Iowa State University
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2005

Yield improvement of chemical mechanical planarization processes

Sutee Eamkajornsiri
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Mechanical Engineering Commons, and the Operations Research, Systems Engineering and Industrial Engineering Commons

Recommended Citation

Eamkajornsiri, Sutee, "Yield improvement of chemical mechanical planarization processes " (2005). *Retrospective Theses and Dissertations*. 3043.
https://lib.dr.iastate.edu/rtd/3043

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

المنارة للاستشارات

www.manaraa.com

Yield improvement of chemical mechanical planarization processes

by

Sutee Eamkajornsiri

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Engineering

Program of Study Committee:
Frank Peters, Co-major Professor
Abhijit Chandra, Co-major Professor
John Jackman
Sigurdur Olafsson
Wallapak Tavanapong

Iowa State University

Ames, Iowa

2005

Copyright © Sutee Eamkajornsiri, 2005. All rights reserved.

UMI Number: 3229146

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

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

UMI®

UMI Microform 3229146

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Sutee Eamkajornsiri

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES x

ABSTRACT xi

CHAPTER 1. INTRODUCTION TO CHEMICAL MECHANICAL PLANARIZATION (CMP) 1

- 1.1. Introduction 1
- 1.2. Chemical Mechanical Planarization Components and Functions 5
  - 1.2.1. Slurry 5
  - 1.2.2. Polishing Pad 6
  - 1.2.3. Wafer Carrier 6
- 1.3. Dissertation Outline 7

CHAPTER 2. BACKGROUND OF CMP PROCESS 9

CHAPTER 3. WAFER-SCALE MODEL SIMULATION 12

- 3.1. Wafer-scale Model 12
- 3.2. Control Formulations 15
  - 3.2.1. Alternative Objective Functions 15
    - a. Number of good sectors within tolerance 15
    - b. Deviation form target surface 16
  - 3.2.2. Alternative Decision Variables 17
    - a. Load as control parameter 17
    - b. Curvature as control parameter 17
- 3.3. Greedy Algorithm for Yield Improvement 18
  - 3.3.1. Objective Function 18
  - 3.3.2. Model with Greedy Algorithm 19
    - a. Switching Logic 19
    - b. Load Control 20
    - c. Curvature Control 21
    - d. Combined Curvature & Load Control 21
  - 3.3.3. Simulation Results – Greedy Algorithm 22
- 3.4. Method Heuristic for Yield Improvement 24
  - 3.4.1. Genetic Algorithms 24
  - 3.4.2. Model with Genetic Algorithm 25
    - a. Coding Scheme 25
    - b. Fitness Value 26

c. Genetic Algorithm Steps for Curvature Control 26
3.4.3. Simulation Results – Genetic Algorithm 32
3.5. Non-linear Programming for Yield Improvement 39
3.5.1. Model with Non-linear Programming 39
a. Objective Function 39
b. Non-linear Programming Model 41
3.5.2. Simulation Results – Non-linear Programming Model 42

CHAPTER 4. FEATURE / DIE-SCALE MODEL SIMULATION 50

4.1. Feature / Die-Scale Model Development 50
4.2. Model Validation 56
4.3. Global Planarization Control Algorithm 58
4.3.1. Objective Function 59
4.3.2. Model of Global Planarization Model 59
a. Spatial Pressure Control: Algorithm 59
b. Spatial and Temporal Pressure Control: Algorithm 62
c. Look-ahead Scheduled Pressure Control: Algorithm 64
4.4 Simulation Results 67

CHAPTER 5. MODEL FOR GENERAL WAFER SURFACE 82

5.1. Simplified Representation of General Wafer Surface Profile 83
5.2. Two-Wave Feature / Die-Scale Model Development 83
5.3. Two-Wave Model Verification and Simulation Results 92
5.3.1 Comparison two-wave model to Fu and Chandra's model 92
5.3.2 Simulation results of two-wave model 95

CHAPTER 6. SUMMARY AND CONCLUSION 105

6.1. Discussion for Wafer-Scale Model 106
6.1.1. Comparison of Control Strategies with the Greedy Algorithm 106
6.1.2. Comparison between the Stochastic, the Greedy Algorithm and the Genetic Algorithm of Curvature Control Strategy 107
6.1.3. Comparison of the Greedy Algorithm and the Non-linear Programming Model with Load Control Strategy 108
6.2 Discuss for Feature / Die-Scale Model 109
6.2.1 Comparison of Control Algorithms using Interface Pressure as Control Parameter 109
6.2.2 Model for General Wafer Surface 110
6.3. Conclusion 110
6.4 Future Work 111

APPENDIX A. LEAST SQUARE CURVE FITTING OF DISCRETE POINT 113

APPENDIX B. THE FILM CURVATURE OF THE WAFER AT ANY TIME 114

APPENDIX C. EXAMPLE PROGRAM OF NON-LINEAR PROGRAMMING USING LINGO SOFTWARE 116

APPENDIX D. GRAPH BETWEEN MODEL PREDICTION AND OUMA'S EXPERIMENTAL DATA 117

APPENDIX E. LOOK-AHEAD PROCEDURE 121

APPENDIX F. FEATURE / DIE-SCALE CONTROL FLOWCHART 123

F.1. Spatial Pressure Control: Flowchart 123
F.2. Spatial and Temporal Pressure Control: Flowchart 124
F.3. Look-ahead Scheduled Pressure Control: Flowchart 125

APPENDIX G. SIMULATION RESULTS BASED ON OUMA'S EXPERIMENTAL DATA 126

G.1. Simulation Result of Line 4 126
G.2. Simulation Result of Line 5 131
G.3. Simulation Result of Line 6 136

APPENDIX H. FEATURE / DIE-SCALE MODEL FOR GENERAL WAFER SURFACE 141

REFERENCES CITED 146

ACKNOWLEDGEMENTS 149

LIST OF FIGURES

Figure 1.1. Illustration of the CMP process 2

Figure 1.2. Planarization defects due to pattern density variations 4

Figure 1.3. Wafer partitioned into dies 4

Figure 1.4. Lists of input and output parameters in CMP process 8

Figure 2.1. FEM model for investigating wafer chamfering (a), and for investigating retainer ring (b) 10

Figure 3.1. Model of pad and wafer contact, a_0 is a rigid body displacement of the wafer due to the down load and $2a$ is the wafer diameter 12

Figure 3.2. Kinematics of the CMP process 14

Figure 3.3. Moment calculation using the upper moment and the lower moment 17

Figure 3.4. Load switching using the moment of deviation objective function 20

Figure 3.5. The initial of population sets with the series of fixtured curvature 27

Figure 3.6. The standard GA diagram 28

Figure 3.7. The fitness values of four series of fixtured curvature on roulette wheel 28

Figure 3.8. The single crossover, mutation, and prime module function 29

Figure 3.9. Flow chart of the genetic algorithm program 31

Figure 3.10. Normalized pressure profile curves shown for seven values of fixtured curvature a_2^1 (m^{-1}) with the indentation depth of 3.678e-04 (m) 32

Figure 3.11. Fixtured curvature variation vs. polishing time in stochastic with three initial film curvatures 35

Figure 3.12. Fixtured curvature variation vs. polishing time in greedy algorithm with three initial film curvatures 36

Figure 3.13. Fixtured curvature variation vs. polishing time in genetic algorithms with three initial film curvatures 37

Figure 3.14. Modified fixtured curvature variation vs. polishing time in genetic algorithm with three initial film curvatures 38

Figure 3.15. Number of good sectors with polishing time for tolerance 400 Å 45

Figure 3.16. Interface pressure variation in greedy algorithm with pressure control for tolerance 400 Å 47

Figure 3.17. Interface pressure variation in non-linear programming model using lingo software with steepest edge strategy for tolerance 400 Å 48

Figure 3.18. Modified interface pressure variation in non-linear programming model using lingo software with steepest edge strategy for tolerance 400 Å 49

Figure 4.1. Schematic diagram of Fu and Chandra [7] step height reduction model 51

Figure 4.2. Schematic diagram of planarization model 54

Figure 4.3. Part of the layout of the characterization mask 56

Figure 4.4. Initial surface profile of line 3 in 5 different zones 68

Figure 4.5. Surface evolution of line 3 vs. die position with no control algorithm 71

Figure 4.6. Surface evolution of line 3 vs. die position with spatial pressure control Algorithm 72

Figure 4.7. Surface evolution of line 3 vs. die position with spatial and temporal pressure control algorithm 73

Figure 4.8. Surface evolution of line 3 vs. die position with look-ahead scheduled pressure control algorithm 74

Figure 4.9. Applied pressure of line 3 vs. polishing time in 5 different zones 76

Figure 4.10. Material removal rate of line 3 vs. polishing time in 5 different zones 78

Figure 5.1. The schematic of wafer surface tranformation 82

Figure 5.2. Pseudo surface profile of two-wave sine and square waveforms 84

Figure 5.3. The illustration of the example 5.3.1. 93

Figure 5.4. Graph between surface height vs. polishing time of two-wave model and Fu's model 94

Figure 5.5. Graph between step height vs. polishing time of two-wave model and Fu's model 95

Figure 5.6. The illustration of the two-wave model. 96

Figure 5.7. Graph between surface heights vs. polishing time of two-wave model in example 1 100

Figure 5.8. Graph between step heights vs. polishing time of two-wave model in example 1 100

Figure 5.9. Surface evolution of two-wave model in example 1 101

Figure 5.10. Graph between surface heights vs. polishing time of two-wave model in example 2 103

Figure 5.11. Graph between step heights vs. polishing time of two-wave model in example 2 103

Figure 5.12. Surface evolution of two-wave model in example 2 104

Figure D.1. Graph between thickness vs. die position of simulation results and experimental data (line 3, 4, 5 and 6) with bending factor of 5.0508e6 N/m^2 119

Figure E.1. Pressure selection loop schematic diagram 121

Figure F.1. Pressure selection loop schematic diagram 123

Figure F.2. Pressure selection loop schematic diagram 124

Figure F.3. Pressure selection loop schematic diagram 125

Figure G.1.1. Initial surface profile of line 4 in 5 different zones 126

Figure G.1.2. Surface evolution of line 4 vs. die position with no control algorithm 127

Figure G.1.3. Surface evolution of line 4 vs. die position with spatial pressure control algorithm 128

Figure G.1.4. Surface evolution of line 4 vs. die position with spatial and temporal pressure control algorithm 129

Figure G.1.5. Surface evolution of line 4 vs. die position with look-ahead scheduled pressure control algorithm 130

Figure G.2.1. Initial surface profile of line 5 in 5 different zones 131

Figure G.2.2. Surface evolution of line 5 vs. die position with no control algorithm 132

Figure G.2.3. Surface evolution of line 5 vs. die position with spatial pressure control algorithm 133

Figure G.2.4. Surface evolution of line 5 vs. die position with spatial and temporal pressure control algorithm 134

Figure G.2.5. Surface evolution of line 5 vs. die position with look-ahead scheduled pressure control algorithm 135

Figure G.3.1. Initial surface profile of line 6 in 5 different zones 136

Figure G.3.2. Surface evolution of line 6 vs. die position with no control algorithm 137

Figure G.3.3. Surface evolution of line 6 vs. die position with spatial pressure control algorithm 138

Figure G.3.4. Surface evolution of line 6 vs. die position with spatial and temporal pressure control algorithm 139

Figure G.3.5. Surface evolution of line 6 vs. die position with look-ahead scheduled pressure control algorithm 140

LIST OF TABLES

Table 3.1. List of variables and notations used 22

Table 3.2. Comparison of control strategies with wafer film thickness at the center of 8000 Å 23

Table 3.3. Comparison of stochastic, greedy algorithm and genetic algorithm for curvature control with wafer film thickness at the center of 10000 Å 33

Table 3.4. Comparison of greedy algorithm and non-linear programming model for load control with wafer film thickness at the center of 10000 Å 43

Table 4.1. List of variables and notations used 57

Table 4.2. Model prediction results for validation 58

Table 4.3. The simulation data of final polished surface for the oxide CMP characterization mask of Ouma et al. [17] 69

Table 5.1. List of variables and notations used in the two-wave model 85

Table 5.2. List of variables used for comparing two-wave model to Fu's model 92

Table 5.3. The results of two-wave model and Fu and Chandra's model 93

Table 5.4. List of variables used for simulation results in two-wave model 96

ABSTRACT

Chemical mechanical polishing (CMP) is a planarization process that produces high quality surfaces both locally and globally. It is one of the key process steps during fabrication of very large scale integrated (VLSI) chips in integrated circuit (IC) manufacturing. High and reliable wafer yield is critical in the CMP process; it is dependent upon uniformity of material removal rate across the entire wafer.

The focus on this research is the development of control algorithm for CMP process. Wafer-scale and die-scale models are the two scales used in this study. To achieve improvement in wafer yield, three control strategies are formulated with greedy algorithm, method heuristic and non-linear programming in this wafer-scale. The simulation results show that average wafer yield from genetic algorithm is improved, compared to greedy algorithm. Moreover, average wafer yield from non-linear programming is also improved, compared to greedy algorithm.

At die-scale, a comprehensive control algorithm is developed based on the MRR equations with interface pressure as a control parameter. The interface pressure is varied spatially and/or temporally across the die. In this concept, three control strategies are developed and studied. The strategies are included spatial pressure control, spatial and temporal pressure control, and look-ahead scheduled pressure control. The simulation results of these three strategies show improvement in the upper surface uniformity; however, look-ahead scheduled pressure control seems to be the promising algorithm.

CHAPTER 1. INTRODUCTION TO CHEMICAL MECHANICAL PLANARIZATION (CMP)

1.1. Introduction

Planarization technology is one of the prime process steps during the fabrication of ultra-large scale integrated / very large-scale integrated (ULSI/VLSI) chips in integrated circuit (IC) manufacturing. The *chemical mechanical planarization (CMP)* process has emerged to be the most promising because of its demonstrated capability to provide better local and global planarization of wafer surfaces (Steigerwald et al. [25]) and its ability to planarize over longer length scales than traditional planarization techniques. Beside inter level dielectric (ILD) and inter-metal dielectric (IMD) planarization (Luo and Dornfeld, [12]), CMP has also found applications in the copper damascene process, and shallow trench isolation (e.g., Kaanta et al. [10], Kranenberg and Woerlee [11]).

CMP consists of a chemical process and a mechanical process being performed together to reduce height variation across a dielectric region. The chemical effects are the chemical reactions between the slurry and the wafer surface, which change the solubility and mechanical properties of the wafer surface, while mechanical processes are affected by the interface pressure, the rotational speed of the pad and the wafer, and viscosity of the slurry (Chen and Lee [3]). The Preston equation (Preston [17]) summarizes the material removal rate (MRR) as

$$\frac{dH}{dt} = K_P P V \qquad [1]$$

Where, dH/dt is the MRR per unit surface area,

P is interface pressure,

V is relative velocity between wafer and polishing pad, and

K_p is the Preston coefficient.

A schematic diagram of the CMP process is shown in Figure 1.1. In general, a CMP machine uses orbital, circular and lapping motions. The wafer with size (diameter) ranging from 4 inches to 12 inches is held on a rotating carrier or wafer carrier, while the face being polished is pressed against a polishing pad attached to a rotating platen disk. Then the slurry with nano-scale abrasive particles and special chemicals that flows between the wafer and the pad is used as the chemical abrasive (Luo and Dornfeld [12]). CMP can be carried out on metals as well as on oxides.

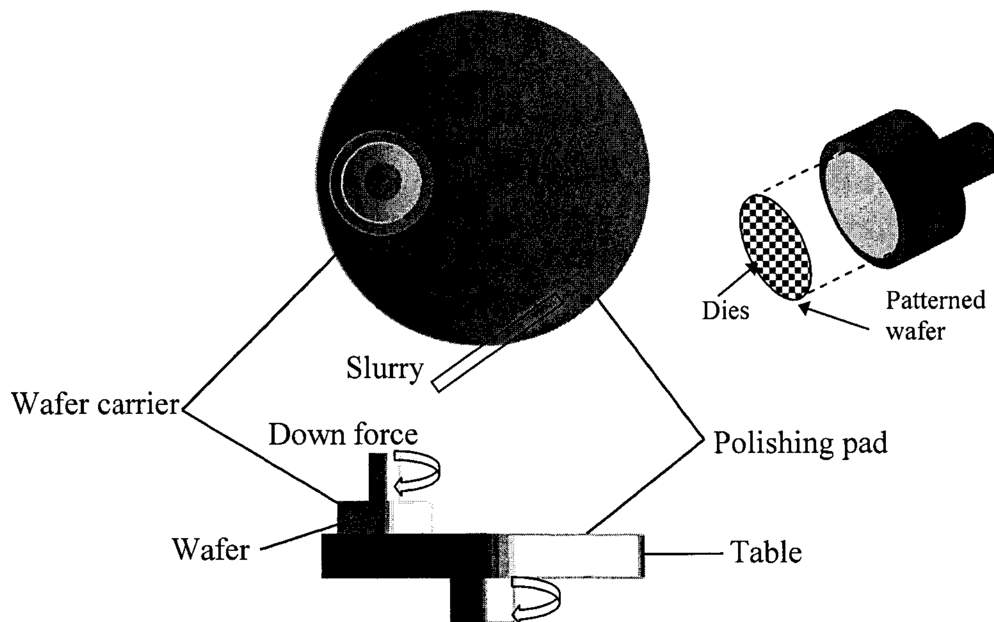


Figure 1.1. Illustration of the CMP process.

Good wafer planarity, both local and global, is essential for the dimensional accuracy required at subsequent lithography stages of wafer manufacture. On a global scale, the

within wafer non-uniformity (WIWNU) is required to be within 200 nm across a 200 mm wafer. Even tighter tolerances become necessary as the wafer size increases and line width decreases (Byrne et al. [2]).

There are two main concerns when considering global planarity of wafers:

1. The edge-ring effect, where sharp variation in removal rate is observed near the edge of the wafer, and
2. The less severe variation in removal rate from the center to the periphery of the wafer.

For a typical 300 mm (diameter) wafer, a 3 mm wide annular ring along the wafer periphery touches about 20% of the chips that represent an annual revenue stream of about $ 2.7B per year for a single IC fabrication facility. Thus the issue of wafer scale uniformity of the MRR has a significant impact on the yield from a wafer and is of critical importance to the IC manufacturing community.

Even though CMP can planarize over longer length scales, pattern density variation across a chip leads to large variation in global thickness across the die. For that reason, CMP removes local steps but generates global steps as illustrated in Figure 1.2. The initial difference in layout pattern density creates a global step height between these two regions due to the difference in removal rates before the local patterns are planarized (Ouma [14]). The global thickness variation still has a serious impact on subsequent process steps such as etching and on multi-level integrated circuits. Because of the depth difference, it is difficult to determine a suitable etch time. The global thickness variation also impacts the circuit performance such as clock skew (Stine et al. [26]). Thus, global step height formation is a significant problem.

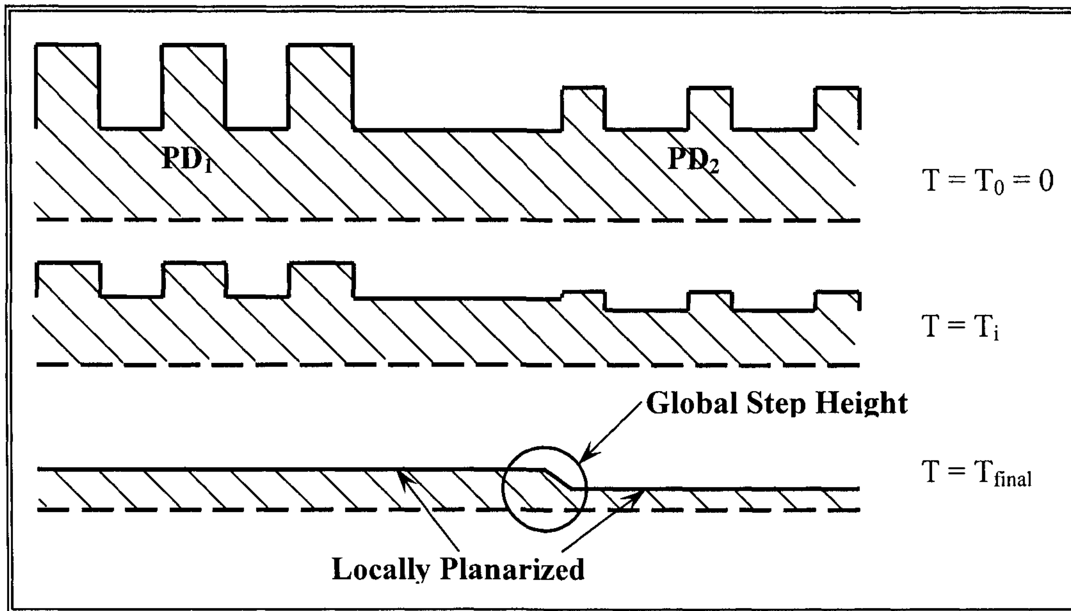


Figure 1.2. Planarization defects due to pattern density variations.

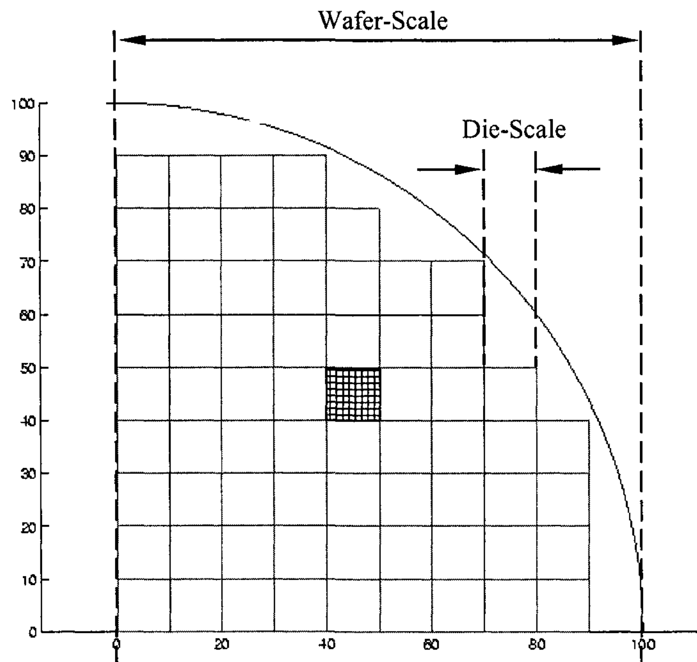


Figure 1.3. Wafer partitioned into dies.

The current studies of CMP process has been ranged into three different scales that are wafer-scale model, feature / die-scale model, and particle-scale model. Where wafer-

scale is in the range of 200 – 300 mm, feature / die-scale is in the range of 10 mm and particle-scale is in the range of nanometers. However, these particular studies are only focused on wafer-scale and die-scale model as shown in Figure 1.3.

1.2. Chemical Mechanical Planarization Components and Functions

In CMP process, slurry, polishing pad and wafer carrier are three main components with different functions. Components and functions of this CMP are explained in details below.

1.2.1. Slurry

Slurry is a liquid mixture of abrasives and chemical solution. The chemicals in slurry react with wafer surface to weaken the material to be easier for removal, where the abrasives act as cutting tools and involving in the repeated plowing action leading to material removal of soften surface (Fu [7]).

In general, abrasives are inorganic oxides. The most commonly used for these abrasives are silica (SiO_2), alumina (Al_2O_3) and ceria (CeO_2). In addition, to be able to achieve consistent polishing, abrasives must be ultra-pure and have nearly uniform size and shape. Size of abrasives is in the range of 50 to 200 nm with very tight distribution. Two main are the number of active abrasives and the size of active abrasives (Luo and Dornfeld [12]).

In chemical solution, water acts as a coolant and also helps removing the abraded material out of the polishing system. As mentioned above, the surface of wafer can be weakened by using a slurry solution to be a hydrolizer in oxide CMP. It changes a Si-O bond to a Si-OH bond that is weaker and easier to remove. Moreover, in metal CMP, the

slurry solution acts as an oxidizer to convert metal into metal oxide which is then removed by abrasives (Fu [7]).

1.2.2. Polishing Pad

Pads have been developed depending on the requirement of CMP. Some are focused on the material removal rate while others might be optimized for surface finish. Based on their structural characteristics, Cook [4] classified polishing pads as:

- Felt and polymer impregnated felts
- Microporous synthetic leathers
- Filled polymer films
- Unfilled textured polymer films

In material removal processes, pad topography and pad properties (Young's modulus, poisson ratio, compressibility and viscoelastic properties) play an important role. It must be able to hold abrasives and transfer the load to abrasives. During this CMP process, pad surface is also planarized in addition to the wafer surface. The abraded pad material, abrasive particles and redeposited wafer surface material fill in the pad pores, causing glaze. The pad becomes smoother and less abrasive, and requires reconditioning.

1.2.3. Wafer Carrier

The wafer carrier is a component used to hold the wafer in a certain kinematics motion while it is polished. It provides a mechanism, such as vacuum, to keep the wafer in place while loading and unloading, and also creates the wafer curvature (fixtured curvature a_2^1) when loading to minimize within wafer non-uniformity (WIWNU).

Furthermore, the wafer carrier prevents the wafer from becoming loose during the polishing process using a retainer ring. A retainer ring can be applied with a separate load to reduce interface pressure near the wafer edge leading to better wafer surface uniformity after the CMP process. In addition, the carrier film at the back of the carrier (backing film) should be able to compensate for small amounts of wafer bow or wafer tilt.

1.3. Dissertation Outline

The first objective of this dissertation is to explain the improvement of global planarity and wafer yield, which leads to accurately prediction of the stopping time. This model is using three control strategies based on various interface pressures and wafer curvatures for a wafer-scale model. It will be developed with practical wafer properties and process specifications, and an assumed uniform pattern density across the entire polish span. These studies are dependent on many input variables and process specifications as shown in Figure 1.4. The yields from these control strategies are studied.

Based on the effective pattern density in a region and the utilization of the step height reduction model developed by Fu et al. [9], the second objective is to explain a control based open loop algorithm to obtain uniformity over the pattern dependant non uniform wafer surfaces in a die-scale model. The die in the wafer surface is assumed to have n number of zones of different heights and different pattern densities. The 2D simulation process is applied to track the amount of removal and current step heights for each step. A two-wave die-scale model is also formulated.

The objectives of this research are to study wafer yield improvement of CMP process in wafer-scale model and to obtain uniformity over pattern dependant, non uniform wafer surfaces in a die-scale model. This dissertation begins with the background of CMP process introduced in chapter 2. Studies on wafer yield improvement in wafer-scale model are then explained in chapter 3. Three control strategies for yield improvement are formulated with greedy algorithm, method heuristic and non-linear programming along with their simulation results.

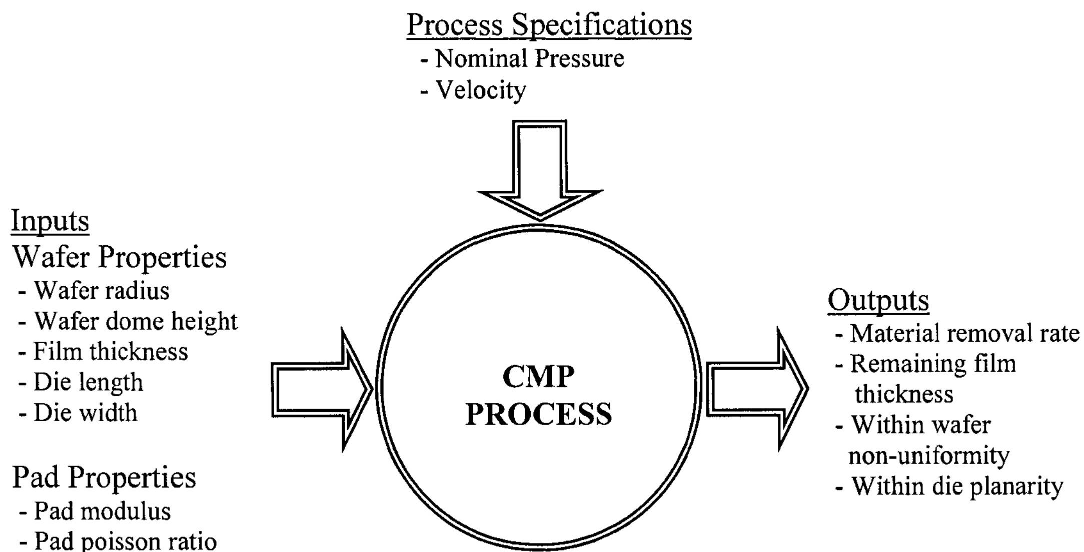


Figure 1.4. Lists of input and output parameters in CMP process.

A step height reduction model and three control strategies of the die-scale model are explained in chapter 4. Comparison results between three control strategies and no control strategy are also discussed in this chapter. Chapter 5 discusses further on two-wave die-scale model, developed from Fu and Chandra [9], along with simulation results. Discussion, conclusion and future works are presented in chapter 6.

CHAPTER 2. BACKGROUND OF CMP PROCESS

Introducing the mechanics of beam bending, Sivaram et al. [22] investigated wafer scale variations on MRR by modifying Preston's equation via varying the contact pressure according to the deflection profile. They assume that the section modulus (E*I) for the polishing pad is the same or higher than that of the platen; which is usually not the case. Runnels and Renteln [21] focus on finite element modeling of the pad deflection and investigated wafer edge effects and wafer curvature effects during polishing. They observed that the normal pressure was uniform under the wafer except within one millimeter from the wafer edge. This does not explain the experimental observations where, a significant increase in MRR is observed further (around 5 mm) from the edge of the wafer.

In a series of papers, Runnels and a colleague ([18], [19] and [20]) have investigated the phenomenon of material removal in a CMP process at the feather scale. They focus primarily on the effects of slurry flow and its associated fluid dynamics. Runnels and Eyman have used the steady-state incompressible Navier-Stoke equation to model the slurry flow at the wafer-polishing pad interface. Their model is physically based, but is constructed on idealized geometry. The wafer surface is assumed to be smooth and spherical with a large radius of curvature, which implies that issues relating to the polishing mechanism and the structure of the wafer surface are neglected. The fluid layer thickness and the angle of attack between the pad and the wafer are obtained through an iterative procedure satisfying force and momentum balance. Motivated by feature scale observations, Runnnels proposes a modification to Preston's equation, where the relative velocity, V, is replaced by the tangential stress, σ_t, on the wafer surface. Runnels's modified equation may be expressed as

$dH/dt = C\sigma_n\sigma_t$. Where, σ_n and σ_t represent the local normal and the tangential stresses, respectively.

Assuming solid-solid contact, Wang [28] and Srinivasa-Murthy et al. [24] have also investigated the effects of various process parameters on the degree of wafer scale nonuniformity. It has been observed that decreasing polishing pad compressibility will improve the uniformity of planarization process across the wafer. Sasaki et al. [23] have conducted a detailed finite element method (FEM) analysis of the pressure distribution under a wafer, and investigated the influences of wafer chamber and retainer ring on the pressure distribution (see Figure 2.1). Fu and Chandra [8] derive an analytical solution for the normalized interface pressure distribution based on an assumption of an elastic / visco-elastic half-space for a pad. It is observed that depending on the wafer curvature and polishing conditions, the interface pressure may exhibit significant variation.

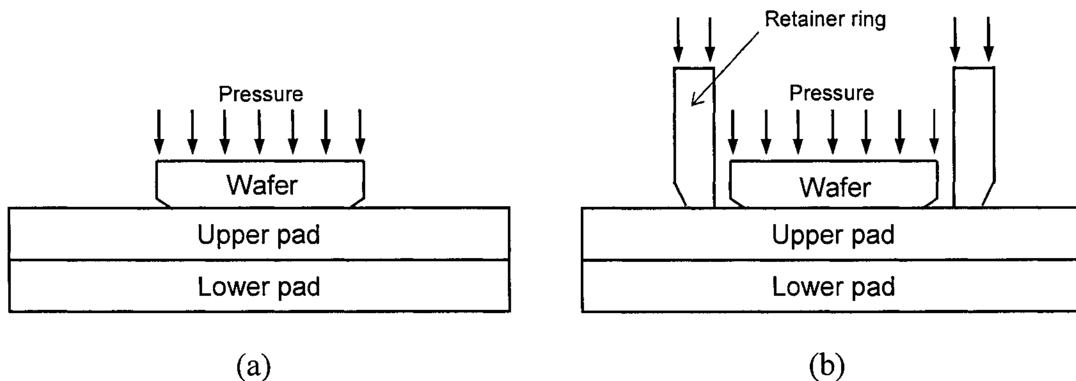


Figure 2.1. FEM model for investigating wafer chamfering (a), and for investigating retainer ring (b).

In the present work, the lower pad stiffness (compared to platen) is accommodated by treating the wafer pad contact as the indentation of a rigid indenter on an elastic half-space (Eamkajornsiri et al. [6]). The spatial variation of the interface pressure profile is

incorporated into the Preston equation to obtain the spatial distribution of the MRR. This model accurately reflects the observed experimental variations in MRR over several millimeters from the wafer edge.

The wafer-scale model is used as a basis for simulating and controlling material removal in the CMP process. The wafer scale model allows the investigation of CMP process parameters such as down load, polishing time, pad properties, pre-existing wafer curvature etc. on the yield. Therefore model based planning and control of the CMP process is possible and can serve as an alternative to the trial and error design procedure that is practiced today.

Tugbawa et al. [27] formulated a density step height based on a die-scale model for abrasive-free copper CMP process. Ouma et al. [16] provides an analytical model using planarization length as a characterization parameter for a particular CMP process. They also use FFT technique to obtain estimates of effective discretized densities across the entire wafer. Fu and Chandra [9] also formulate an analytical model for a dishing and step height model by assuming force redistribution is proportional to dishing height. Their model introduces an alpha value as the characterization parameter for a particular CMP process similar to planarization length in the Ouma et al. [16] model. Based on Fu and Chandra [9], dishing and step height reduction depend on various parameters such as Preston's constant, pad stiffness, pad bending, pressure, and velocity.

CHAPTER 3. WAFER-SCALE MODEL SIMULATION

In this chapter, the wafer-scale model simulation is modified from earlier work by Eamkajornsiri [5]. The wafer curvature (a_2) obtained in [5] was described as an oxide thickness profile. Therefore, to be more realistic, the wafer curvature in this chapter is defined by the wafer curvature that occurred from the CMP machine during the polishing process (a_2^1) and the nominal thickness of the oxide (a_2^2). First, the wafer-scale model is explained. Then control strategies for yield improvement are formulated with greedy algorithm, method heuristic and non-linear programming and their example results are also presented.

3.1. Wafer-scale Model

Pre-existing wafer curvature is modeled as a quadratic function a_2r^2 due to effect of loading wafer into wafer carrier and type of material of film (oxide, copper, etc.). As shown in Figure 3.1, the wafer is subjected to a rigid body displacement of a_0 due to the down load.

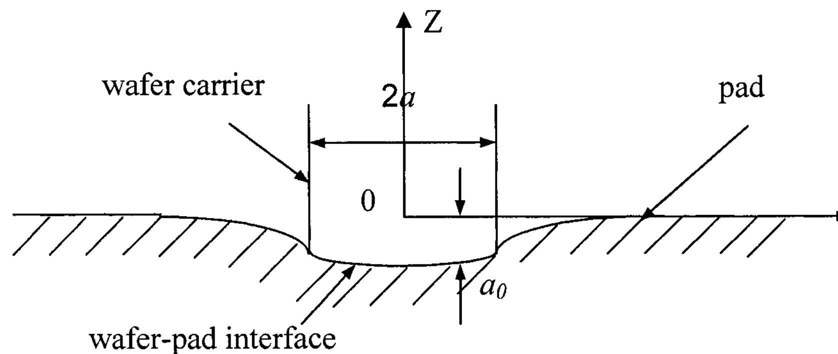


Figure 3.1. Model of pad and wafer contact, a_0 is a rigid body displacement of the wafer due to the down load and $2a$ is the wafer diameter.

The displacement field right under the wafer can therefore be expressed as:

$$f(r) = a_0 + a_2 r^2 \qquad [2]$$

Where, r measures the radial distance from the center of the wafer,

$2a_2$ is the wafer curvature caused by preexisting wafer bow, and

a_0 is the vertical displacement of the wafer (depth of penetration) into pad.

The pressure distribution on the contact area by Fu and Chandra [8] can be expressed as:

$$P(r) = K \frac{4a_2 r^2 + \left(a_0 - 2a_2 a^2\right)}{\left(a_0 - 2a_2 a^2\right)\sqrt{1 - \left(\frac{r}{a}\right)^2}} \qquad [3]$$

Where, r is any arbitrary radius,

a is the radius of the wafer, and

K is a constant that depends on the pad properties.

Based on the kinematics as depicted in Figure 3.2, the magnitude of the polishing velocity, which is the relative velocity between the pad and the wafer, (Chen and Lee [3]) can be written as:

$$\left|V_p\right| = \left|\dot{\theta}_1\right|\left(b_1^2 + r^2(k_v - 1)^2 - 2b_1 r(k_v - 1)\cos\theta_2\right)^{1/2} \qquad [4]$$

Where, $\dot{\theta}_1$ is the angular velocity of the polishing pad,

b_1 is the offset distance between the axes of the pad and the wafer,

r is an arbitrary radius on the wafer, and

k_v is the ratio of the angular velocity of the wafer to the angular velocity of the pad.

The angular velocity of the wafer θ_2 can be expressed as

$$\theta_2 = (\dot{\theta}_2 - \dot{\theta}_1) \cdot t \quad [5]$$

Where, t is the processing time.

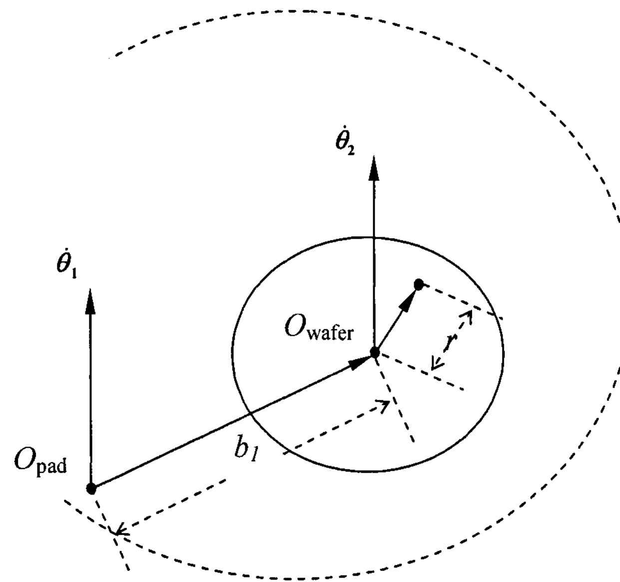


Figure 3.2. Kinematics of the CMP process.

The material removal rate (MRR) can now be calculated at any point in space and time by using Equation 6. The height removed may be obtained by integrating dH/dt over time.

$$\frac{dH}{dt} = K_P PV \quad [6]$$

Where, dH/dt is the material removal rate per unit surface area,

P is the interface pressure,

V is the relative velocity between the wafer and the polishing pad, and

K_p is the Preston's coefficient.

This model had been compared against experimental observations of Srinivasa-Murthy et al. [24] and the model predictions compare very well with the experimental data.

3.2. Control Formulations

In order to improve the wafer yield, control schemes with load (applied pressure) and/or fixtured wafer curvature as control parameters are introduced.

3.2.1. Alternative Objective Functions

a. *Number of good sectors within tolerance*

This objective function can be formulated by considering number of good sectors together with polishing time at the end of CMP process, and it can be written as:

$$\text{Objective} = w \times \textit{number of good sectors}\ (NGS) - \textit{polishing time}\ (t) \qquad [7]$$

Where, NGS is maximum number of good sectors occurred during the polishing process at the polishing time step t (in second), and

w is the arbitrary weights value.

In addition, a good sector is defined as a sector that all four corners fall within specified tolerance.

b. *Deviation from target surface*

It is assumed that by minimizing a deviation from a desired surface will lead to a better yield improvement. Thus, this objective function can be formulated by simultaneously considering:

- The curvature of the wafer and
- The amount of material available to be subsequently removed.

A moment function comprised of an upper moment and a lower moment is used to capture this information. Figure 3.3 illustrates a wafer profile and the upper and lower moments. The wafer radius is discretized to N intervals (in this work N = 10). The lower moment of deviation (LMD) is used to measure curvatures. LMD is a summation of the product of radius (r_i) and height difference $(h_C - F(r_i))$ between oxide profile at the center of the wafer and the radius r_i, as explained in Equation 8. The upper moment of deviation (UMD) is a summation of the product of radius (r_i) and the height difference $(F(r_i) - h_A)$ between oxide profile and desired surface of the oxide, shown in Equation 9. In each case, the moment amplifies the height difference by the radius. Thus, the height difference at the edge of the wafer is weighted more. This is justified because for any annular region; the area $(2\pi r dr)$ at the edge of the wafer is greater than the center of the wafer. The upper and lower moments are combined together by a weighting term α that lies in the interval [0, 1], called the combined moment of deviation (MOD) shown in Equation 10.

$$LMD = \sum_{i=1}^{N} \{(h_C - F(r_i)) * r_i\} \qquad [8]$$

$$UMD = \sum_{i=1}^{N} \{(F(r_i) - h_A) * r_i\} \qquad [9]$$

$$MOD = \alpha(LMD) + (1-\alpha)(UMD) \qquad [10]$$

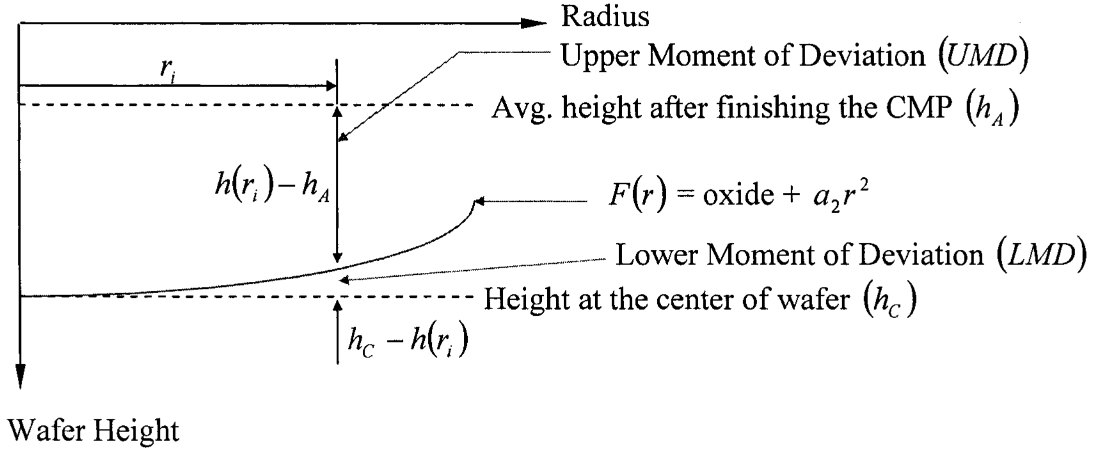


Figure 3.3. Moment calculation using the upper moment and the lower moment.

