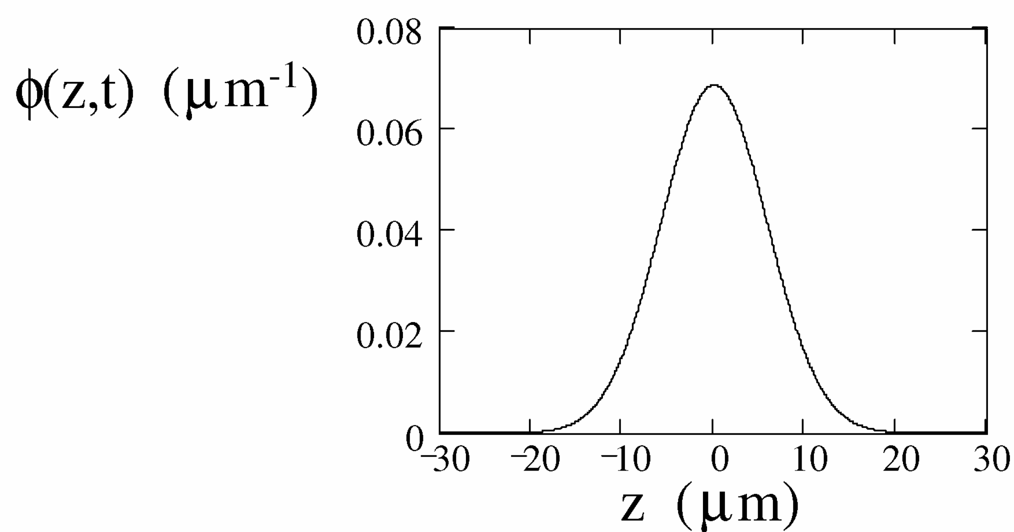
Figure 6.4 Steady-state CCDF with an exponential diamond cutting depth distribution and the corresponding PDF. (a) Steady-state CCDF with an exponential diamond cutting depth distribution. The standard deviation of exponential distribution is 5 μm. The solid line shows the case with separation distance between the rough diamond surface and the pad, d = 9.5 μm and the dotted line shows the case with d = 28.9 μm. The dashed line shows the fitting of the two cases very well by complementary error function (erfc). The erfc function has a standard deviation of 5.8 μm. (b) The corresponding PDF of a normal distribution with standard deviation of 5.8 μm.

المنارة للاستشارات

www.manaraa.com



(a)



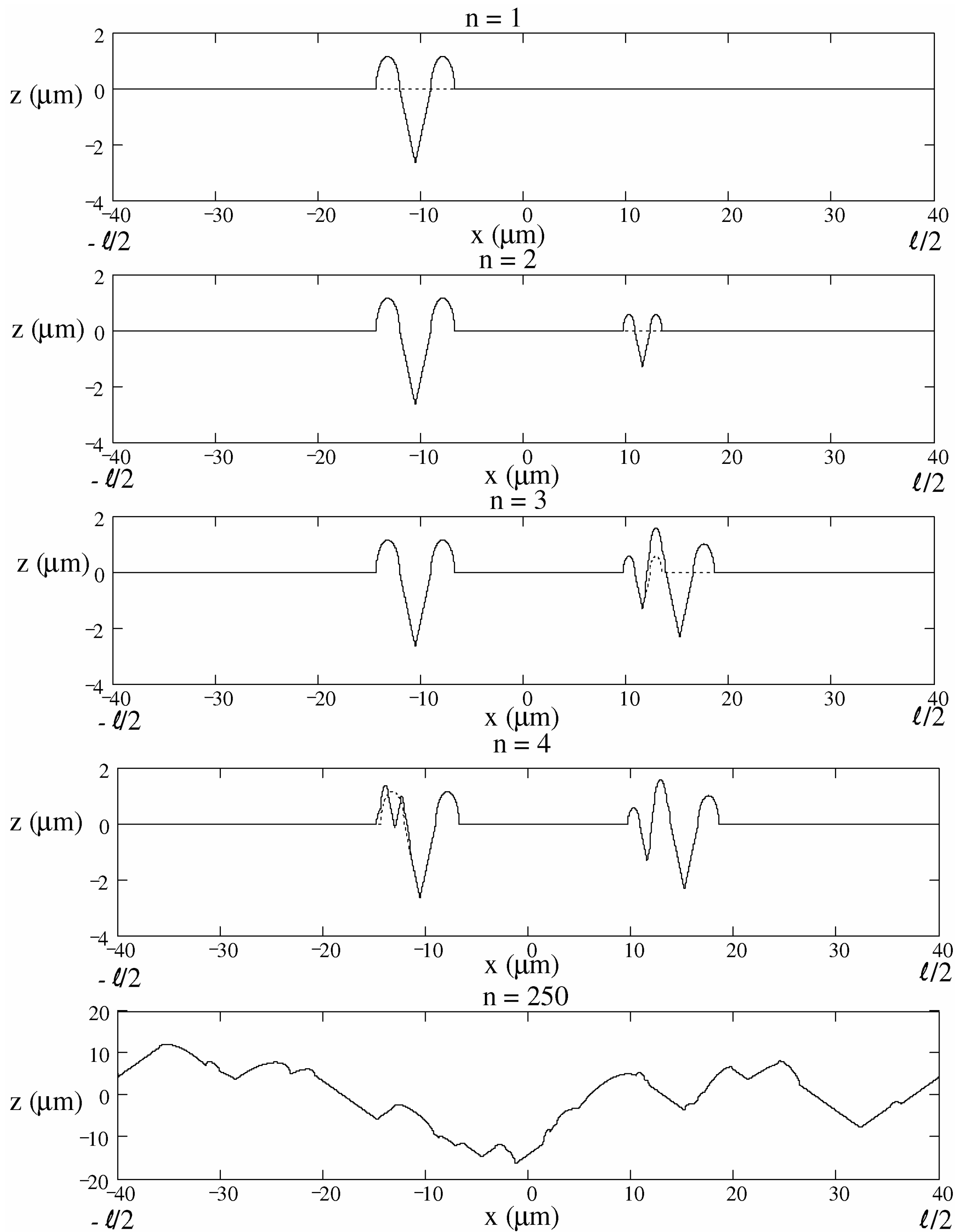
(b)

المنارة للاستشارات
www.manaraa.com

Figure 6.5 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process for n = 1, 2, 3, 4, 250 time steps with initially flat surface and variable diamond cutting depths. The dotted line represents the surface in the previous step. The penetration depth for each cut is the diamond protrusion length as determined from the exponential distribution subtracting the conditioner-pad separation d. The parameters are $\alpha = \pi/3$, $\ell = 80$ μm, $f_{add} = 1$ and $\sigma = 5$ μm (diamond height distribution).