3.2.2. Alternative Decision Variables

To improve the wafer yield, these control schemes are based on two control parameters, load and curvature, shown as:

a. *Load as control parameter*

A uniform applied pressure (P) is used as a decision variable and it can be continuously varied in the range of $P_{Min} \leq P \leq P_{Max}$, depended on machine capability.

b. *Curvature as control parameter*

A fixtured wafer curvature (a_2^1) is used as a decision variable and it can be varied continuously in the range of $a_{21}^{Min} \leq a_2^1 \leq a_{21}^{Max}$, also depended capability of the CMP machine.

3.3. Greedy Algorithm for Yield Improvement

Three control strategies with greedy algorithm are explored. The first strategy varies the interface pressure to control wafer yield. The second strategy controls the curvature of the wafer. It is assumed that the wafer is set in a fixture that can control the curvature. In curvature control, the applied load is held a constant. In the third strategy, both the curvature and the applied load are controlled. Each of these three strategies is discussed in detail below and objective function of model with greedy algorithm is explained first.

3.3.1. Objective Function

The objective function of this model shown in 3.2.1.b, *deviation from the target surface*, can be written as:

$$MOD = \alpha \cdot (LMD) + (1-\alpha) \cdot (UMD)$$

Where, α is a weighting term that lies in the interval [0, 1],

LMD is referred in equation 8, and

UMD is referred in equation 9.

From equation 8 and 9, the objective function can be rewritten as:

$$MOD = \alpha . \sum_{i=1}^{N} \{[h_C - F(r_i)]r_i\} + (1-\alpha) \sum_{i=1}^{N} \{[F(r_i) - h_A]r_i\}$$

Where, r_i is a distance from the center of the wafer,

h_C, h_A and $F(r_i)$ are referred in Figure 3.3, and

N is number of points.

3.3.2. Model with Greedy Algorithm

To improve the wafer yield, the objective function shown in 3.3.1. is minimized with the decision variables shown is 3.2.2. Thus, the model with the greedy algorithm can be obtained as:

$$Min \left(\alpha . \sum_{i=1}^{N} \{ [h_C - F(r_i)] r_i \} + (1-\alpha) . \sum_{i=1}^{N} \{ [F(r_i) - h_A] r_i \} \right) \quad \text{at every time step}$$

Subject to $P_{Min} \leq P \leq P_{Max}$ for Load Control

Or

Subject to $a_2^{Min} \leq a_2^1 \leq a_2^{Max}$ for Curvature Control

Or

Subject to $P_{Min} \leq P \leq P_{Max}$
and $a_2^{Min} \leq a_2^1 \leq a_2^{Max}$ for Combined Curvature and Load Control

In order to vary the decision variables every time step, next, switching logic is implemented and discussed.

a. *Switching Logic*

The moment of deviation (MOD) objective function is used for control. Figure 3.4 shows two curves (P_1, P_2) of moment of deviation vs. polishing time. At polishing time t_1, pressure P_1 is selected because it gives the steeper slope curve compared to pressure P_2. This pressure is applied until it reaches polishing time t_2 where applying pressure P_2 gives the steeper slope curve. Again, this pressure is presented until it reaches the next switching point t_3 where another pressure gives the steeper slope curve. This pressure is applied until it stops.

Thus, in switching logic, the steepest slope curve is chosen at all times as it helps to improve the wafer surface in the fastest possible manner. Therefore, the control algorithms discussed next used the moment of deviation as the objective function and used the steepest descending moment curves to select the values of the load and curvature to obtain the best possible wafer yield in the fastest possible time. The algorithm is a "greedy" algorithm and simulation results indicate that they improve the wafer yield substantially.

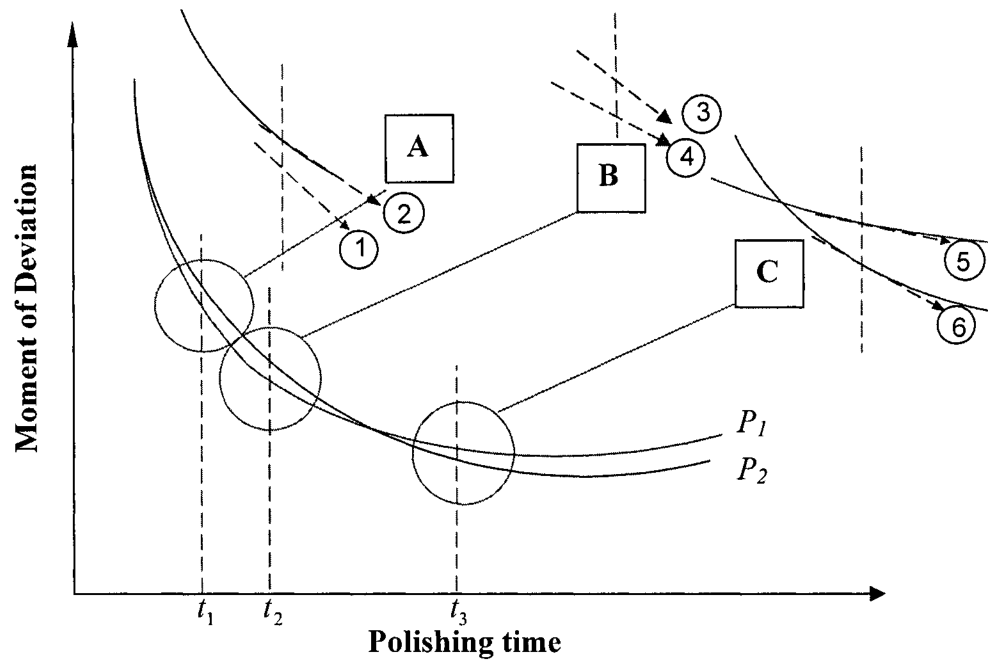


Figure 3.4. Load switching using the moment of deviation objective function.

b. *Load Control*

In load control, the interface pressure is controlled to improve the wafer yield. The initial wafer profile is defined by the wafer curvature and the nominal thickness of the oxide. The wafer curvature a_2 is computed from

$$a_2 = a_2^1 + a_2^2$$

Where, a_2^1 is the fixtured curvature and a_2^2 is the film curvature.

The chosen interface pressure at each CMP step is the one that minimizes the moment of deviation. Once the appropriate load for a particular CMP step is chosen, polishing is carried out for that time step. At the end of the time step, the new wafer curvature is computed by performing a least squares fit of the polished wafer profile. The number of good die sectors is computed next. A die is declared good if all four corners of the die are within the tolerance limits. The wafer yield is tracked through the simulation, and the CMP process should be stopped when the yield reaches its maxima.

c. *Curvature Control*

In curvature control, the curvature is controlled to improve the wafer yield for a fixed and constant value of the interface pressure. The initial wafer profile is defined by the film thickness at the center and at the edge of the wafer. The fixtured curvature at each CMP step is the one that minimizes the moment of deviation. Once the fixtured curvature for a particular CMP step is chosen, polishing is carried out for that time step. At the end of the time step, the new film wafer curvature a_2^2 is computed by performing a least squares fit of the wafer profile. The number of good die sectors is computed next. The wafer yield is tracked through the simulation, and the CMP process should be stopped when the yield reaches its maxima.

d. *Combined Curvature and Load Control*

In combined curvature & load control, the fixtured curvature is calculated by minimizing the moment of deviation for the determined value of a_0. At the end of the

time step, the new film wafer curvature a_2^2 is computed by performing a least squares fit of the wafer profile. The value of a_0 in the next time step is reduced by the amount of material polished at the wafer center. The fixtured curvature for the next time step is then computed by minimizing the MOD. The number of good die sectors is computed next. The wafer yield is tracked through the simulation, and the CMP process should be stopped when the yield reaches its maxima. While both the load and curvature are varied at each time step; it is not a true curvature and load control since the load adjustment is not selected independently.

3.3.3. Simulation Results – Greedy Algorithm

In this example, the three control strategies above are simulated. Three values of initial film curvature a_2^2 (-10e-6 m^{-1}, -15e-6 m^{-1} and -20e-6 m^{-1}) are chosen. The simulation parameters used are listed in Table 3.1 below:

Table 3.1. List of variables and notations used.

Variable and Notation	Value	Unit
Pad angular velocity	35	rpm
Wafer angular velocity	20	rpm
Wafer diameter	200	mm
The offset between the axes of pad and wafer	170	mm
Film thickness at the center	8000	Å
Desired final surface height (after finishing CMP process)	1000	Å
Tolerance	200	Å
The fixed curvature (a_2^1)	-20e-04	m^{-1}

The results in Table 3.2 show that by applying the load control strategy, there are 216 to 224 good dies out of possible 268 dies with polishing times between 369 to 387 seconds. There are 216 to 240 good dies out of possible 268 dies with polishing time between 106 to 108 seconds by applying curvature control strategy. The combined curvature and load control strategy yields the same result. Because, in both strategies, curvature is the primary control where load is kept constant in the curvature control and load is varied only to keep the wafer holder constantly that make it easier to control in the combined curvature and load control. In addition, the wafer yield is 81% to 84% by applying load control, but the polishing time is excessive. By running either with curvature control or combined curvature and load control, the wafer yield is still 81% to 90%, but the polishing time is shortened. The next chapter discusses the wafer-scale model with genetic algorithms.

Table 3.2. Comparison of control strategies with wafer film thickness at the center of 8000 Å.

	Low Film Curvature			Medium Film Curvature			High Film Curvature		
	# of good dies	CMP time (sec)	α	# of good dies	CMP time (sec)	α	# of good dies	CMP time (sec)	α
Load Control	216	369	0.55	224	387	0.58	224	376	0.54
Curvature Control	216	108	0.24	224	108	0.34	240	106	0.34
Combined Curvature & Load Control	216	108	0.26	224	108	0.34	240	106	0.34

3.4. Method Heuristic for Yield Improvement

In the previous section, the wafer yield is improved by implementing a greedy algorithm to the three control strategies for the wafer-scale model. To investigate better yield improvement, the curvature control with Genetic Algorithms (GAs), as one of the techniques of method heuristic, is introduced. First, an overview of genetic algorithms is explained. Then genetic algorithm for yield improvement and the simulation results are presented.

3.4.1. Genetic Algorithms

Genetic algorithms are adaptive methods, which may be used to solve search and optimization problems (Busetti [1]). They are search algorithms based on the mechanics of natural selection and natural genetics. GAs combine survival of the fittest among the string structures with a structured yet randomized information exchange to form a search algorithm with some innovative flair. In every generation, a new set of artificial creatures (strings) is created using bits and pieces of the fittest of the old; an occasional new part is tried for good measure. GAs efficiently exploit historical information to consider new search points with expected improved performance.

The genetic algorithms are different from the conventional optimization techniques in the following ways.

(i) The search is based on the selection from a population.

(ii) It is a blind search.

(iii) The search uses stochastic operators, not deterministic rules.

Moreover, the advantages of these algorithms are following:

(i) It is useful when analytical solutions cannot be found due to the complexity of the problem.

(ii) There is less risk of getting 'stuck' at local optima as GAs give continual improved performance.

(iii) Constant monitoring is not needed if programmed well.

(iv) Ideal for problems involving large number of parameters; especially since this can be now easily accomplished with decreasing cost of computation power.

3.4.2. Model with Genetic Algorithm

This problem consists of several parameters, where the interface pressure and the fixtured curvature are the two main variables. Because there is not an analytical solution to this problem, a genetic algorithm is used to find a better result. In doing this, the best fitness value is calculated using a series of either interface pressure or fixtured curvature to control the CMP process. This fitness value will then be used to determine the wafer yield improvement. The genetic algorithm steps are discussed in detail below. The coding scheme is first explained, followed by an explanation of the fitness value used in this algorithm.

a. *Coding Scheme*

Before running a GA, a proper coding for the problem must be devised (Busetti [1]). It is assumed that a potential solution to a problem may be represented as a series of either interface pressure (P) or fixtured curvature (a_2^1).

A series of interface pressures can be illustrated as:

$$P_1 \; P_2 \; P_3 \; P_4 \; P_5 \; P_6 \; P_7 \; P_8 \; ..... \; P_{n-2} \; P_{n-1} \; P_n .$$

Where, P_i represents an interface pressure of the ith processing time.

And a series of fixtured curvature values can be illustrated as:

$$a_{21}^{1}\ a_{21}^{2}\ a_{21}^{3}\ a_{21}^{4}\ a_{21}^{5}\ a_{21}^{6}\ a_{21}^{7}\ a_{21}^{8}\ \ldots..\ a_{21}^{n-2}\ a_{21}^{n-1}\ a_{21}^{n}.$$

Where, a_{21}^{i} represents a fixtured curvature of processing time i.

b. *Fitness Value*

Besides the coding scheme used, finding the fitness value is the most complicated and the most important step in GA. The search seeks to find a better wafer yield through various GA operations. The fitness is the most important value in the genetic algorithm search. In addition, for this particular problem, the fitness value shown in 3.2.1.a, *number of good sectors within tolerance*, can be written as:

Fitness Value = w x *number of good sectors* (NGS) - *polishing time* (t)

Where, NGS is maximum number of good sectors occurred during the polishing process at the polishing time step t, and

w is an arbitrary weight's value of 100 to differentiate fitness values when there are two or more sample sets with the same number of good sectors.

c. *Genetic Algorithm Steps for Curvature Control*

First, the model with the genetic algorithm is expressed as:

Min $(w$ x *number of good sectors* (NGS) - *polishing time* $(t))$

Subject to $a_2^{Min} \le a_2^1 \le a_2^{Max}$ for Curvature Control

The flowchart of this program (refer to Figure 3.9) and the steps are shown below:

(i) Generate one hundred initial population sets (refer to Figure 3.5). The series of fixtured curvature values in each population set is created using random generator in visual basic software. In addition, the randomization in this process varies upon the range determined by manufacturer $\left(a_{21}^{\min} \text{ to } a_{21}^{\max}\right)$.

(ii) Run the CMP polishing process simulation with each sample set using the wafer-scale model until reaching the maximum number of good sectors at the target surface, then stop.

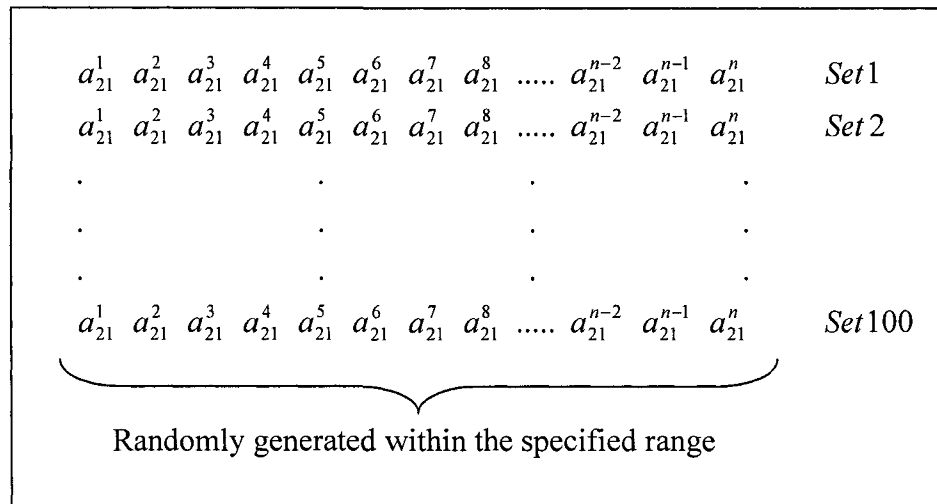


Figure 3.5. The initial of population sets with the series of fixtured curvature.

(iii) Calculate the fitness values of 100 sample sets after running the CMP polishing process using equation 7.

(iv) Find the new generation (refer to Figure 3.6).

a) Select the best N fitness values and save the sample sets to the next generation.

b) Use the genetic algorithm operations to get the rest of the new generation by

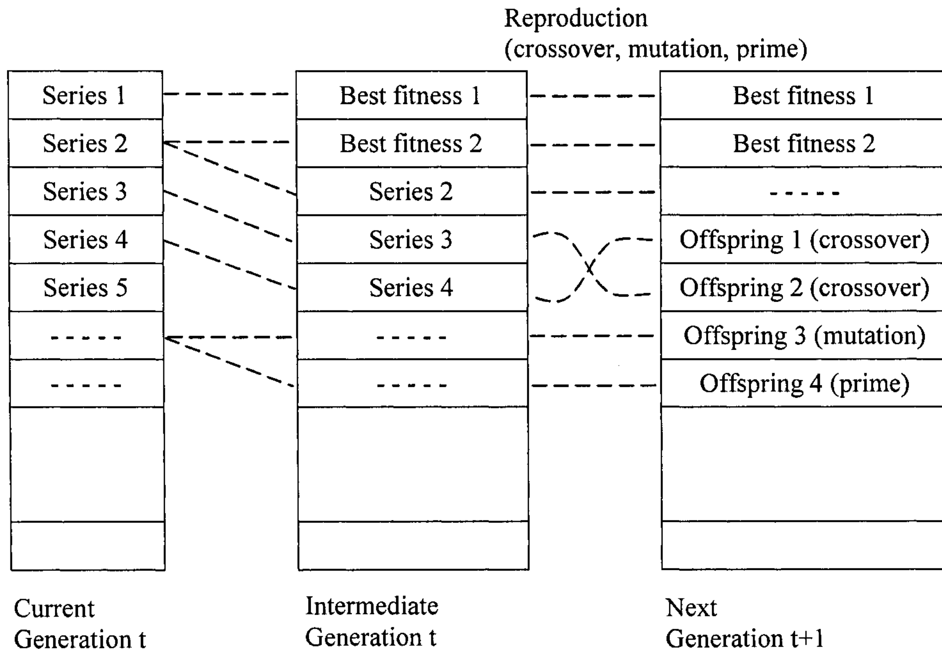


Figure 3.6. The standard GA diagram.

i. Select the next generation with the probability of 0.2 using the roulette wheel selection (refer to Figure 3.7). The size of the section in the roulette wheel is proportional to the fitness value of every series of fixtured curvature. Clearly, from the Figure 3.7, the series with the larger fitness value will be selected more often (Obitko [13]).

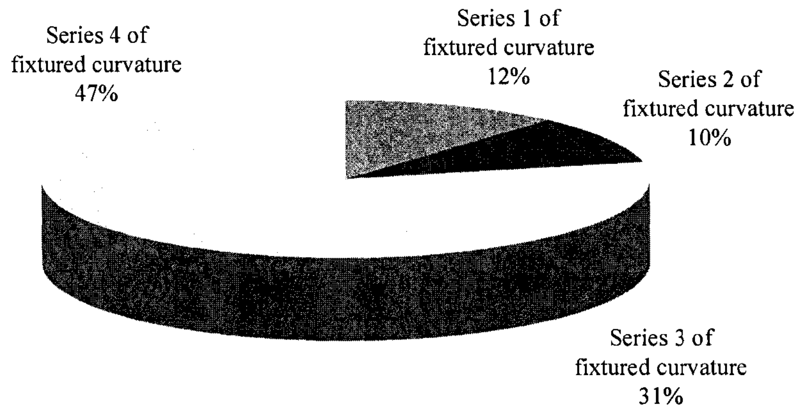


Figure 3.7. The fitness values of four series of fixtured curvature on roulette wheel.

ii. Apply one of the following operations (refer to Figure 3.8) with the respective probability.

1. *Single crossover with P(0.5).* If a random number (RN) is smaller than or equal to 0.5 $(RN \leq 0.5)$, two series of either fixtured curvature called Mom and Dad are chosen. Then these two series are crossed over at a selected crossover point.

2. *Mutation with P(0.05) and P(0.05) for each step.* If RN is greater than 0.5 and smaller than or equal to 0.55 $(0.5 < RN \leq 0.55)$, there is a chance that the selected series of fixtured curvature will be mutated. Each step of the series will be mutated with a probability of 0.05. The new step is selected randomly between $a_{21}^{\min}$ and $a_{21}^{\max}$. Where, $a_{21}^{\min}$ is the minimum fixtured curvature that can be applied, and $a_{21}^{\max}$ is the maximum fixtured curvature that can be applied.

Crossover (| is the crossover point)

Series 1	$a_{21}^{11}\ a_{21}^{12}\ a_{21}^{13} \mid a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17}\ a_{21}^{18}\ a_{21}^{19}$
Series 2	$a_{21}^{21}\ a_{21}^{22}\ a_{21}^{23} \mid a_{21}^{24}\ a_{21}^{25}\ a_{21}^{26}\ a_{21}^{27}\ a_{21}^{28}\ a_{21}^{29}$
Offspring 1	$a_{21}^{11}\ a_{21}^{12}\ a_{21}^{13} \mid a_{21}^{24}\ a_{21}^{25}\ a_{21}^{26}\ a_{21}^{27}\ a_{21}^{28}\ a_{21}^{29}$
Offspring 2	$a_{21}^{21}\ a_{21}^{22}\ a_{21}^{23} \mid a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17}\ a_{21}^{18}\ a_{21}^{19}$

Mutation (3rd position is randomly picked to mutate)

Series 3	$a_{21}^{11}\ a_{21}^{12}\ a_{21}^{13}\ a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17}\ a_{21}^{18}\ a_{21}^{19}$
Offspring 3	$a_{21}^{11}\ a_{21}^{12}\ a_{21}^{13_new}\ a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17}\ a_{21}^{18}\ a_{21}^{19}$

Prime (1st and 8th positions are randomly picked to prime)

Series 4	$a_{21}^{11}\ a_{21}^{12}\ a_{21}^{13}\ a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17}\ a_{21}^{18}\ a_{21}^{19}$
Offspring 4	$a_{21}^{11_new}\ a_{21}^{12}\ a_{21}^{13}\ a_{21}^{14}\ a_{21}^{15}\ a_{21}^{16}\ a_{21}^{17_new}\ a_{21}^{18}\ a_{21}^{19}$

Figure 3.8. The single crossover, mutation, and prime module function.

3. *Prime with P(0.1).* If RN is greater than 0.55 and smaller than or equal to 0.65, each of the steps in the series is computed as

$$a_{21}^{i_new} = a_{21}^{\max} - a_{21}^{i}.$$

Where, $a_{21}^{i_new}$ is the new fixtured curvature at i^{th} step, and a_{21}^{i} is current fixtured curvature at i^{th} step.

c) Finally, the new generation is composed of

i. series from the first five fitness values of the previous generation.

ii. ~ 47 to 48 series with crossover.

iii. ~ 4 to 5 series with mutation.

iv. ~ 9 to 10 series with prime.

v. And roughly 32 to 35 series from the previous generation.

(v) Keep running for 1000 generations from step (ii) to step (iv) then STOP.

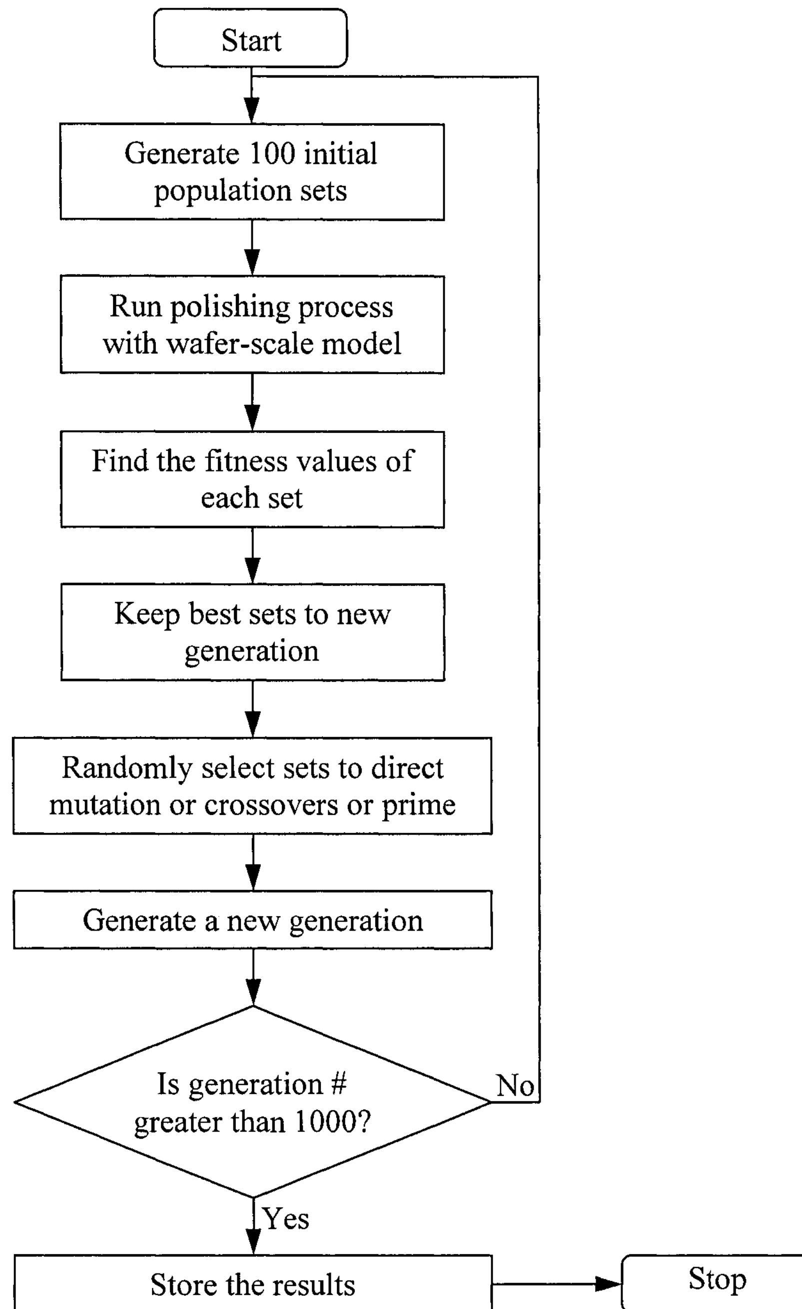


Figure 3.9. Flow chart of the genetic algorithm program.

3.4.3. Simulation Results – Genetic Algorithm

In this example, the genetic algorithm for curvature control is applied and simulated. Three values of initial film curvature a_2^2 (-5e-6 m^{-1}, -10e-6 m^{-1} and -15e-6 m^{-1}) are chosen. The simulation parameters used in this example are the same as those in 3.4, except the target film thickness is at 10000 Å. The range of fixtured curvature a_2^1 is from -2.0e-02 to -2.0e-03 m^{-1}. Figure 3.10 shows normalized pressure profile variation with wafer radius for different fixtured curvatures.

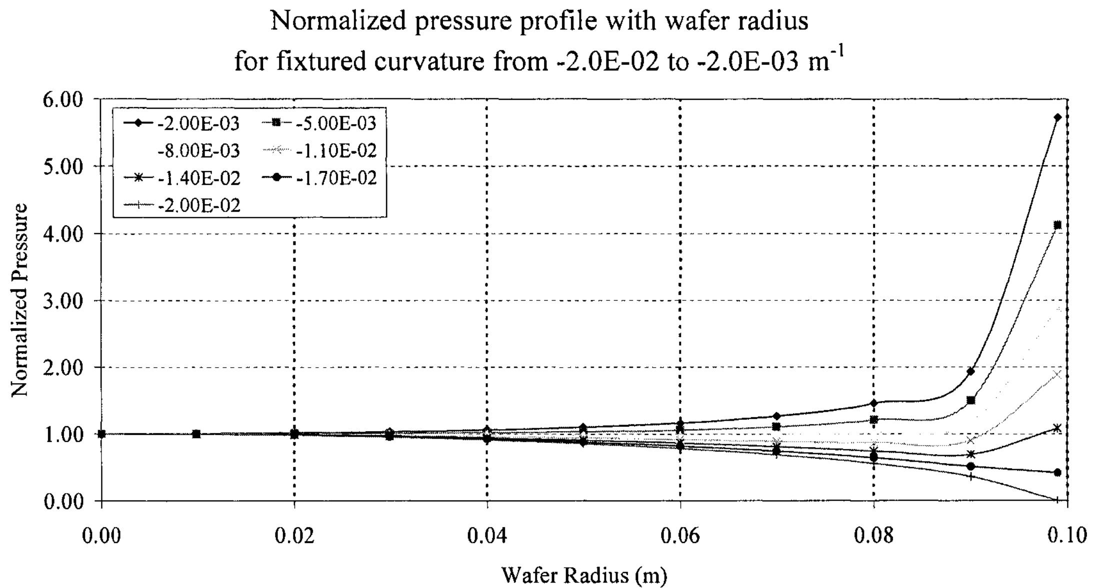


Figure 3.10. Normalized pressure profile curves shown for seven values of fixtured curvature a_2^1 (m^{-1}) with the indentation depth of 3.678e-04 (m).

The comparison of the wafer yield prediction among stochastic, greedy algorithm, genetic algorithm, and the genetic algorithm with modified fixtured curvature for curvature control is shown in Table 3.3. The results of stochastic curvature control are run and selected from 1000 sample sets. Unlike stochastic curvature control, the results

of best wafer yield with shortest time are chosen by running curvature control with the greedy algorithm for alpha values from 0 to 1. In the genetic algorithm for curvature control, the results are selected after running 1000 generations. Lastly, the series of fixtured curvature are sorted from GA results sequentially from high curvature to low curvature or in the reverse direction. Afterward, the sorted series are used to compute the wafer yield.

Table 3.3. Comparison of stochastic, greedy algorithm and genetic algorithm for curvature control with wafer film thickness at the center of 10000 Å.

	Low Film Curvature		High Film Curvature		Very High Film Curvature	
	# of good dies	CMP time (sec)	# of good dies	CMP time (sec)	# of good dies	CMP time (sec)
Stochastic[1]	208	138	208	135	124	133
Greedy Algorithm	188	135	208	134	188	133
Genetic Algorithm with 1000 generations[1]	208	138	224	138	240	134
GA with modified fixtured curvature[1]	208	138	216	138	240	134

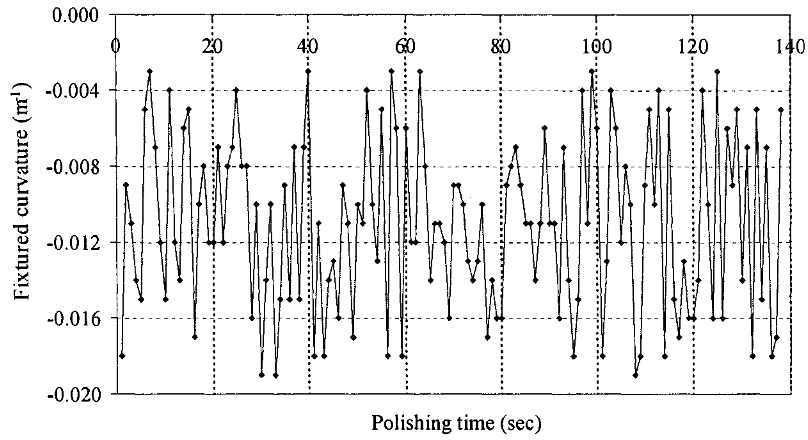
[1] average results based on their simulations with standard deviation of zero, except one of stochastic with initial very high film curvature.

The results from Table 3.3 show that by applying the genetic algorithm through curvature control, the estimated wafer yields are increased by 8 to 28 percent, comparing to the results from greedy algorithm. Similarly, the improvement of 8 to 28 percent from the greedy algorithm is still presented when comparing the results to the modified

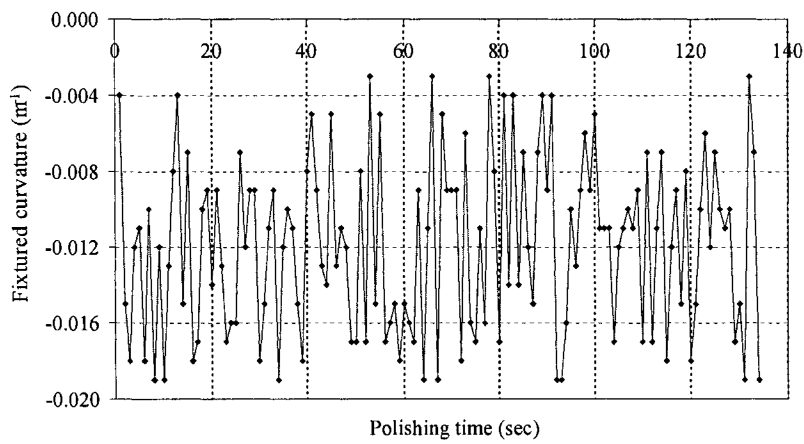
fixtured curvature of GA. However, the results from GA with modified fixtured curvature are not as good as those from the pure genetic algorithm.

Figures 3.11 to 3.14, shown below, illustrate the series of fixtured curvature control in relation to polishing time using different approaches. Figure 3.11 shows that the model curvatures fluctuated during the entire period of polishing process in the stochastic. With the greedy algorithm, shown in Figure 3.12, curvatures vary only 3 to 4 times during polishing process. Similar to the results from Figure 3.11, Figure 3.13 shows the variation of curvature for all polishing times using genetic algorithms. In Figure 3.14, curvatures after modifying are varied 15 to 20 times during 138 seconds of polishing time. These results indicate that the greedy algorithm and the GA with the modified fixtured curvature are more realistic and easier to control than the stochastic and pure genetic algorithm. In addition, the GA with the modified fixtured curvature offers a better wafer yield.

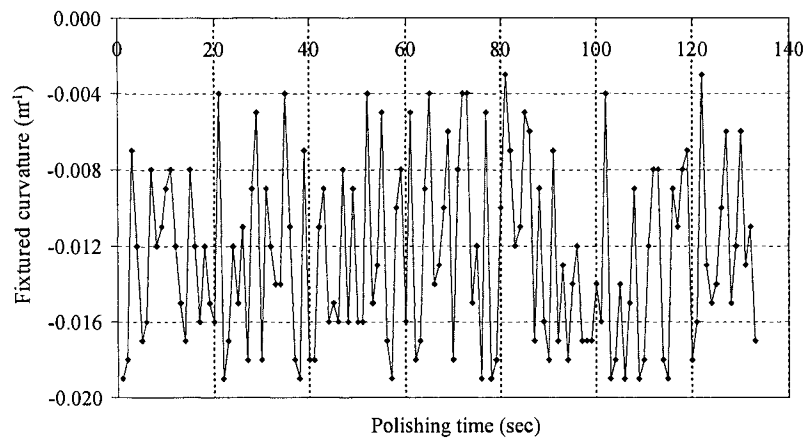
Even though the GA with modified fixtured curvature provides the most appropriate results, it is not yet the best wafer yield. In order to find the best possible wafer yield, the more effective method will be discussed.



(a) with initial low film curvature of -1e-05 m^{-1}

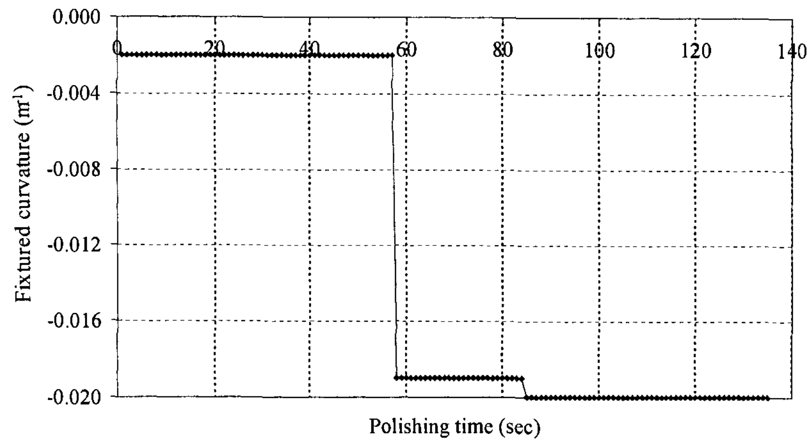


(b) with initial high film curvature of -2e-05 m^{-1}

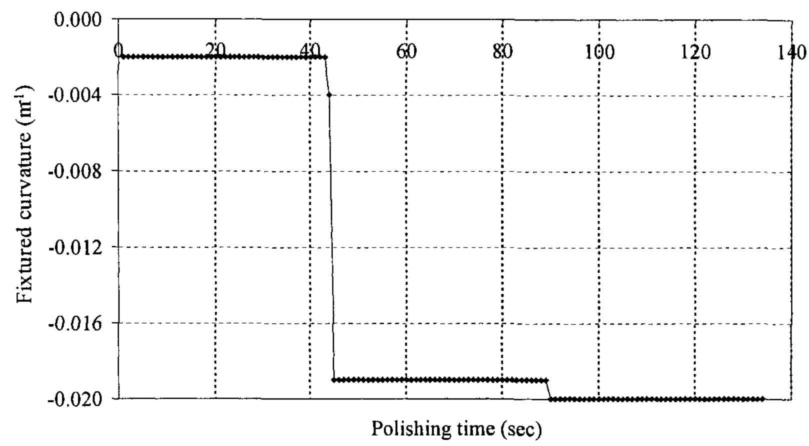


(c) with initial very high film curvature of -3e-05 m^{-1}

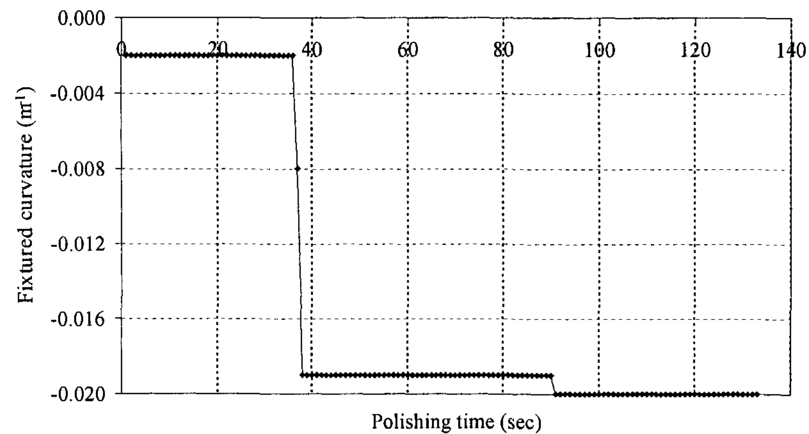
Figure 3.11. Fixtured curvature variation vs. polishing time in stochastic with three initial film curvatures.



(a) with initial low film curvature of -1e-05 m^{-1}

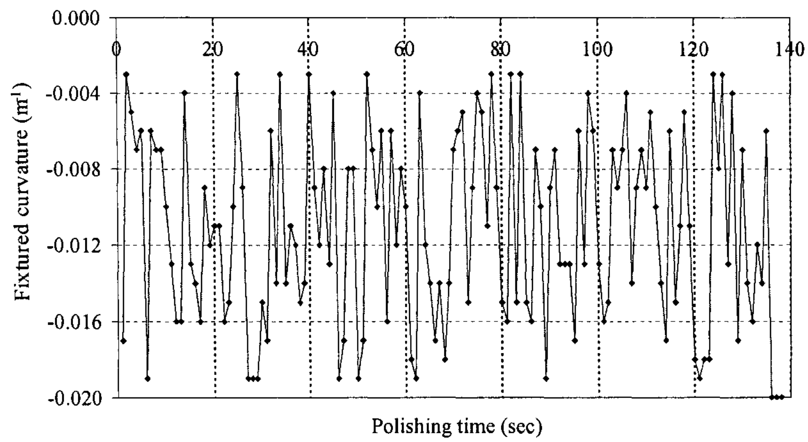


(b) with initial high film curvature of -2e-05 m^{-1}

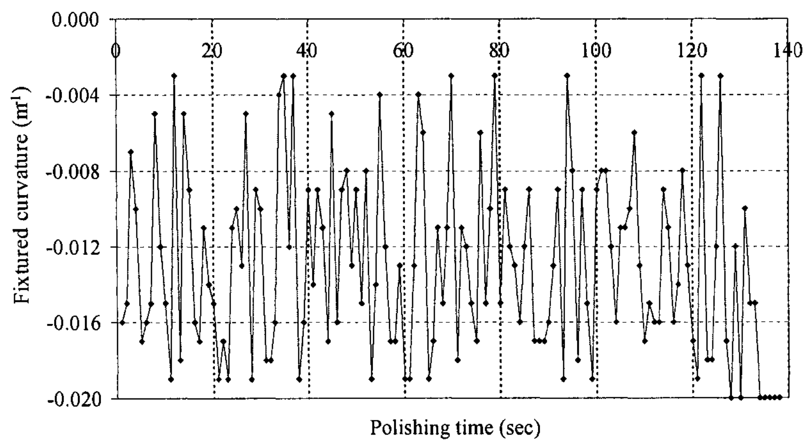


(c) with initial very high film curvature of -3e-05 m^{-1}

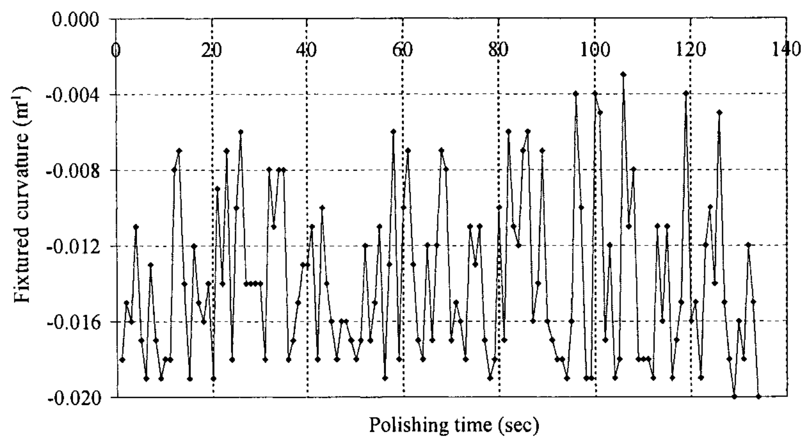
Figure 3.12. Fixtured curvature variation vs. polishing time in greedy algorithm with three initial film curvatures.