المنارة للاستشارات

www.manaraa.com

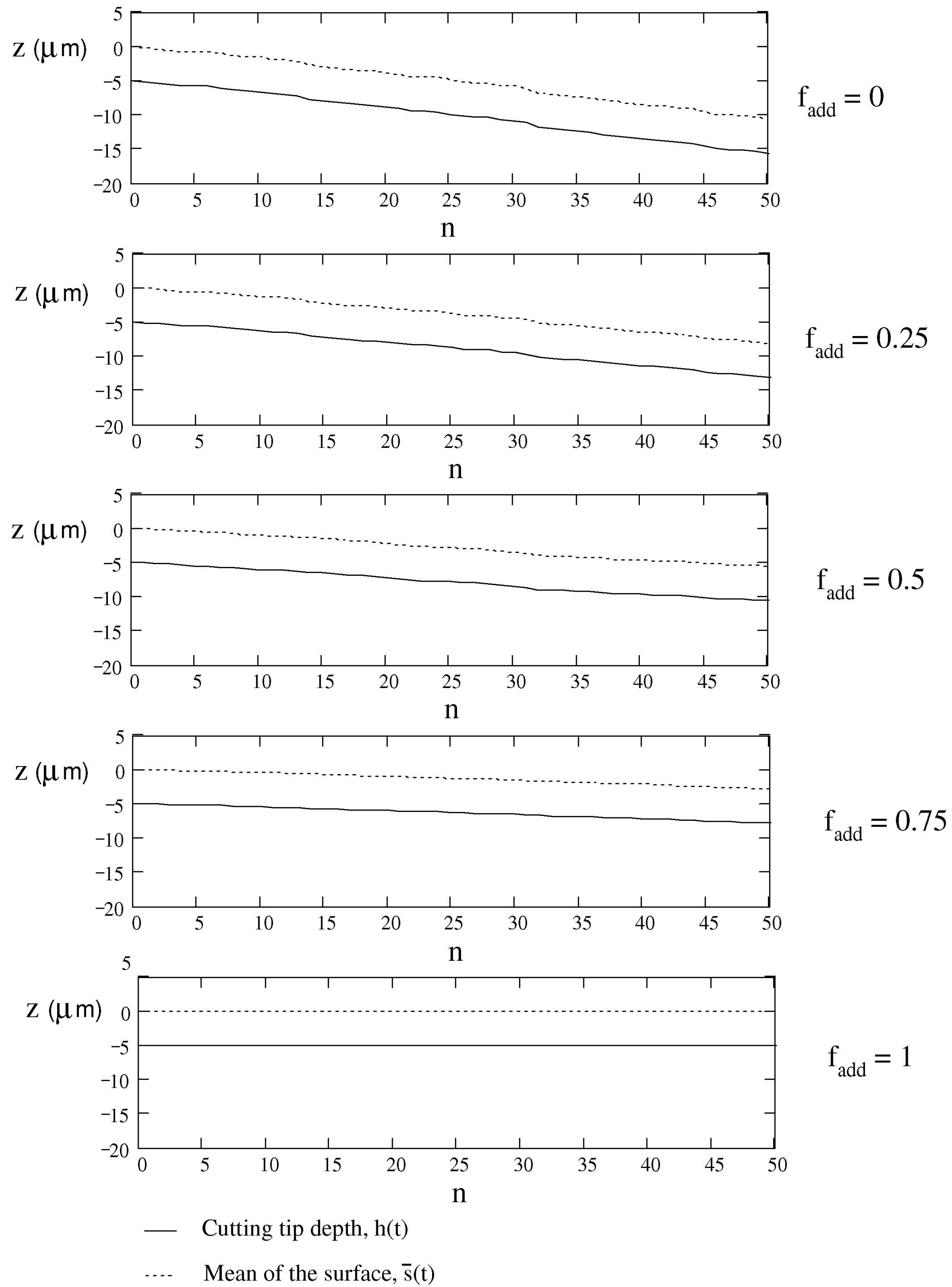


www.manaraa.com

Figure 6.6 Monte Carlo simulations for the changing of the average surface height (mean) $\bar{s}(t)$ and cutting depth h(t) for different chosen f_{add} values. The program calculates the new mean after each cut and moves down the cutting tip relative to the new mean to ensure that the cutting depth is always h_0 relative to the mean at each cut. The mean $\bar{s}(t)$ lags the cutting depth h(t) by h_0 for all the cases. The slope of $\bar{s}(t)$ or h(t) is the cut-rate. Lager f_{add} value gives smaller cut-rate. At $f_{add} = 1$, the average surface height remains at z = 0 (pure plastic deformation). At $f_{add} = 0$, the average surface height moves down fastest (pure brittle fracture) which is the case in [10]. In these simulations, $\ell = 80$ μm, $\alpha = \pi/3$, and $h_0 = -5$ μm [9].

المنارة للاستشارات

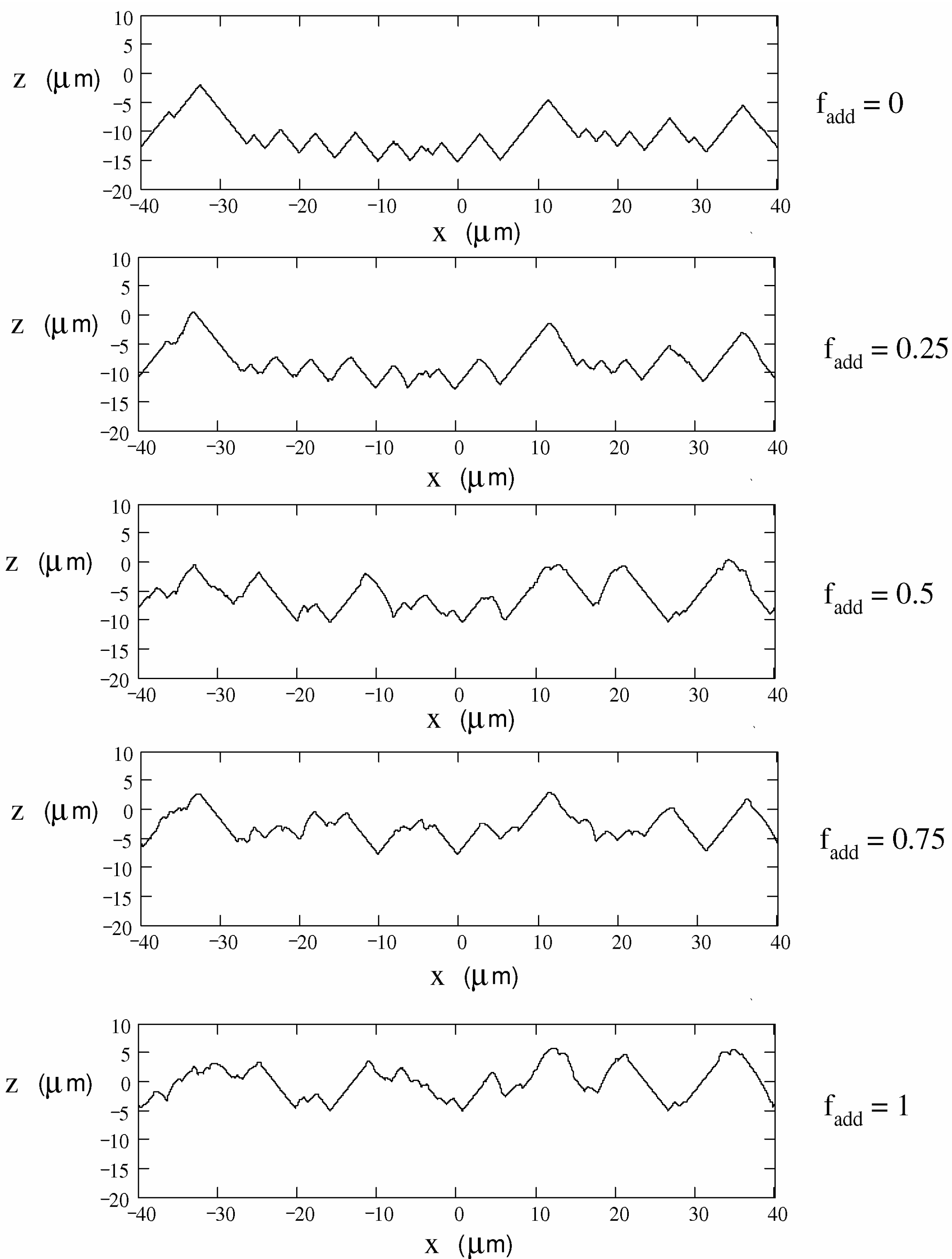
www.manaraa.com



المنارة للاستشارات
www.manaraa.com

Figure 6.7 The surface profile s(x,t) after 50 time steps for different f_{add} values. The profile moves down slower with increasing f_{add} values. At f_{add} = 1, the average surface height remains at z = 0 (pure plastic deformation). At f_{add} = 0, the average surface height moves down fastest (pure brittle fracture) which is the case in [10]. In these simulations, ℓ = 80 μm, $\alpha = \pi/3$, and h_0 = -5 μm [9].

المنارة للاستشارات
www.manaraa.com



المنارة للاستشارات
www.manaraa.com

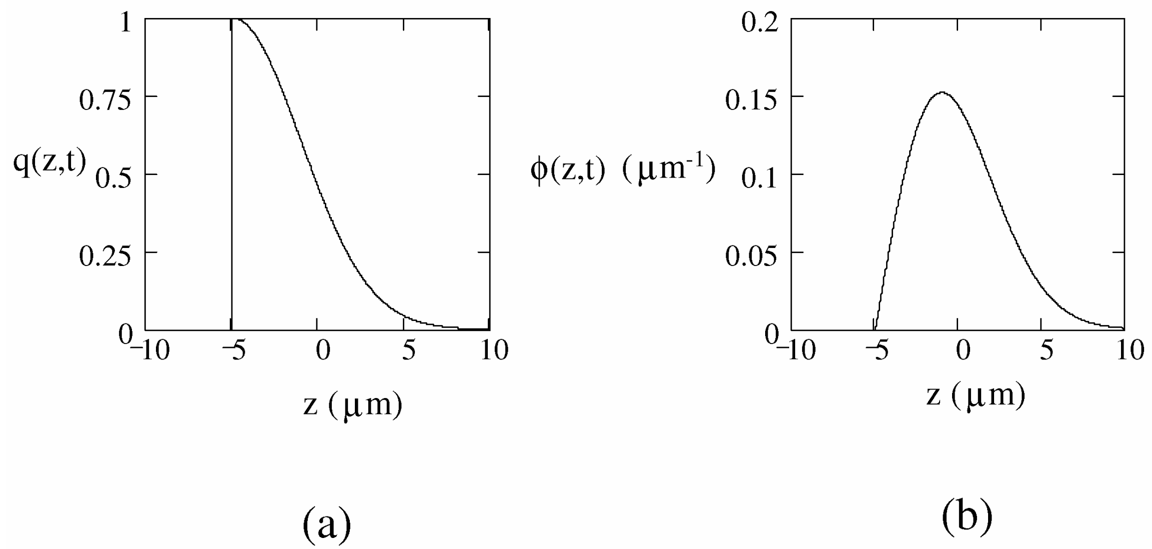


Figure 6.8 Analytical solutions for CCDF and PDF, Equation (6.12) and Equation (6.13) respectively, with single diamond cutting depth, h_0 = -5 μm [9], [10].