(a) with initial low film curvature of -1e-05 m^{-1}

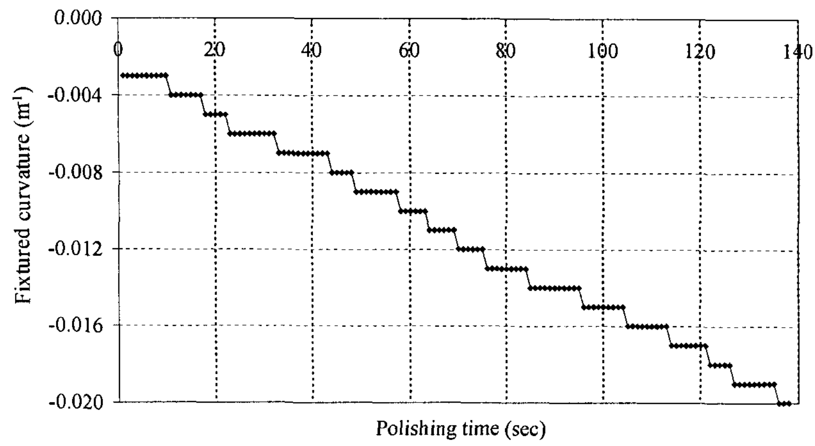


(b) with initial high film curvature of -2e-05 m^{-1}

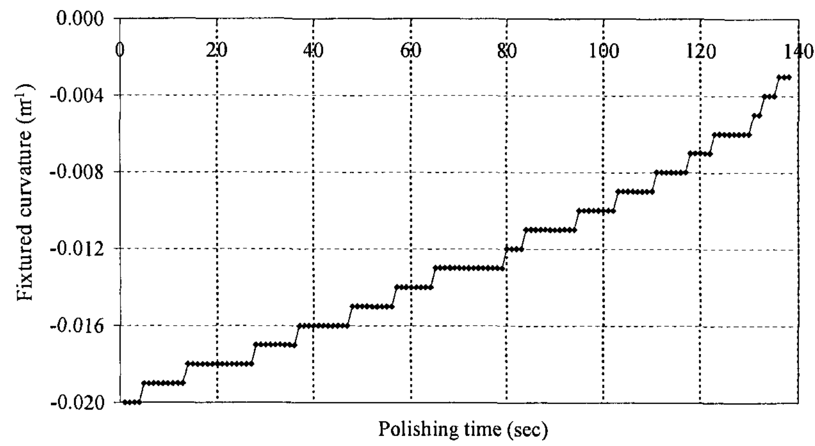


(c) with initial very high film curvature of -3e-05 m^{-1}

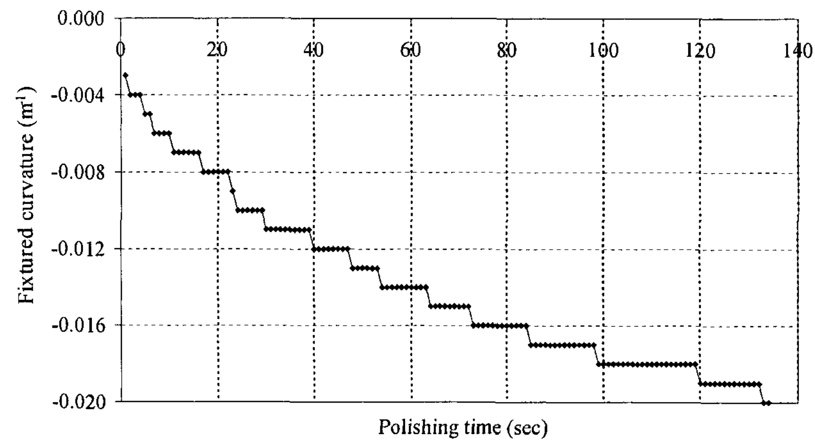
Figure 3.13. Fixtured curvature variation vs. polishing time in genetic algorithms with three initial film curvatures.



(a) with initial low film curvature of -1e-05 m^{-1}



(b) with initial high film curvature of -2e-05 m^{-1}



(c) with initial very high film curvature of -3e-05 m^{-1}

Figure 3.14. Modified fixtured curvature variation vs. polishing time in genetic algorithm with three initial film curvatures.

3.5. Non-Linear Programming for Yield Improvement

The wafer yields of the genetic algorithm are considerably improved with respect to the yields of the greedy algorithm but it is not the best solution for one time-dependent parameter. In this section, the optimized wafer yield for the wafer-scale model will be formulated and simulated. This is based on the objective function from Eamkajornsiri [5] using Lingo software.

3.5.1. Model with Non-linear Programming

Based on the switching logic from [4], the logic can be called the local optimization because the steepest slope curve is chosen at all times as it helps improve the wafer surface in the fastest possible manner at each instant. To achieve the global optimization, Lingo software is used to obtain the minimum objective function that provides us with a series of control parameters.

a. *Objective Function*

Similar to the objective function explained in 3.2.1.b, the combined moment of deviation (MOD) is comprised of the upper moment (UMD) and the lower moment (LMD) deviations. The upper moment is formalized by considering an amount of material available to be removed, whereas the lower moment implies the current curvature of the wafer. According to equation 10, this objective function can be written as:

$$MOD = \alpha \cdot (LMD) + (1 - \alpha) \cdot (UMD)$$

Where, α is a weighting term that lies in the interval [0, 1],

LMD is referred to in equation 8, and

UMD is referred to in equation 9.

From equation 8 and 9, the objective function can be rewritten as:

$$MOD = \alpha.\sum_{i=1}^{N}\{[h_C - F(r_i)]r_i\} + (1-\alpha).\sum_{i=1}^{N}\{[F(r_i) - h_A]r_i\}$$

Where, r_i is a distance from the center of the wafer,

h_C, h_A and $F(r_i)$ are referred to in Figure 3.3, and

N is number of points.

Since $F(r_i) = h_C + a_2^2 r^2$, substituting this $F(r_i)$ into the first term of the above equation, the MOD's equation may be expressed as:

$$MOD = -\alpha.\sum_{i=1}^{N}\left(a_2^2 r_i^3\right) + (1-\alpha).\left[\sum_{i=1}^{N}\{F(r_i).r_i\} - \sum_{i=1}^{N}(h_A.r_i)\right]$$

By multiplying $(1-\alpha)$ into the second term of the equation above, it may be expressed as:

$$MOD = -\alpha.\sum_{i=1}^{N}\left(a_2^2 r_i^3\right) + (1-\alpha).\sum_{i=1}^{N}\{F(r_i).r_i\} - (1-\alpha).\sum_{i=1}^{N}(h_A.r_i)$$

Furthermore, the equation shown above can be rewritten in terms of the current time step t as:

$$MOD(t) = -\alpha.\sum_{i=1}^{N}\left(a_{22}^t r_i^3\right) + (1-\alpha).\sum_{i=1}^{N}\{F(r_i,t).r_i\} - (1-\alpha).\sum_{i=1}^{N}(h_A.r_i) \qquad [11]$$

Where, the film curvature $\left(a_{22}^t\right)$ and the film profile $(F(r_i,t))$ are the functions of polishing time, t. The film profile at any time step may be expressed as:

$$F(r_i,t) = h_C(t) + a_{22}^t r_i^2$$

In addition, the height at the center of the wafer $(h_C(t))$ at time t can be calculated by subtracting the removed material from initial height at the center of the wafer (h_0). The amount of removed material depends on the pressure applied to the surface from the beginning of the polishing process until it reaches time t $\left(K_P KV \cdot \sum_{j=1}^{t} P_j\right)$. Thus, the film profile may be expressed as:

$$F(r_i, t) = h_O + a_{22}^t r_i^2 - K_P KV \cdot \sum_{j=1}^{t} P_j \qquad [12]$$

By substituting equation 12 into equation 11, we have

$$MOD = \begin{bmatrix} -\alpha . \sum_{i=1}^{N} \left(a_{22}^t r_i^3\right) + \\ (1-\alpha) . \sum_{i=1}^{N} \left\{ \left[h_O + a_{22}^t r_i^2 - K_P KV . \sum_{j=1}^{t} P_j \right] . r_i \right\} - (1-\alpha) . \sum_{i=1}^{N} (h_A r_i) \end{bmatrix} \qquad [13]$$

b. *Non-linear Programming Model*

To obtain the non-linear programming model, the objective function in equation 13 must be minimized. Thus, the model can be obtained as:

$$Min \quad MOD$$

$$Min \quad \begin{pmatrix} -\alpha . \sum_{i=1}^{N} \left(a_{22}^t r_i^3\right) + \\ (1-\alpha) . \sum_{i=1}^{N} \left\{ \left[h_O + a_{22}^t r_i^2 - K_P KV . \sum_{j=1}^{t} P_j \right] . r_i \right\} - (1-\alpha) . \sum_{i=1}^{N} (h_A r_i) \end{pmatrix}$$

According to this equation, summation of the second and third terms represents the upper moment of deviation (UMD). The UMD should be greater than or equal to zero, so the polishing process is stopped at the desired surface. In order to

maintain this condition, the absolute value of this summation is required. The non-linear programming model can be expressed as:

$$Min \begin{pmatrix} -\alpha.\sum_{i=1}^{N}\left(a_{22}^{t} r_i^3\right) + \\ Abs\left((1-\alpha).\sum_{i=1}^{N}\left\{\left[h_O + a_{22}^{t} r_i^2 - K_P KV.\sum_{j=1}^{t} P_j \right].r_i\right\} - (1-\alpha).\sum_{i=1}^{N}\left(h_A r_i\right)\right) \end{pmatrix} \quad [14]$$

Subject to $P_{Min} \leq P \leq P_{Max}$ for Pressure Control

Or

Subject to $a_2^{Min} \leq a_2 \leq a_2^{Max}$ for Curvature Control

Where, $a_{22}^{t} = \dfrac{\left(N.\sum_{i=1}^{N}\left\{Z^t(r_i).r_i^2\right\} - \sum_{i=1}^{N} Z^t(r_i).\sum_{i=1}^{N} r_i^2 \right)}{\left(N.\sum_{i=1}^{N} r_i^4 - \left[\sum_{i=1}^{N} r_i^2\right]^2 \right)}$ referred to Appendix B.

3.5.2. Simulation Results – Non-linear Programming (NLP)

In this example, the non-linear programming model with pressure control is applied and computed for wafer yields using Lingo software. In this software, there is a limitation of 800 nonlinear variables and 20 global variables that can be declared. To compare this non-linear programming model to the greedy algorithm from [4], some changes need to be made due to the limitation of the software. These changes include adjusting the parameters to obtain the best possible results.

Three values of initial film curvature a_2^2 (-5e-6 m^{-1}, -10e-6 m^{-1} and -15e-6 m^{-1}) are chosen. The parameters used in this example are the same as those in 3.3, except

Film thickness at the center is 10000 Å.

Tolerances are 450, 400, and 300 Å.

Desired final surface height is 8000 Å (after finishing CMP).

The range of interface pressure is from 2 to 6 psi.

The comparison of the wafer yield prediction among the greedy algorithm, the non-linear programming model, and the model with modified interface pressure for pressure control are shown in Table 3.4. The results of the greedy algorithm are run based on the parameters shown above with the pressure control. In the non-linear programming model for pressure control, the best wafer yields are chosen from three results using Lingo software. The result of the first run is computed based on the SLP directions strategy (Type I). To obtain the second result, Steepest Edge strategy (Type II) is applied. By running both strategies (Type III) together, the third result is computed. Furthermore, the series of interface pressures are sorted from non-linear programming results, sequentially from high pressure to low pressure, in order to make the CMP process easier to control. These sorted series are then used to calculate the wafer yield.

Table 3.4. Comparison of greedy algorithm and non-linear programming model for load control with wafer film thickness at the center of 10000 Å.

	Number of good sectors								
	Medium-Low Film Curvature			Low Film Curvature			Very Low Film Curvature		
Tolerance (Å)	450	400	350	450	400	350	450	400	350
Greedy Algorithm	144	120	112	164	152	144	188	184	164
NLP Type I	156	144	120	188	180	160	216	208	208
NPL Type II	156	148	120	192	180	164	208	208	192

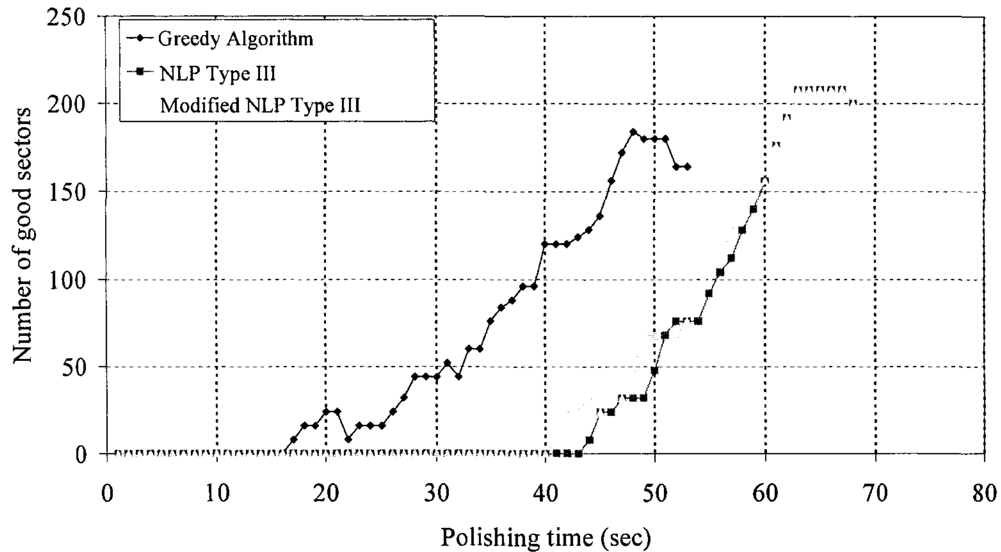
Table 3.4. (Continued).

	Number of good sectors								
	Medium-Low Film Curvature			Low Film Curvature			Very Low Film Curvature		
Tolerance (Å)	450	400	350	450	400	350	450	400	350
NLP Type III	156	148	120	188	180	160	216	208	208
Modified NLP	The wafer yield results are the same for all three types.								

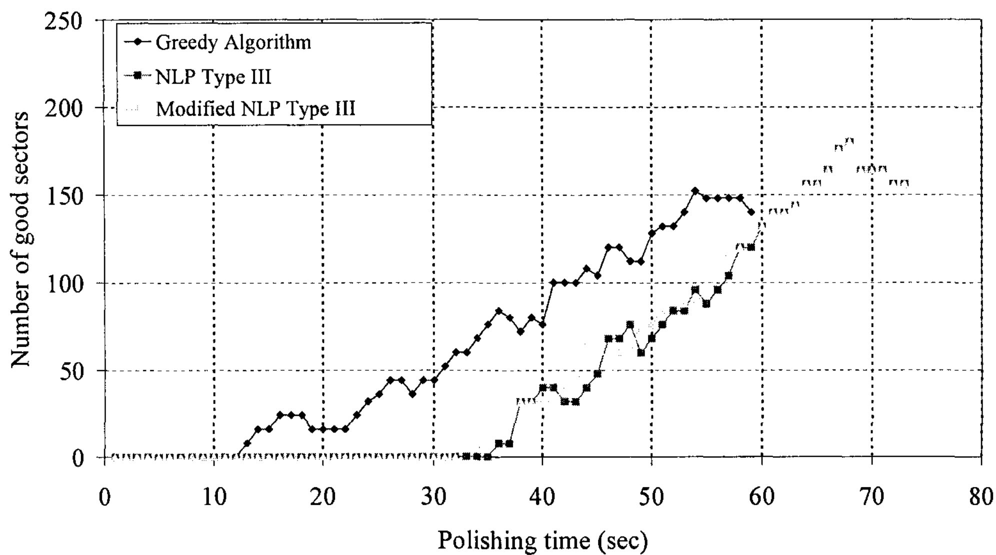
The CMP yields of three different approaches are shown in Figure 3.15. For initial very low film curvature of 5e-6 m^{-1}, the best yield of 208 occurs at 63 seconds for both the non-linear programming and modified non-linear programming, and the yield of 184 occurs at 48 seconds for greedy algorithm (Figure 3.15(a)). The results in Figure 3.15(b) show wafer yields for initial low film curvature of 10e-6 m^{-1}. This figure explains that the NLP or modified NLP give the same yield of 180 at 78 seconds, while greedy algorithm provide the yield of 152 at 54 seconds. For initial medium-low film curvature of 15e-6 m^{-1}, shown in Figure 3.15(c), wafer yield of 120 at 61 seconds is the result from the greedy algorithm. The better yields of 148 at 72 seconds are represented for both the NLP and the modified NLP.

Compared with the greedy algorithm, wafer yields from the non-linear programming models with pressure control are increased 7 to 23 percent for initial medium-low film curvature. An improvement of 11 to 18 percent is presented for initial low film curvature, and 11 to 27 percent for initial very low film curvature. Even though the wafer yields are improved in the NLP models, polishing times are also increased by

18 to 44 percent when comparing to the greedy algorithm. This increment in polishing time implies the longer planarization process to obtain the better wafer yield.

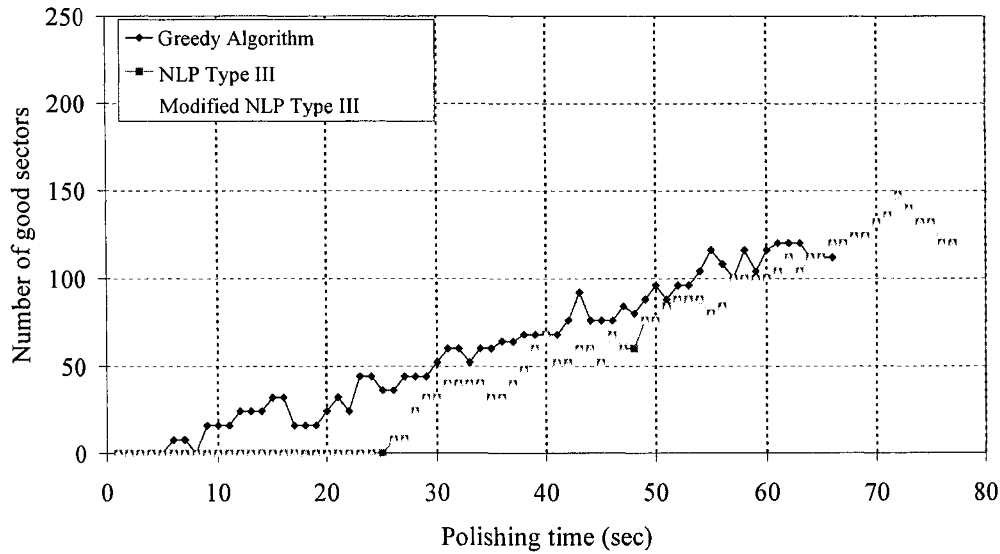


(a) with initial very low film curvature of -0.5e-05 m^{-1}



(b) with initial low film curvature of -1e-05 m^{-1}

Figure 3.15. Number of good sectors with polishing time for tolerance 400 Å.

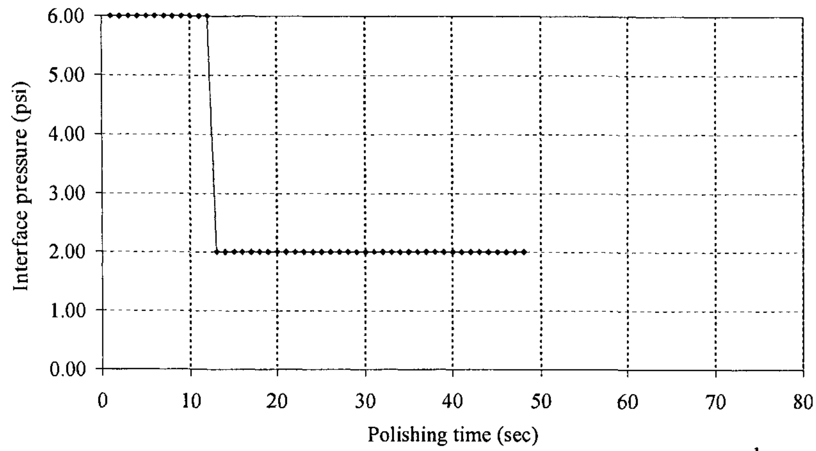


(c) with initial medium-low film curvature of -1.5e-05 m^{-1}

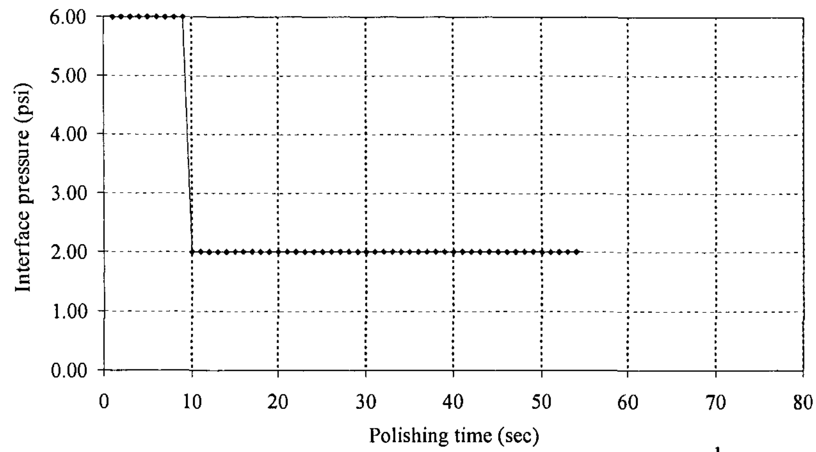
Figure 3.15. (Continued).

Moreover, Figures 3.16-3.18 illustrate the series of interface pressure controls in relation to polishing time using different approaches. These results indicate that the greedy algorithm and non-linear programming with modified interface pressure are more realistic than pure non-linear programming. In addition, non-linear programming with modified interface pressure offers a better wafer yield.

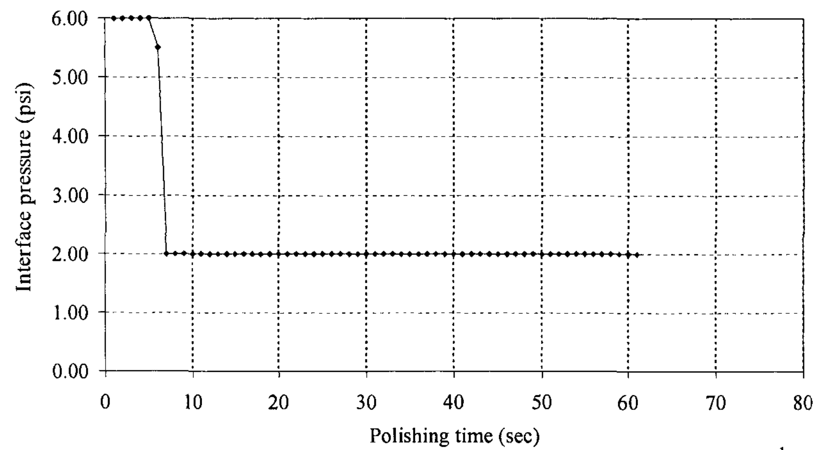
To this point, we have found the best possible wafer yield using the greedy algorithm, genetic algorithm, and non-linear programming model for wafer-scale. The interface pressure and/or fixtured curvature are controlled to manipulate the normalized pressure throughout the wafer. In the next chapter, we will start focusing on another scale called feature/die scale, which will show the improvement in upper surface uniformity.



(a) with initial very low film curvature of -0.5e-05 m^{-1}

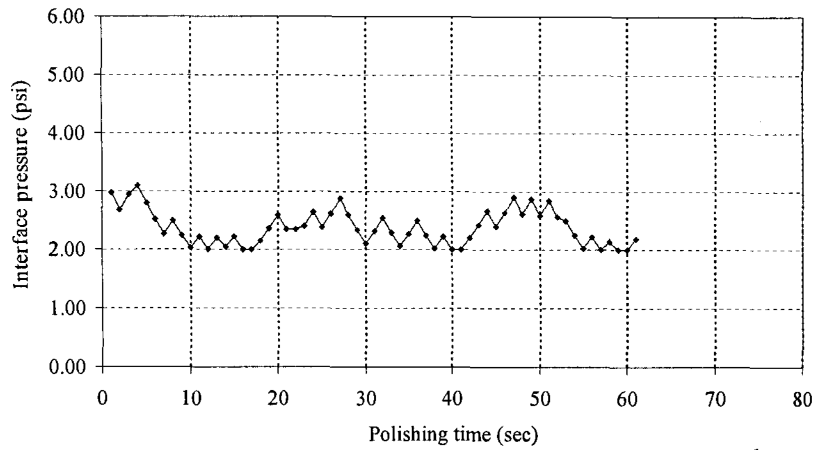


(b) with initial low film curvature of -1e-05 m^{-1}

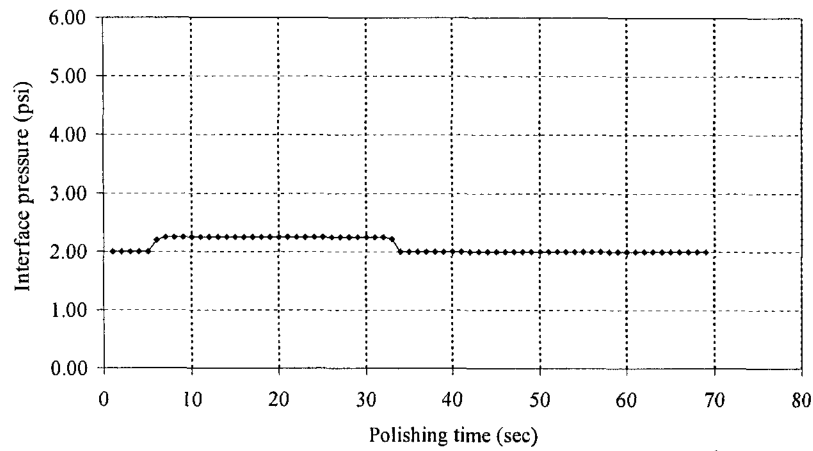


(c) with initial medium-low film curvature of -1.5e-05 m^{-1}

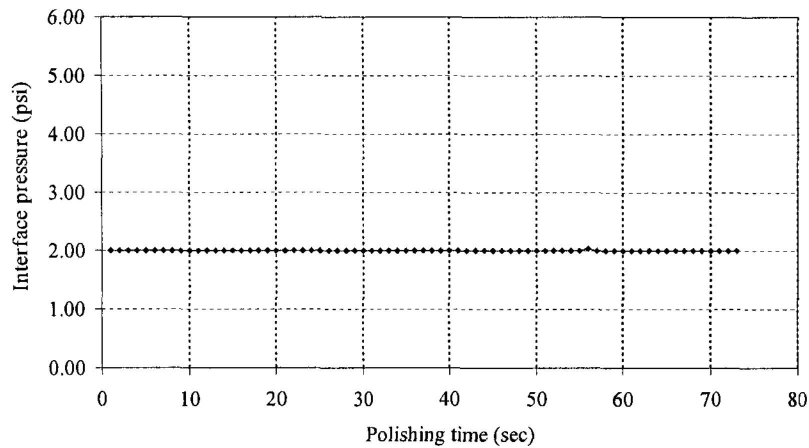
Figure 3.16. Interface pressure variation in greedy algorithm with pressure control for tolerance 400 Å.



(a) with initial very low film curvature of -0.5e-05 m^{-1}

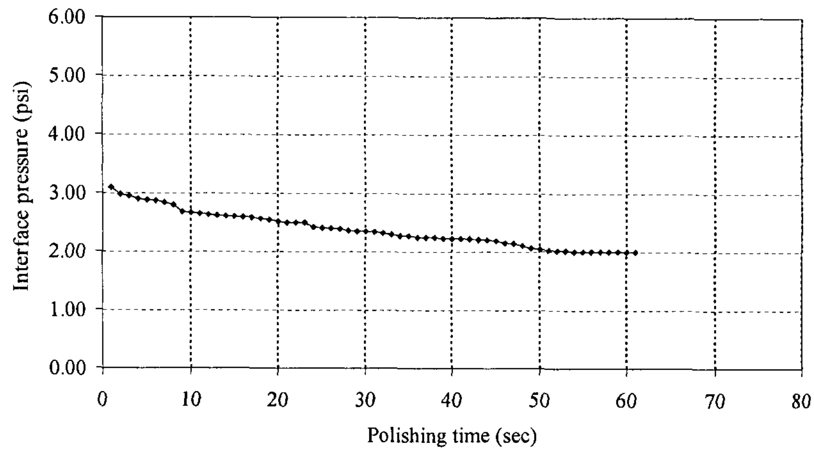


(b) with initial low film curvature of -1e-05 m^{-1}

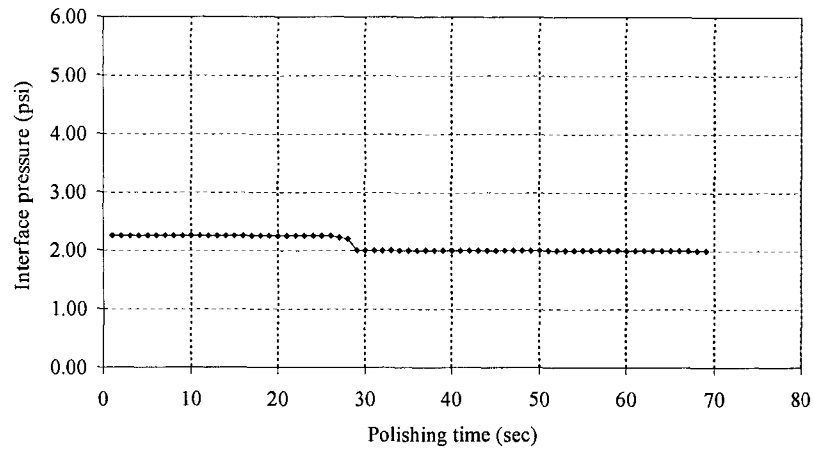


(c) with initial medium-low film curvature of -1.5e-05 m^{-1}

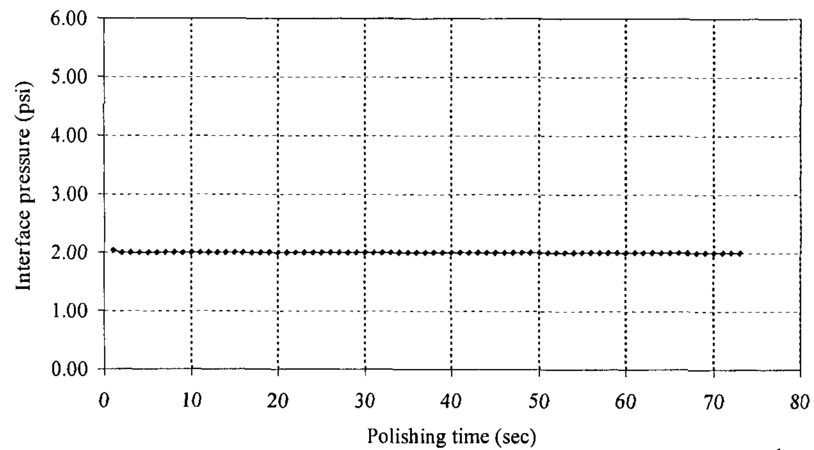
Figure 3.17. Interface pressure variation in non-linear programming model using lingo software with steepest edge strategy for tolerance 400 Å.



(a) with initial very low film curvature of -0.5e-05 m^{-1}



(b) with initial low film curvature of -1e-05 m^{-1}



(c) with initial medium-low film curvature of -1.5e-05 m^{-1}

Figure 3.18. Modified interface pressure variation in non-linear programming model using lingo software with steepest edge strategy for tolerance 400 Å.

CHAPTER 4. FEATURE / DIE-SCALE MODEL SIMULATION

This chapter discusses the feature / die-scale model development, the model validation, three control strategies based on the applied pressure for global surface improvement, and example simulation results. These control strategies are under the provisional patent, ISURF #03193. This simulation is developed based on the model from Fu and Chandra [9] and experimental data by Ouma et al. [16]. The Fu and Chandra [9], model predicted that at the feature-scale, high areas of wafer experience higher pressure than low areas. Due to this assumption, force redistribution is introduced with the fitting parameter α.

4.1. Feature / Die-Scale Model Development

Based on the understanding from the analytical step height model developed by Fu and Chandra [9], it is determined that the surface evolution mainly depends on the applied pressure and relative velocity between the pad and the wafer. The paper by Fu and Chandra has the following assumptions:

- The pad is assumed to deform like an elastic foundation.
- Force redistribution due to pad bending is proportional to the current step height.
- Wafer and pad are in contact at any point on the interface.

By following Preston's equation, this model (refer Figure 4.1) may be expressed as,

$$\frac{dY_{di}}{dt} = K_{di}\{k[H-(Y-Y_{di})]\}V$$

$$\frac{dY_{met}}{dt} = K_{met}\{k[H-(Y-Y_{met})]\}V .$$

And the feature-scale force equilibrium may be expressed as,

$$\{k[H-(Y-Y_{di})]\}\cdot[(b-1)\cdot 1]+\{k[H-(Y-Y_{met})]\}\cdot(a\cdot 1)=P\cdot(b\cdot 1).$$

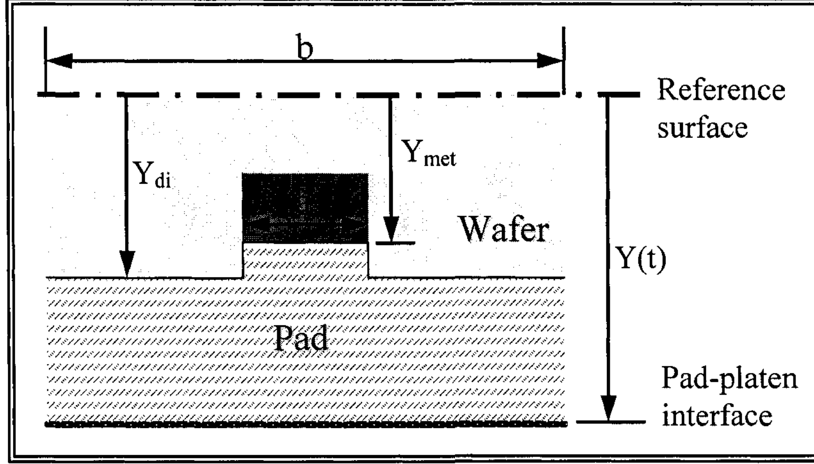


Figure 4.1. Schematic diagram of Fu and Chandra [9] step height reduction model.

Due to the pad bending, there is a force redistribution. The three modified governing equations including pad bending effects are

$$\frac{dY_{di}}{dt}=K_{di}\left\{k[H-(Y-Y_{di})]+\frac{\Delta F}{b-a}\right\}V$$

$$\frac{dY_{met}}{dt}=K_{met}\left\{k[H-(Y-Y_{met})]-\frac{\Delta F}{a}\right\}V$$

$$\{k[H-(Y-Y_{di})]\cdot[(b-a)\cdot 1]+\Delta F\}+\{k[H-(Y-Y_{met})]\cdot(a\cdot 1)-\Delta F\}=P\cdot(b\cdot 1).$$

From the second assumption, the force redistribution due to the pad bending can be expressed as

$$\Delta F=\alpha(Y_{di}-Y_{met}).$$

where α is the pad bending parameter. This is similar to planarization length (Ouma [15]). If the planarization length is zero, then the α value is zero as well.

By solving these differential equations, the metal region and dielectric region can be expressed as

$$Y_{met}(t)=\frac{\left\{K_{met}Vk\left(-\frac{P}{k}\right)+K_{met}Vk\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{di}(0)-K_{di}Vk\left(-\frac{a}{b}-\frac{\alpha}{kc}\right)Y_{met}(0)-\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}\right\}}{\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vk}$$

$$+\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}t$$

$$+\left(Y_{met}(0)-\frac{\left\{K_{met}Vk\left(-\frac{P}{k}\right)+K_{met}Vk\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{di}(0)-K_{di}Vk\left(-\frac{a}{b}-\frac{\alpha}{kc}\right)Y_{met}(0)-\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}\right\}}{\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vk}\right)$$

$$\cdot\exp\left\{-\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vkt\right\}$$

$$Y_{di}(t)=\frac{\left\{K_{di}Vk\left(-\frac{P}{k}\right)+K_{di}Vk\left(-\frac{a}{b}-\frac{\alpha}{kc}\right)Y_{met}(0)-K_{met}Vk\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{di}(0)-\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}\right\}}{\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vk}$$

$$+\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}t$$

$$+\left(Y_{di}(0)-\frac{\left\{K_{di}Vk\left(-\frac{P}{k}\right)+K_{di}Vk\left(-\frac{a}{b}-\frac{\alpha}{kc}\right)Y_{met}(0)-K_{met}Vk\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{di}(0)-\frac{K_{di}K_{met}(Vk)\left(-\frac{P}{k}\right)\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]}{K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)}\right\}}{\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vk}\right)$$

$$\cdot\exp\left\{-\left[K_{di}\left(\frac{a}{b}+\frac{\alpha}{kc}\right)+K_{met}\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\right]Vkt\right\}.$$

With these two equations, the local dishing model as a function of time can be written as

$$D(0) = Y_{di}(0) - Y_{met}(0)$$

$$D(t) = \frac{K_{met} - K_{di}}{K_{di}\left(\frac{a}{b} + \frac{\alpha}{kc}\right) + K_{met}\left(\frac{c}{b} + \frac{\alpha}{ka}\right)} \cdot \frac{P}{k} + \left\{ [Y_{di}(0) - Y_{met}(0)] - \frac{K_{met} - K_{di}}{K_{di}\left(\frac{a}{b} + \frac{\alpha}{kc}\right) + K_{met}\left(\frac{c}{b} + \frac{\alpha}{ka}\right)} \cdot \frac{P}{k} \right\} \cdot \exp\left\{ -\left[K_{di}\left(\frac{a}{b} + \frac{\alpha}{kc}\right) + K_{met}\left(\frac{c}{b} + \frac{\alpha}{ka}\right)\right] Vkt \right\}.$$

Where, $D(t)$ is the height difference between the dielectric surface and the metal surface at any polishing time, t.

According to Fu and Chandra [9], the metal region, dielectric region and local dishing equations stated above are classified into three cases.

Case 1: When $Y_{met}(0) = Y_{di}(0)$ and $K_{met} = K_{di} = K$; this situation can be explained as involving a homogeneous film with an absolutely flat initial surface. This indicates that the height of metal region is equal to the height of dielectric region, and a nonselective slurry or a slurry with selectivity of unity is used.

Case 2: When $Y_{met}(0) = Y_{di}(0)$ and $K_{met} \neq K_{di}$; this is a typical scenario in a copper polishing process, where the metal region and dielectric region are flushed against each other, and a selective slurry is being used. Thus, in this case, the initial surface is started with the flat surface and a slurry with selectivity of unity.

Case 3: When $Y_{met}(0) \neq Y_{di}(0)$ and $K_{met} = K_{di} = K$; this situation represents planarization of a patterned wafer with a nonselective slurry. It also represents the case where there is only one material (dielectric or metal) on the surface.

The model discussed in this chapter focuses on the step height on one type of surface, which is relevant to the third case according to Fu and Chandra [9]. Figure 4.2 shows the schematic diagram of this planarization model. Since the model assumes that there is only one material on the surface, the Preston's constant K_{met} and K_{di} will be noted as K. In addition, the notation of dielectric surface (Y_{di}) and metal surface (Y_{met}) will be changed to Y_{upper} and Y_{lower} respectively. By substituting these variables, an equation for the step height as a function of time can be expressed as

$$D(t) = \left[Y_{upper}(0) - Y_{lower}(0)\right] \exp\left\{-K\left[1 + \frac{\alpha}{k}\left(\frac{1}{a} + \frac{1}{c}\right)\right]Vkt\right\}. \qquad [15]$$

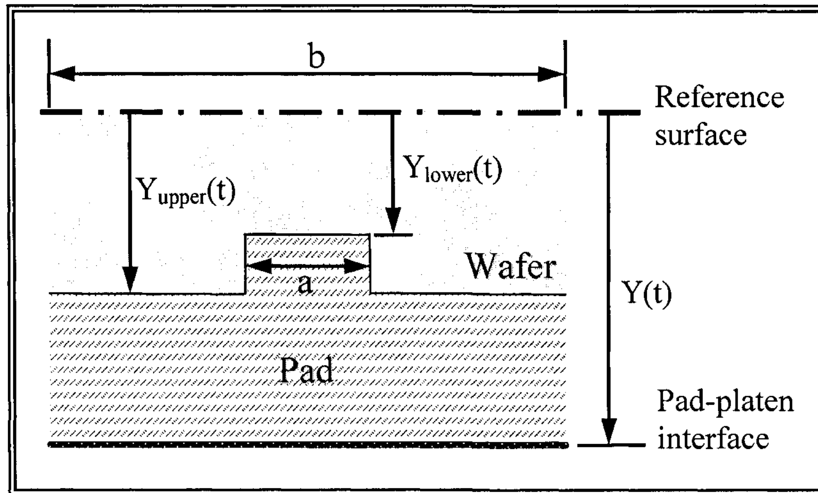


Figure 4.2. Schematic diagram of planarization model.

The heights of the upper surface and lower surface after polishing for a given period of polishing time t are expressed as

$$Y_{upper}(t)=\frac{\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{upper}(0)+\left(\frac{a}{b}+\frac{\alpha}{kc}\right)Y_{lower}(0)}{1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)}-KPVt$$
$$+\frac{\frac{a}{b}+\frac{\alpha}{kc}}{1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)}\left[Y_{upper}(0)-Y_{lower}(0)\right]\exp\left\{-K\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]Vkt\right\} \quad [16]$$

$$Y_{lower}(t)=\frac{\left(\frac{c}{b}+\frac{\alpha}{ka}\right)Y_{upper}(0)+\left(\frac{a}{b}+\frac{\alpha}{kc}\right)Y_{lower}(0)}{1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)}-KPVt$$
$$+\frac{\frac{c}{b}+\frac{\alpha}{ka}}{1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)}\left[Y_{lower}(0)-Y_{upper}(0)\right]\exp\left\{-K\left[1+\frac{\alpha}{k}\left(\frac{1}{a}+\frac{1}{c}\right)\right]Vkt\right\}. \quad [17]$$

And the removal rate equations are

$$\frac{dY_{upper}}{dt}=KVk\left[\left(-\frac{a}{b}-\frac{\alpha}{kc}\right)\left(Y_{upper}-Y_{lower}\right)-\frac{P}{k}\right] \quad [18]$$

$$\frac{dY_{lower}}{dt}=KVk\left[\left(\frac{c}{b}+\frac{\alpha}{ka}\right)\left(Y_{upper}-Y_{lower}\right)-\frac{P}{k}\right] \quad [19]$$

The equations 16 and 17 are terminal equations, meaning the heights of the surface are the final heights after polishing for a given period of time, t. On the other hand, equations 18 and 19 are intermediate equations, meaning the removal rate changes for every time step, dt.

4.2. Model Validation

To compare model predictions against experimental observations, the data from Ouma et al. [16] is used. Their test wafer consists of 1 μm TEOS (Tetra-Ethyl-Ortho-Silicate) oxide deposition, metal deposition and patterning using the characteristic mask, and the final deposition of a 2.5 μm TEOS layer. The characterization mask (refer Figure 4.3) is constructed of four sets of structures with constant pitch but different in densities across the die or in the opposite way, where density is constant and pitches are varied. Each density structure in Line 3, Line 4 and Line 5 consist of vertical lines and spaces of 100 μm pitch, in opposition, Line 6 has a constant density of 0.5 with different pitch and linewidth. The size of the mask is 20 x 20 mm and it is divided into 4 x 4 mm blocks. Their experimental data was obtained by polishing two sets of three wafers at the polishing time of 29 seconds and 88 seconds. The wafers were polished with a single head of a five-head SpeedFarm polisher to eliminate head-to-head variability. Line 3, Line 4, Line 5 and Line 6 of their experimental data are used to verify our simulation model (refer Appendix D).