المنارة للاستشارات
www.manaraa.com

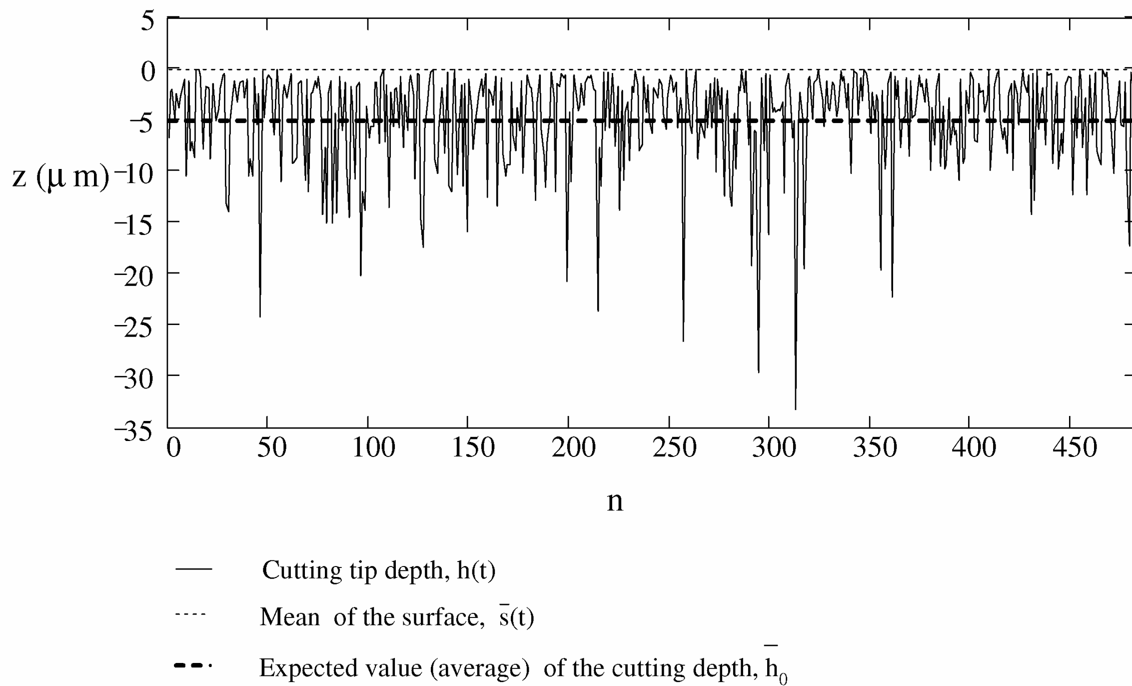


Figure 6.9 Monte Carlo simulations for the changing of the average surface height (mean), $\bar{s}(t)$ and cutting depth h(t) for the simulation in Figure 6.5. The mean stays at z = 0 because of pure plastic deformation ($f_{add} = 1$). The mean is behind the cutting depth h(t) by the expected value of h_0, which is 5.107 μm for exponential distribution of diamond heights with standard deviation σ = 5 μm.

المنارة للاستشارات
www.manaraa.com

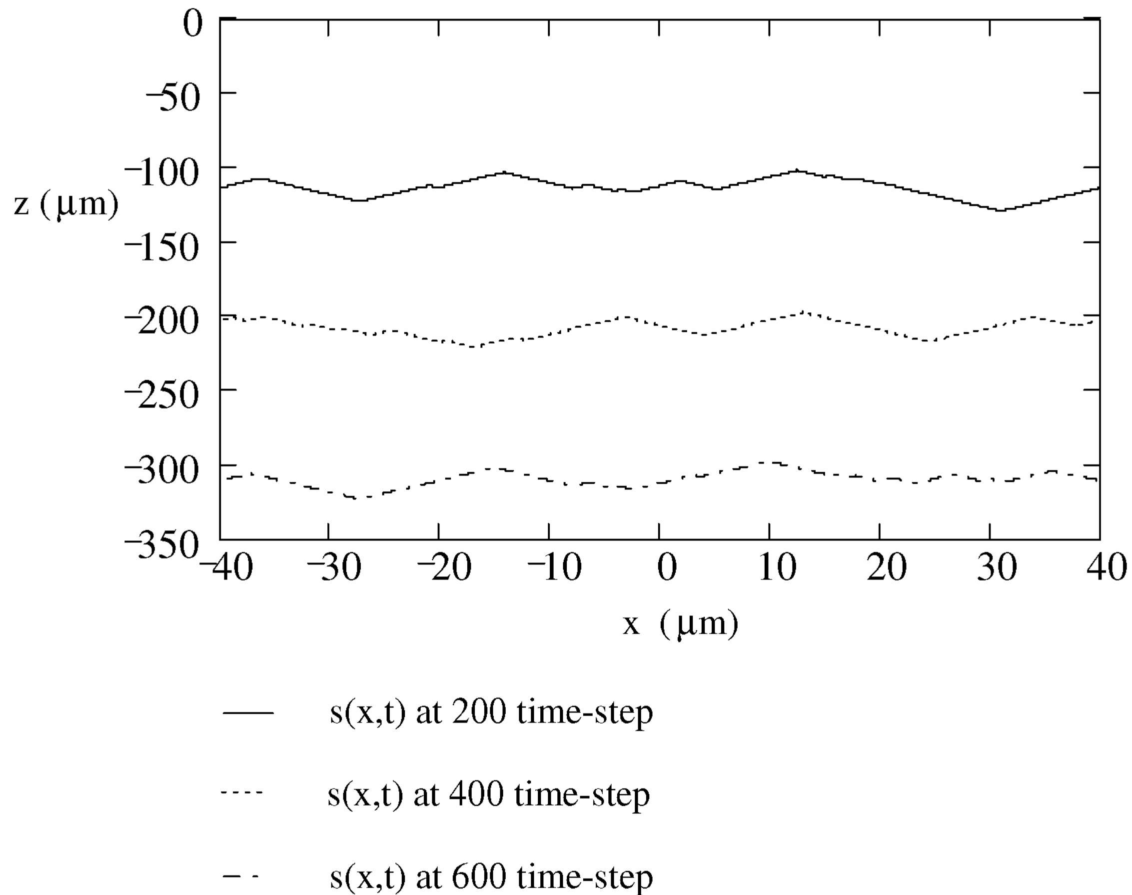


Figure 6.10 An example of the sequence of solid pad surface profiles after successive steps of discrete conditioning process for n = 200, 400, and 600 time steps (pad rotations) with initially flat surface and variable diamond cutting depth. The parameters are $\alpha = \pi/3$, $\ell = 80$ µm, $f_{add} = 0$ and $\sigma = 5$ µm (exponential diamond height distribution). The cut-rate is -0.49 µm per pad rotation. So after 600 time-steps (pad rotations), the mean of the surface has traveled to about -294 µm.