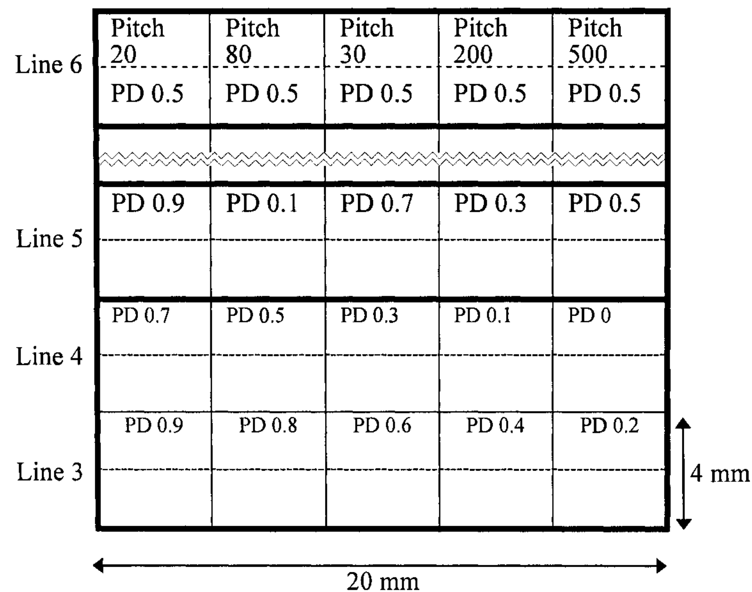


Figure 4.3. Part of the layout of the characterization mask.

To compare between the model prediction and the experimental data, it is necessary that a parameter α is identified a priori. In doing this, a set of experimental data in Line 5, is used as a reference. This set of data is constructed with five different pattern densities of 0.9, 0.8, 0.6, 0.4 and 0.2, respectively. In order to obtain an estimation of parameter α, the least root mean square (RMS) error is calculated by fitting model predictions to data from Line 5. After finding the best fit for the parameter α of 5.0508e6 N/m^2, the model predictions of other cases are simulated using the parameters shown in Table 4.1.

Table 4.1. List of variables and notations used.

Variable and Notation	Value	Unit	
Pressure (P)	3.4474e+04	N/m^2	
Relative Velocity (V)	0.65	m/s	
Bending factor (α)	5.0508e+06	N/m^2	
Preston's constant (K)	2.1075e-13	N/m^2	
Pad stiffness (k)	5.2130e+09	N/m^3	
Pattern density (PD)	a/b		for Line 3, 4, and 5
Linewidth (a)	from exp. data	μm	for Line 3, 4, and 5
Pitch (b)	100	μm	
Pattern density (PD)	0.5		for Line 6 only
Linewidth (a)	$b \cdot PD$	μm	for Line 6 only
Pitch (b)	from exp. data	μm	

According to Table 4.2 shown below, the root means square errors are calculated for each line, with two different regions, upper and lower surfaces, and polishing times of 29 and 88 seconds. At 29 seconds, the average RMS prediction error for the upper surface areas is 582 Å and the prediction error of the lower surface areas is 305 Å. At 88 seconds, the average errors of 518 Å and 308 Å are presented for upper and lower surface areas respectively.

Table 4.2. Model prediction results for validation.

	RMS Prediction Error (Å)			
	Upper surface area		Lower surface area	
Polishing Time	29 sec	88 sec	29 sec	88 sec
Line 3	780	417	397	298
Line 4	466	689	175	260
Line 5	499	449	342	365
Line 6	RMS of this line is not computed			
Average	582	518	305	308

4.3. Global Planarization Control Algorithm

As seen from the previous section, even with the initial uniform surface, the surfaces after polishing 88 seconds are shown non-uniformities across the die in all three different regions. In order to obtain uniformity over the pattern dependencies in a die-scale model, three control strategies are explored. The first strategy controls the interface pressure in

space called spatial pressure control. The second and third strategies control the interface pressure both in space and in time. The second strategy is called spatial and temporal pressure control. The last strategy is called look-ahead scheduled pressure control. A 2D simulation is developed using visual basic software to track current step heights for every time step. Each of these three strategies is discussed in detail below. The spatial pressure control strategy is explained first.

4.3.1. Objective Function

To obtain uniformity over the pattern dependencies in a die-scale model, the objective function can be formulated by considering the deviation from the target surface and it can be written as:

$$\text{Objective Function} = \sum_{i=1}^{n} \left(Y_{upper}(i) - Y_{desired} \right)$$

Where, n is number of zones.

This objective function will be applied through out these three control strategies.

4.3.2. Model of Global Planarization Control

The model of global planarization is expressed as

$$Min \quad \left(\sum_{i=1}^{n} \left(Y_{upper}(i) - Y_{desired} \right) \right)$$

Subject to $\quad P_{Min} \leq P_i \leq P_{Max} \quad$, for i = 1 to n (# of zones)

a. *Spatial Pressure Control: Algorithm*

The principle idea behind this pressure control is to planarize the upper surface of each zone, with different initial surface topography, down to a specific

target surface at the best possible time. In order to achieve this goal, maximum pressure capability for a specific CMP machine will be applied to calculate the polishing time needed for each zone. This process allows us to specify time required to planarize every zone down to the same level. Applied interface pressures will then be calculated based on specified time in the earlier process. To achieve the specific target surface, the calculated pressures will be applied simultaneously throughout the entire period of polishing time. This strategy is calculated using the following algorithm for each of n zones. The general structure of this algorithm (refer Appendix F.1 for flowchart) is shown in the following steps:

Step 1: Calculated Total Material

Calculate *the total material* (Mat_Total) to be removed in all zones together. This step and the next step are used together to find when the polishing process will finish.

$$Mat_Total = \sum_{i=1}^{n} \left(Y_{upper}(i) - Y_{desired} \right)$$

Where, $Y_{desired}$ is the desired height, and

Y_{upper} is the initial upper surface height.

Step 2: Calculated Time Needed

Calculate *the polishing time needed* for each zone (T_{zone}) to reach the target or desired surface with the maximum interface pressure (the maximum pressure that the user would like to apply) using equation 16 by following Newton-Raphson method. Where $f(t)$ is $Y_{upper}(t)$ in equation 16.

$t_{i+1} = t_i - f(t)/f'(t)$, until $t_{i+1} - t_i < 1e-8$

Step 3: Calculated Applied Pressure

Compare the polishing time for all n zones and find *the maximum polishing time needed* to have all applied interface pressure values of all zones to be less than or equal to maximum pressure.

$T_{\max} = Max(T_{zone})$, for zone = 1 to n

With polishing time as the $T_{\max}$, calculate *the applied interface pressure* for each zone by using equation 16

Step 4: Calculate Step Height and Check

4.1: Calculate Step Height

Calculate *the new upper and lower surface* of each zone using the removal rate equation 18 $\left(Y'_{upper}(t)\right)$ and 19 $\left(Y'_{lower}(t)\right)$, respectively for the upper and lower surfaces.

$Y_{upper}(i)^{new} = Y_{upper}(i)^{old} - Y'_{upper}(i).\Delta t$, for $i = 1$ to n (# of zones)

$Y_{lower}(i)^{new} = Y_{lower}(i)^{old} - Y'_{lower}(i).\Delta t$, for $i = 1$ to n (# of zones)

Where, $\Delta t = 0.1$ seconds.

4.2: Check

Compare the total material left with the previous step until it reaches the least total material left. If it is not, go back to step 4.1, and continue polishing and recalculate the new upper and lower surface.

Step 5: Calculate Error

The errors of upper surface of each zone are calculated as

$$Error_{upper} = \left(Y_{upper}(final)_i - Y_{desired}\right) / \left(Y_{upper}(0)_i - Y_{desireed}\right) x100$$

b. Spatial and Temporal Pressure Control: Algorithm

This control is divided into two phases. In the first phase, the surface is polished using low interface pressure for controlling the local step height. By using this low pressure, only the upper surface is polished, while the lower surface remains the same. After the height difference between upper and lower surface reaches its limitation point, depending on the surface topography and the pad properties, this phase will no longer exist. In order to control the global step height, the second phase is presented. The applied interface pressures are calculated using spatial pressure control for each of n zones based on the present upper surface evolution from the previous phase. The following algorithm explains the procedure for spatial and temporal pressure control (refer Appendix F.2 for flowchart).

Step 1: Calculate Minimum Step Height

From the machine specifications, the minimum interface pressure capability is specified. With this pressure, *the smallest step height* achievable (the only upper surface is polished) for each zone $\left(SH_i^{\min}\right)$ is calculated using equation 19.

Find $SH_i^{\min}$, for $i = 1$ to n (# of zones)

such that $Y'_{lower}\left(P_{\min}\right) = 0$

Where, $P_{\min}$ is the minimum pressure capability for a specific CMP machine.

Step 2: Calculate Max Pressure

With the respective step heights of each zone, *the maximum pressure* that can be applied is calculated for each zone $\left(P_i^{\max}\right)$ using equation 19.

Find $P_i^{\max}$, for $i = 1$ to n (# of zones)

such that $Y'_{lower}(SH_i) = 0$

Where, SH_i is the present step height of i-th zone.

Step 3: Calculate Interface Pressure for each zone of the First Phase

Calculate material removal rate on the upper surface of each zone $\overline{Y'_{upper_i}}$ with $P_i^{\max}$ from step 2 using equation 18.

Calculate the material need to be removed of each zone (Mat_i) by setting

$$Mat_i = SH_i - Max\left(SH_i^{\min}\right)$$

Calculate the ratio (R_i) by setting

$$R_i = \frac{Mat_i}{\overline{Y'_{upper_i}}}$$

Assuming the relation between step height and time to be linear, calculate the material removal rate on the upper surface

$$Y'_{upper_i} = \frac{Mat_i}{Max(R_i)}$$

Calculate *the interface pressure* (P_i) using equation 18 and material removal rate on the lower surface $\left(Y'_{lower}(t)\right)$ using equation 19.

Step 4: Polish the surface and Check

Calculate *the new upper and lower surface* of each zone

$$Y_{upper}(i)^{new} = Y_{upper}(i)^{old} - Y'_{upper_i}.\Delta t \qquad \text{, for } i = 1 \text{ to } n \text{ (\# of zones)}$$

$$Y_{lower}(i)^{new} = Y_{lower}(i)^{old} - Y'_{lower\,i}.\Delta t \qquad \text{, for } i = 1 \text{ to } n \text{ (\# of zones)}$$

Where, $\Delta t = 1$ sec.

Step 5: *Check*

Repeat step 2 to step 4 until the following condition is satisfied. This condition helps determine whether the surface has reached the least step height $\left(SH_i^{\min}\right)$

$$\exists\left(Mat_i < Max\left(SH_i^{\min}\right)\right) \qquad \text{, for } i = 1 \text{ to } n \text{ (\# of zones)}$$

Step 6: Calculate Interface Pressure for each zone of the Second Phase

After reaching the stipulated step height, the spatial pressure control is applied to attain the target surface.

Step 7: Calculate Error

Same calculation as in *the spatial pressure control*

c. *Look-ahead Scheduled Pressure Control: Algorithm*

There are two phases in this control. The first phase is similar to spatial and temporal pressure control, where the surface is polished using low interface pressure to control the local step height. In addition, this control strategy views the step height ahead of the process and modifies the pressure based on the desired step height with indicated time. Similarly, after the height difference between upper and lower surface reaches its limitation point, the first phase will no longer exist. The second phase is then applied in order to obtain the global step height. With this process, the interface

pressures are calculated using spatial pressure control for each of n zones based on the present upper surface evolution from the previous phase. The following algorithm explains the procedure for the look-ahead scheduled pressure control (refer Appendix F.3 for flowchart).

Step 1: Calculate Minimum Step Height

Same calculation as in *the spatial and temporal pressure control*

Step 2: Calculate Max Pressure

Same calculation as in *the spatial and temporal pressure control*

Step 3: Calculate Interface Pressure for each zone of the First Phase

Calculate the material need to be removed from each zone (Mat_i) by setting

$$Mat_i = SH_i - Max(SH_i^{\min})$$

With $P_{\min}$ and $P_i^{\max}$ as the inputs for each zone, the minimum possible step height left $(MSH_i^{\min})$ is identified in each zone after a specific period of time using *look-ahead procedure* (refer Appendix E for procedure).

$$MSH_i^{\min} = Look_ahead(t_{step}, P_{\min}, P_i^{\max}) \quad \text{, for } i = 1 \text{ to } n \text{ (\# of zones)}$$

Where, t_{step} is the specified time step by user.

The step height needed to be removed is calculated after the specified time step using the following equation:

$$RSH_i = SH_i - MSH_i^{\min}$$

The ratio is calculated as follows:

$$R_i = \frac{Mat_i}{RSH_i}$$

Calculate the material to be removed from each zone, based on the zonal ratio, which should be occurred by setting

$$LSH_i = \frac{Mat_i}{Max(R_i)}$$

Find *the interface pressure* of each zone using the *look-ahead procedure* (refer Appendix E for procedure) for MSH_i^{actual} (the step height to be left after the prescribed time step)

$$MSH_i^{actual} = SH_i - LSH_i$$

$$P_i = Look_ahead(t_{step}, P_{\min}, P_i^{\max}, MSH_i) \quad \text{, for } i = 1 \text{ to } n \text{ (\# of zones)}$$

Step 4: Polish the surface

Calculate the material removal rate of the upper surface and lower surface using equations 18 and 19. The polishing is carried out on the wafer surface.

Step 5: Check

Same check as in *the spatial and temporal pressure control*

Step 6: Calculate Interface Pressure for each zone of Second Phase

Same calculation as in *the spatial and temporal pressure control*

Step 7: Calculate Error

Same calculation as in *the spatial pressure control*

4.4. Simulation Results

In order to aid the understanding of the control algorithms shown above, the simulation example based on the experimental data from Ouma et al. [16] is provided. According to this example shown in Figure 4.4, initial upper and lower surfaces are set with different pattern densities, at 2500 nm and 1900 nm, respectively. The pattern densities of each line in each zone are set based on the MIT mask shown in Figure 4.3. The first simulation is run by applying uniform pressure of 5 psi (3.4474e+04 N/m^2) across the die to polish for 88 seconds. Unlike the first simulation, the other two control algorithms vary pressure spatially and/or temporally across the die. These two algorithms enable us to obtain the prediction values of the final upper surface, which is an average value of 1714 nm, at approximately 88 seconds of polishing time, according to the experimental data.

Table 4.3 shows the simulation data of the final polished surface, including polishing time, upper surface percent error, and standard deviation (*Stdev*). The following percent error equation allows us to evaluate the accuracy of the upper surface prediction from the three algorithms explained previously. *Stdev* presented in Table 4.3 represents standard deviation of the upper surface areas or of step heights in each zone.

$$\%Error_{upper} = \sum_{i=1}^{n} \left[\left(Y_{upper}(final)_i - Y_{desired} \right) / \left(Y_{upper}(0)_i - Y_{desireed} \right) \right] x100 .$$

The objective of this model is to polish the surface using various pattern densities to control the uniformity of final upper surface. The simulation results from three different regions evidently show that there is a significant improvement in the uniformity of the upper surface when pressure across the die is controlled spatially (spatial pressure control). Moreover, further improvement of the uniformity of the upper surface is presented when

pressure is varied in time besides being controlled spatially (spatial and temporal pressure control / look-ahead scheduled pressure control). These similar results are obtained for all three regions, including Line 3, 4, and 5. At the end of this chapter, graphs of series of surface evolution, applied pressure, and material removal rate of Line 3 are shown in Figures 4.5 to 4.10.

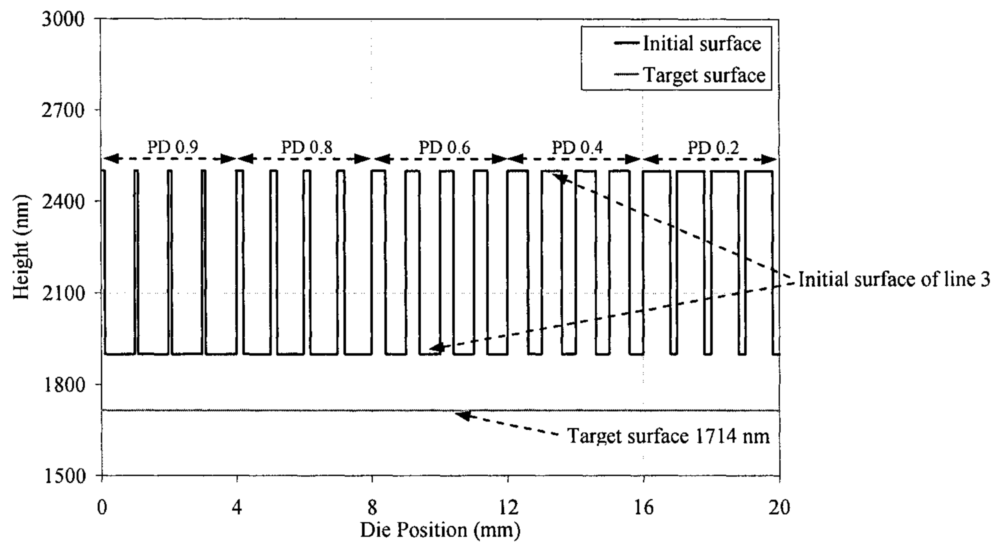


Figure 4.4. Initial surface profile of line 3 in 5 different zones.

Table 4.3. The simulation data of final polished surface for the oxide CMP characterization mask of Ouma et al. [16].

Zone	*No control*			*Spatial pressure control*			*Spatial and temporal pressure control*			*Look-ahead scheduled pressure control*		
	Final Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)
For Line 3												
1	1545	1544	1	1714	1714	0	1714	1713	1	1714	1713	1
2	1613	1601	12	1716	1702	14	1714	1695	19	1714	1696	18
3	1741	1696	45	1720	1672	48	1714	1660	54	1714	1663	51
4	1828	1773	55	1713	1663	50	1714	1660	54	1714	1663	51
5	1882	1843	39	1686	1664	22	1713	1688	25	1714	1688	26
Time		88 seconds			86.4 seconds			84.9 seconds			86.6 seconds	
%Error	-	-	-	0.94	w/ Max 8 psi		0.03	w/ Max 8 psi		~0	w/ Max 8 psi	
Stdev	141.8	-	22.9	13.57	-	21.75	0.45	-	23.11	~0	-	21.67
For Line 4												
1	1682	1651	31	1872	1837	35	1871	1828	43	1870	1826	44
2	1789	1736	53	1885	1825	60	1871	1824	47	1871	1821	50
3	1858	1810	48	1876	1824	52	1871	1826	45	1870	1823	47
4	1900	1879	21	1841	1828	13	1871	1839	32	1870	1836	34
5	1914	1900	14	1817	1817	0	1871	1870	1	1870	1869	1
Time		88 seconds			87.8 seconds			89.3 seconds			87.6 seconds	
%Error	-	-	-	3.28	w/ Max 6.5 psi		0.16	w/ Max 12 psi		0.29	w/ Max 12.5 psi	
Stdev	95.3	-	16.8	28.37	-	25.39	~0	-	19.13	0.45	-	20.04

Table 4.3. (Continued)

Zone	No control			Spatial pressure control			Spatial and temporal pressure control			Look-ahead scheduled pressure control		
	Final Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)	*Final* Y_{upper} (nm)	*Final* Y_{lower} (nm)	*Final* *SH* (nm)
For Line 5												
1	1545	1544	1	1766	1766	0	1765	1764	1	1765	1764	1
2	1900	1879	21	1724	1719	5	1765	1749	16	1758	1751	7
3	1682	1651	31	1775	1745	30	1765	1725	40	1765	1725	40
4	1858	1810	48	1764	1725	39	1765	1723	42	1765	1722	43
5	1789	1736	53	1778	1727	51	1765	1709	56	1765	1710	55
Time	88 seconds			89.2 seconds			88.7 seconds			89.1 seconds		
%Error	-	-	-	1.80	w/ Max 7.5 psi		~0	w/ Max 7.2 psi		0.19	w/ Max 7.2 psi	
Stdev	143.4	-	21.1	21.72	-	21.92	~0	-	22.09	3.13	-	23.77

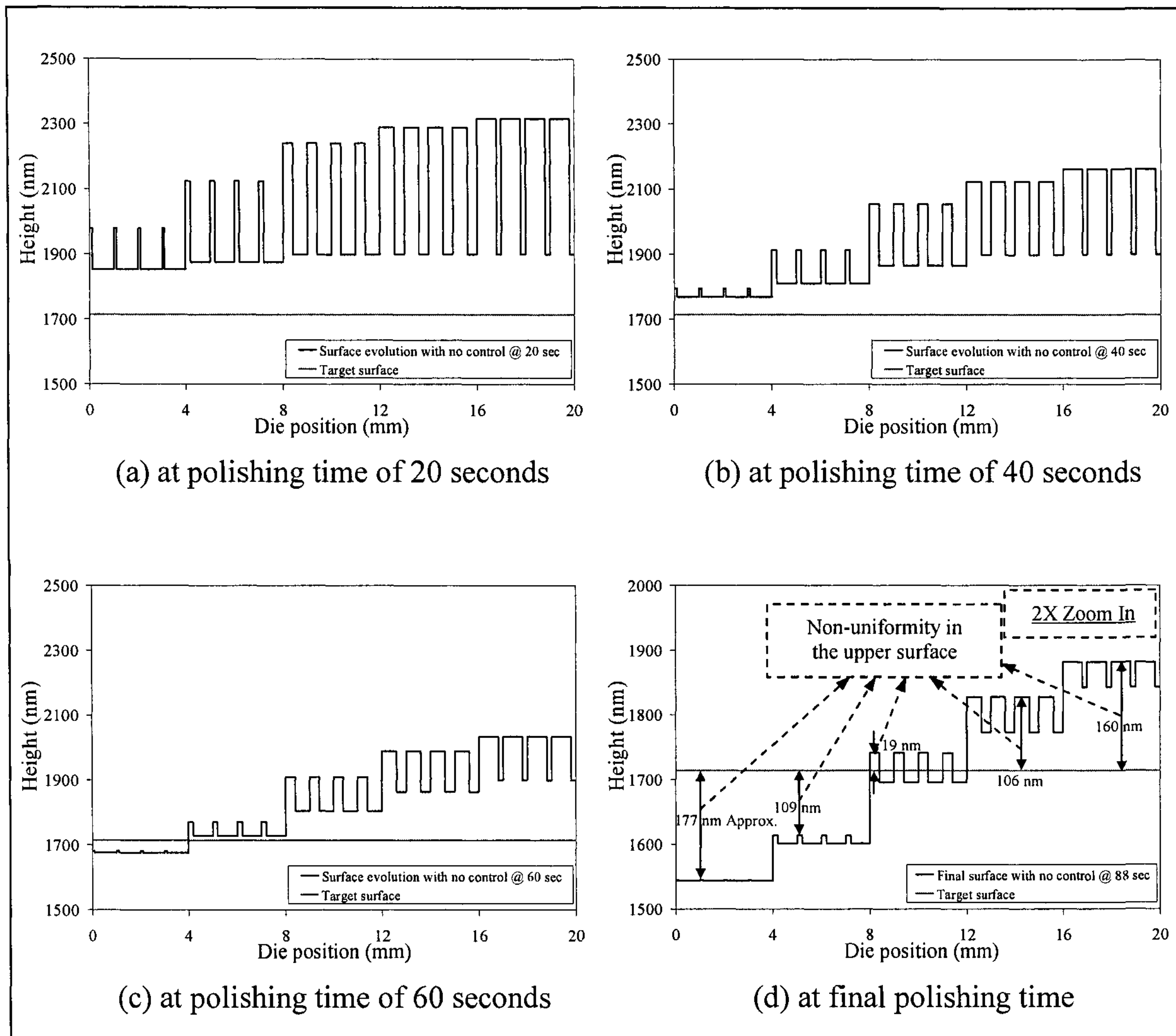


Figure 4.5. Surface evolution of line 3 vs. die position with no control algorithm.

According to Figure 4.5 (a), the pressure of 5 psi is applied for 20 seconds, resulted in non-uniformity of the upper surfaces across the die. In addition, it shows that the upper surfaces are polished faster with higher pattern densities; therefore, the lower surface regions with low pattern densities remain unpolished. As the surfaces continue polishing, the variation of the upper surfaces and the step heights are existed as shown in Figures 4.5 (b) and (c). Moreover, the non-uniformity of the upper surfaces of 340 nm is noticeable in Figure 4.5 (d).

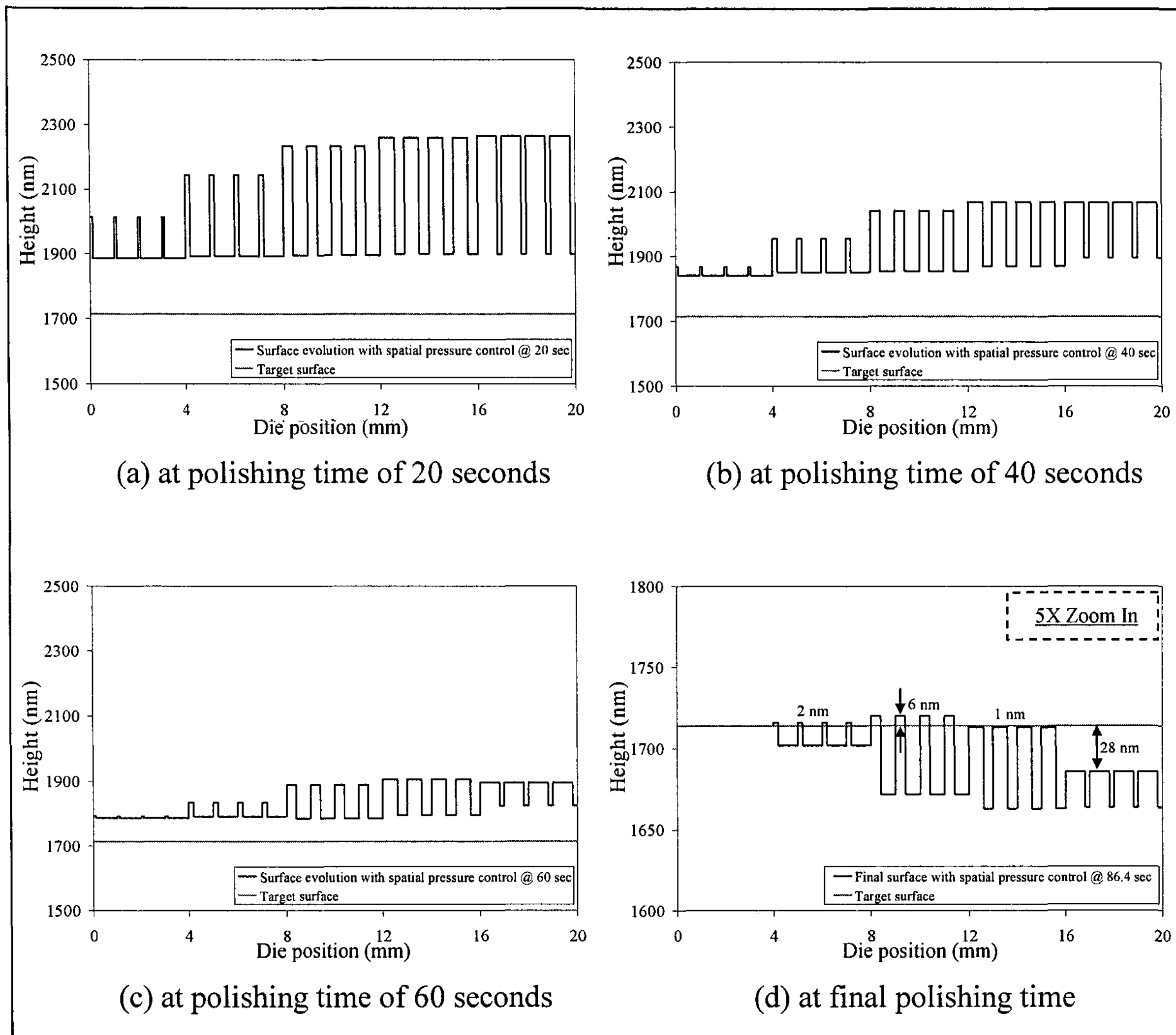


Figure 4.6. Surface evolution of line 3 vs. die position with spatial pressure control algorithm.

By varying the applied pressure across the die, at polishing time of 20 seconds, non-uniformity of the upper surfaces occur. It is shown that the variation of the upper surfaces in Figure 4.6 (a) is smaller than the variation of those in Figure 4.5 (a). In addition, most of the lower regions in Figure 4.6 (a) are still unpolished.

According to Figure 4.6 (b) and (c), a significant improvement of the upper surface variation appears as the polishing process is longer. In this algorithm, the simulation is

stopped at a polishing time of 86.4 seconds at the target upper surface of 1714 nm. In addition, the variation of the upper surface is approximately 34 nm shown in Figure 4.6 (d). In comparison to the uniformity of the previous simulation results, Figure 4.6 (d) illustrates a major improvement.

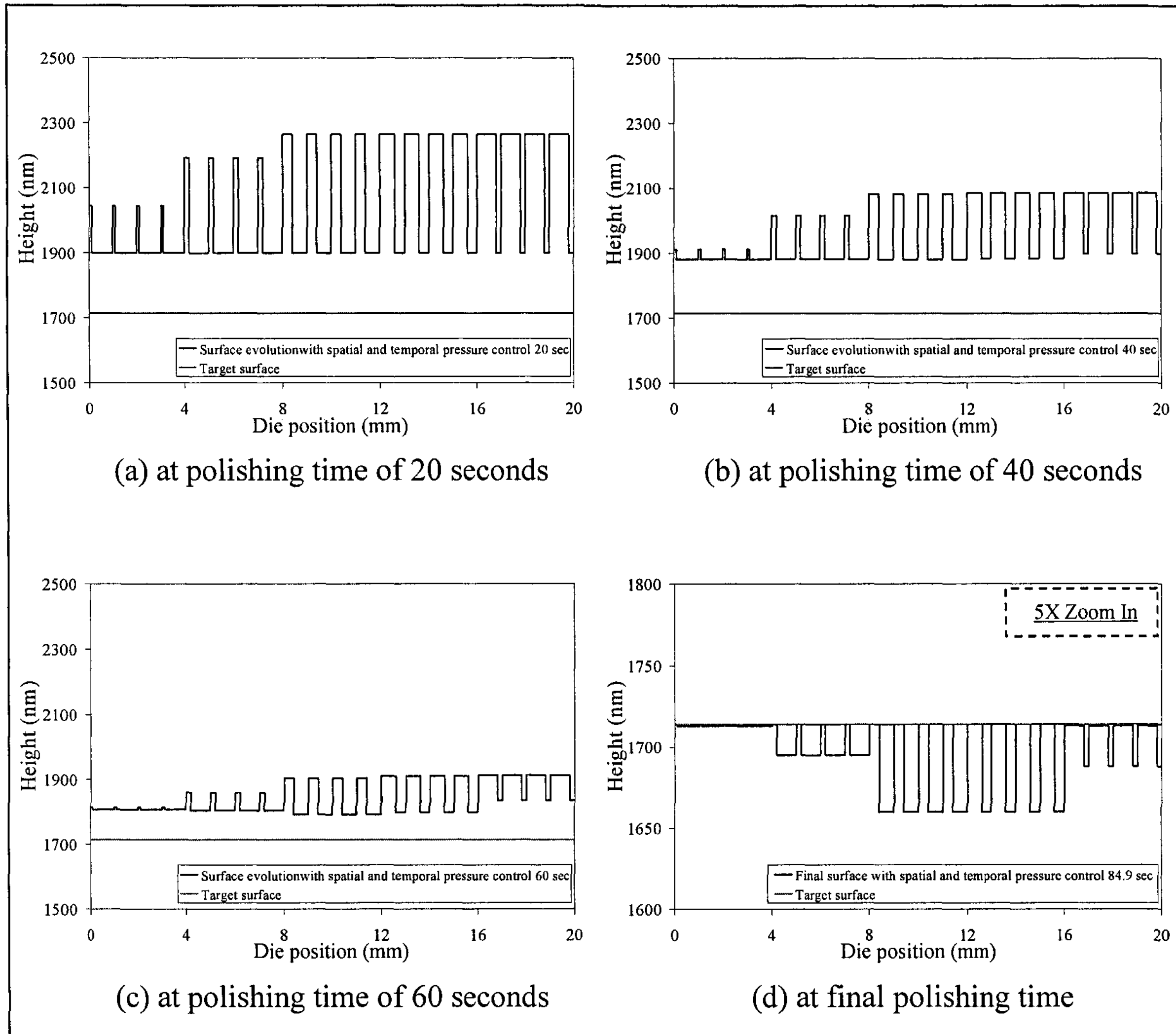
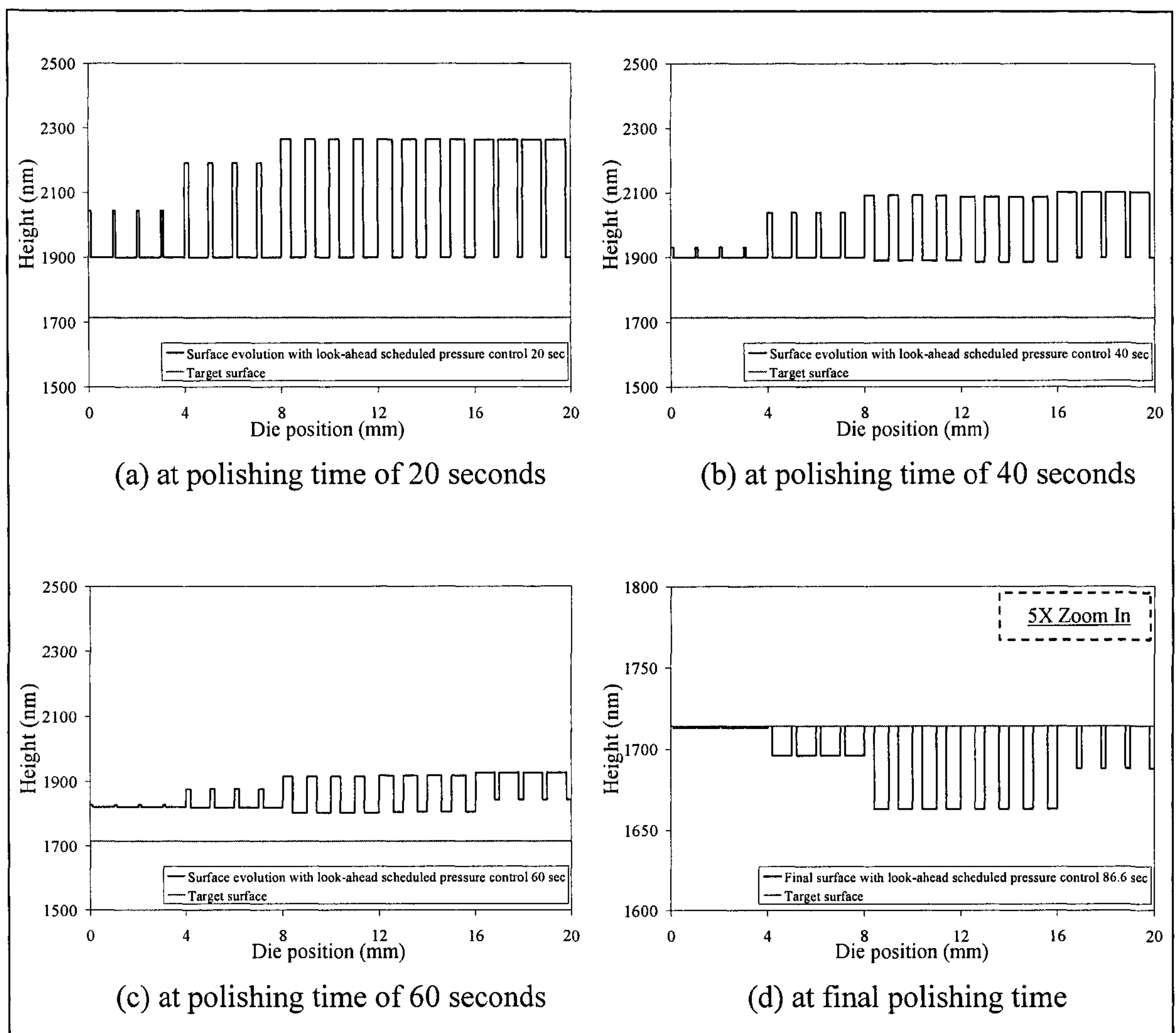


Figure 4.7. Surface evolution of line 3 vs. die position with spatial and temporal pressure control algorithm.

Figure 4.7 (a) shows that a low pressure is applied at the beginning of the polishing process. After polishing for 20 seconds, the entire lower regions are unpolished while it still

generates the non-uniformity in the upper surfaces across the die, as found in earlier simulation results. At 40 seconds, the variation in the upper surface heights is found concurrently while the lower regions begin the polishing process, shown in Figure 4.7 (b). After 40 seconds of polishing time, the spatial pressure control is applied, resulting in better uniformity of the upper regions. Finally, the simulation is stopped at the 84.9 seconds, with excellent uniformity in the upper surfaces as shown in Figure 4.7 (d).



(a) at polishing time of 20 seconds

(b) at polishing time of 40 seconds

(c) at polishing time of 60 seconds

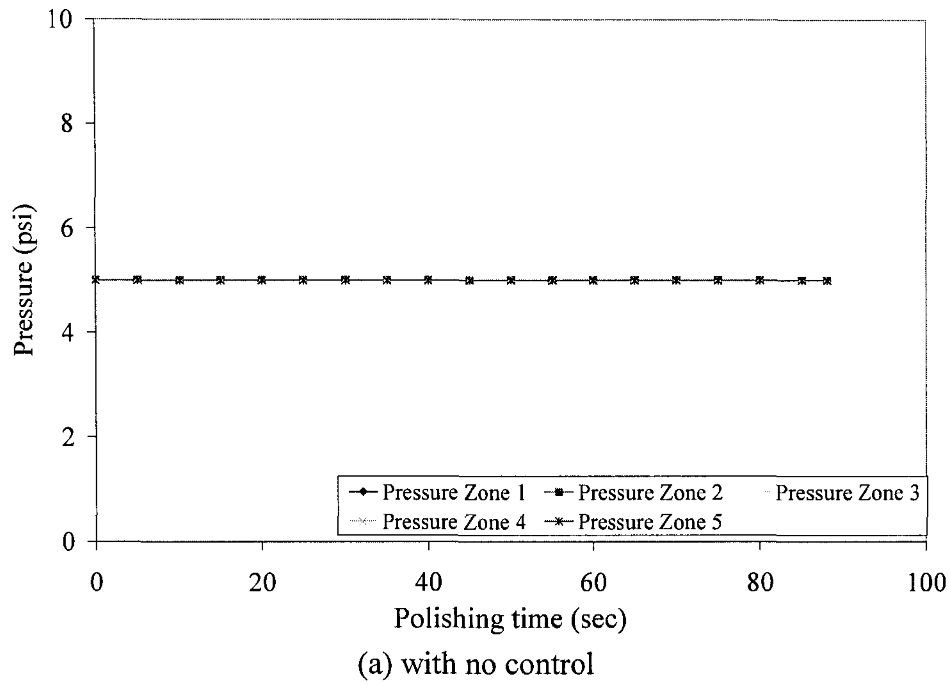
(d) at final polishing time

Figure 4.8. Surface evolution of line 3 vs. die position with look-ahead scheduled pressure control algorithm.

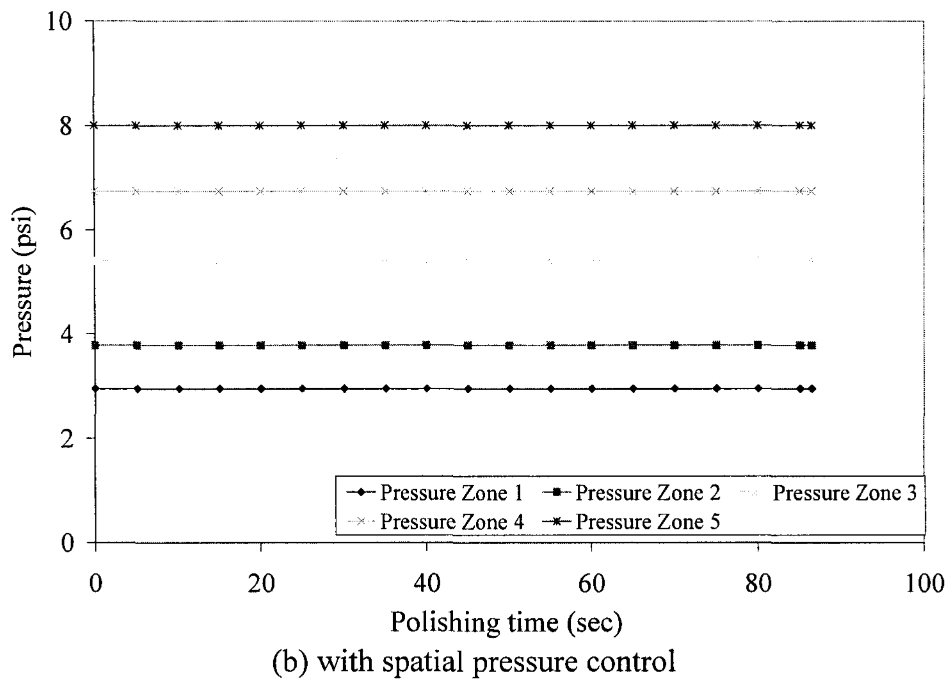
Figure 4.8 shows the similar results to the spatial and temporal pressure control simulation. The final surface evolution presents with the outstanding upper region uniformity at 86.6 seconds of polishing time.

In conclusion, the series of graphs in Figures 4.5 to 4.8 show substantial improvement in the uniformity of the upper regions between the various control algorithms. Furthermore, the applied pressure of each control algorithm and the material removal rate variation in upper and lower surfaces for various controls are shown in Figures 4.9 and 4.10, respectively. Two-wave feature/die scale model will be discussed in the next chapter.

The results of surface evolution of Line 4, 5, and 6 are shown in Appendix G.