www.manaraa.com

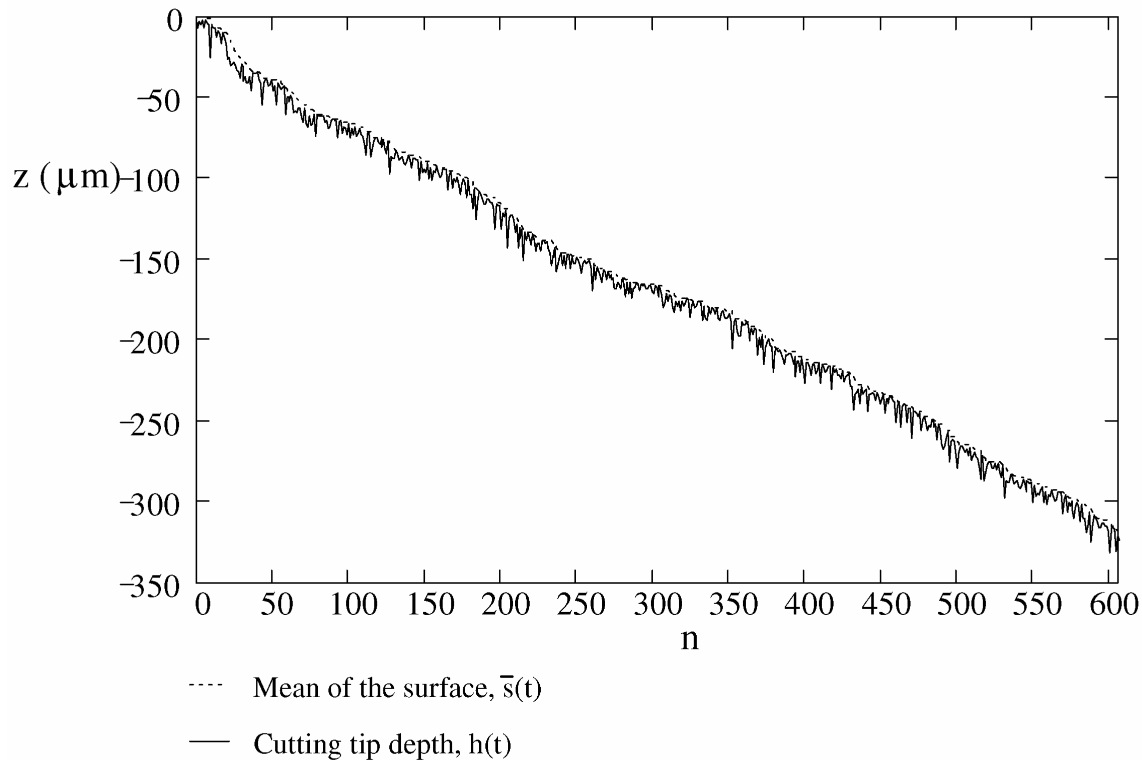


Figure 6.11 Monte Carlo simulations for the changing of the average surface height (mean), $\bar{s}(t)$ and cutting depth h(t) for the simulation in Figure 6.10. The dotted line is the mean of the surface traveling down the z-axis. The solid line is the cutting depth h(t) also traveling down the z-axis. The program calculates the new mean after each cut and moves down the cutting tip relative to the new mean. The cutting depth relative to the new mean at each cut is h_0 which is the diamond height taken randomly from the exponential distribution PDF subtracting the conditioner-pad separation d. Linear fittings of the cutting depth h(t) and mean $\bar{s}(t)$ show that the mean lags (offsets) the cutting depth h(t) by about 5.3 μm for an exponential distribution with standard deviation σ = 5 μm. The slope (cut-rate) of the mean and h(t) is -0.49 μm.

المنارة للاستشارات
www.manaraa.com

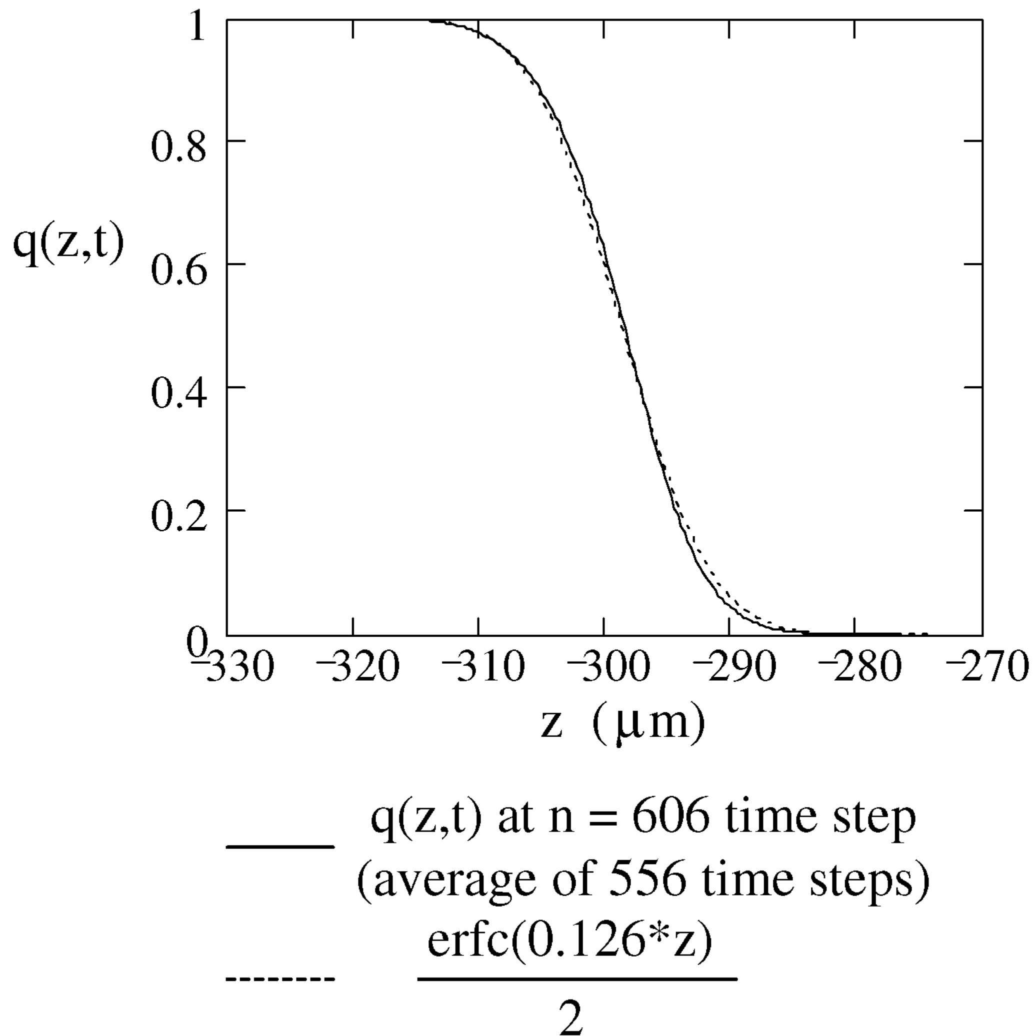


Figure 6.12 Steady-state CCDF with an exponential diamond cutting depth distribution for the simulation in Figure 6.4. The standard deviation of exponential distribution is 5 μm. The solid line shows CCDF at n = 606 time step with the average of 556 time steps (discard the first 50 time steps). The mean of the steady-state CCDF has traveled to about -294 μm after 606 time steps. The dashed line shows the fitting of the CCDF very well by complementary error function (erfc). The erfc function has a standard deviation of 5.612 μm.