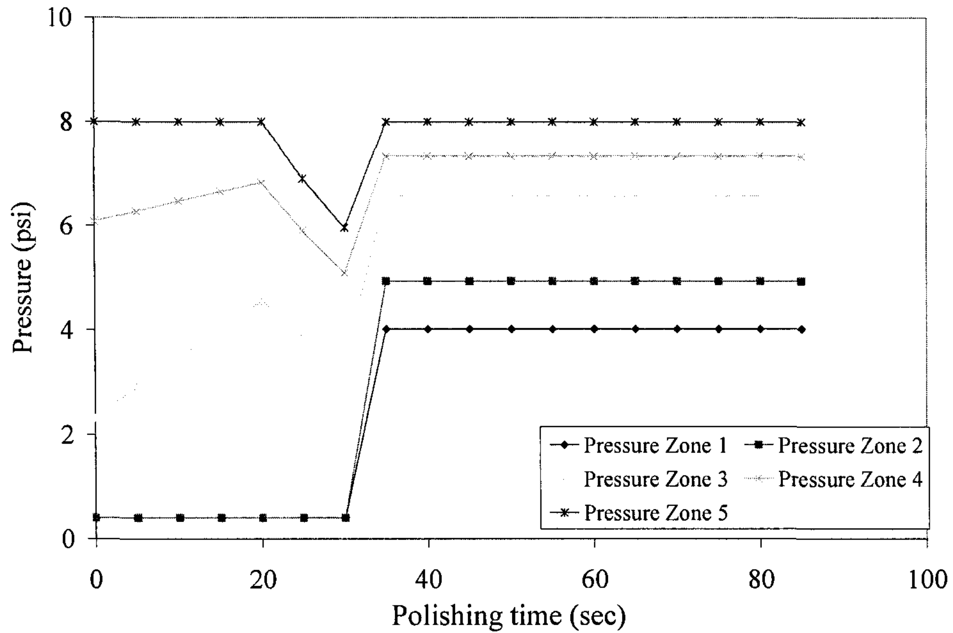


(a) with no control

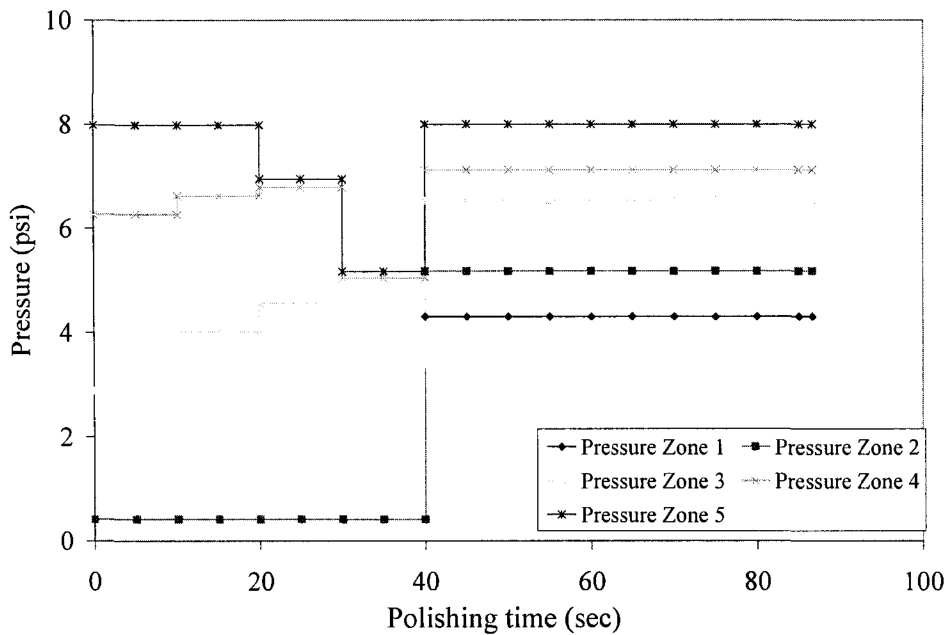


(b) with spatial pressure control

Figure 4.9. Applied pressure of line 3 vs. polishing time in 5 different zones.

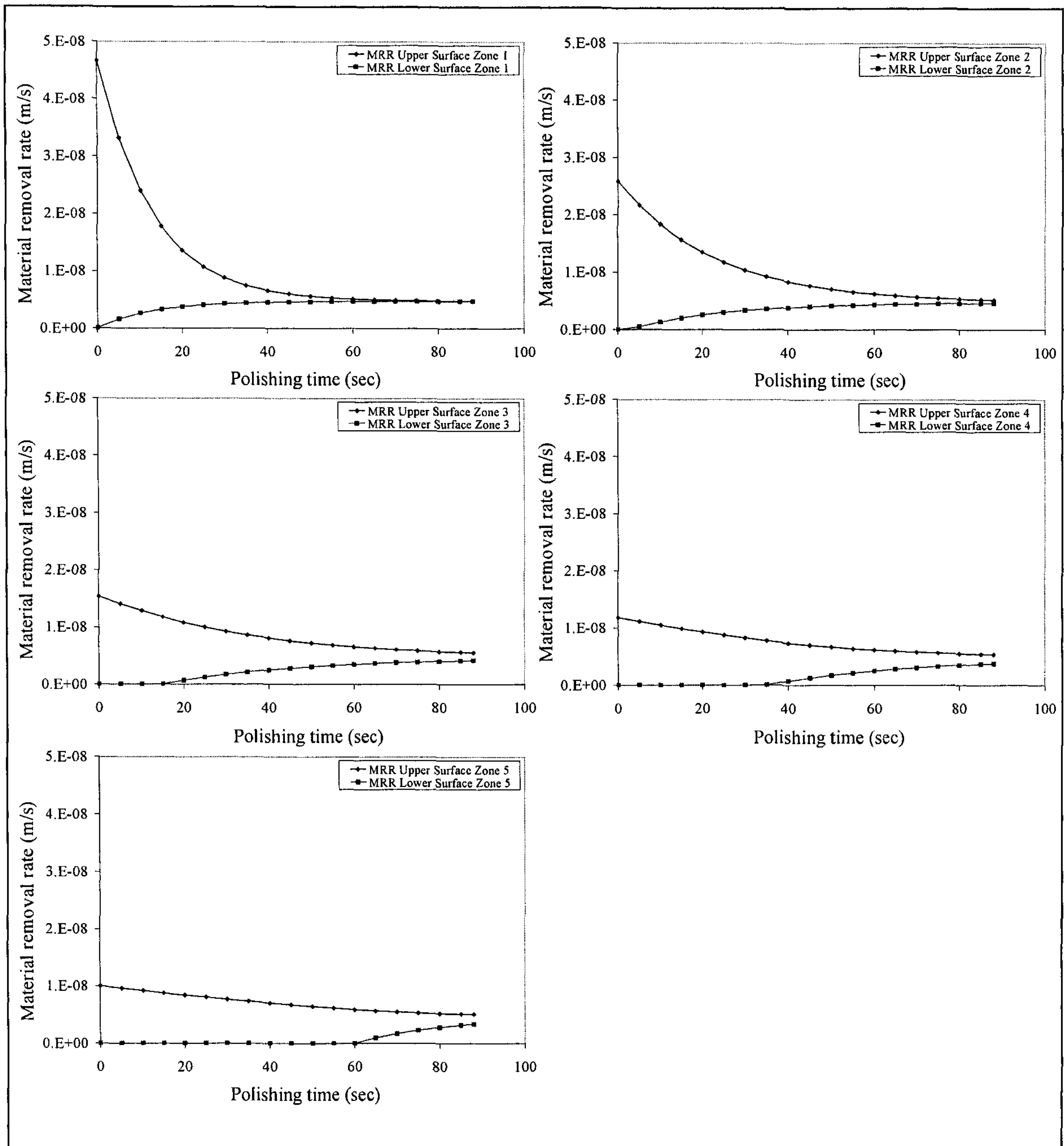


(c) with spatial and temporal pressure control



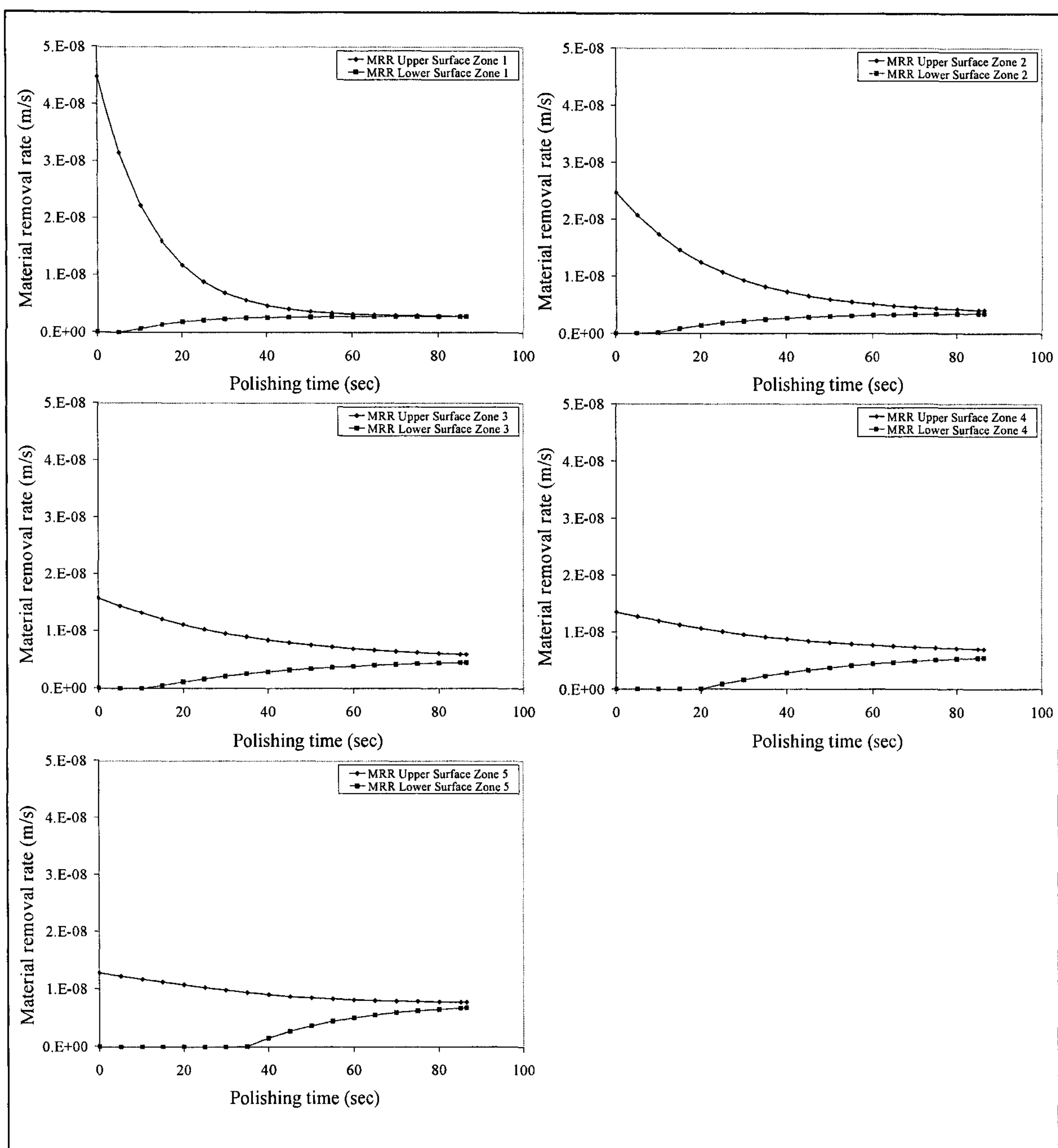
(d) with look-ahead scheduled pressure control

Figure 4.9. (Continued).



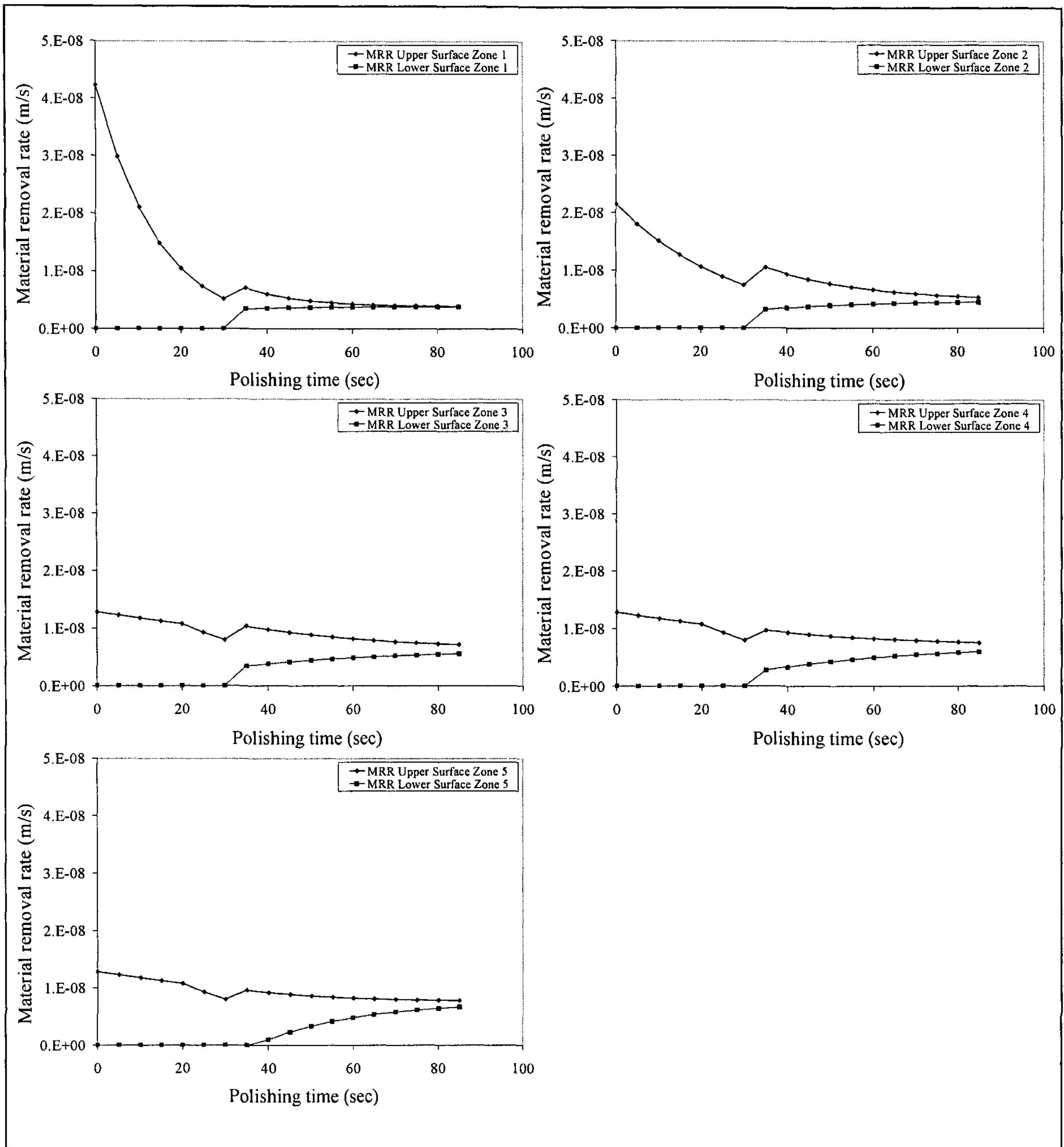
(a) with no control

Figure 4.10. Material removal rate of line 3 vs. polishing time in 5 different zones.



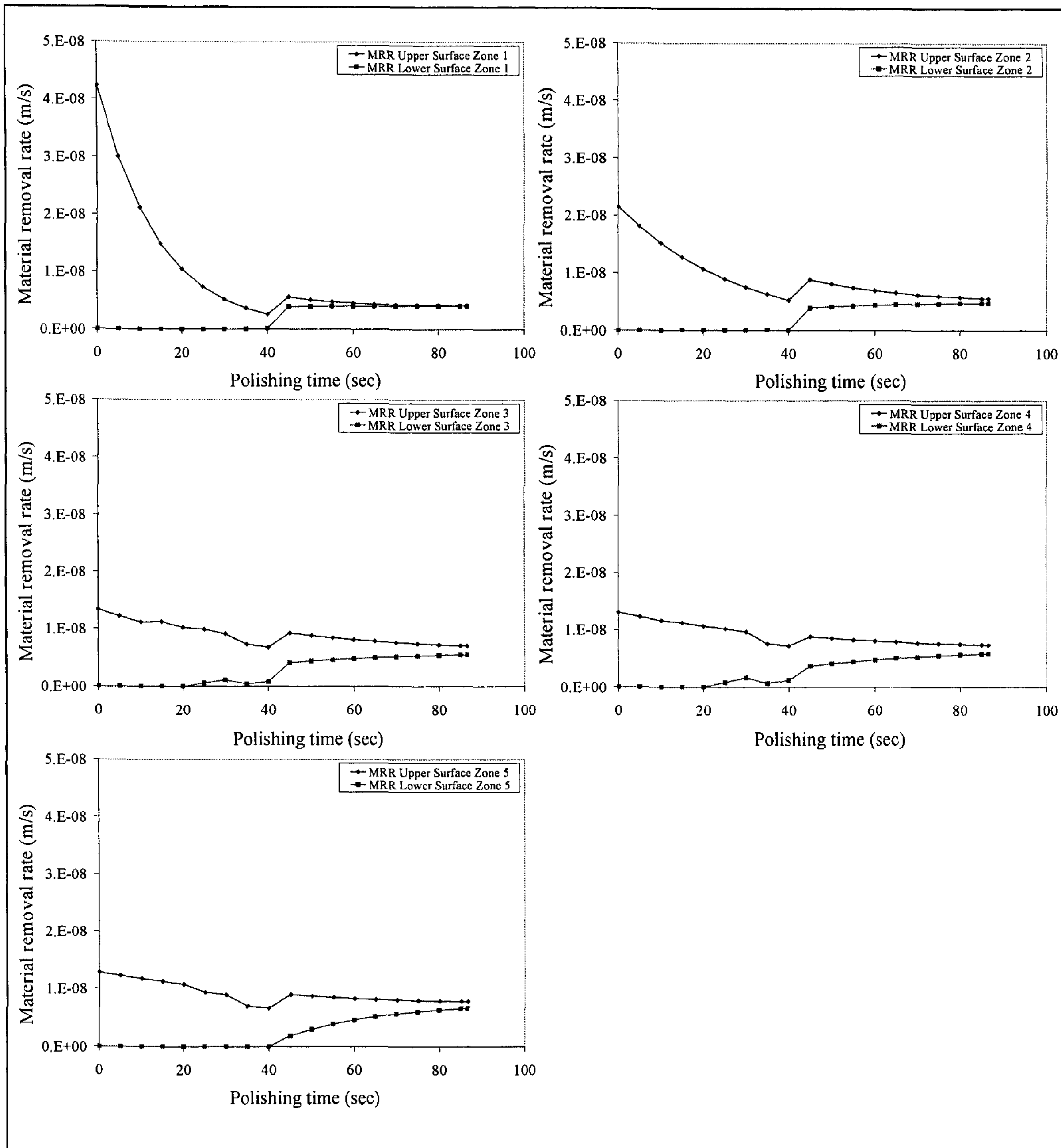
(b) with spatial pressure control

Figure 4.10. (Continued).



(c) with spatial and temporal pressure control

Figure 4.10. (Continued).



(d) with look-ahead scheduled pressure control

Figure 4.10. (Continued).

CHAPTER 5. MODEL FOR GENERAL WAFER SURFACE

This chapter discusses feature / die-scale model for general wafer surface. The two-wave model is attempted first and then verified based on the model from Fu and Chandra [9]. Examples and results of this model will be shown at the end of this chapter.

In this model, surface obtained from experimental data is imported and converted to a standard form of general wafer surface. This standard form is then taken Fast Fourier Transform (FFT) to capture dominated frequencies with their amplitudes. A series of sine waveforms is formulated based on these frequencies and amplitudes. Lastly, this series of sine waveforms is simplified to a series of square waveforms. The schematic of this manner is shown in Figure 5.1.

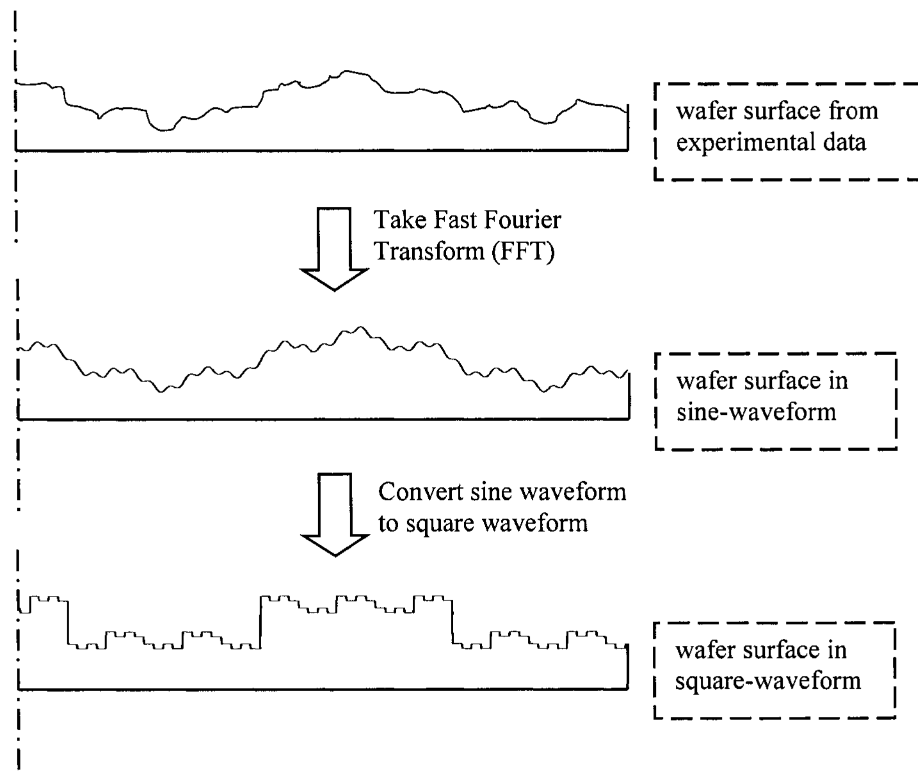


Figure 5.1. The schematic of wafer surface tranformation.

5.1. Simplified Representation of General Wafer Surface Profile

To develop a multi-wave feature / die-scale model, the general wafer surface profile is first formulated. The entire wafer surface obtained from the experimental data can be expressed as:

$$Y(x) = Y_0 + A_x$$

Where, Y_0 is an average wafer surface from the experimental data,

A_x is the film amplitudes from an average wafer surface (Y_0),

By taking the Fast Fourier Transform (FFT), a series of sine waveforms is formulated as shown below.

$$Y(x) = Y_0 + A_1 \sin(f_1 x) + A_2 \sin(f_2 x) + \ldots + A_n \sin(f_n x)$$

Where, $A_1, A_2, \ldots, A_n$ are the film amplitudes,

$f_1, f_2, \ldots, f_n$ are the frequencies,

x is the distance from the center of the wafer, and

n is the number of waves.

Moreover, for simplicity, this series of sine waves is represented with a series of square waves where the square waveform amplitudes are two-thirds of sine waveform amplitudes.

5.2. Two-Wave Feature / Die-Scale Model Development

In this part, the feature / die-scale model is considered to contain only two waves. According to Figure 5.2, the entire wafer surface is discretized, where Line 1 represents the wafer surface, expressed as a summation of two sinusoids.

$$Y = Y_0 + A_1 \sin(f_1 x) + A_2 \sin(f_2 x).$$

By simplifying Line 1, the wafer surface is then written in a square waveform, represented by Line 2.

$$
\begin{aligned}
Y &= Y_0 + (0.67)A_1 + (0.67)A_2 && ,if\ 0 \le x - 2m\pi \le \pi/f_1\ and \\
&&& \quad 0 \le x - 2m\pi \le \pi/f_2 \\
&= Y_0 + (0.67)A_1 - (0.67)A_2 && ,if\ \pi/f_1 \le x - 2m\pi \le 2\pi/f_1\ and \\
&&& \quad \pi/f_2 \le x - 2m\pi \le 2\pi/f_2 \\
&= Y_0 - (0.67)A_1 + (0.67)A_2 && ,if\ 0 \le x - 2m\pi \le \pi/f_1\ and \\
&&& \quad 0 \le x - 2m\pi \le \pi/f_2 \\
&= Y_0 - (0.67)A_1 - (0.67)A_2 && ,otherwise
\end{aligned}
$$

As shown in figure below, wafer surface Y is measured from the same reference plane on the wafer.

Figure 5.2. Pseudo surface profile of two-wave sine and square waveforms.

Based on the paper of Fu and Chandra [9], major assumptions are:

- The pad is assumed to deform like an elastic foundation (a set of linear elastic springs);

- Force redistribution due to pad bending is proportional to step height;
- Wafer and pad are in contact at any point of the interface;
- It is a feature scale model accounting for local feature patterns.

Additional assumptions in this model are:

- Wafer surface is approximated by Fast Fourier Transform (FFT) and is simplified to a series of sine waves and then into a series of square waves.
- The force redistribution is calculated only on the step height at the particular point.
- This model only considers a two-wave step height.

Moreover, the notations used in this chapter are shown in Table 5.1.

Table 5.1. List of variables and notations used in the two-wave model.

Variable	Notation
$Y(t)$	Current height
$D(t)$	Step height
P	Interface pressure
V	Relative velocity
K	Preston's constant
k	Pad stiffness
a_1	Linewidth of low frequency
b_1	Pitch of low frequency
c_1	$b_1 - a_1$
a_2	Linewidth of high frequency
b_2	Pitch of high frequency
c_2	$b_2 - a_2$
α	Bending factor
a/b	Pattern density
s	Variable in Laplace space
t	Time

From Preston's equation (Preston [17]), the material removal rate of four different regions can be expressed as:

$$\frac{dY_1}{dt} = K\{k[H-(Y-Y_1)]\}V \qquad [19]$$

$$\frac{dY_2}{dt} = K\{k[H-(Y-Y_2)]\}V \qquad [20]$$

$$\frac{dY_3}{dt} = K\{k[H-(Y-Y_3)]\}V \qquad [21]$$

$$\frac{dY_4}{dt} = K\{k[H-(Y-Y_4)]\}V \qquad [22]$$

From Fu and Chandra [9], the pad thickness in the highest region at time t is $Y(t)-Y_1(t)$. At this moment, the deformation of the pad corresponding to its original undeformed thickness H is $H-(Y-Y_1)$. Thus, the pressure of this deformation should be $k[H-(Y-Y_1)]$. The total area in this region with unit thickness (length into the page) is $(b_2-a_2).[(b_1-a_1)/b_2]$. Because the pressure on this region is the same, the total force is $\{k[H-(Y-Y_1)]\}.\{(b_2-a_2).[(b_1-a_1)/b_2]\}$. Similarly, the total force on the other regions are $\{k[H-(Y-Y_1)]\}.\{(a_2).[(b_1-a_1)/b_2]\}$, $\{k[H-(Y-Y_1)]\}.\{(b_2-a_2).(a_1/b_2)\}$, and $\{k[H-(Y-Y_1)]\}.\{(a_2).(a_1/b_2)\}$.

Assuming a constant downforce, the force equilibrium equation can be expressed as:

$$\begin{aligned}
&\{k\cdot[H-(Y-Y_1)]\cdot(b_2-a_2)\cdot[(b_1-a_1)/b_2]\}+\\
&\{k\cdot[H-(Y-Y_2)]\cdot(a_2)\cdot[(b_1-a_1)/b_2]\}+\\
&\{k\cdot[H-(Y-Y_3)]\cdot(b_2-a_2)\cdot(a_1/b_2)\}+\\
&\{k\cdot[H-(Y-Y_4)]\cdot(a_2)\cdot(a_1/b_2)\}\\
&=P\cdot(b_1\cdot 1)
\end{aligned} \qquad [23]$$

Due to pad bending, force redistribution on the wafer surface is presented. Since there is waviness on the surface, it causes contact force on the lower region to drop. In order to maintain overall force balance between low and high regions, there is a corresponding rise in the contact force on the higher region. Modified governing equations, including pad bending effect are shown below in equations 24 to 28.

$$\frac{dY_1}{dt} = K\left\{k[H-(Y-Y_1)] + \frac{\Delta F_1}{b_1 - a_1} + \Delta F_2 / \{(b_2 - a_2)\cdot[(b_1 - a_1)/b_2]\}\right\}V \qquad [24]$$

$$\frac{dY_2}{dt} = K\left\{k[H-(Y-Y_2)] + \frac{\Delta F_1}{b_1 - a_1} - \Delta F_2 / \{a_2 \cdot[(b_1 - a_1)/b_2]\}\right\}V \qquad [25]$$

$$\frac{dY_3}{dt} = K\left\{k[H-(Y-Y_3)] - \frac{\Delta F_1}{a_1} + \Delta F_2 / \{(b_2 - a_2)\cdot[a_1/b_2]\}\right\}V \qquad [26]$$

$$\frac{dY_4}{dt} = K\left\{k[H-(Y-Y_4)] - \frac{\Delta F_1}{a_1} - \Delta F_2 / \{a_2 \cdot[a_1/b_2]\}\right\}V \qquad [27]$$

and

$$\begin{aligned}
&\{k\cdot[H-(Y-Y_1)]\cdot(b_2-a_2)\cdot[(b_1-a_1)/b_2] + \Delta F_1 + \Delta F_2\} + \\
&\{k\cdot[H-(Y-Y_2)]\cdot(a_2)\cdot[(b_1-a_1)/b_2] + \Delta F_1 - \Delta F_2\} + \\
&\{k\cdot[H-(Y-Y_3)]\cdot(b_2-a_2)\cdot(a_1/b_2) - \Delta F_1 + \Delta F_2\} + \\
&\{k\cdot[H-(Y-Y_4)]\cdot(a_2)\cdot(a_1/b_2) - \Delta F_1 - \Delta F_2\} \\
&= P\cdot(b_1\cdot 1)
\end{aligned} \qquad [28]$$

By rewriting equation 28 for *Y(t),* it can be expressed as:

$$Y(t) = \left(\frac{c_1 c_2}{b_1 b_2}\right)Y_1 + \left(\frac{c_1 a_2}{b_1 b_2}\right)Y_2 + \left(\frac{a_1 c_2}{b_1 b_2}\right)Y_3 + \left(\frac{a_1 a_2}{b_1 b_2}\right)Y_4 + \left(H - \frac{P}{k}\right) \qquad [29]$$

Substituting (29) into (24) to (27), the material removal rate equations can be expressed as:

$$\frac{dY_1}{dt} = KVk\begin{Bmatrix} \left(\frac{c_1c_2 - b_1b_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{1}{k}\left(P + \left(\frac{1}{c_1}\right)\Delta F_1 + \left(\frac{b_2}{c_1c_2}\right)\Delta F_2\right) \end{Bmatrix} \quad [30]$$

$$\frac{dY_2}{dt} = KVk\begin{Bmatrix} \left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2 - b_1b_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{1}{k}\left(P + \left(\frac{1}{c_1}\right)\Delta F_1 - \left(\frac{b_2}{c_1a_2}\right)\Delta F_2\right) \end{Bmatrix} \quad [31]$$

$$\frac{dY_3}{dt} = KVk\begin{Bmatrix} \left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2 - b_1b_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{1}{k}\left(P - \left(\frac{1}{a_1}\right)\Delta F_1 + \left(\frac{b_2}{a_1c_2}\right)\Delta F_2\right) \end{Bmatrix} \quad [32]$$

$$\frac{dY_4}{dt} = KVk\begin{Bmatrix} \left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2 - b_1b_2}{b_1b_2}\right)Y_4 \\ -\frac{1}{k}\left(P - \left(\frac{1}{a_1}\right)\Delta F_1 - \left(\frac{b_2}{a_1a_2}\right)\Delta F_2\right) \end{Bmatrix} \quad [33]$$

From the second assumption, force redistributions due to the pad bending can be expressed as:

$$\begin{aligned} \Delta F_1 &= \alpha(Y_1 - Y_3)\ or\ \alpha(Y_2 - Y_4) \\ \Delta F_2 &= \alpha(Y_1 - Y_2)\ or\ \alpha(Y_3 - Y_4) \end{aligned} \quad [34]$$

Where, α is the bending factor and it is zero if there is no pad bending.

It should be noted that α has no co-relation to pad stiffness k.

ΔF_1 is the force redistribution on the surface profile of amplitude A_1, and

ΔF_2 is the force redistribution on the surface profile of amplitude A_2.

Substituting (34) into (30) to (33) corresponding to the sixth assumption, the material removal rate equations can be expressed as:

$$\frac{dY_1}{dt} = KVk\left\{\begin{array}{l}\left(\frac{c_1c_2 - b_1b_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{\alpha}{k}\left(\frac{1}{c_1}\cdot(Y_1 - Y_3) + \left(\frac{b_2}{c_1c_2}\right)\cdot(Y_1 - Y_2)\right) - \frac{P}{k}\end{array}\right\} \quad [35]$$

$$\frac{dY_2}{dt} = KVk\left\{\begin{array}{l}\left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2 - b_1b_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{\alpha}{k}\left(\frac{1}{c_1}\cdot(Y_2 - Y_4) - \left(\frac{b_2}{c_1a_2}\right)\cdot(Y_1 - Y_2)\right) - \frac{P}{k}\end{array}\right\} \quad [36]$$

$$\frac{dY_3}{dt} = KVk\left\{\begin{array}{l}\left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2 - b_1b_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2}{b_1b_2}\right)Y_4 \\ -\frac{\alpha}{k}\left(-\frac{1}{a_1}\cdot(Y_1 - Y_3) + \left(\frac{b_2}{a_1c_2}\right)\cdot(Y_3 - Y_4)\right) - \frac{P}{k}\end{array}\right\} \quad [37]$$

$$\frac{dY_4}{dt} = KVk\left\{\begin{array}{l}\left(\frac{c_1c_2}{b_1b_2}\right)Y_1 + \left(\frac{c_1a_2}{b_1b_2}\right)Y_2 + \left(\frac{a_1c_2}{b_1b_2}\right)Y_3 + \left(\frac{a_1a_2 - b_1b_2}{b_1b_2}\right)Y_4 \\ -\frac{\alpha}{k}\left(-\frac{1}{a_1}\cdot(Y_2 - Y_4) - \frac{b_2}{a_1a_2}\cdot(Y_3 - Y_4)\right) - \frac{P}{k}\end{array}\right\} \quad [38]$$

Where,

$$c_1 = (b_1 - a_1)$$
$$c_2 = (b_2 - a_2)$$

Rewriting (35) to (38), they are expressed as:

$$\frac{dY_1}{dt} = C_{11}Y_1 + C_{12}Y_2 + C_{13}Y_3 + C_{14}Y_4 + C_{15} \quad [39]$$

$$\frac{dY_2}{dt} = C_{21}Y_1 + C_{22}Y_2 + C_{23}Y_3 + C_{24}Y_4 + C_{25} \quad [40]$$

$$\frac{dY_3}{dt} = C_{31}Y_1 + C_{32}Y_2 + C_{33}Y_3 + C_{34}Y_4 + C_{35} \quad [41]$$

$$\frac{dY_4}{dt} = C_{41}Y_1 + C_{42}Y_2 + C_{43}Y_3 + C_{44}Y_4 + C_{45} \quad [42]$$

Where,

$$C_{11} = KVk\left(\frac{c_1c_2 - b_1b_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{c_2 + b_2}{c_1c_2}\right)\right),\ C_{12} = KVk\left(\frac{c_1a_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-b_2}{c_1c_2}\right)\right),$$

$$C_{13} = KVk\left(\frac{a_1c_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-c_2}{c_1c_2}\right)\right),\ C_{14} = KVk\left(\frac{a_1a_2}{b_1b_2}\right),\ C_{15} = KVk\left(-\frac{P}{k}\right),$$

$$C_{21} = KVk\left(\frac{c_1c_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-b_2}{c_1a_2}\right)\right),\ C_{22} = KVk\left(\frac{c_1a_2 - b_1b_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{a_2 + b_2}{c_1a_2}\right)\right),$$

$$C_{23} = KVk\left(\frac{a_1c_2}{b_1b_2}\right),\ C_{24} = KVk\left(\frac{a_1a_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-a_2}{c_1a_2}\right)\right),\ C_{25} = C_{15},$$

$$C_{31} = KVk\left(\frac{c_1c_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-c_2}{a_1c_2}\right)\right),\ C_{32} = KVk\left(\frac{c_1a_2}{b_1b_2}\right),$$

$$C_{33} = KVk\left(\frac{a_1c_2 - b_1b_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{c_2 + b_2}{a_1c_2}\right)\right),\ C_{34} = KVk\left(\frac{a_1a_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-b_2}{a_1c_2}\right)\right),\ C_{35} = C_{15},$$

$$C_{41} = KVk\left(\frac{c_1c_2}{b_1b_2}\right),\ C_{42} = KVk\left(\frac{c_1a_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{-a_2}{a_1a_2}\right)\right),$$

$$C_{43} = KVk\left(\frac{a_1c_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{a_2 - b_2}{a_1a_2}\right)\right),\ C_{44} = KVk\left(\frac{a_1a_2 - b_1b_2}{b_1b_2} - \frac{\alpha}{k}\left(\frac{b_2}{a_1a_2}\right)\right),\ C_{45} = C_{15}$$

Using a Laplace transform on (39) to (40), we have

$$s\widetilde{Y}_1 - Y_1(0) = C_{11}\widetilde{Y}_1 + C_{12}\widetilde{Y}_2 + C_{13}\widetilde{Y}_3 + C_{14}\widetilde{Y}_4 + \frac{C_{15}}{s} \quad [43]$$

$$s\widetilde{Y}_2 - Y_2(0) = C_{21}\widetilde{Y}_1 + C_{22}\widetilde{Y}_2 + C_{23}\widetilde{Y}_3 + C_{24}\widetilde{Y}_4 + \frac{C_{25}}{s} \quad [44]$$

$$s\widetilde{Y}_3 - Y_3(0) = C_{31}\widetilde{Y}_1 + C_{32}\widetilde{Y}_2 + C_{33}\widetilde{Y}_3 + C_{34}\widetilde{Y}_4 + \frac{C_{35}}{s} \quad [45]$$

$$s\widetilde{Y}_4 - Y_4(0) = C_{41}\widetilde{Y}_1 + C_{42}\widetilde{Y}_2 + C_{43}\widetilde{Y}_3 + C_{44}\widetilde{Y}_4 + \frac{C_{45}}{s} \quad [46]$$

According to equations 43 to 46, the unknowns are $\widetilde{Y}_1,\ \widetilde{Y}_2,\ \widetilde{Y}_3,\ and\ \widetilde{Y}_4$. Those four equations can be rewritten in matrix form as:

$$\begin{bmatrix} s-C_{11} & -C_{12} & -C_{13} & -C_{14} \\ -C_{21} & s-C_{22} & -C_{23} & -C_{24} \\ -C_{31} & -C_{32} & s-C_{33} & -C_{34} \\ -C_{41} & -C_{42} & -C_{43} & s-C_{44} \end{bmatrix} \begin{bmatrix} \widetilde{Y}_1 \\ \widetilde{Y}_2 \\ \widetilde{Y}_3 \\ \widetilde{Y}_4 \end{bmatrix} = \begin{bmatrix} \frac{C_{15}}{s} + Y_1(0) \\ \frac{C_{25}}{s} + Y_2(0) \\ \frac{C_{35}}{s} + Y_3(0) \\ \frac{C_{45}}{s} + Y_4(0) \end{bmatrix} \quad [47]$$

Solving the matrix in equation 47, we have

$$\begin{bmatrix} \widetilde{Y}_1 \\ \widetilde{Y}_2 \\ \widetilde{Y}_3 \\ \widetilde{Y}_4 \end{bmatrix} = \begin{bmatrix} s-C_{11} & -C_{12} & -C_{13} & -C_{14} \\ -C_{21} & s-C_{22} & -C_{23} & -C_{24} \\ -C_{31} & -C_{32} & s-C_{33} & -C_{34} \\ -C_{41} & -C_{42} & -C_{43} & s-C_{44} \end{bmatrix}^{-1} \begin{bmatrix} \frac{C_{15}}{s} + Y_1(0) \\ \frac{C_{25}}{s} + Y_2(0) \\ \frac{C_{35}}{s} + Y_3(0) \\ \frac{C_{45}}{s} + Y_4(0) \end{bmatrix} \quad [48]$$

This inverse Laplace transform can be solved using any powerful mathematic software such as the “Mathematica”, “Maple” or “Matlab”.

5.3. Two-Wave Model Verification and Simulation Results

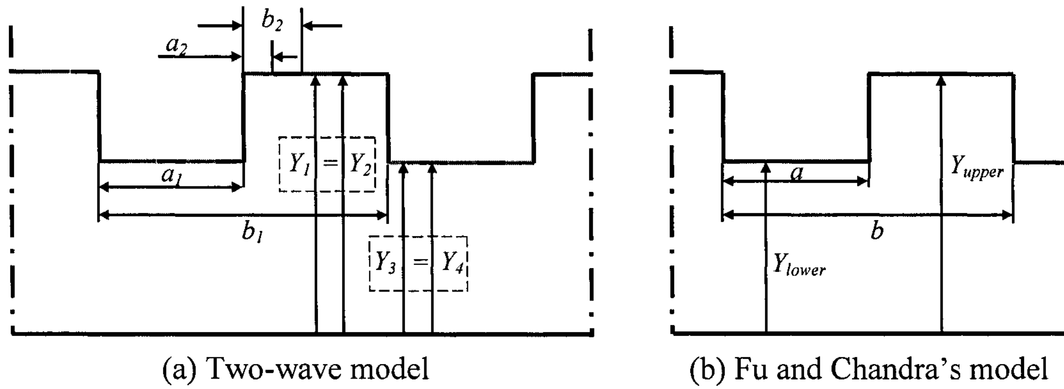
To verify the two-wave model, it is first compared to Fu and Chandra's single wave model. Then, example results of two different initial surface topographies are simulated.

5.3.1. Comparison two-wave model to Fu and Chandra's (single wave) model

Table 5.2 shows variables which are taken into consideration in verifying this two-wave model. The model is assumed that the surface contains one frequency of a square wave, similar to Fu and Chandra's model. Thus, heights of Y_1 and Y_2 in the two-wave model are the same as Y_{upper} in Fu and Chandra's. Similarly, heights of Y_3 and Y_4 are equal to Y_{lower}. The illustration of this example is shown in Figure 5.3.

Table 5.2. List of variables used for comparing two-wave model to Fu's model.

Two-wave model			Fu and Chandra's model		
a_1	50e-06	m	a	50e-06	m
a_2	10e-06	m			
b_1	100e-06	m	b	100e-06	m
b_2	20e-06	m			
Y_1	1650	Å	Y_{upper}	1650	Å
Y_2	1650	Å			
Y_3	1050	Å	Y_{lower}	1050	Å
Y_4	1050	Å			
P	21.9	kPa			
V	0.65	m/s			
K	4.683e-13	N/m^2			
k	5.213e09	N/m^3			
α	1.403e06	N/m^2			



(a) Two-wave model

(b) Fu and Chandra's model

Figure 5.3. The illustration of the example 5.3.1.

Table 5.3. The results of two-wave model and Fu and Chandra's model.

Time (sec)	Two-Wave model					Fu's model
	Y_1	Y_2	Y_3	Y_4	Step Height	Step Height
0	1.6500E-06	1.6500E-06	1.0500E-06	1.0500E-06	6.0000E-07	6.0000E-07
10	1.5322E-06	1.5322E-06	1.0344E-06	1.0344E-06	4.9782E-07	4.9782E-07
20	1.4232E-06	1.4232E-06	1.0102E-06	1.0102E-06	4.1304E-07	4.1304E-07
30	1.3214E-06	1.3214E-06	9.7866E-07	9.7866E-07	3.4270E-07	3.4270E-07
40	1.2255E-06	1.2255E-06	9.4118E-07	9.4118E-07	2.8433E-07	2.8433E-07
50	1.1346E-06	1.1346E-06	8.9873E-07	8.9873E-07	2.3591E-07	2.3591E-07
60	1.0479E-06	1.0479E-06	8.5216E-07	8.5216E-07	1.9574E-07	1.9574E-07
70	9.6456E-07	9.6456E-07	8.0216E-07	8.0216E-07	1.6240E-07	1.6240E-07
80	8.8407E-07	8.8407E-07	7.4933E-07	7.4933E-07	1.3474E-07	1.3474E-07
90	8.0594E-07	8.0594E-07	6.9414E-07	6.9414E-07	1.1180E-07	1.1180E-07
100	7.2975E-07	7.2975E-07	6.3700E-07	6.3700E-07	9.2757E-08	9.2757E-08
110	6.5519E-07	6.5519E-07	5.7823E-07	5.7823E-07	7.6960E-08	7.6960E-08
120	5.8198E-07	5.8198E-07	5.1812E-07	5.1812E-07	6.3854E-08	6.3854E-08

Table 5.3 compares the results of both models after applying the variables. This result is then used to plot the graphs, shown in Figures 5.4 and 5.5. Figure 5.4 shows the graph between the surface height and polishing time, where surface heights are decreased as polishing time increases. While the upper surface heights (Y_1 and Y_2) are decreased with

faster rate at the beginning of the process, lower surface heights (Y_3 and Y_4) are decreased slowly. Similarly, upper surface heights (Y_1 and Y_2) are decreased at a slower rate toward the end of the process, where lower surface heights (Y_3 and Y_4) are decreased faster. Furthermore, surface heights in both areas are slightly different by the end of this process. As the polishing process continues, slopes of both surface areas will become equal. Planarization efficiency will then drop to zero.

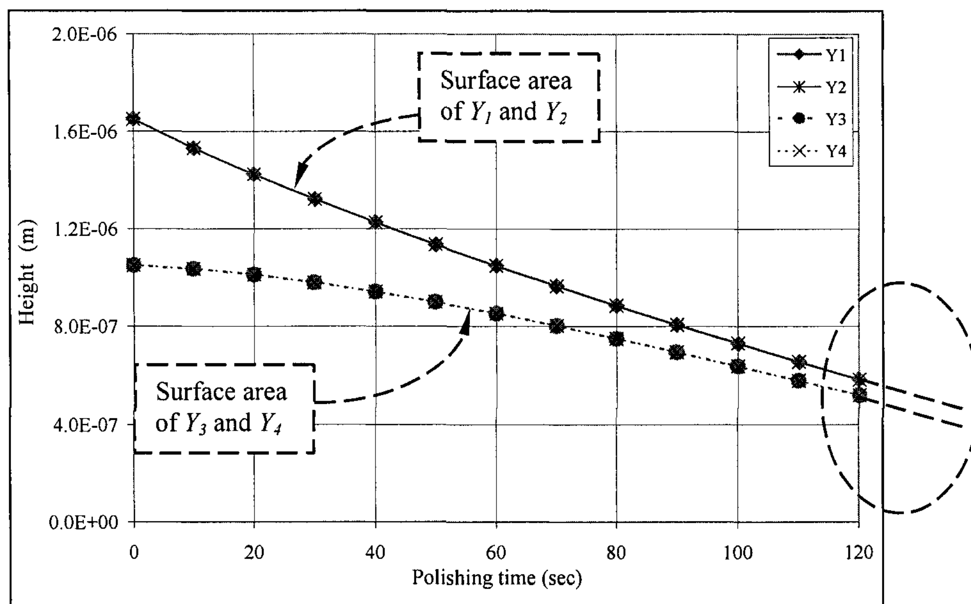


Figure 5.4. Graph between surface height vs. polishing time of two-wave model and Fu's model.

Figure 5.5 illustrates the relationship between step height and polishing time. It illustrates that step height curves in both models are reduced as polishing time increases. In addition, step height will reach a limit and cannot reduce indefinitely, called an asymptotic function. More importantly, the result from both models appears to be exceptionally similar, where the graph shows one curve on top of another.

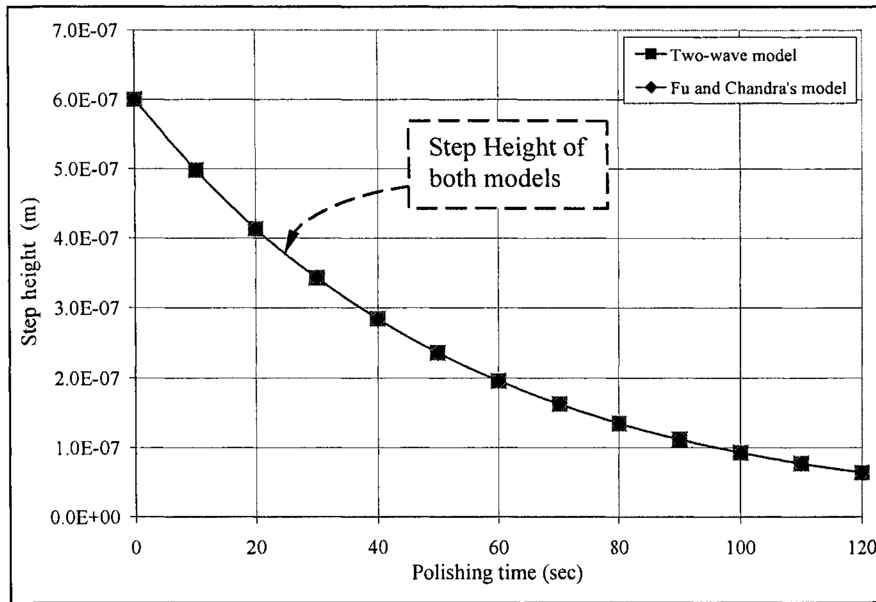


Figure 5.5. Graph between step height vs. polishing time of two-wave model and Fu's model.

5.3.2. Simulation results of two-wave model

The simulation of examples 1 and 2 assumed that the wafer surface comprises of two frequencies of a square waveform. In example 1, amplitudes of low frequency (A_1) and high frequency (A_2) are 300 Å and 50 Å, respectively, and average wafer surface is 1300 Å. Example 2 is assigned different frequencies and amplitudes to compare with the result of the first example. Where, amplitude of low frequency (A_1) is 50 Å, amplitude of high frequency (A_2) is 200 Å, and average wafer surface is 1400 Å. Square waveform behavior of these examples illustrates that pattern densities of 0.5 are presented in both low and high frequencies, shown in Figure 5.6. The variables used in this simulation are shown below in Table 5.4.

Table 5.4. List of variables used for simulation results in two-wave model.

Notation	Ex 1	Ex 2	Unit	Notation	Ex 1	Ex 2	Unit
P	21.9		kPa	V	0.65		m/s
K	4.683e-13		N/m^2	k	5.213e09		N/m^3
α	1.403e06		N/m^2				
a_1	50	50	nm	a_2	10	10	nm
b_1	100	100	nm	b_2	20	20	nm
Y_1	1650	1650	Å	Y_2	1550	1250	Å
Y_3	1050	1550	Å	Y_4	950	1150	Å

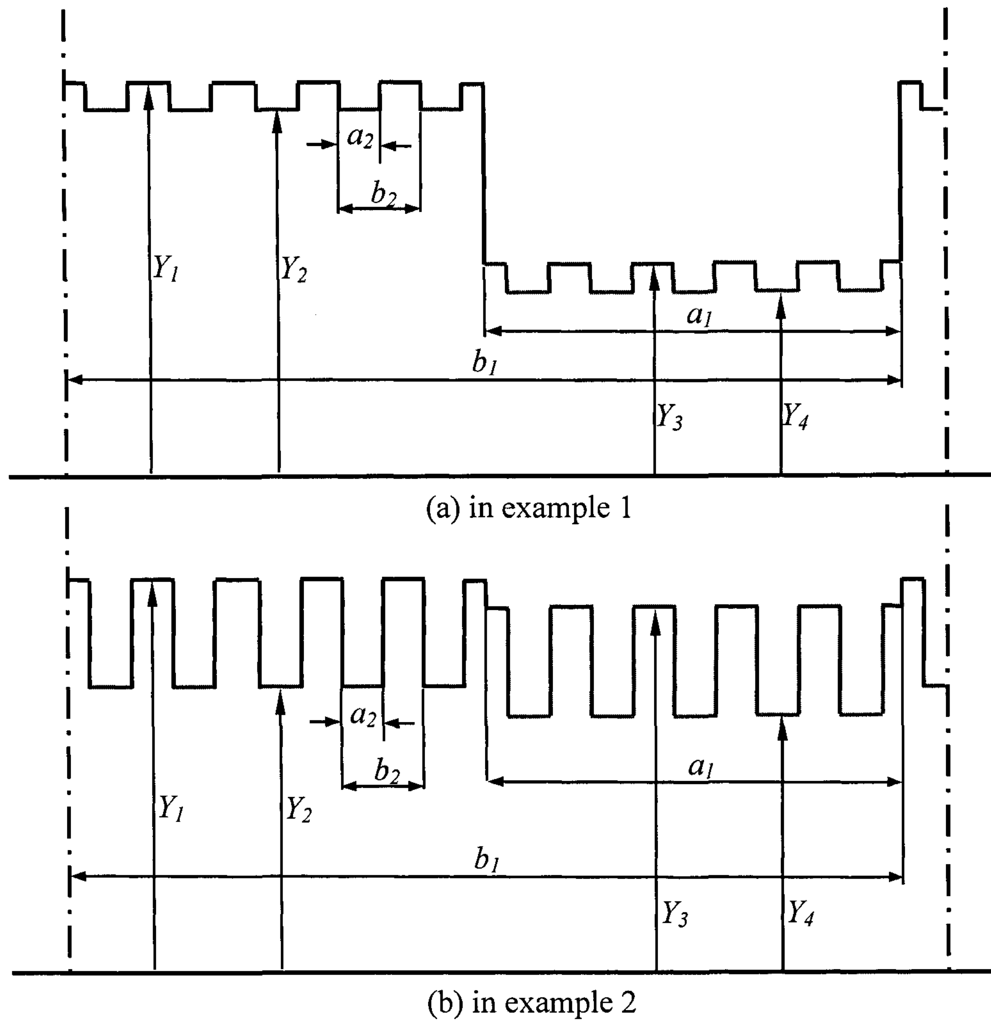


Figure 5.6. The illustration of the two-wave model.

This simulation illustrates wafer surface profiles of the two-wave model, based on variables discussed above. In order to solve this model, Mathematica software is used. In example 1, wafer surface profiles as function of polishing time, t, are shown in equations 49 to 52, and wafer surface profiles in example 2 are expressed in equations 53 to 56.

$$Y_1(t)^{Ex1} = \frac{105364428831}{81517873945000000} + \frac{e^{-\frac{3733891707 \cdot t}{2E+11}}}{3125000} + \frac{\left(-338025971435 + 93678490167\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} - \frac{\left(338025971435 + 93678490167\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [49]$$

$$Y_2(t)^{Ex1} = \frac{105364428831}{81517873945000000} + \frac{e^{-\frac{3733891707 \cdot t}{2E+11}}}{3125000} - \frac{\left(-392853380125 + 53898029033\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} + \frac{\left(392853380125 + 53898029033\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [50]$$

$$Y_3(t)^{Ex1} = \frac{105364428831}{81517873945000000} - \frac{e^{-\frac{3733891707 \cdot t}{2E+11}}}{3125000} - \frac{\left(42872933035 + 198673250457\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} - \frac{\left(-42872933035 + 198673250457\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [51]$$

$$Y_4(t)^{Ex1} = \frac{105364428831}{81517873945000000} - \frac{e^{-\frac{3733891707 \cdot t}{2E+11}}}{3125000} + \frac{\left(-46374555255 + 92978165723\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} + \frac{\left(46374555255 + 92978165723\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [52]$$

$$Y_1(t)^{Ex2} = \frac{111689794333}{81517873945000000} + \frac{26 \cdot e^{-\frac{3733891707 \cdot t}{2E+11}}}{2E+8} + \frac{\left(-1352103885740 + 374713960668\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} + \frac{\left(1352103885740 + 374713960668\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [53]$$

$$Y_2(t)^{Ex2} = \frac{111689794333}{81517873945000000} + \frac{26 \cdot e^{-\frac{3733891707 \cdot t}{2E+11}}}{2E+8} - \frac{\left(-1571413520500 + 215592116132\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} - \frac{\left(1571413520500 + 215592116132\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [54]$$

$$Y_3(t)^{Ex2} = \frac{111689794333}{81517873945000000} - \frac{26 \cdot e^{-\frac{3733891707 \cdot t}{2E+11}}}{2E+8} - \frac{\left(171491732140 + 794693001828\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5+\sqrt{5})} - \frac{\left(-171491732140 + 794693001828\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489+14030\sqrt{5}\right) t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5+\sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \qquad [55]$$

$$Y_4(t)^{Ex2} = \frac{111689794333}{81517873945000000} - \frac{26 \cdot e^{-\frac{3733891707 \cdot t}{2E+11}}}{2E+8} + \frac{\left(-185498221020 + 371912662892\sqrt{5}\right) \cdot e^{\frac{60879 \cdot \left(-131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 - 14030\sqrt{5}\right)^2 \cdot (-5 + \sqrt{5})} + \frac{\left(185498221020 + 371912662892\sqrt{5}\right) \cdot e^{-\frac{60879 \cdot \left(131489 + 14030\sqrt{5}\right)t}{2E+11}}}{(2E+8) \cdot \left(131483 + 14030\sqrt{5}\right)^2 \cdot (5 + \sqrt{5})} - \frac{13332501 \cdot t}{2E+15} \quad [56]$$

The result of surface height of examples 1 is plotted in Figure 5.7. Comparing Y_1 to Y_2 and Y_3 to Y_4, the surface areas of Y_1 and Y_3 are polished at a faster rate than Y_2 and Y_4, because of their initial surface heights. The graph also shows that Y_1 and Y_2 are nearly the same height toward the end of the process. This similar result is also presented when comparing Y_3 to Y_4. Moreover, surface heights of Y_1 to Y_4 will converse as the polishing process continues.

Figure 5.8 illustrates the relationship between step heights and polishing time of the two-wave model. According to this graph, step heights of low amplitudes are reduced to 50 to 57 nm at a polishing time of 20 seconds, while those of high amplitudes are reduced to 415 to 422 nm. At a polishing time of 60 seconds, step heights of low amplitudes are reduced to 13 to 18 nm and these of high amplitudes are reduced to 200 to 205 nm. After 80 seconds of polishing time, step heights of low amplitudes are exceptionally small and can be disregarded.

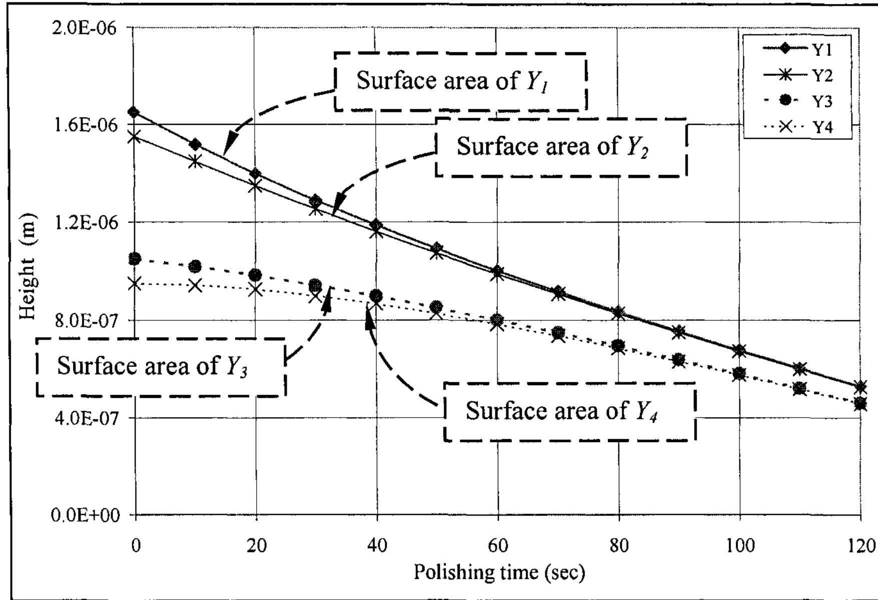


Figure 5.7. Graph between surface heights vs. polishing time of two-wave model in example 1.

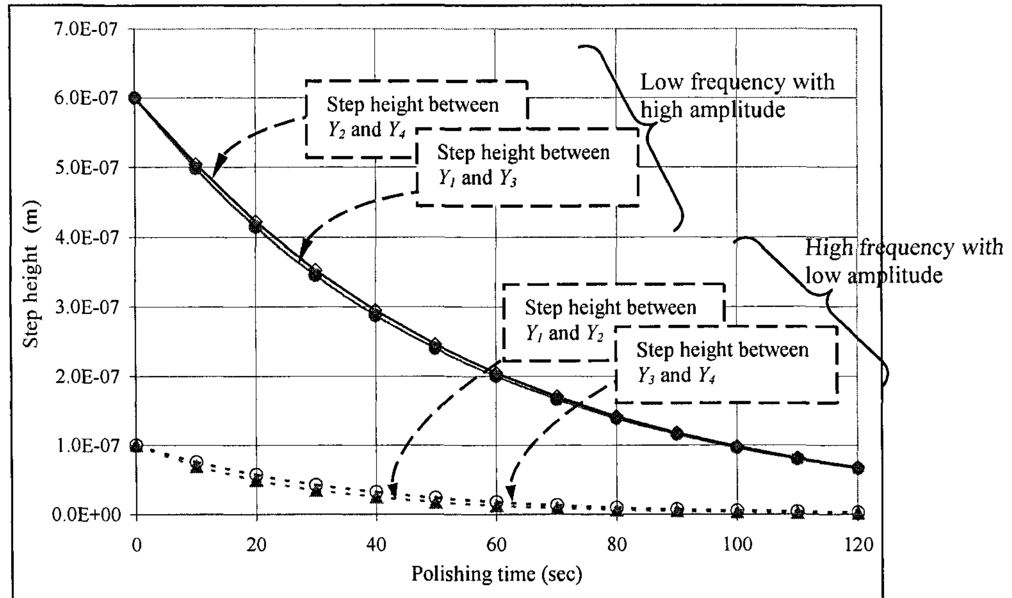


Figure 5.8. Graph between step heights vs. polishing time of two-wave model in example 1.

Surface evolution of example 1 is plotted in Figure 5.9, where the surface heights are shown at the beginning of the polishing process, and at polishing time of 40, 80 and

120 seconds. It shows that step heights of low amplitudes with high frequency are reduced and can be overlooked at 80 seconds of polishing time. Step heights of high amplitude with low frequency also show sign of improvement after 120 seconds; however, the planarization process is still necessary. In addition, step heights with high frequency will reach their limitation before step heights with low frequency. Once the step heights are at their limits, they cannot be reduced indefinitely.

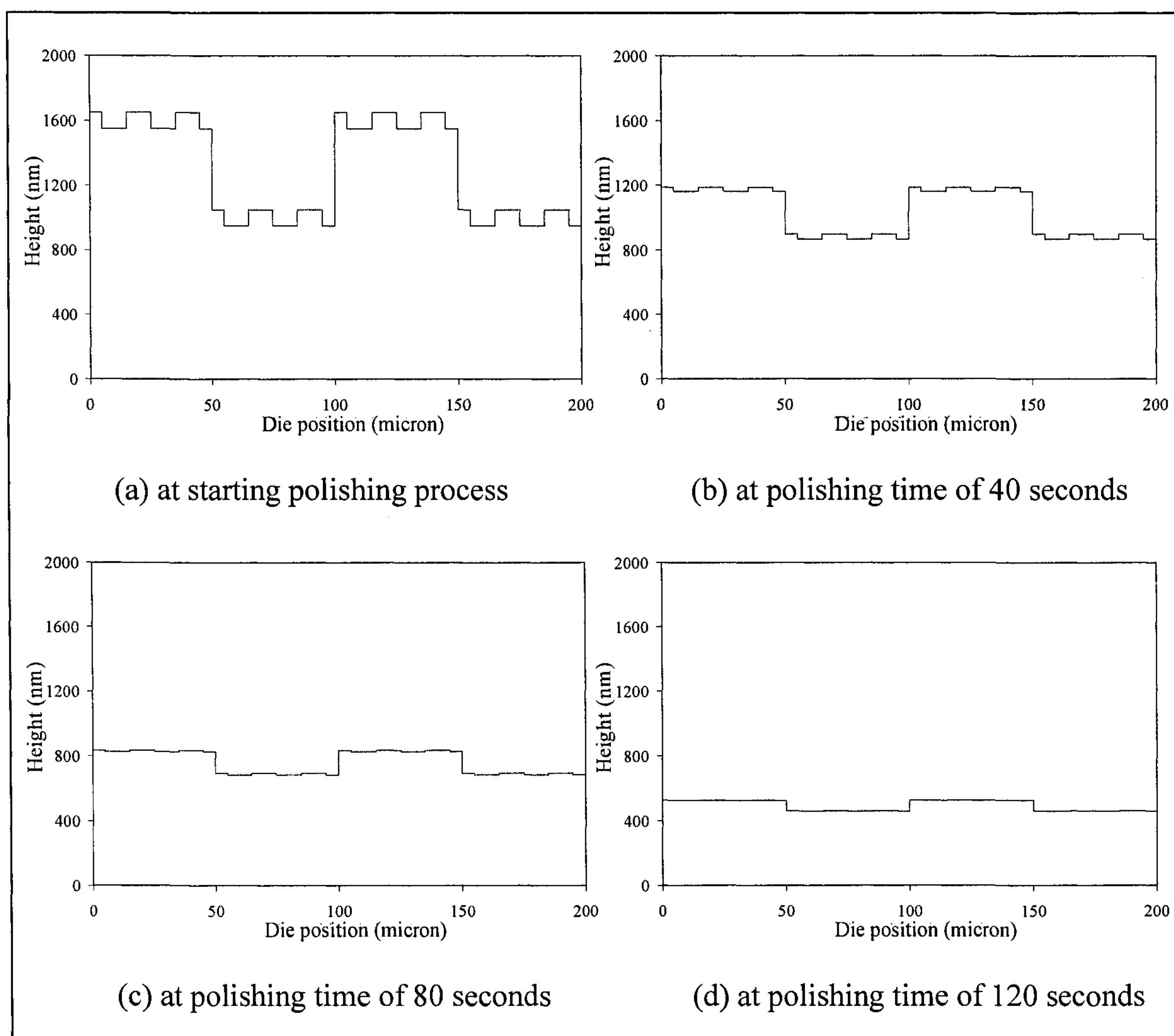


Figure 5.9. Surface evolution of two-wave model in example 1.

Figure 5.10 shows results of example 2, plotted between surface heights and polishing time. The graph shows that the surface heights of Y_1 and Y_2 are fairly similar toward the end the process. Similarly, this result is also presented when comparing Y_3 to Y_4. At a polishing time of 120 seconds, heights of Y_1 and Y_2 are slightly higher than heights of Y_3 and Y_4. Furthermore, as polishing process continues further, surface heights of Y_1 to Y_4 will get closer together.

Figure 5.11 illustrates the relationship between the step height and polishing time of two-wave model. Unlike the results in example 1, step heights with both frequencies tend to get exceptionally small toward the end of the polishing process at the same time. In addition, step heights with high frequency are smaller than the heights with low frequency.

Surface evolution in example 2 is plotted in Figure 5.12, showing at the initial surface, and at 40, 80 and 120 seconds of polishing process. It shows that step heights of high amplitude with high frequency are reduced at faster rate compared to step heights of low amplitude with low frequency. At 120 seconds of polishing time, step heights are improved and surface heights are close together.

In addition, the feature / die-scale model for general wafer surface profile is described in Appendix G.

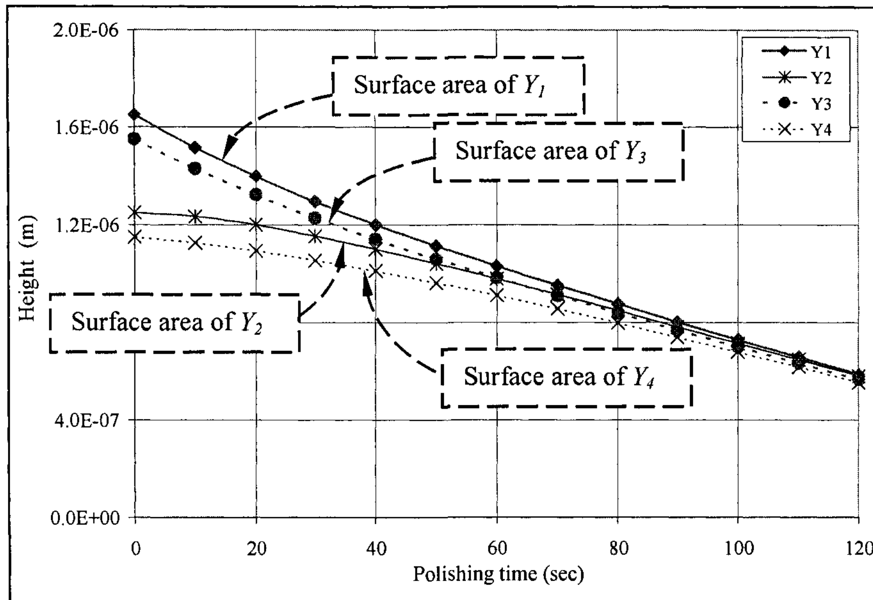


Figure 5.10. Graph between surface heights vs. polishing time of two-wave model in example 2.

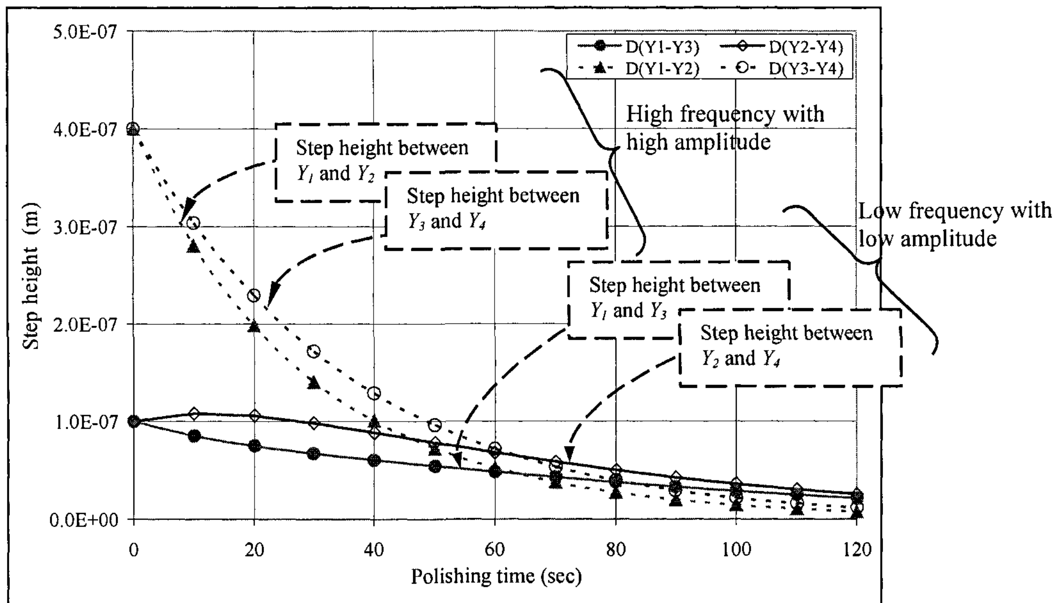


Figure 5.11. Graph between step heights vs. polishing time of two-wave model in example 2.

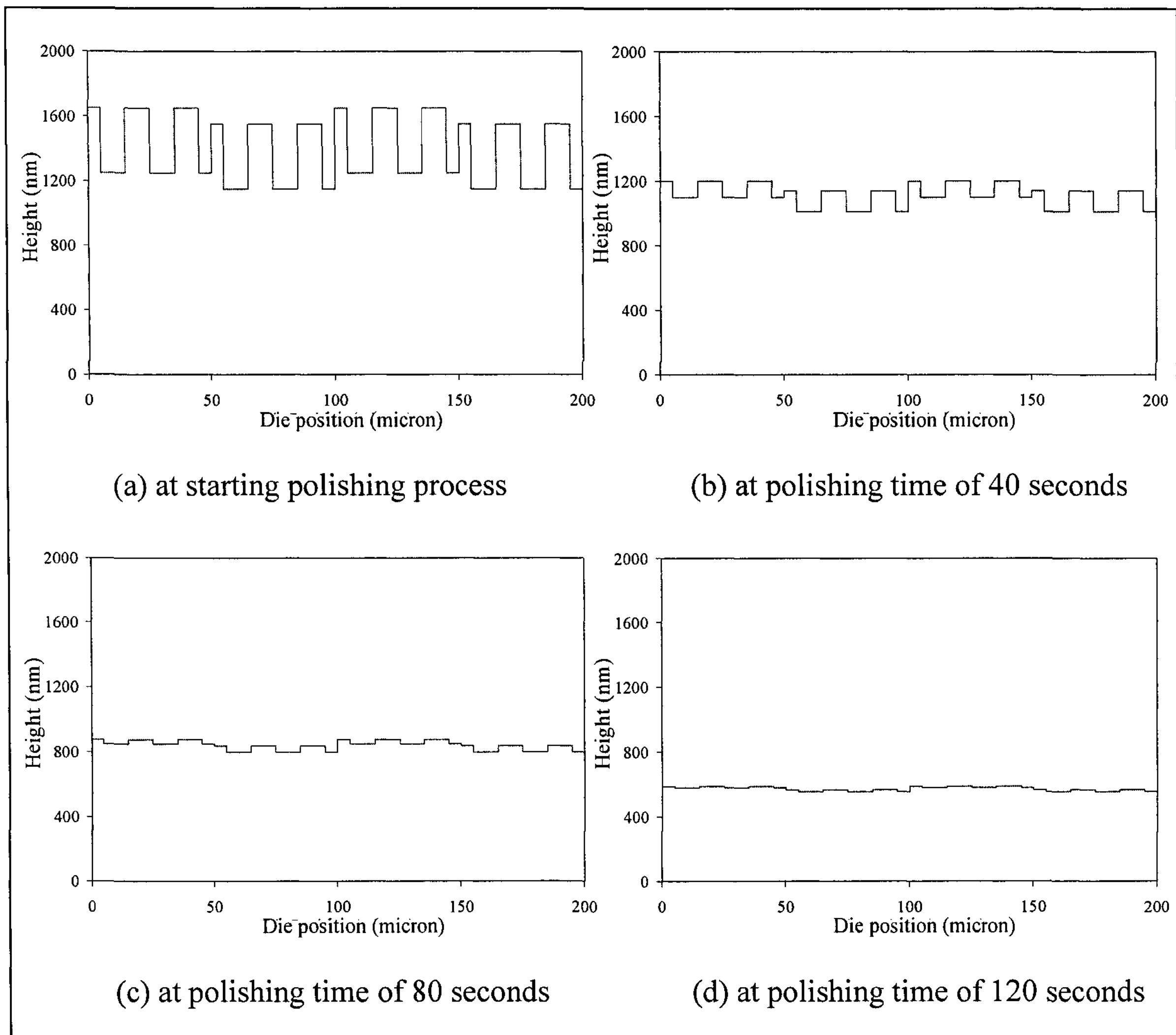


Figure 5.12. Surface evolution of two-wave model in example 2.

CHAPTER 6. SUMMARY AND CONCLUSION

In both the local and global scales, good wafer planarity is essential for dimensional accuracy of the final upper wafer surface. On a global scale, within wafer non-uniformity (WIWNU) in material removal rate (MRR) is a critical parameter in determining the quality of a wafer planarized by a CMP process. In addition, large variation in global thickness across the die has a serious impact on subsequent process steps due to pattern density variation across a chip.

To improve global planarity and wafer yield in a CMP process, this dissertation presents three control strategies based on various interface pressure and wafer curvature for wafer-scale model. This model can be used as a CMP design tool, where it is assumed a uniform pattern density across the entire polish span and a solid-solid contact between wafer and pad. The pad is represented as an elastic half space, indented by a rigid wafer. Interface pressure and fixtured wafer curvature (a_2^1) can be chosen to increase the wafer yield. Furthermore, a genetic algorithm or non-linear programming model shows better wafer yield improvement, when compared with the results of the greedy algorithm.

This dissertation also presents three open loop algorithms using interface pressure as the control parameter to control polishing at feature / die-scale. By applying these algorithms, uniformity over the pattern dependant non-uniformity wafer surfaces in a die-scale model is obtained. Although a spatial pressure control algorithm is able to improve uniformity across the die, better surface uniformity is presented with spatial and temporal pressure control and look-ahead scheduled pressure control algorithms. Due to the difficulty

in varying interface pressure every second in spatial and temporal pressure control, look-ahead scheduled pressure control seems to be the promising algorithm.

6.1. Discussion for Wafer-Scale Model

Based on wafer-scale model of removal rate in Fu and Chandra [8], the performance of three different control strategies in the greedy algorithm is first evaluated. Then the genetic algorithm of curvature control is compared to the stochastic and the greedy algorithm. Lastly, the performance of the non-linear programming of load control is compared to the greedy algorithm.

6.1.1. Comparison of Control Strategies with the Greedy Algorithm

The control schemes in the greedy algorithm are quite sensitive to the value of parameter α chosen in the moment function (MOD). To obtain the best wafer yield in load control, α values of 0.55, 0.58 and 0.54 are selected for initial low, medium and high film curvatures respectively. It means that 54 to 58 percent of the moment function is weighted on wafer curvature, resulting in rapid decrement of interface pressure and correspondingly, the best wafer yield occurs at a delayed time. Based on simulation results in 3.3, a wafer yield of 216 is obtained for an initial low film curvature at a polishing time of 369 seconds, and yields of 224 are obtained for initial medium and high film curvatures at 387 and 376 seconds, respectively.

The selection process of α values is different for the curvature control model. The parameter α of 0.24 is chosen to obtain the best wafer yield for initial low film curvature and α value of 0.34 is selected for initial medium and high film curvatures. Similarly, an α value in combined curvature and load control is 0.34 on initial medium

and high film curvatures, while an α value of 0.26 is selected for initial low film curvature. Moreover, higher wafer yields are found with high initial wafer curvature in both control schemes. According to the simulation results, the best yield of 216 occurs at 108 seconds on initial curvature of -10 x 10^{-6} m^{-1}, a yield of 224 is found on initial curvature of -15 x 10^{-6} m^{-1} at 108 seconds, and when initial curvature is -20 x 10^{-6} m^{-1} the best yield is 240 at 106 seconds.

By comparing the performance of these three control strategies, the wafer yields of curvature control, and combined curvature and load control are slightly better than results from load control. It also shows that their processing times are three times faster. In addition, the better wafer yields are found in the higher initial film curvature, comparing to low initial film curvature. These results are consistent across different control strategies.

6.1.2. Comparison between the Stochastic, the Greedy Algorithm and the Genetic Algorithm of Curvature Control Strategy

Different algorithms in curvature control strategy are compared based on their performances. Table 4.1 shows an application of the strategy on a wafer with a nominal film thickness of 10000 Å. In the greedy algorithm, the best wafer yields of 188, 208 and 288 are shown for initial film curvature of -5e-6 m^{-1}, -10e-6 m^{-1} and -15e-6 m^{-1}, respectively. Comparing to the stochastic method, wafer yield improves from 188 to 208 on initial low film curvature, reduced from 188 to 148 on initial very high film curvature, and no sign of improvement is presented on initial high film curvature.

By comparing the genetic algorithm to the stochastic method and greedy algorithm, it presents better results on any initial film curvatures. Wafer yield of the genetic algorithm is improved from 188 to 208 for initial low film curvature; however, processing time is slightly increased from 135 to 138 seconds, compared to the greedy algorithm. For initial high film curvature, the wafer yield improves from 208 to 224. Similarly, the processing time is increased by 4 seconds. For initial very high film curvature, yield is significantly improved form 188 to 240, while the polishing time is barely increased. Moreover, comparing genetic algorithm to stochastic, yields are improved in initial film curvature of -10e-6 m^{-1} and -15e-6 m^{-1}.

6.1.3. Comparison of the Greedy Algorithm and the Non-Linear Programming Model with Load Control Strategy

Application of load control on a wafer with a nominal film thickness of 10000 Å is presented. For initial very low film curvature, average wafer yield from non-linear programming model is improved from 180 to 208, compared to the greedy algorithm; however, average processing time is increased from 48 to 63 seconds. For initial low film curvature, average yield is also improved from 153 to 177, and average polishing time is increased from 54 to 78 seconds. Moreover, average yield of non-linear programming model is improved from 125 to 141, and average time is increased from 61 to 72 seconds.

In modified non-linear programming model, yield improvements are the same as the original results. The advantage of this modified model over the non-linear

programming model is the ability to more easily control interface pressure in the CMP process.

6.2. Discussion for Feature / Die-Scale Model

Based on the step height model from Fu and Chandra [9], an analytical model on one type of surface material is compared to the experimental data by Ouma et al [16]. Their results show that this model fits fairly well with experimental data. It was also shown that variations in pattern density across a die surface are a major cause for global non-planarities. Therefore, three control algorithms using interface pressure as a control parameter are introduced to improve this circumstance.

6.2.1. Comparison of Control Algorithms using Interface Pressure as Control Parameter

This comparison concentrates on controlling zonal pressure across the die, using a Zonal Process Controller (ZPC). This controller is a pixel-based control to vary local pressure at feature / die-scale levels.

According to Table 6.3, uniformity of the upper surface with different pattern densities is improved significantly by applying spatial pressure control algorithm. An average error of the final upper surface is 2 percent of the desired target upper surface. By gradually varying pressure spatially and temporally, the upper surface uniformity can be controlled effectively, where an average error is shown at 0.06 percent. Furthermore, to be able to apply the second algorithms to the controller unit, the look-ahead scheduled pressure control is derived. In this scheme, zonal pressure is varied with a specified time step for a period of time, based on the desired step height and is then varied spatially

using the spatial pressure control. It is observed from simulation that the upper surface uniformity can also be effectively controlled with an average error of 0.16%. However, by varying the zonal pressure both in space and in time, improving the local step height is still questionable.

6.2.2. Model for General Wafer Surface

To formulate the feature / die-scale model for general wafer surface profiles, a simplified representation with two highest amplitudes is first attempted. It is observed that this two-wave model fits accurately to Fu's single-wave model. In addition, simulation results show that when applying the pressure for a long period of time, the step height will eventually reach its limitation and cannot reduce indefinitely.

Example 1 of a two-wave model in 5.3.2 illustrates that when the amplitude of high frequency is lower than the amplitude of low frequency, step heights of high frequency are reduced and reach their limitations, while step heights of low frequency are improved and planarization process is needed. In opposition, example 2 shows that when the amplitude of high frequency is higher than the amplitude of low frequency, step heights of both frequencies are exceptionally small toward the end of polishing process and finally reach their limitations at the same time.

6.3. Conclusion

Control algorithms for yield improvement in CMP are developed using interface pressure and fixtured wafer curvature as control parameters. Greedy algorithms, genetic algorithms, and optimization models are studied. The greedy algorithm is based on minimizing a moment function that incorporates curvature of the film layer on the wafer

surface, as well as the final desired oxide layer thickness. At each step of the CMP process, the moment function is minimized. The genetic algorithm calculates the best fitness value that incorporates the number of good sectors and polishing time. This algorithm is used for a series of fixtured wafer curvatures to control the CMP process. Lastly, the optimization model is derived, based on minimizing a moment function. Unlike the greedy algorithm, a moment function is minimized only at the end of the CMP process. The best wafer yield is shown with the optimization model, when comparing the results among these algorithms.

The feature / die-scale model explains the effects of applied pressure and pattern density on material removal rate. It is obvious that the upper surface height can be controlled as a function of applied pressure in pattern density variation. In this concept, three control strategies are developed and studied. The strategies are included spatial pressure control, spatial and temporal pressure control, and look-ahead scheduled pressure control. These mechanisms are developed based on modifying the applied pressure across the die over different pattern densities, resulting in improvement of the final surface uniformity. The simulation results of these three strategies show improvement in the upper surface uniformity; however, look-ahead scheduled pressure control seems to be the promising algorithm.

6.4. Future Work

Although better yield improvement is presented in the non-linear programming model, there is one more parameter needed to be considered. It is the weighting term in the objective function of the non-linear programming model that can be varied when considering yield improvement corresponding to polishing time.

The control parameters, interface pressure and fixtured curvature, are required to obtain better wafer yield; however, the actual details of implementation of the curvature control still needs to be work out. To control curvature, a fixturing device or spatially varying load distribution to alter this curvature is needed.

Finally, the featured / die-scale model control strategies are simulated; the results prove that there is a significant improvement in surface uniformity. Therefore, these strategies should be applied in the realistic environment, such as on a MIT mask with Zonal Process Controller (ZPC) to further improve the model as well as control strategies.

APPENDIX A. LEAST SQUARE CURVE FITTING OF DISCRETE POINT

$F(r_i) = a_0 + a_2 r_i^2$

a_0 is the vertical displacement of the wafer (depth of penetration).

a_2 is the wafer curvature caused by the preexisting wafer bow.

r_i measures the radial distance from the center of the wafer.

n is the number of discrete points.

$$\begin{bmatrix} n & \sum_{i=1}^{n} r_i^2 \\ \sum_{i=1}^{n} r_i^2 & \sum_{i=1}^{n} r_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} F(r_i) \\ \sum_{i=1}^{n} \left(r_i^2 F(r_i)\right) \end{bmatrix}$$

$$\begin{bmatrix} a_0 \\ a_2 \end{bmatrix} = inv\left(\begin{bmatrix} n & \sum_{i=1}^{n} r_i^2 \\ \sum_{i=1}^{n} r_i^2 & \sum_{i=1}^{n} r_i^4 \end{bmatrix}\right) \cdot \begin{bmatrix} \sum_{i=1}^{n} F(r_i) \\ \sum_{i=1}^{n} \left(r_i^2 F(r_i)\right) \end{bmatrix}$$

APPENDIX B. THE FILM CURVATURE OF THE WAFER AT ANY TIME

$F(r_i,t) = a_0^t + a_{22}^t r_i^2$

Where, a_0^t is the vertical displacement of the wafer (depth of penetration to pad).

a_{22}^t is the film curvature at any time, t.

r_i is the distance measured from the center of the wafer.

By applying the least square curve fitting of discrete points (refer Appendix A.), we have

$$\begin{bmatrix} a_0^t \\ a_{22}^t \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} r_i^2 \\ \sum_{i=1}^{N} r_i^2 & \sum_{i=1}^{N} r_i^4 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{i=1}^{N} Z(r_i) \\ \sum_{i=1}^{N} \{Z(r_i).r_i^2\} \end{bmatrix}$$

$$\begin{bmatrix} a_0^t \\ a_{22}^t \end{bmatrix} = \frac{1}{\left(N.\sum_{i=1}^{N} r_i^4 - \left[\sum_{i=1}^{N} r_i^2\right]^2\right)} \cdot \begin{bmatrix} \sum_{i=1}^{N} r_i^4 & -\sum_{i=1}^{N} r_i^2 \\ -\sum_{i=1}^{N} r_i^2 & N \end{bmatrix} \cdot \begin{bmatrix} \sum_{i=1}^{N} Z(r_i) \\ \sum_{i=1}^{N} \{Z(r_i).r_i^2\} \end{bmatrix}$$

$$a_{22}^t = \frac{\left(N.\sum_{i=1}^{N} \{Z^t(r_i).r_i^2\} - \sum_{i=1}^{N} Z^t(r_i).\sum_{i=1}^{N} r_i^2\right)}{\left(N.\sum_{i=1}^{N} r_i^4 - \left[\sum_{i=1}^{N} r_i^2\right]^2\right)} \qquad [B.1]$$

Where, N is the number of discrete points.