المنارة للاستشارات
www.manaraa.com

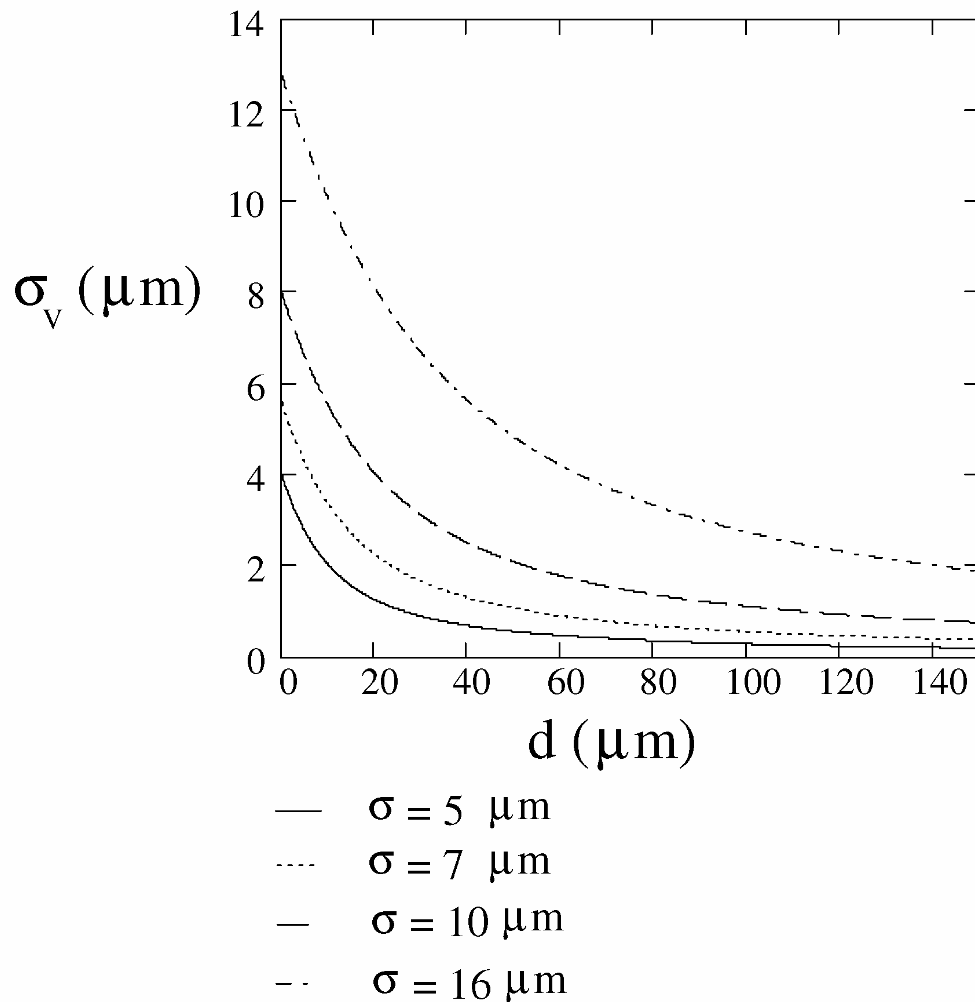


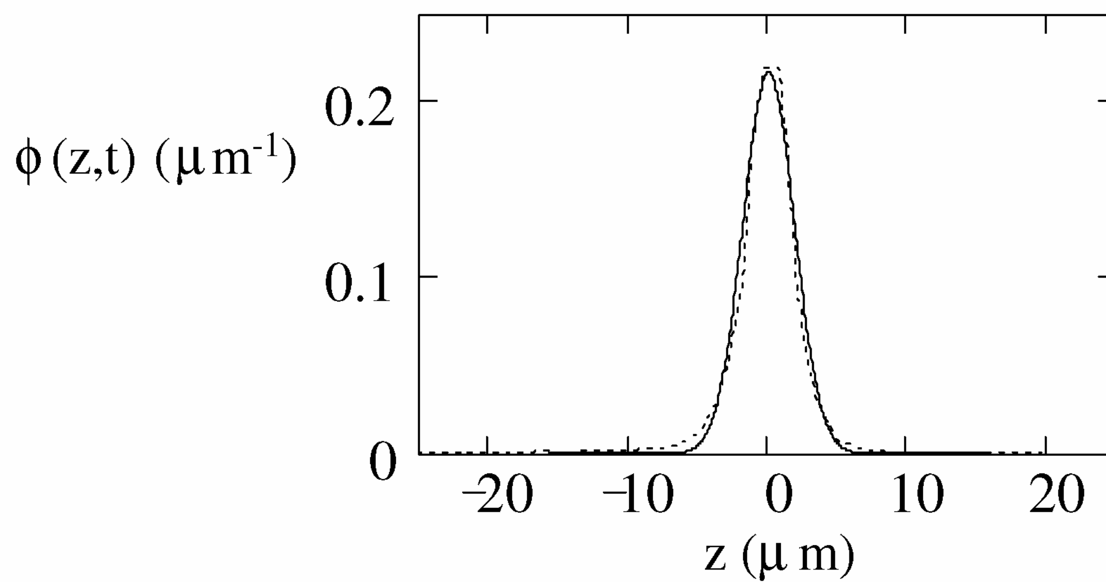
Figure 6.13 The changing of the standard deviation of the conditioned pad asperity PDF with d (separation distance between conditioner diamond surface and pad) and σ (standard deviation of the diamond height PDF with a Gaussian distribution) by using Equation (6.28) and (6.29).

المنارة للاستشارات
www.manaraa.com

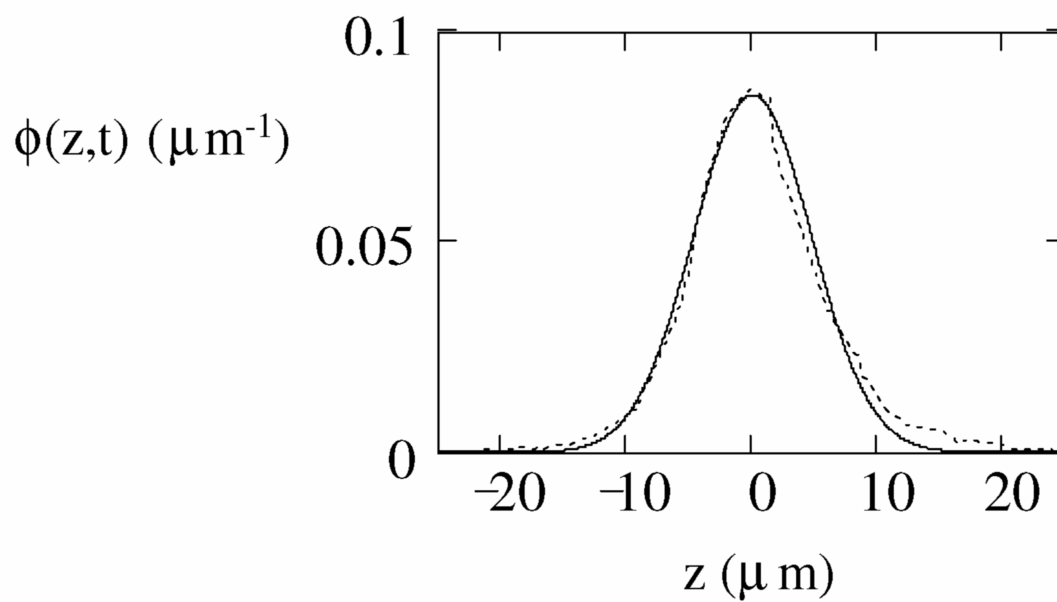
Figure 6.14 The measured conditioned pad asperity PDF at two different loads for 100 grit conditioner and the best fits of the data by Gaussian functions with different standard deviation. The dotted lines are the experimental measurements from [12]. The solid lines are the fittings of the experimental data by Gaussian function. (a) The measured PDF at 3.6 lb load. The standard deviation of the Gaussian function is 1.84 μm. (b) The measured PDF at 8.0 lb load. The standard deviation of the Gaussian function is 4.7 μm.

المنارة للاستشارات

www.manaraa.com



(a)



(b)

المنارة للاستشارات
www.manaraa.com

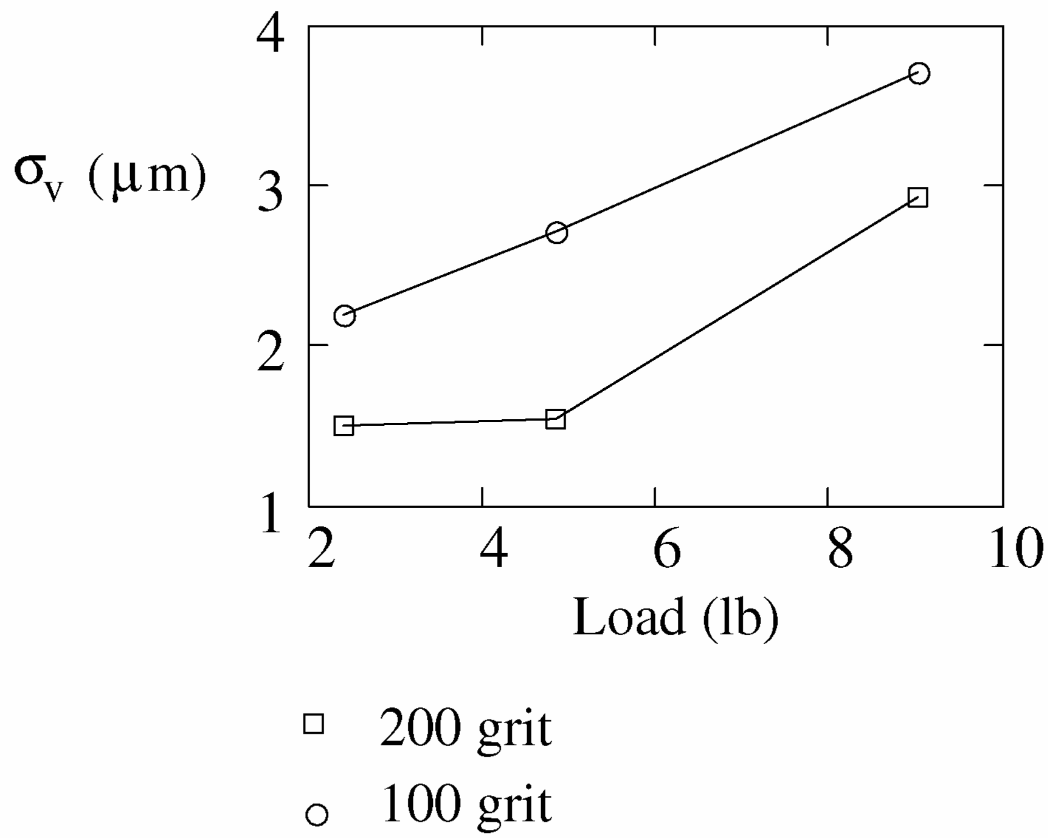


Figure 6.15 The measured standard deviations of the conditioned pad asperity PDF as a function of the load on the conditioner [13]. Conditioners with 100 grit and 200 grit sizes are used.