$Z^t(r_i) = Z^{t-1}(r_i) - K_p KVP_t NP(a_0, a_2, r_i)$.

$$Z^t(r_i) = Z^0(r_i) - K_p KV \sum_{j=1}^{t} \{P_j NP(a_0^j, a_2^j, r_i)\}. \qquad [B.2]$$

$$NP\left(a_0^j, a_2^j, r_i\right) = \frac{4a_2^j r_i^2 + \left(a_0^j - 2a_2^j a^2\right)}{\left(a_0^j - 2a_2^j a^2\right)\sqrt{1 - \left(\frac{r_i}{a}\right)^2}}$$, referred to equation 3 (normalized distribution pressure).

$$P_j = \left(\frac{1}{\pi} \cdot \frac{E}{1 - \nu^2} \cdot \frac{a_0^j - 2.a_2^j.a^2}{a}\right).$$

$$a_0^j = 2.a_2^j.a^2 + \frac{a.\pi.\left(1 - \nu^2\right)}{E}.P_j.$$

Where, E is the pad Young's modulus, and

ν is poisson ratio.

$$NP\left(a_0^j, a_2^j, r_i\right) = \frac{4a_2^j r_i^2 + \left(\frac{a.\pi.\left(1 - \nu^2\right)}{E}.P_j\right)}{\left(\frac{a.\pi.\left(1 - \nu^2\right)}{E}.P_j\right)\sqrt{1 - \left(\frac{r_i}{a}\right)^2}}$$

$$\begin{aligned} a_2^j &= a_{21} + a_{22}^{t-1} \\ &\approx a_{21} \end{aligned}$$

$$NP\left(a_0^j, a_2^j, r_i\right) = \frac{4a_{21} r_i^2 + \left(\frac{a.\pi.\left(1 - \nu^2\right)}{E}.P_j\right)}{\left(\frac{a.\pi.\left(1 - \nu^2\right)}{E}.P_j\right)\sqrt{1 - \left(\frac{r_i}{a}\right)^2}} \qquad [B.3]$$

Thus, the film curvature of the wafer at any time can be expressed as equation B.1, together with equation B.2 and equation B.3.

APPENDIX C. EXAMPLE PROGRAM OF NON-LINEAR PROGRAMMING USING LINGO SOFTWARE

```
!Wafer-Scale Non-Linear Programming Model for Pressure Control;

model:
sets:
    point/ 1 .. 10 /:r, y0;
    time/ 1 .. 61 /:p, a22, MOD;
    pair(time,point):NP;
end sets

data:
    N = 10;                     !Number of Point;
    alpha = 0.3;                        !Alpha;
    Kp = 0.00000000112;         !Preston Constant;
    K = 1;                      !Constant;
    V = 35;                     !Velocity (rpm);
    E = 29700000;               !Pad Modulus (Pa);
    poisson = 0.3;              !Pad Poisson Ratio;
    a = 0.1;                    !Wafer Radius (m);
    Pmin = 1;                   !Minimum Pressure (psi);
    Pmax = 6;                   !Maximum Pressure (psi);
    Pi = 3.14159265359;                 !Pi Value;
    offset = 0.17;                      !Offset Value b/w two centers (m);
    conv = 6894.76;             !Pa from 1 psi;
    hO = 0.000001;              !Initial Surface at the center (m);
    hA = 0.0000008;             !Target Surface (m);
    a21 = -0.002;                       !Fixtured Curvature (m^-1);
    a220 = -0.00002;            !Initial Oxide Curvature (m^-1);
end data
```

```
Min =  @min(time(j): MOD(j));

!Here is interface pressure control take place;
@for(time(j):
        @BND(Pmin, p(j), Pmax));

!Hers is to make the pressure control smooth;
@for(time(j)|j #LE# 60:
        @ABS((p(j+1)-p(j))/p(j)) <= 0.1);

!Calculation;
@for(time(j):
        MOD(j) = @ABS((1-alpha)*(@sum(point(i):(hO+a22(j)*@sqr(r(i))-
                  CA*@sum(time(jj)|jj #LE# j:p(jj)))*r(i))-@sum(point(i):hA*r(i)))) -
                  alpha*a22(j)*@sum(point(i):@pow(r(i),3));
        a22(j) = (N*(@sum(point(i):(y0(i)-CA*@sum(time(jj)|jj
                 #LE#j:p(jj)*NP(jj,i)))*@sqr(r(i))))-(@sum(point(i):(y0(i)-
                 CA*@sum(time(jj)|jj
                 #LE#j:p(jj)*NP(jj,i)))))*@sum(point(i):@sqr(r(i))))/(N*@sum(point(i):@p
                 ow(r(i),4)-@sqr(@sum(point(i):@sqr(r(i)))));
        @free(MOD);
        @free(a22));

@for(pair(j,i):
        NP(j,i) = ((4*a21*@sqr(r(i)))+(CB*p(j)*conv))/(CB*p(j)*conv*@sqrt(1-
                 @sqr(r(i)/a))));

        CA = Kp*K*(2*Pi/60*V*(offset+a));
        CB = a*Pi*(1-@sqr(poisson))/E;
```

```
@for(point(i):
    r(i) = a*(i-1)/N;
    y0(i) = hO+a220*@sqr(r(i)));

end
```

APPENDIX D. GRAPH BETWEEN MODEL PREDICTION AND OUMA'S EXPERIMENTAL DATA

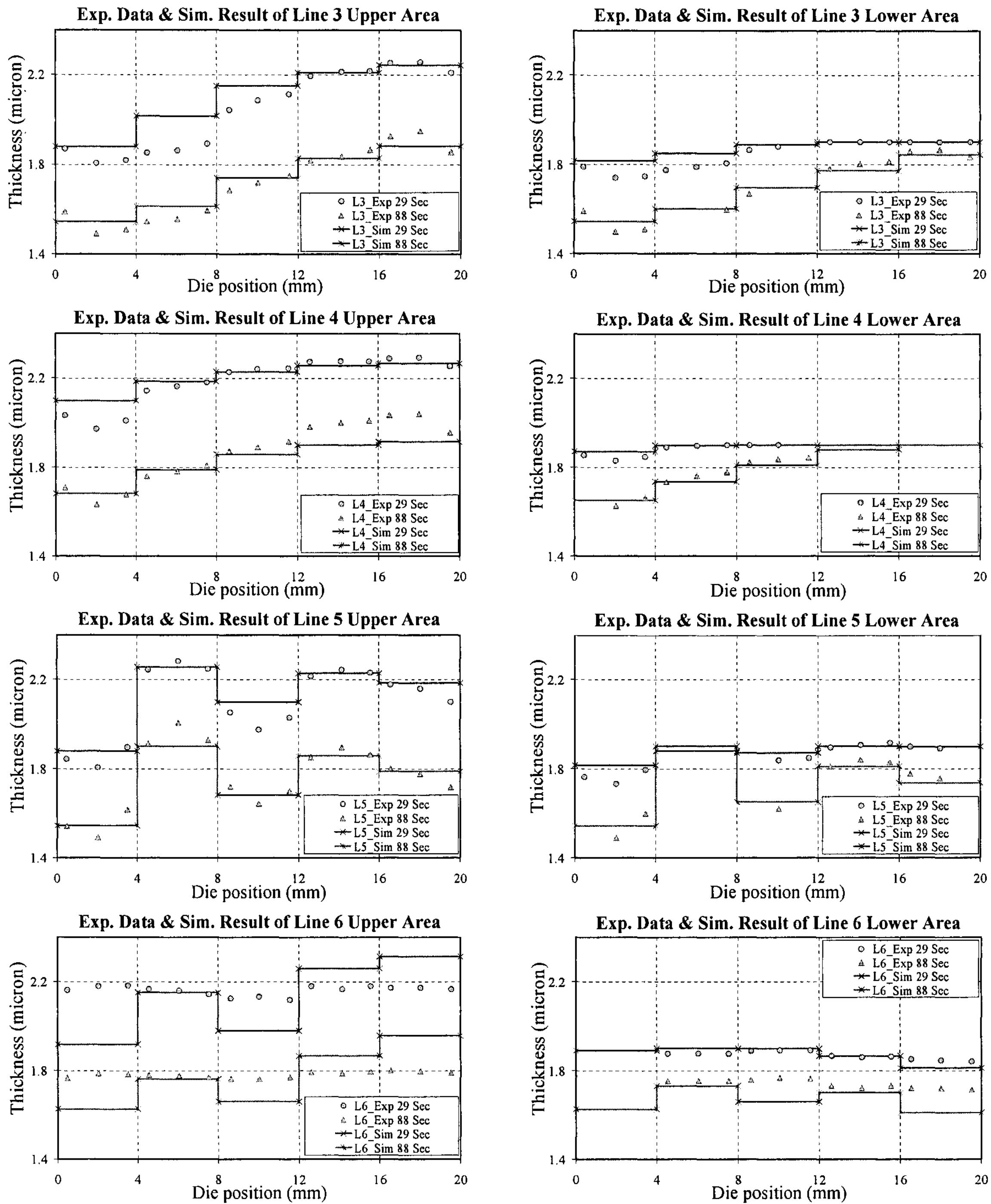


Figure D.1. Graph between thickness vs. die position of simulation results and experimental data (line 3, 4, 5 and 6) with bending factor of 5.0508e6 N/m^2.

The model predictions and experimental data are then plotted on surface height and die position graphs, both in the upper region (a) and lower region (b), shown in Figure D.1. The upper solid line in each graph represents the model prediction after polishing the surface for 29 seconds, where the model prediction after 88 seconds of polishing time is shown in the lower solid line. In comparison to model predictions, the experimental data is shown. Graphs of Line 3, 4 and 5 indicate that the model predictions are consistent with the experimental data in both upper region and lower region. Unlike Line 3, 4 and 5, the model predictions in Line 6 are not corresponding to the experimental data due to the structure of characterization mask, where this set of data represents five different pitches with the pattern densities of 0.5. This variation in pitch affects material removal rate in both the upper and lower regions of the prediction model; however, the effects are not presented in polishing rate for experimental data by Ouma [16]. According to Warnock [29], the material removal rate is also affected by the size of features (pitches) with the overall average pattern factor still remains the same.

APPENDIX E. LOOK-AHEAD PROCEDURE

Procedure $Look-ahead\left(t_{step}, P_{\min}, P_i^{\max}\right)$

The pressure selection procedure is explained below. And the schematic diagram of this procedure is shown in Figure A1.

- Calculate the step heights $\left(SH_{P1}, SH_{P2}\right)$ after the specific time t_{step} for two interface pressures $\left(P_1, P_2\right)$.

- Calculate another step height $\left(SH_{half}\right)$ after the specific time for interface pressure $\left(P_1 + P_2\right)/2$.

- Compare the step height form step 2 to step 1 and substitute the pressure from step 2 for the pressure of step 1 to get the new (P_1, P_2).

If $SH_{half} < SH_{P1}$ then $P_1 = \left(P_1 + P_2\right)/2$

Otherwise, $P_2 = \left(P_1 + P_2\right)/2$.

- Keep doing step 2 and 3 until $P_2 - P_1 < 0.1 * P_{\min}$ for getting *the minimum possible step height left* $\left(MSH_i^{\min}\right)$.

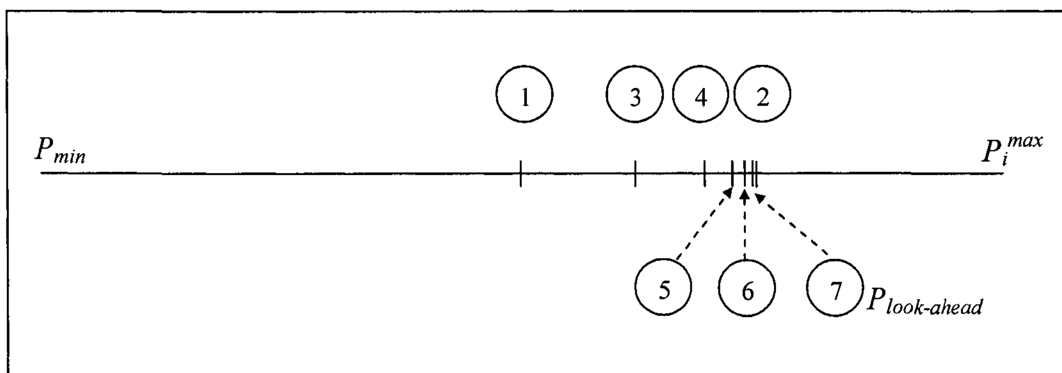


Figure E.1. Pressure selection loop schematic diagram.

Procedure $Look - ahead\left(t_{step}, P_{\min}, P_l^{\max}, MSH_i\right)$

The pressure selection procedure with specific minimum step height is explained below.

- Calculate the step heights $\left(SH_{P1}, SH_{P2}\right)$ after the specific time t_{step} for two interface pressures $\left(P_1, P_2\right)$.
- Calculate another step height $\left(SH_{half}\right)$ after the specific time for interface pressure $\left(P_1 + P_2\right)/2$.
- Compare the step height form step 2 to step 1 and substitute the pressure from step 2 for the pressure of step 1 to get the new (P_1, P_2).

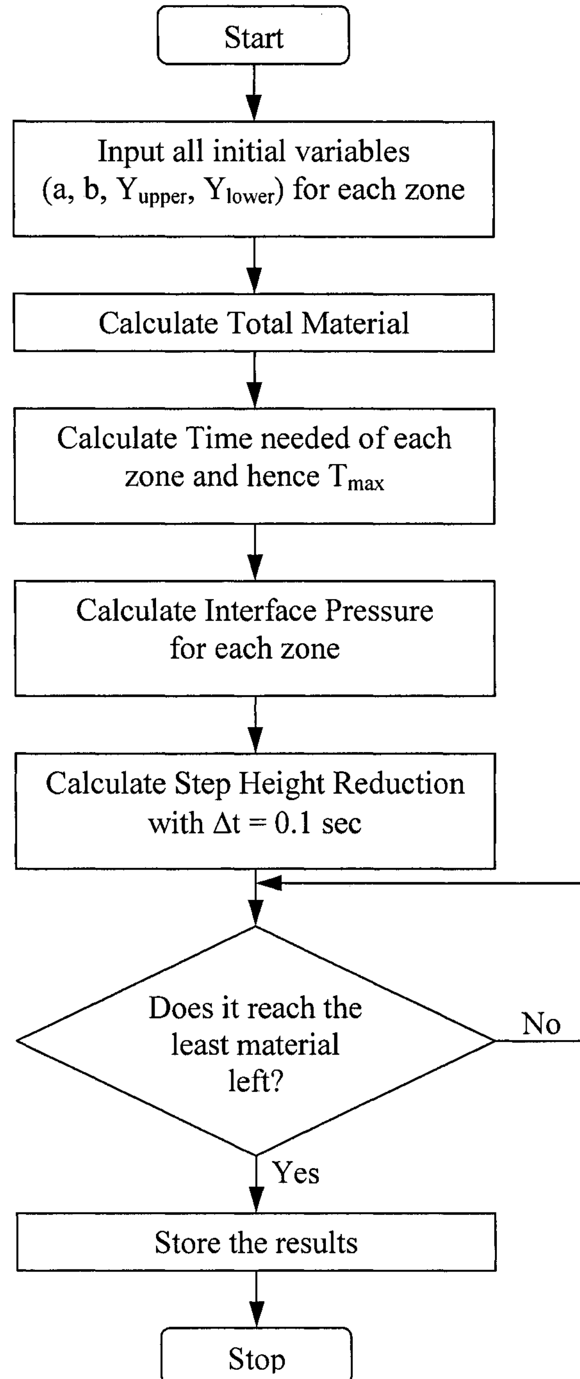
If $SH_{P1} \le MSH_i < SH_{half}$ then $P_1 = \left(P_1 + P_2\right)/2$

Otherwise, $P_2 = \left(P_1 + P_2\right)/2$.

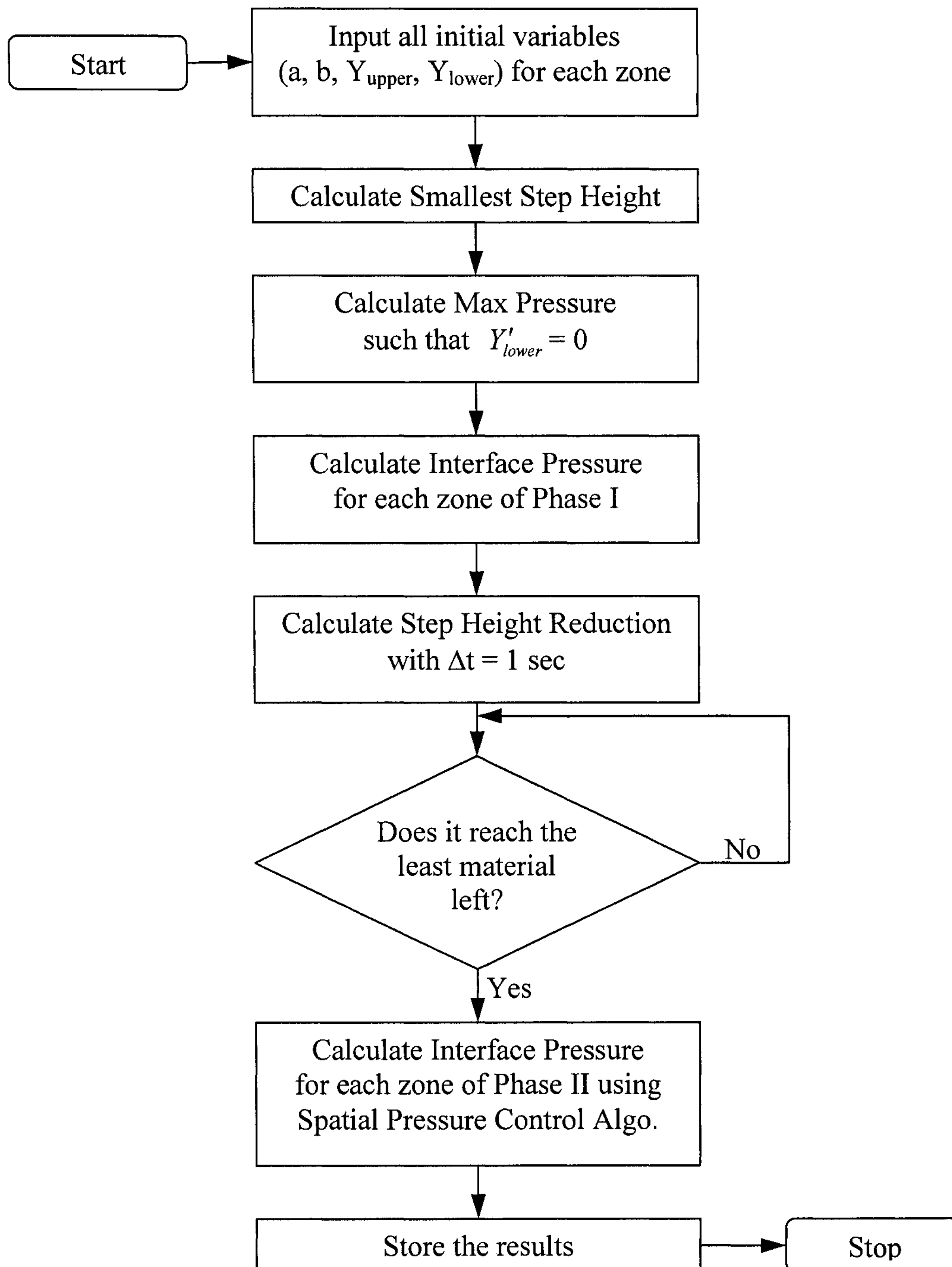
- Keep doing step 2 and 3 until $P_2 - P_1 < 0.1 * P_{\min}$ for getting *the specific step height left* $\left(MSH_i\right)$.

APPENDIX F. FEATURE / DIE-SCALE CONTROL FLOWCHART

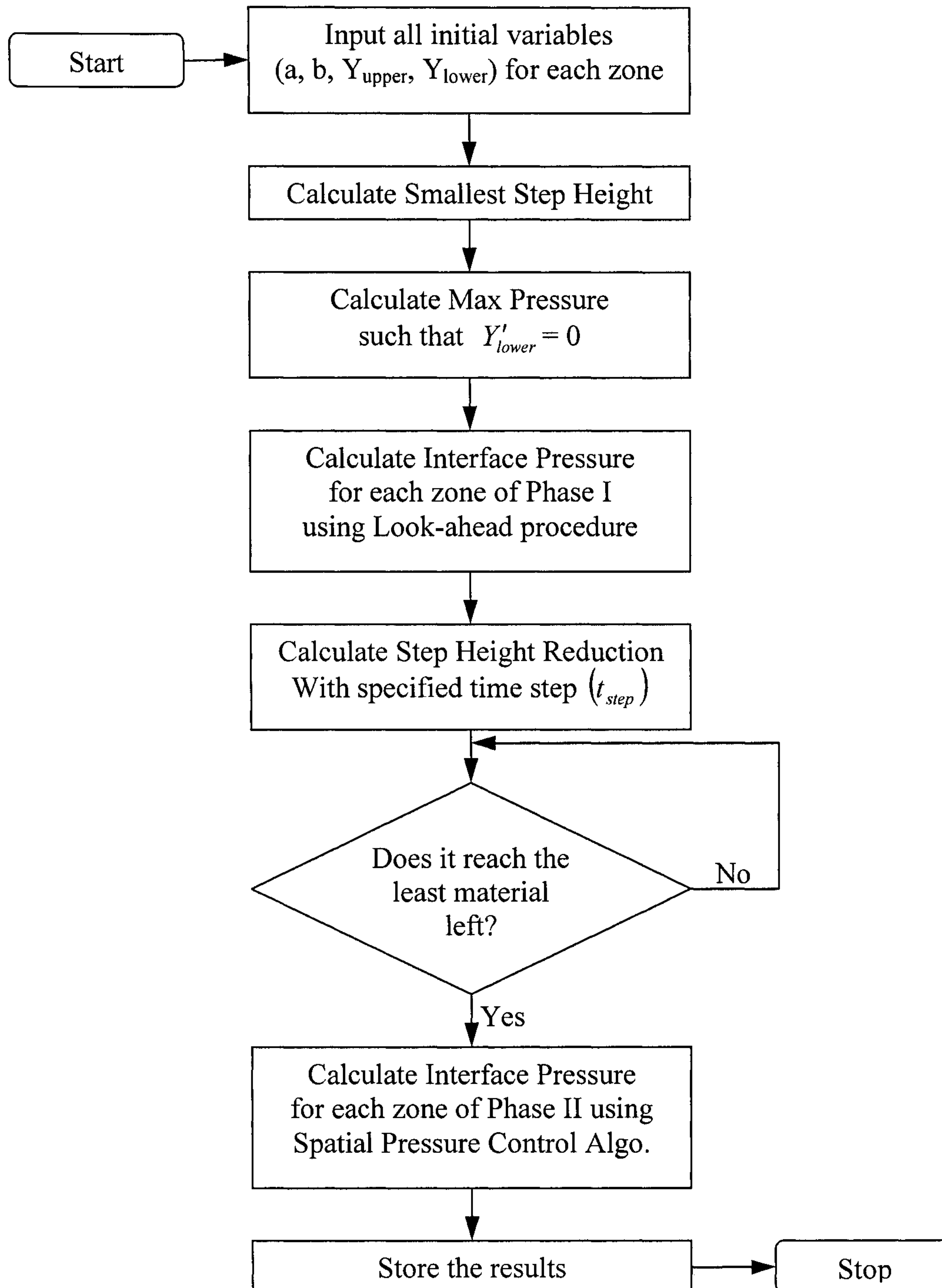
F.1. Spatial Pressure Control : Flowchart



F.2. Spatial and Temporal Pressure Control : Flowchart



F.3. Look-ahead Scheduled Pressure Control : Flowchart



APPENDIX G. SIMULATION RESULTS BASED ON OUMA'S EXPERIMENTAL DATA

G.1. Simulation Result of line 4

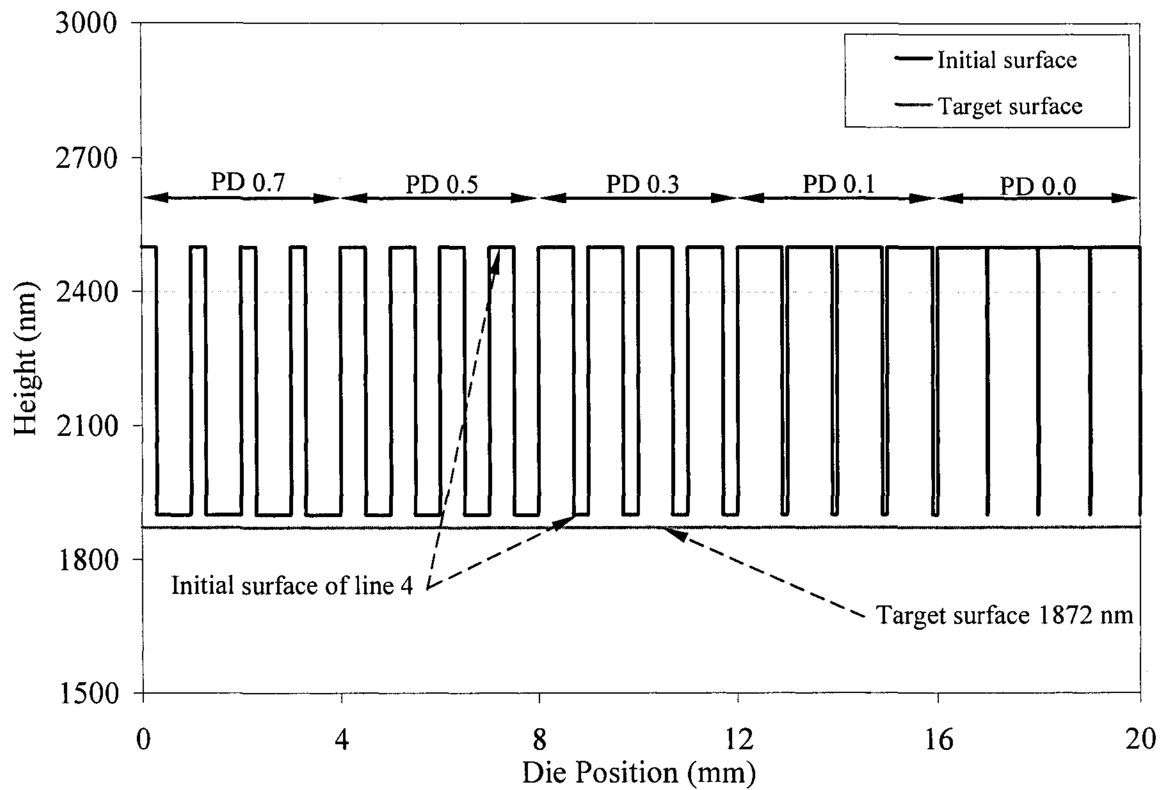


Figure G.1.1. Initial surface profile of line 4 in 5 different zones.

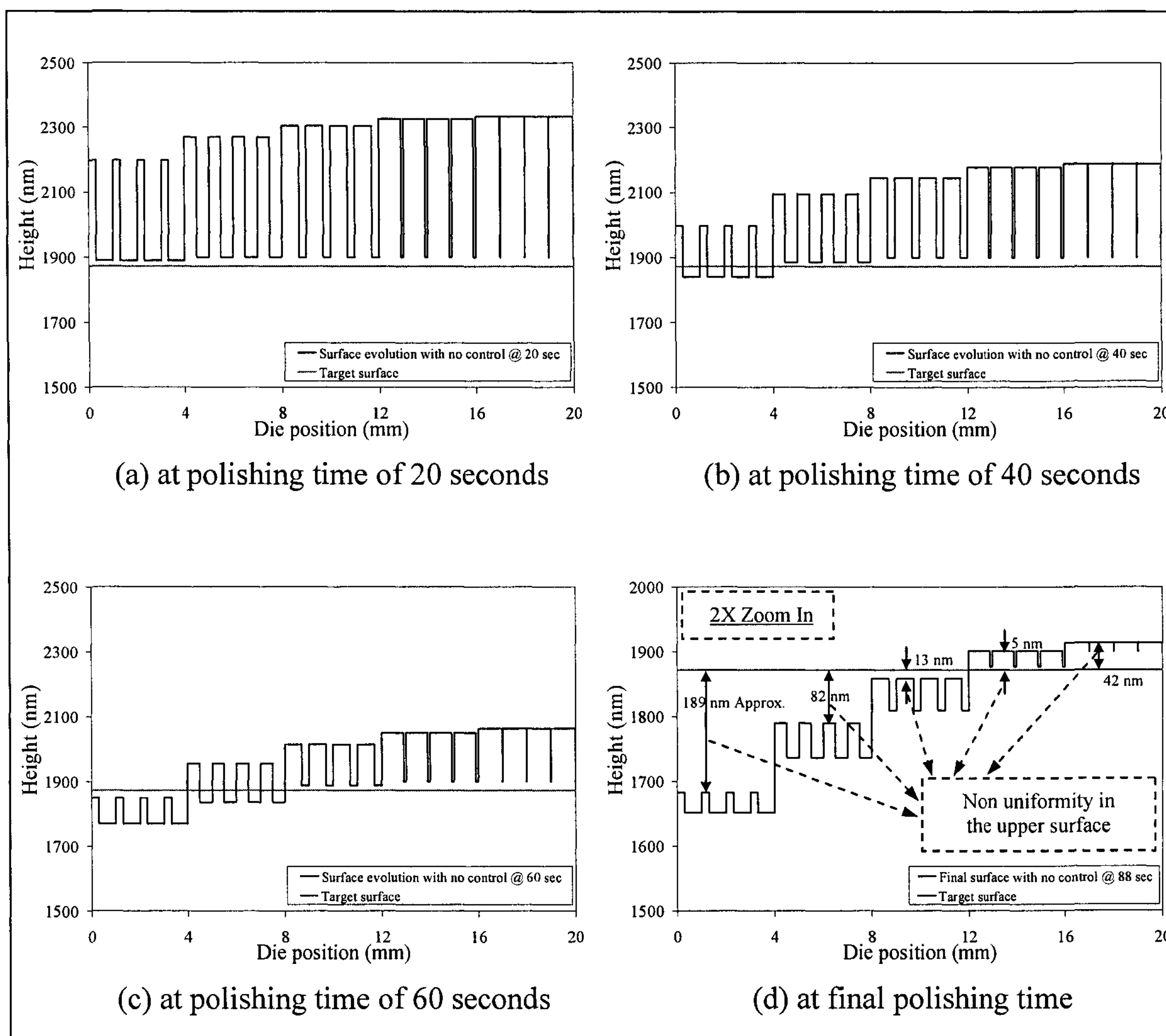


Figure G.1.2. Surface evolution of line 4 vs. die position with no control algorithm.

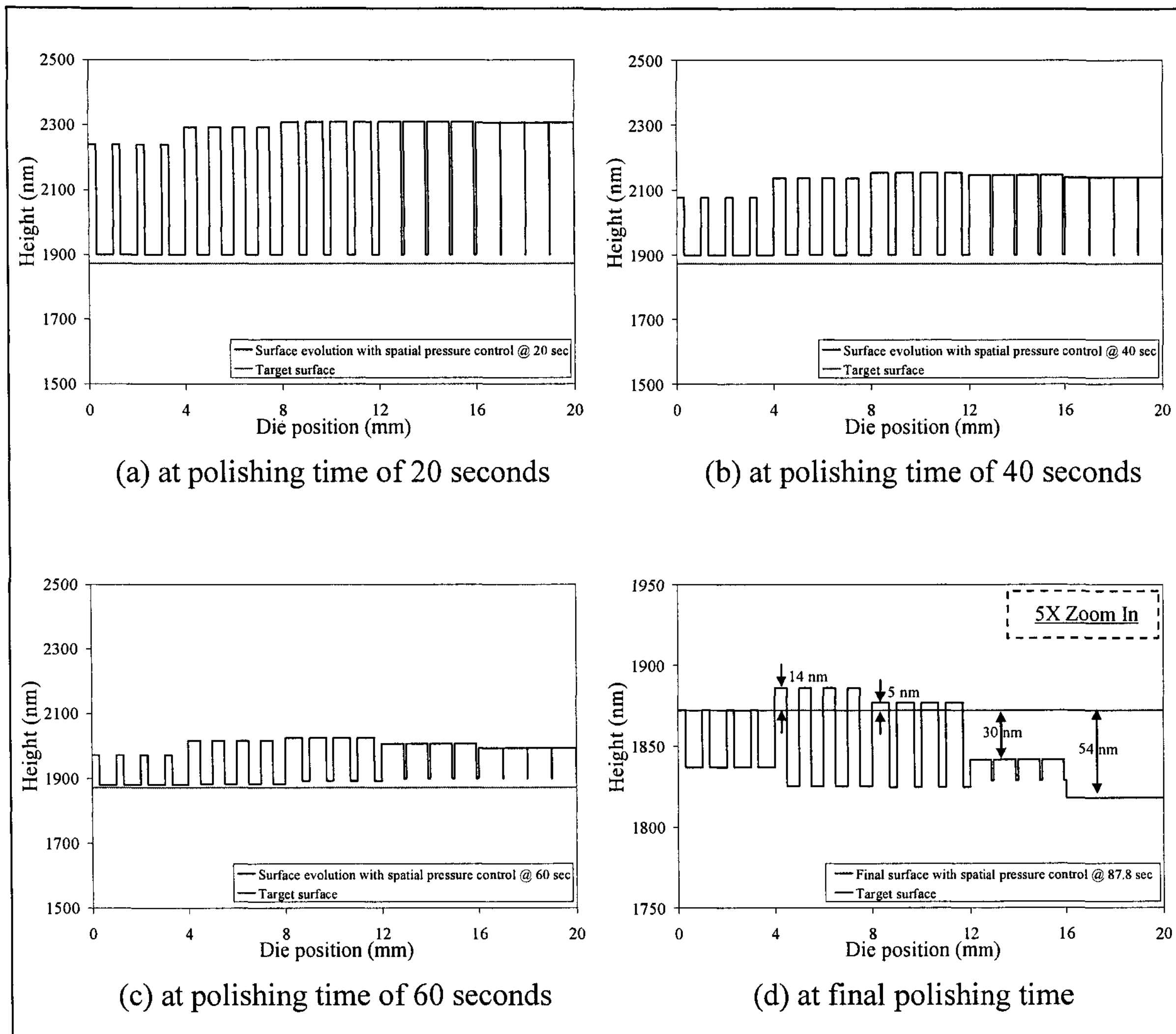


Figure G.1.3. Surface evolution of line 4 vs. die position with spatial pressure control algorithm.

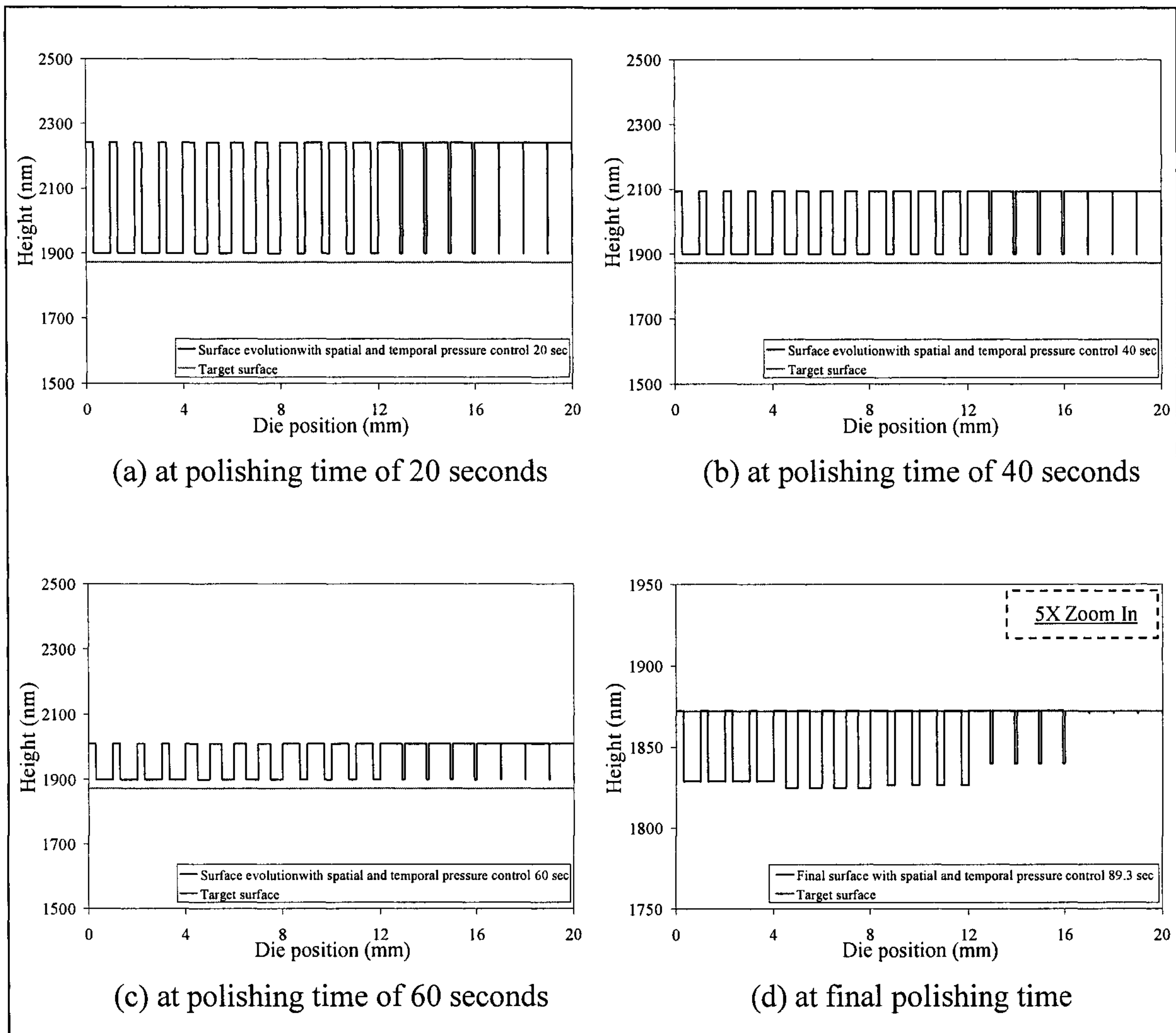


Figure G.1.4. Surface evolution of line 4 vs. die position with spatial and temporal pressure control algorithm.

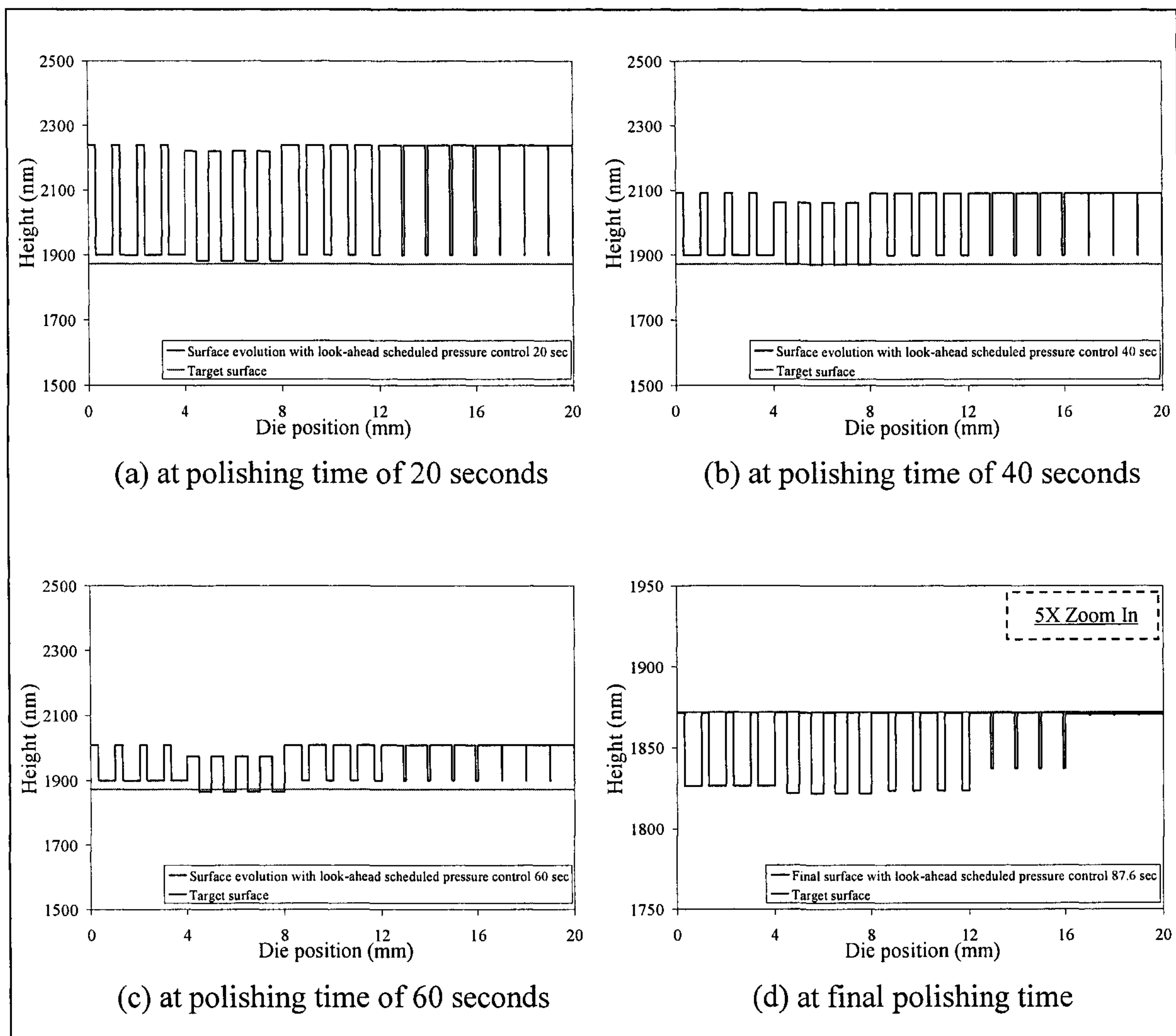


Figure G.1.5. Surface evolution of line 4 vs. die position with look-ahead scheduled pressure control algorithm.