المنارة للاستشارات
www.manaraa.com

CHAPTER 7

CONCLUSION AND FUTURE WORK

المنارة للاستشارات

www.manaraa.com

Conclusion

In this study, a population balance model (PBE) for pad asperities is used to model the pad wear and pad conditioning processes. Monte Carlo numerical simulations for pad conditioning are also performed and found to match the analytical solutions to the population balance model.

Pad wear

An analytical solution to the population balance for the heights of pad asperities worn down during chemical-mechanical polishing of silicon wafers is employed to calculate the chemical-mechanical removal rate decay. This analytical solution is compared with experimental data from Stein and the model results agree with the experimental removal rate decay results. The polishing-rate decay predicted by the analytical solution to asperity population balance is similar to the numerical prediction done by Borucki in previous work. The analytical solution matches well with the experimental data by Stein and the RMS value of the fitting is better than the numerical method. The advantage of the analytical solution is that it is more accurate and numerically faster to compute.

In the next work, fluid mechanics is incorporated into the pad wear model by assuming positive pressure developed in the pad wafer interface. Modeling results with and without fluid effects are compared. Polish-rate model results are compared with experimental data by Stein, et al. Results agree with experimental results for both cases by using different wear rate coefficients to fit the experimental data. A smaller wear rate coefficient must be used to fit Stein's data for fluid case compared to the case without fluid. The wear rate of the pad is calculated from the rate of change of the pad-wafer

المنارة للاستشارات
www.manaraa.com

separation distance during polishing because only asperities above this distance will be in contact and worn down by the wafer and that portion of pad asperities will pile up at the pad-wafer separation distance on the PDF curve of the pad asperities. The results with fluid show much less pad wear because the decrease of pad-wafer separation distance is only 36% of that without fluid after 45 minutes of cumulative wear time at Stein's experimental conditions. During this time, wear has driven the CMP operating regime from boundary-lubrication to hydrodynamic lubrication by shifting more fraction of the load to P_f. This is reflected in the pad-wafer separation distance and contact area fraction, both of which have shown increasing difference comparing the cases without fluid and with fluid during 45 minutes of wear time at Stein's experimental condition. Another important result from our model is that if fluid is neglected, pad-wafer separation distance will continue to drop until it is zero. While for the fluid case, it will reach a dynamic steady-state with no pad wear eventually, indicating that entirely hydrodynamic regime has reached where only chemical removal mechanism is left which is the observed physical situation. The closer the process condition is to hydrodynamic regime, the faster the transition to dynamic steady-state d occurs with wear. This is especially important for copper polishing where small P_{app} or faster velocity is generally used and wear quickly pushes the operating regime from mixed-lubrication to hydrodynamic lubrication in which only chemical removal mechanism (mass transfer) is left.

Polish-rate is a sensitive function of pad asperity PDF. Accurately estimating the interface gap (pad wear) is critical for controlling both mechanical and chemical removal rate. As we have shown in this work, the population balance equation needs to be solved with fluid mechanics to accurately predict the interface gap and pad wear. For stable

المنارة للاستشارات
www.manaraa.com

removal rate (both chemical and mechanical), constant interface gap resulting from constant pad surface needs to be maintained. Accurately estimating pad wear is also important to decide when the pad needs to be changed. Pad wear and pad conditioning work together to maintain a constant pad surface. Models without fluid effects overestimate the pad wear. It will result in over-conditioning of the pad, and making the task of maintaining a constant pad surface difficult.

Pad conditioning

In this pad conditioning work, a population balance model (PBE) for pad asperities is used to model the pad conditioning process. Pad conditioning for a single diamond height with plastic deformation is developed first and then that with variable diamond height is developed. Monte Carlo numerical simulations for both cases match analytical solutions to the PBE for pad asperities. Plastic deformation and variable diamond height are not done in the previous works.

Previous conditioner model developed by Borucki, et al. assumes constant diamond height and brittle fracture in which all of the material being cut is removed from the pad surface. That will overestimate the pad cut-rate. In the Borucki's PBE conditioner model, only a "death" term is used. Pad cut-rate is important to estimate the aggressiveness of the pad conditioning, decide the time needed to bring the pad surface into steady-state in the pad "break-in" process and the time to change the pad. Plastic deformation is important in modeling pad conditioning because the raised ridges along the sides of the grooves are the high points (asperities) that make contacts with the wafer. These asperities act as pressure enhancers for the abrasive particles used to polish the wafer. In this work, plastic deformation is modeled by adding a "birth" term by mass balance to the

المنارة للاستشارات
www.manaraa.com

population balance model to take account of the material pileups (plastic deformation) during pad conditioning. This material pileup is modeled by a f_{add} parameter to represent the fraction of the material removed that is added back to the pad surface at the edge of the groove cut by the diamond. The derived steady-state pad surface conditioning profile only depends on the diamond cutting depth into the pad, but the cut-rate derived from the model depends on the diamond cutting depth, the number of the cutting diamonds and f_{add}. The dependence of the cut-rate on the number of the cutting diamonds is through ℓ. Larger f_{add} will give a smaller cut-rate. The model can be used to model a series of pad materials with plastic properties because f_{add} is a variable parameter. The analytical solutions to the PBE derived for the pad surface conditioning profile and cut-rate agree with the Monte Carlo simulations.

Current conditioner disks are made by attaching numerous similar size diamond grits to a metal substrate. The tips of the diamonds are not leveled to the same height resulting in a Gaussian diamond height distribution. This height distribution can induce variance into the conditioned pad asperity PDF. Hence, in this work, the model with single diamond height is extended to incorporate variable diamond height. Monte Carlo simulations with variable diamond height and plastic deformation are performed. Analytical solutions are derived and match Monte Carlo numerical results. The model results show that the conditioned pad asperity PDF is Gaussian in form and is obtained by integrating the average of the conditioned pad asperity PDF for single diamond height. The diamond height PDF only changes the standard deviation of the conditioned pad asperity PDF and it does not change its functional form which is Gaussian. The conditioned pad asperity PDF being Gaussian is consistent with the experimental results

المنارة للاستشارات
www.manaraa.com

in the literature. Our model results show that while the variance of the conditioned pad asperity PDF is determined solely by the variance of the heights of the cutting diamonds (active diamonds), the cut-rate depends on f_{add}, the variance of the heights of the active diamonds and the number of the active diamonds (or the contact area between the pad and conditioner). The contact form of the pad asperities is important to determine the contact area and interface gap between the rough surface of the pad and the wafer. In this study, we did not perform the simultaneous conditioning and wear together to show if the aggressiveness of the conditioning can overcome the wear of the pad asperities due to polishing. If the aggressiveness of the conditioning can overcome the wear of the pad asperities due to polishing, then the steady-state conditioning profile obtained from the model results for steady-state conditioning in this work is the contact form of the pad asperities with the wafer at steady-state polishing. The conditioning model with variable diamond height has changed the pad asperity PDF compared with the conditioning model with single diamond height done previously. This model is more realistic and it is confirmed by experimental data in the literature.

When the conditioner diamond height PDF is Gaussian as reported in the literature, our model results with variable diamond height show that a different interface gap between the conditioner and pad will give a different standard deviation of the resulting conditioned pad asperity PDF. This is consistent with the experimental findings that with increasing load, the roughness of the conditioned pad asperity PDF increases. Model results in this work show that it is caused by the decreasing of the interface gap between the rough surface of the conditioner diamonds and the pad surface which results in the increase of the roughness of the active diamonds population. When the roughness of the

المنارة للاستشارات
www.manaraa.com

active diamonds population is independent of the conditioner-pad separation (as in the case for exponential diamond height distribution), the roughness of the conditioned pad asperity PDF does not depend on the conditioner-pad separation. For Gaussian diamond height distribution, the roughness of the active diamonds increases with decreasing conditioner-pad separation. The roughness of the conditioned pad asperity PDF is determined only by the roughness of the active diamonds, and it increases as the roughness of the active diamond increases. The load on the conditioner, the rotation speed between the conditioner and the pad and the viscosity of the polishing slurry decide the interface gap between the conditioner and the pad. Our model results show that the diamond height PDF and how its roughness changes with the conditioner-pad separation (interface gap) are very important in pad conditioning for CMP. The conditioner-pad interface gap is very important in determining the roughness of the conditioned pad asperity PDF. The roughness of the conditioned pad asperity PDF increases as the roughness of the diamond height PDF increases, however if the gap is too large, it may increase a little bit or may even decrease. When the diamond grit size increases (e.g., the roughness of the diamond height PDF increases), the gap will also increase under the same process conditions because of the size of the diamonds. This is also observed in the experimental results in which the roughness of the conditioned pad asperity increases only a small amount when the diamond grit size increases a large amount. Our model has successfully explained these experimental findings and the trend shown by the model results is consistent with the trend observed from the experiments.

المنارة للاستشارات
www.manaraa.com

Future work

In the pad wear model, for the micro-contact mechanics, we use Greenwood and Williamson for elastic contacts. However, asperities could deform plastically when the applied load exceeds about half of the yield point load [1]. Only plastically deformed asperities will be worn down. Elastically deformed asperities will not be worn down. We assume all asperities are deformed elastically and worn down. Future work can refine model in this respect. The wear rate of the pad is expected to be reduced because the number of plastically deformed asperities is only a fraction of the total asperities.

Future study can include the study of the slurry hydrodynamics of the grooved pad. There are many groove designs made into CMP pads. Grooves are expected to supply slurry evenly, and to remove wasted particles. These grooves have impact on slurry hydrodynamics. So future study can include groove effect to see how it influences the wear and conditioning of the pad.

In CMP pad conditioning process, each active diamond is cutting while rotating with the conditioner and this rotation may cause one side of the diamond tip cuts more material and plows more material to the top of the groove than the other side depending on the angle of rotation. In this thesis, this effect is not included. Future study can include this effect to see how it influences the conditioned pad asperity PDF. It should increase the variance of the conditioned pad asperity PDF.

Future study on pad conditioning can also compare different diamond grit arrangement on the conditioner, e.g., random, clustered, and grid, to see what's the difference in cut-rate and conditioned pad asperity PDF for these different arrangement.

CMP is a dynamic process. Pad wear flattens the pad surface and pad conditioning

المنارة للاستشارات
www.manaraa.com

roughens the pad surface. Pad wear and pad conditioning work together to maintain a constant pad surface. In this study, we did not model pad wear and pad conditioning simultaneously to see the transient process, i.e., the pad break-in process. During polishing, the aggressiveness of the conditioning should overcome the pad wear effect in order to sustain a stable polish-rate. The pad cut-rate derived using steady-state conditioning in this work is also valid for this transient process. This cut-rate can be used to study how quickly the pad can reach steady-state during the pad break-in process. This may be done in the future work.

References

[1] H. J. Tsai, Y. R. Jeng, and P. Y. Huang, "An improved model considering elastic-plastic contact and partial hydrodynamic lubrication for chemical-mechanical polishing," *Proc. IMechE, Part J: J. Eng. Tribol.*, vol. 222, pp. 761-770, 2008.

المنارة للاستشارات
www.manaraa.com

APPENDIX

SENSITIVITY TEST FOR FITTING PARAMETERS, C_a AND C_w

المنارة للاستشارات
www.manaraa.com

In order to test the sensitivity of the MRR to the fitting parameters. We performed the sensitivity test by using different c_a and c_w parameters. The results are shown in Figure A.1 to Figure A.4. The best fits for Stein's data are c_a = 2.4e(-16) m/s/Pa and c_w = 1.85e(-16) m/s/Pa for no fluid case, and c_a = 2.1e(-16) m/s/Pa and c_w = 1.88e(-16) m/s/Pa for fluid case.

For comparison, c_a = 2.1e(-16) m/s/Pa and c_w = 1.85e(-16) m/s/Pa are used for no fluid case and the results are shown in Figure A.1. The RMS of the fitting to the experimental data by using these parameters is 29.1 nm compared to 19.1 nm for the best fit. For comparison, c_a = 2.4e(-16) m/s/Pa and c_w = 1.88e(-16) m/sPa are used for fluid case and the results are shown in Figure A.2. The RMS of the fitting to the experimental data by using these parameters is 32.3 nm compared to 21.1 nm for the best fit. Figure A.1 and A.2 show that the MRR is very sensitive to the fitting parameter c_a.

For comparison, c_a = 2.4e(-16) m/s/Pa and c_w = 1.95e(-16) m/s/Pa are used for no fluid case and the results are shown in Figure A.3. The RMS of the fitting to the experimental data by using these parameters is 43.0 nm compared to 19.1 nm for the best fit. For comparison, c_a = 2.1e(-16) m/s/Pa and c_w = 1.95e(-16) m/s/Pa are used for fluid case and the results are shown in Figure A.4. The RMS of the fitting to the experimental data by using these parameters is 35.4 nm compared to 21.1 nm for the best fit. Figure A.3 and A.4 show that the MRR is very sensitive to the fitting parameters c_w.

المنارة للاستشارات
www.manaraa.com

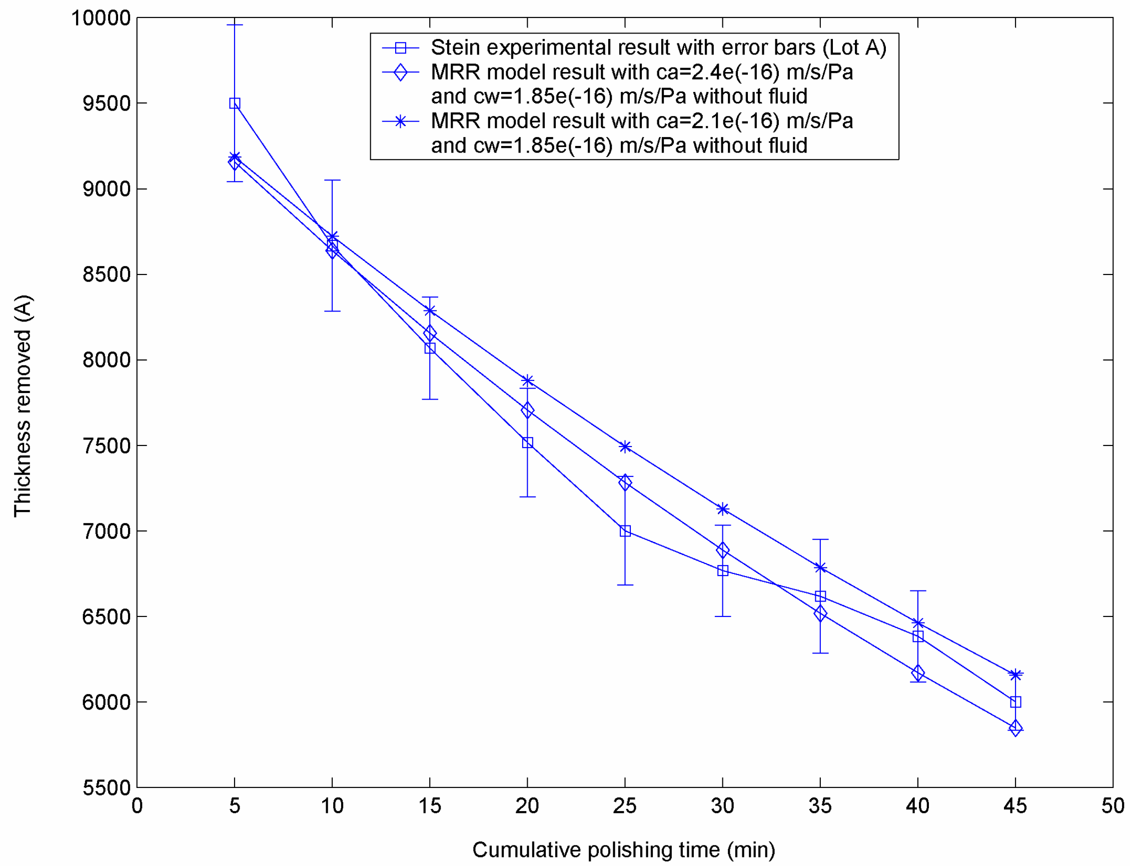


Figure A.1 Removed thickness results calculated with different c_a values for the case without fluid.

المنارة للاستشارات

www.manaraa.com

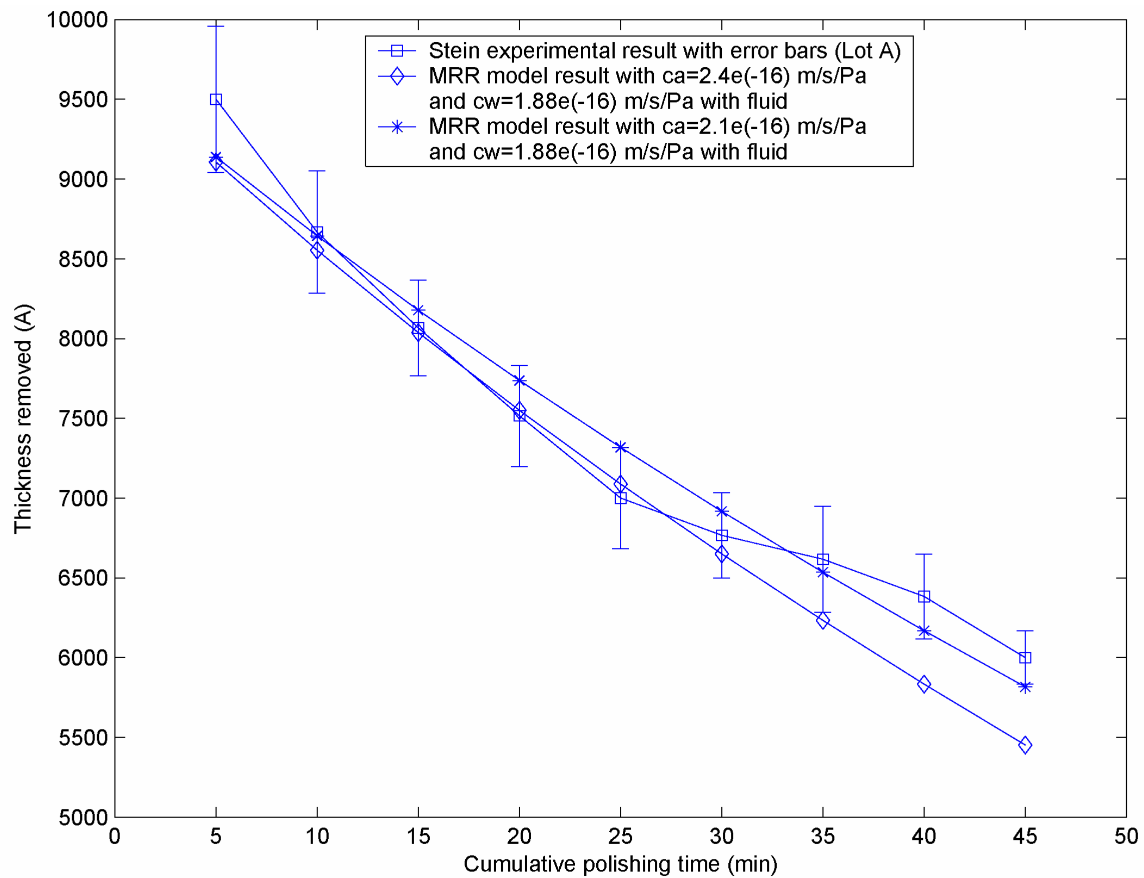


Figure A.2 Removed thickness results calculated with different c_a values for the case with fluid.

المنارة للاستشارات

www.manaraa.com

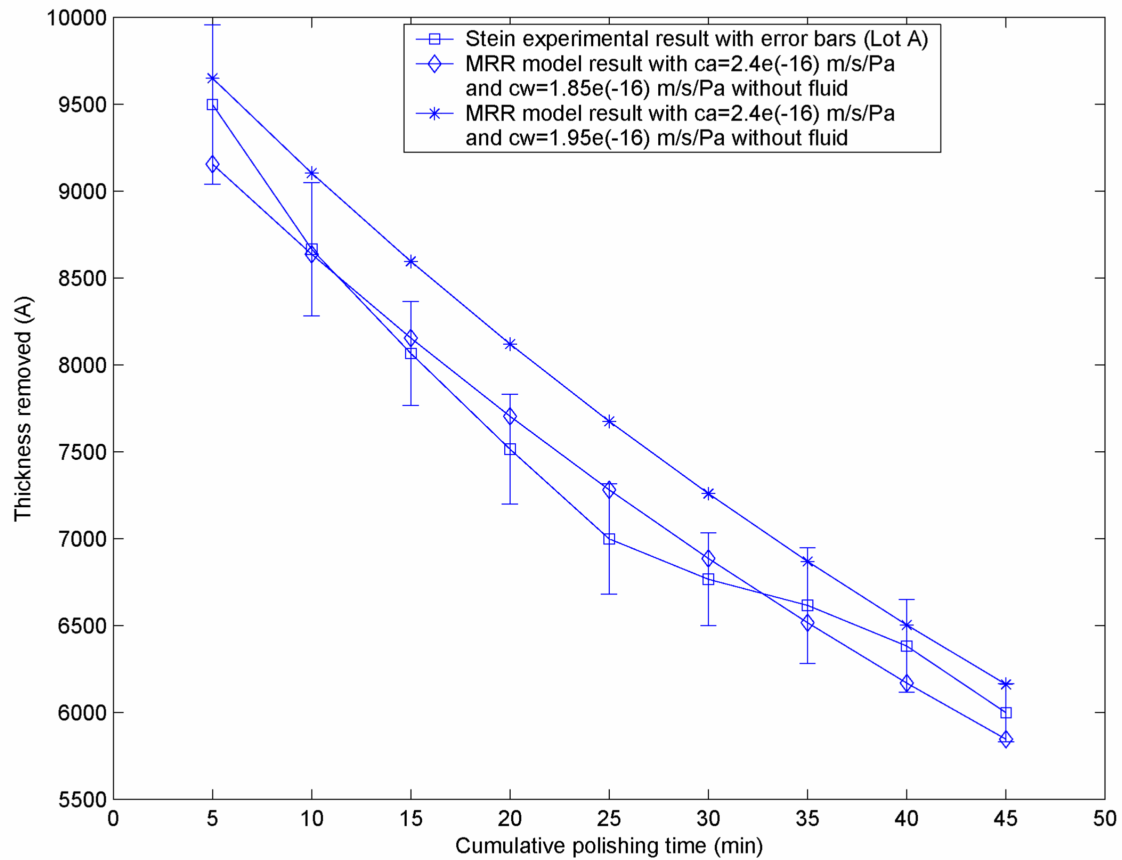


Figure A.3 Removed thickness results calculated with different c_w values for the case without fluid.

المنارة للاستشارات
www.manaraa.com

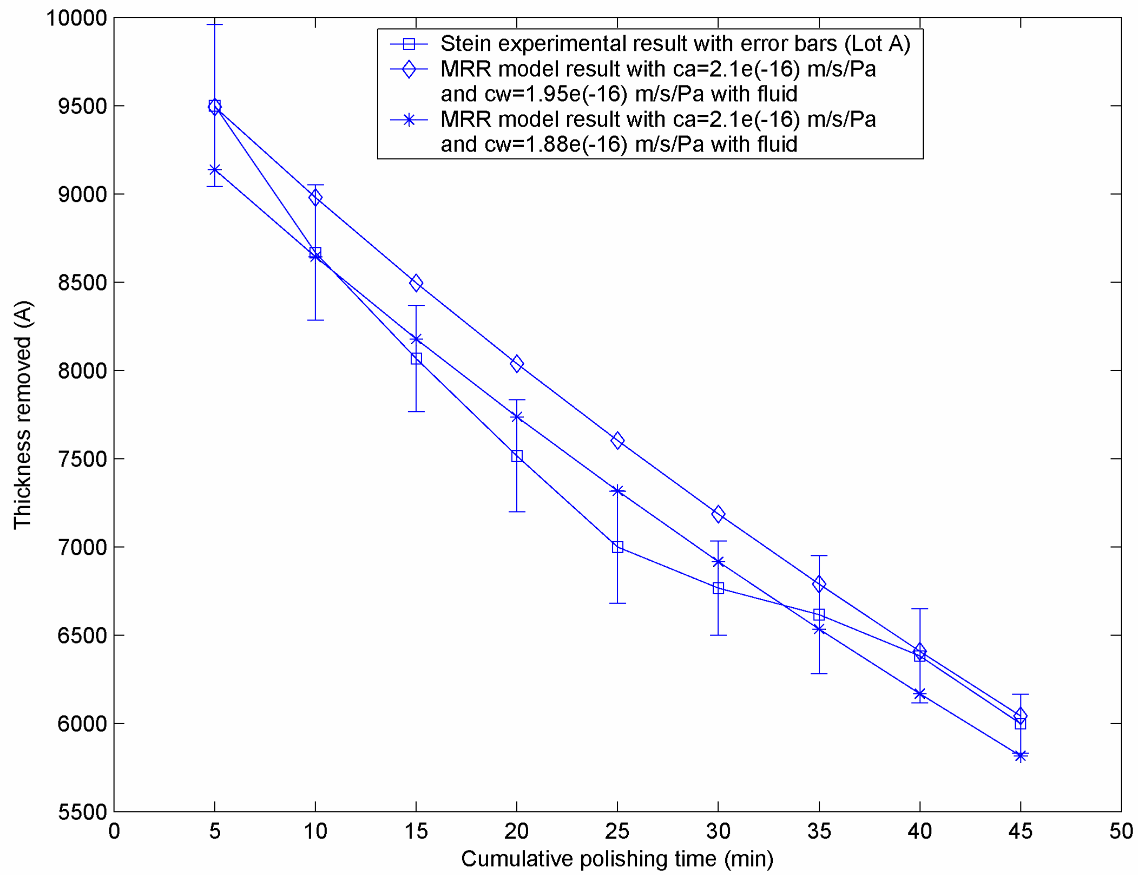


Figure A.4 Removed thickness results calculated with different c_w values for the case with fluid.

المنارة للاستشارات
www.manaraa.com