G.2. Simulation Result of line 5

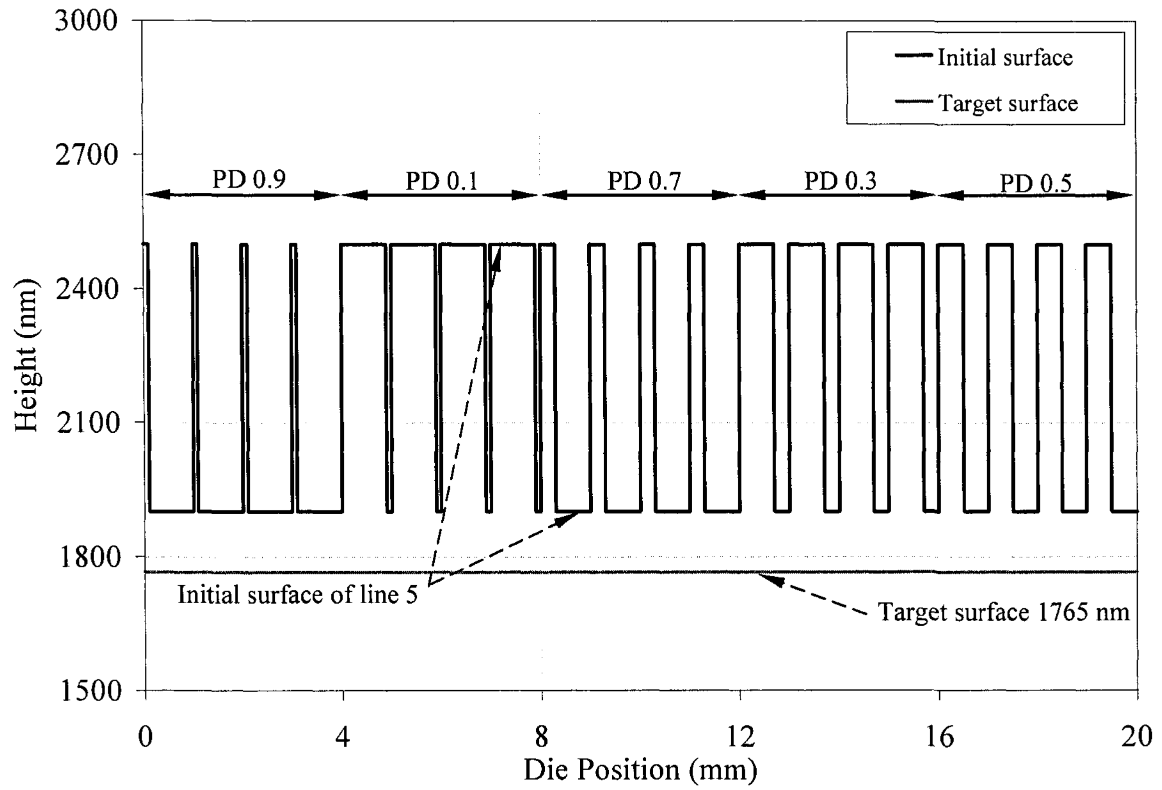


Figure G.2.1. Initial surface profile of line 5 in 5 different zones.

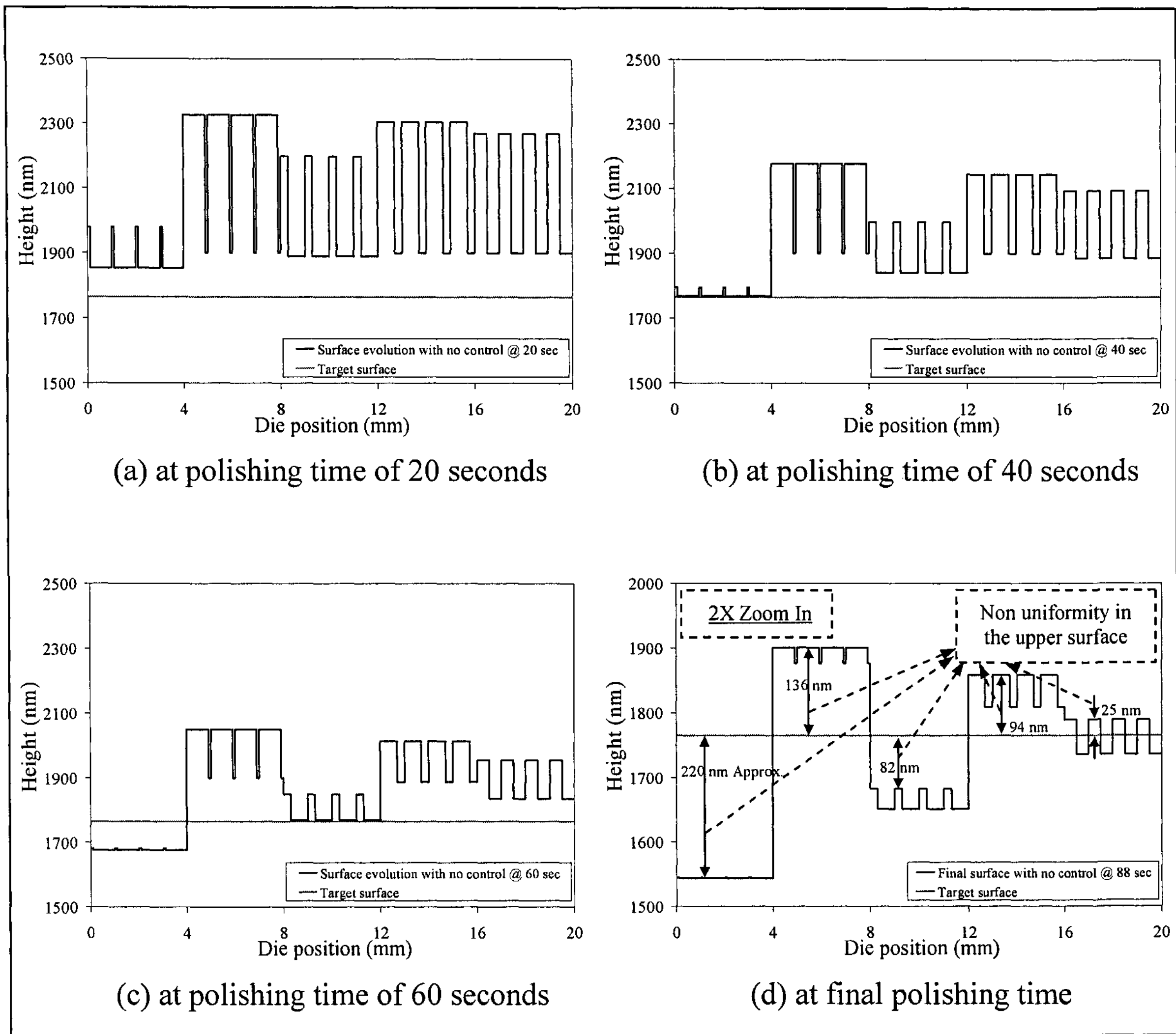


Figure G.2.2. Surface evolution of line 5 vs. die position with no control algorithm.

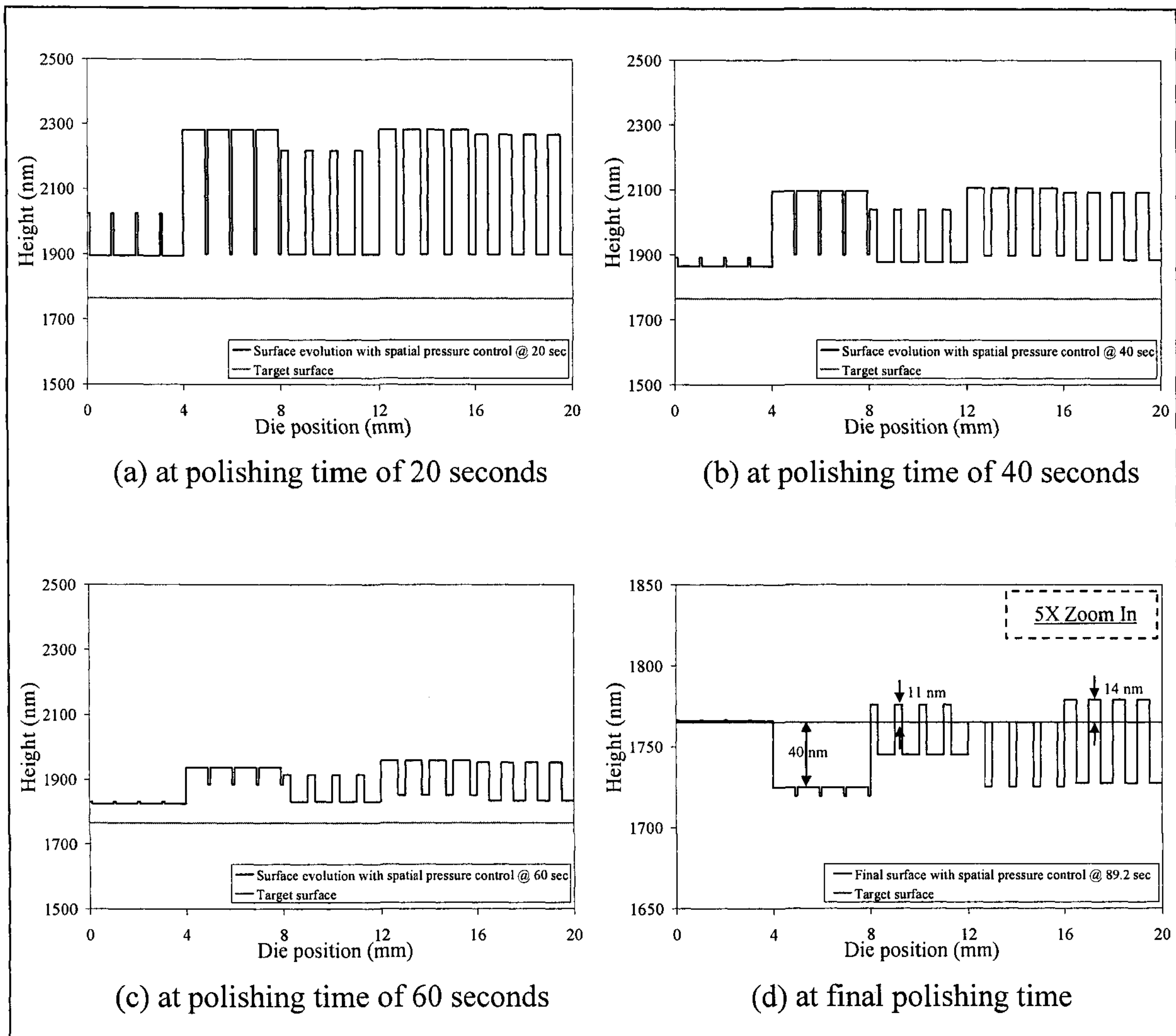


Figure G.2.3. Surface evolution of line 5 vs. die position with spatial pressure control algorithm.

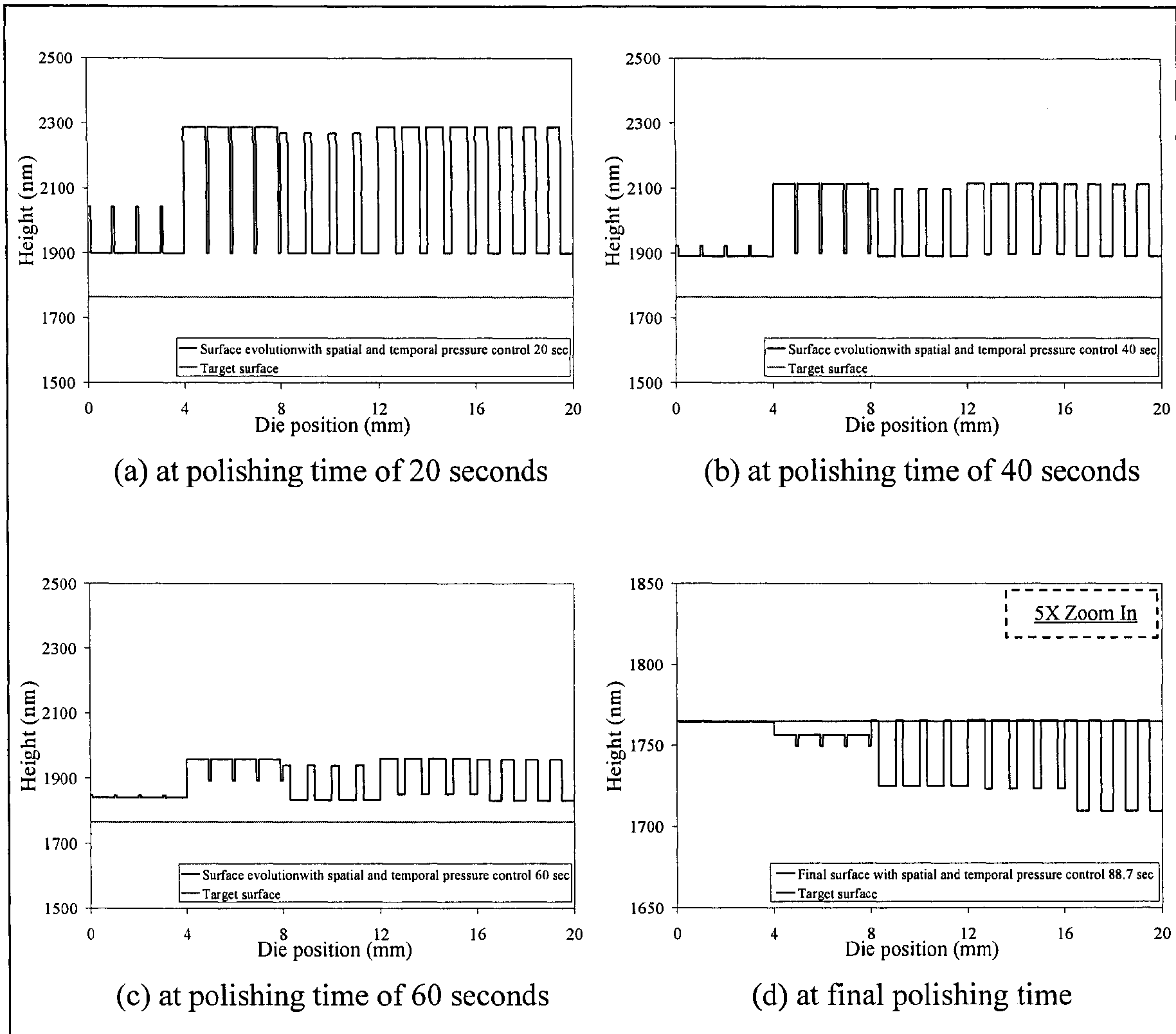


Figure G.2.4. Surface evolution of line 5 vs. die position with spatial and temporal pressure control algorithm.

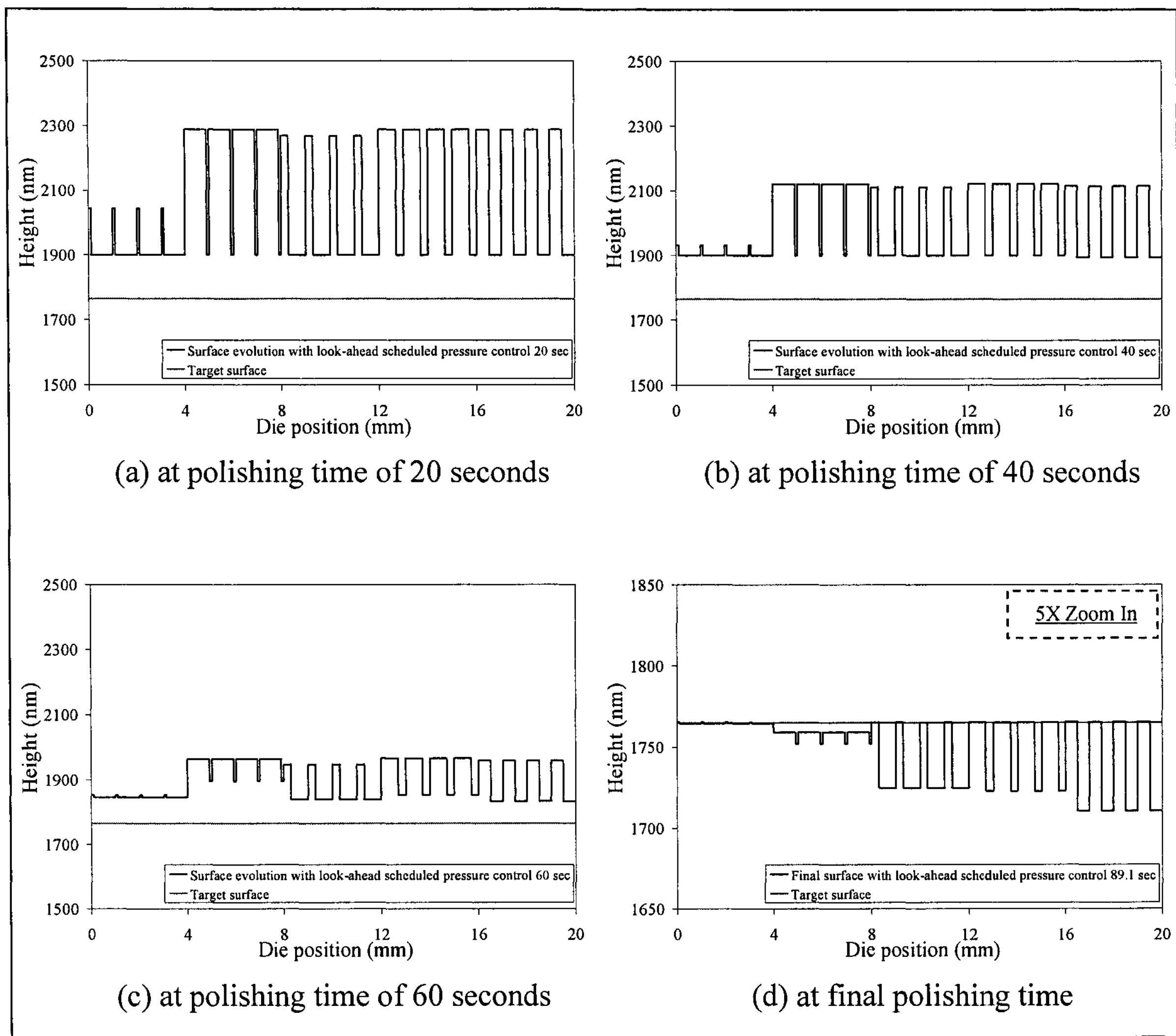


Figure G.2.5. Surface evolution of line 5 vs. die position with look-ahead scheduled pressure control algorithm.

G.3. Simulation Result of line 6

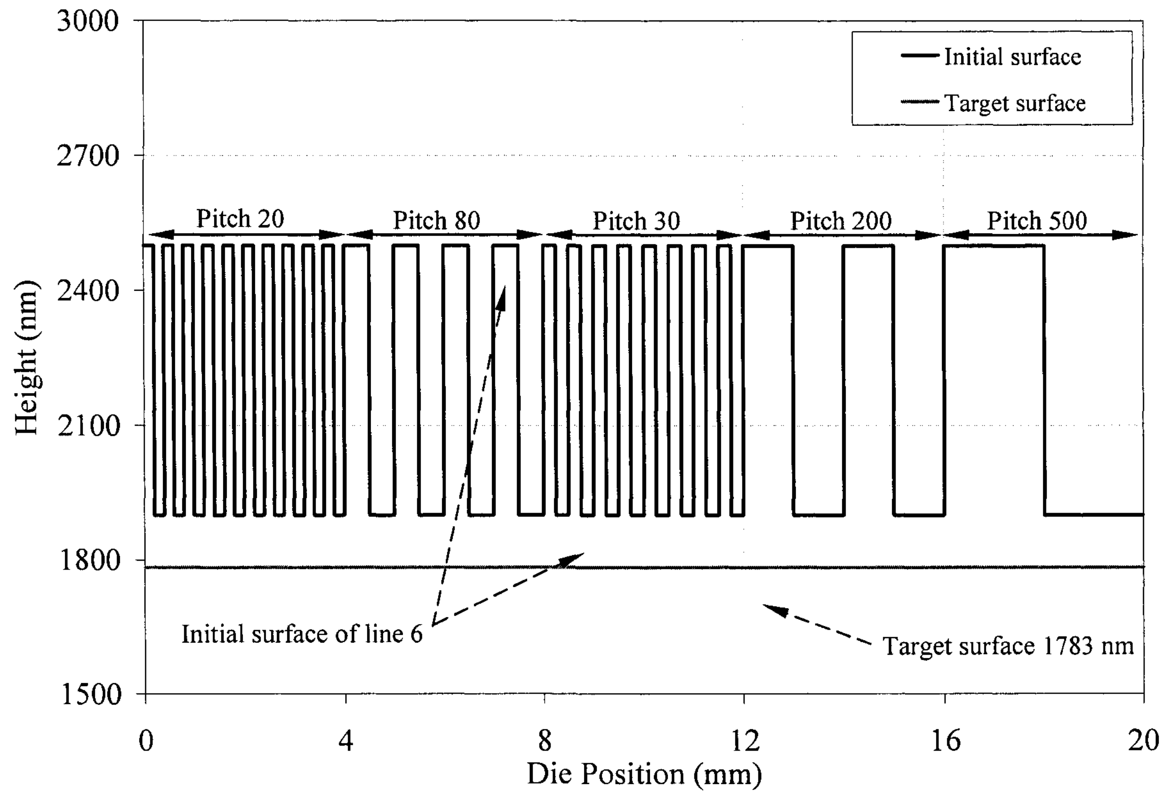


Figure G.2.1. Initial surface profile of line 6 in 5 different zones.

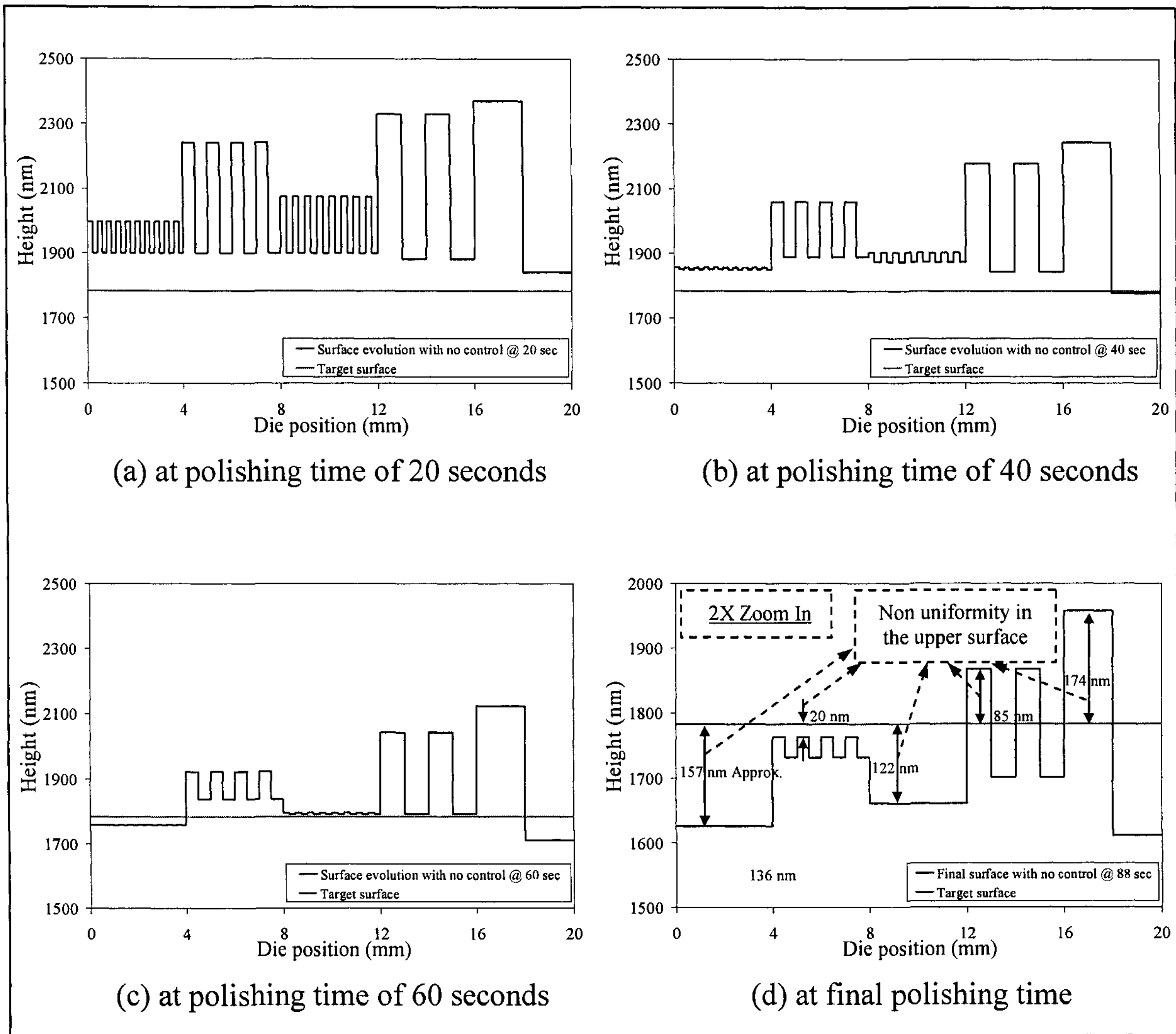


Figure G.3.2. Surface evolution of line 6 vs. die position with no control algorithm.

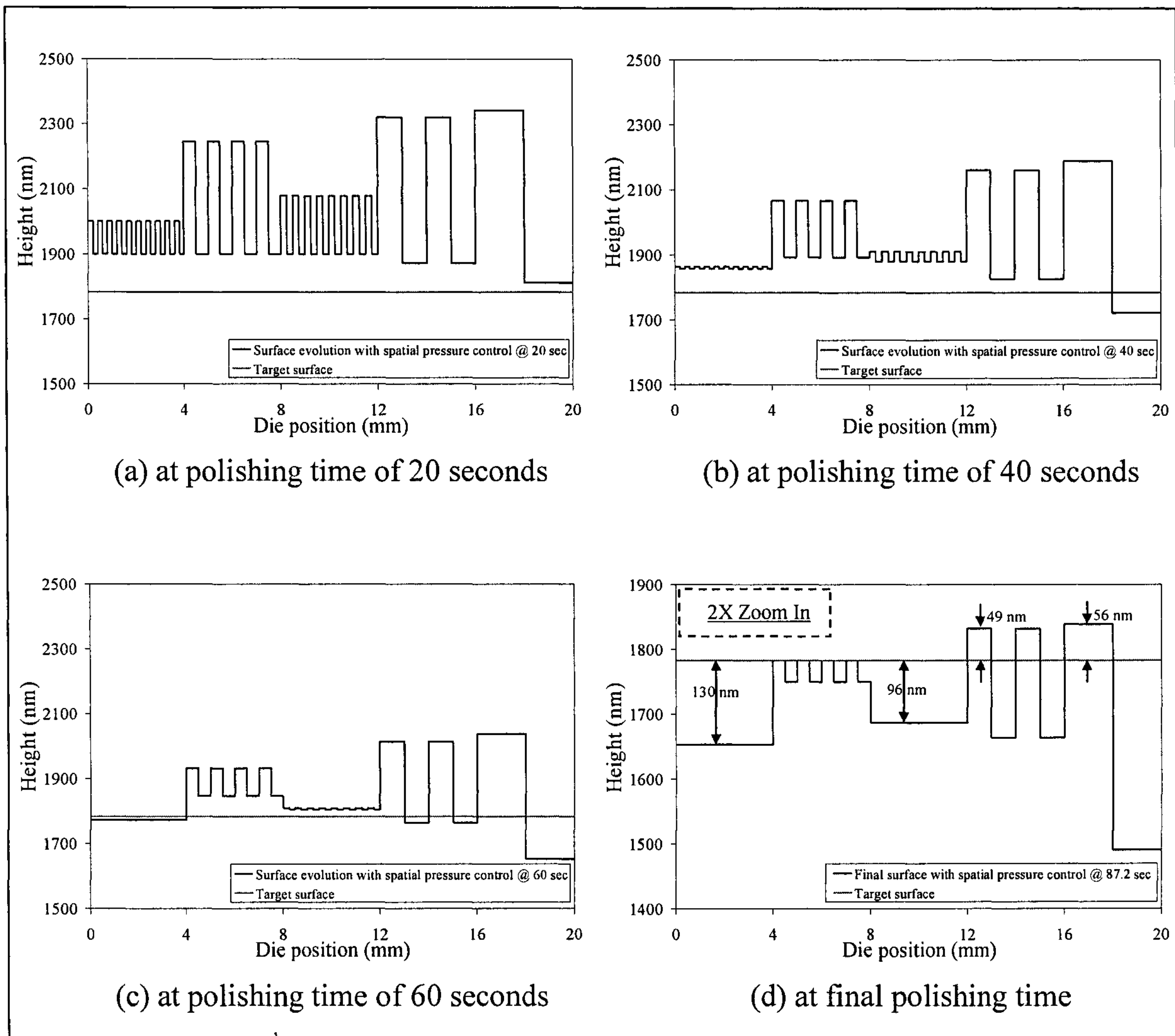


Figure G.3.3. Surface evolution of line 6 vs. die position with spatial pressure control algorithm.

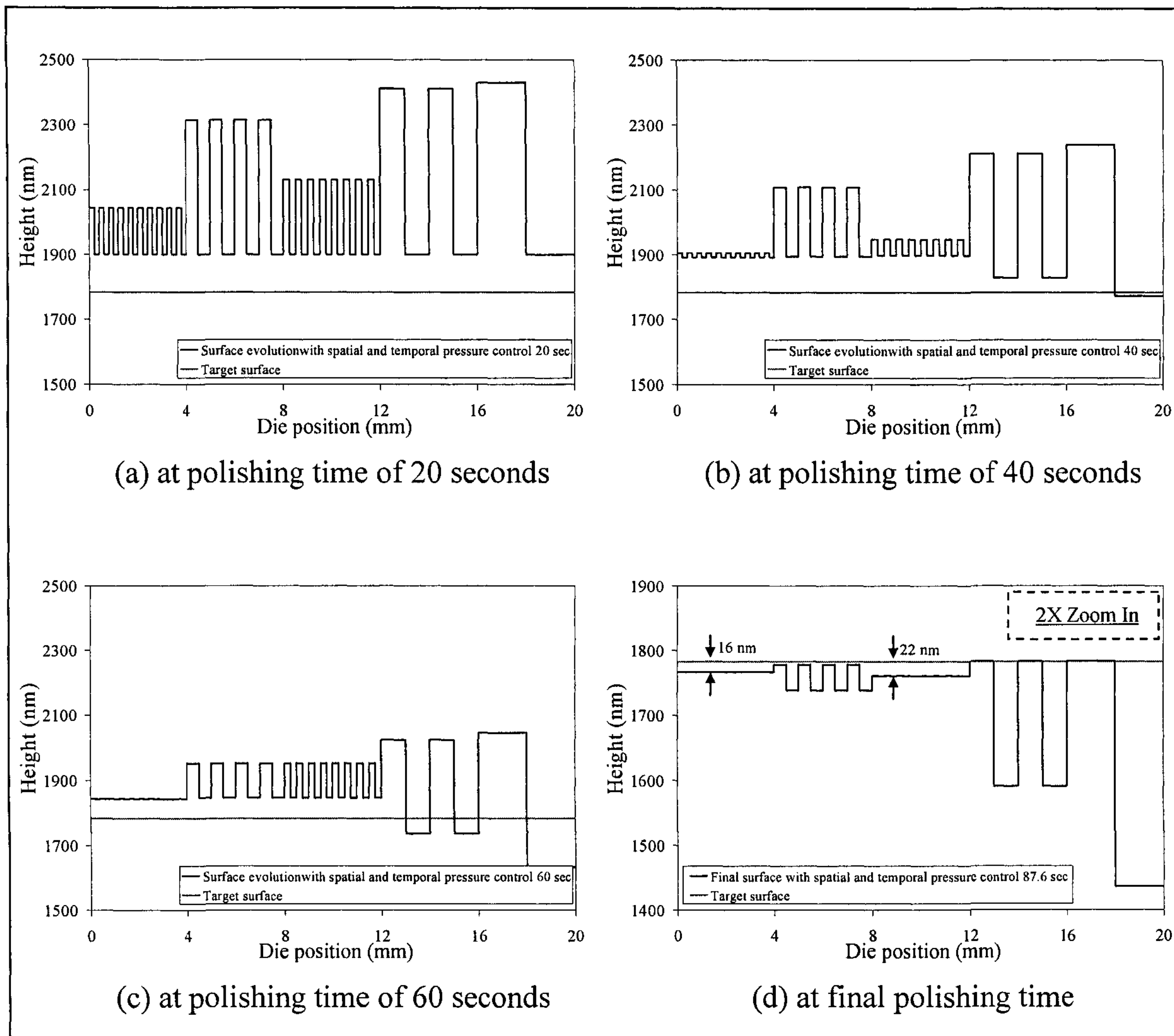


Figure G.3.4. Surface evolution of line 6 vs. die position with spatial and temporal pressure control algorithm.

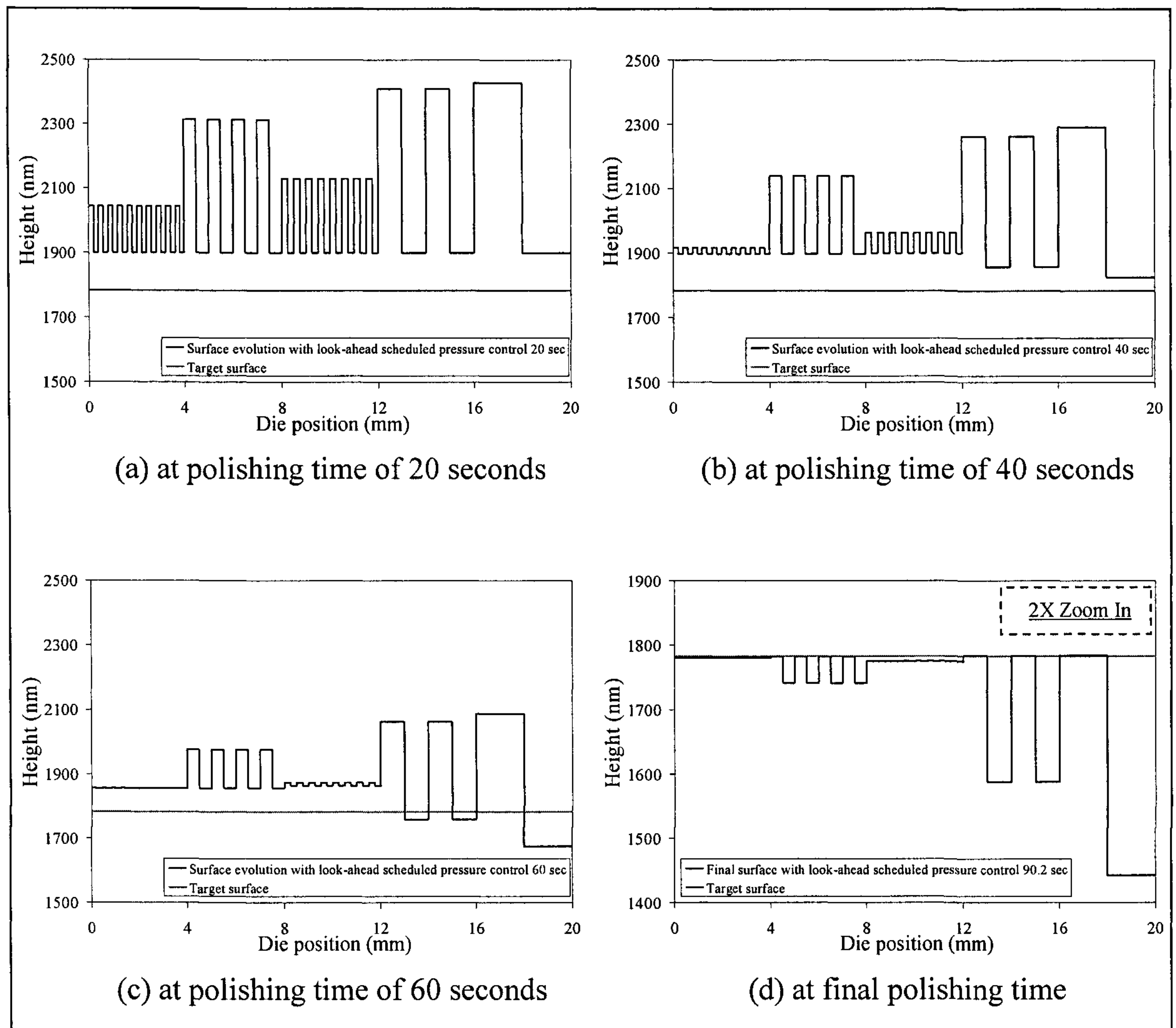


Figure G.3.5. Surface evolution of line 6 vs. die position with look-ahead scheduled pressure control algorithm.

APPENDIX H. FEATURE / DIE-SCALE MODEL FOR GENERAL WAFER SURFACE

General Wafer Surface Form

The entire wafer surface is setup into the series of sine waves by Fast Fourier Transform (FFT) as:

$$Y = Y_0 + A_1 \sin(f_1 x) + A_2 \sin(f_2 x) + \ldots + A_n \sin(f_n x)$$

Where, Y_0 is an average wafer surface from the experimental data,

$A_1, A_2, \ldots, A_n$ are the film amplitudes,

$f_1, f_2, \ldots, f_n$ are the frequencies,

x is the distance from the center of the wafer, and

n is the number of waves.

Moreover, for simplicity, this series of sine waves is represented with a series of square waves where square waveform amplitudes are two-thirds of the sine waveform amplitudes.

From Preston's equation and the force redistribution with the assumptions shown in Chapter 5, the material removal rate equation can be expressed as:

$$\frac{dY_i}{dt} = K\left\{k[H - (Y - Y_i)] + \frac{\Delta F_1}{D_1} + \frac{\Delta F_2}{D_2} + \ldots + \frac{\Delta F_n}{D_n}\right\}V \qquad \text{[H.1]}$$

Where, ΔF_i is the force redistribution on the surface profile of amplitude A_i.

$i = 1 \; to \; 2^n$,

$B2 = Base2(i)$,

$$
\begin{aligned}
D_1 &= c_1 && ,if\ B2 = 0xxxx.... \\
&= -a_1 && ,otherwise \\
D_2 &= c_1 c_2 / b_2 && ,if\ B2 = 00xxx.... \\
&= -c_1 a_2 / b_2 && ,if\ B2 = 01xxx.... \\
&= a_1 c_2 / b_2 && ,if\ B2 = 10xxx.... \\
&= -a_1 a_2 / b_2 && ,otherwise
\end{aligned}
$$

$$
\begin{aligned}
D_3 &= c_1 c_2 c_3 / b_2 b_3 && ,if\ B2 = 000xx.... \\
&= -c_1 c_2 a_3 / b_2 b_3 && ,if\ B2 = 001xx.... \\
&= c_1 a_2 c_3 / b_2 b_3 && ,if\ B2 = 010xx.... \\
&= -c_1 a_2 a_3 / b_2 b_3 && ,if\ B2 = 011xx.... \\
&= a_1 c_2 c_3 / b_2 b_3 && ,if\ B2 = 100xx.... \\
&= -a_1 c_2 a_3 / b_2 b_3 && ,if\ B2 = 101xx.... \\
&= a_1 a_2 c_3 / b_2 b_3 && ,if\ B2 = 110xx.... \\
&= -a_1 a_2 a_3 / b_2 b_3 && ,otherwise
\end{aligned}
$$

.

.

.

$$
\begin{aligned}
D_n &= c_1 c_2 ... c_{n-1} c_n / b_2 b_3 ... b_n && ,if\ B2 = 000...00 \\
&= -c_1 c_2 ... c_{n-1} a_n / b_2 b_3 ... b_n && ,if\ B2 = 000...01 \\
& ... \\
&= a_1 a_2 ... a_{n-1} c_n / b_2 b_3 ... b_n && ,if\ B2 = 111...10 \\
&= -a_1 a_2 ... a_{n-1} a_n / b_2 b_3 ... b_n && ,otherwise
\end{aligned}
$$

And also, from the force balance equation, we have

$$\begin{aligned}
&\left\{k\cdot\left[H-\left(Y-Y_1\right)\right]\cdot\frac{c_1c_2\ldots c_{n-1}c_n}{b_2b_3\ldots b_n}+\Delta F_1+\Delta F_2+\ldots+\Delta F_{n-1}+\Delta F_n\right\}+\\
&\left\{k\cdot\left[H-\left(Y-Y_2\right)\right]\cdot\frac{c_1c_2\ldots c_{n-1}a_n}{b_2b_3\ldots b_n}+\Delta F_1+\Delta F_2+\ldots+\Delta F_{n-1}-\Delta F_n\right\}+\\
&\left\{k\cdot\left[H-\left(Y-Y_3\right)\right]\cdot\frac{c_1c_2\ldots a_{n-1}c_n}{b_2b_3\ldots b_n}+\Delta F_1+\Delta F_2+\ldots-\Delta F_{n-1}+\Delta F_n\right\}+\\
&\ldots+\\
&\left\{k\cdot\left[H-\left(Y-Y_n\right)\right]\cdot\frac{a_1a_2\ldots a_{n-1}a_n}{b_2b_3\ldots b_n}-\Delta F_1-\Delta F_2-\ldots-\Delta F_{n-1}-\Delta F_n\right\}\\
&=P\cdot\left(b_1\cdot 1\right)
\end{aligned} \qquad [H.2]$$

By solving the above equation for *Y(t)*, it can be expressed as:

$$\begin{aligned}
Y(t) = &\left(\frac{c_1c_2\ldots c_{n-1}c_n}{b_1b_2\ldots b_n}\right)Y_1+\left(\frac{c_1c_2\ldots c_{n-1}a_n}{b_1b_2\ldots b_n}\right)Y_2+\ldots+\\
&\left(\frac{a_1a_2\ldots a_{n-1}c_n}{b_1b_2\ldots b_n}\right)Y_{n-1}+\left(\frac{a_1a_2\ldots a_{n-1}a_n}{b_1b_2\ldots b_n}\right)Y_n+\left(H-\frac{P}{k}\right)
\end{aligned} \qquad [H.3]$$

Substituting (H.3) into (H.1), the material removal rate equations can be expressed as:

$$\frac{dY_i}{dt}=KVk\left\{\begin{aligned}&\left(\frac{c_1c_2\ldots c_{n-1}c_n}{b_1b_2\ldots b_n}\right)Y_1+\left(\frac{c_1c_2\ldots c_{n-1}a_n}{b_1b_2\ldots b_n}\right)Y_2+\ldots+\left(\frac{a_1a_2\ldots a_{n-1}c_n}{b_1b_2\ldots b_n}\right)Y_{n-1}\\
&+\left(\frac{a_1a_2\ldots a_{n-1}a_n}{b_1b_2\ldots b_n}\right)Y_n-Y_i-\frac{1}{k}\left(P+\frac{\Delta F_1^i}{D_1}+\frac{\Delta F_2^i}{D_2}+\ \ldots\ +\frac{\Delta F_n^i}{D_n}\right)\end{aligned}\right\} \qquad [H.4]$$

From the second assumption, we have

$$\Delta F_1^i = \alpha(Y_i - Y_{2^{n-1}+i}) \qquad i = 1 \text{ to } 2^{n-1}$$
$$\Delta F_2^i = \alpha(Y_i - Y_{2^{n-2}+i}) \qquad i = 1 \text{ to } 2^{n-2} \text{ and } 2^{n-1}+1 \text{ to } 2^n$$

.

. [H.5]

.

$$\Delta F_{n-1}^i = \alpha(Y_i - Y_{2^1+i}) \qquad i = 1,2,5,6,9,10,\dots,2^n-3,2^n-2$$
$$\Delta F_n^i = \alpha(Y_i - Y_{2^0+i}) \qquad i = 1 \text{ to } 2^n \quad \text{for odd number}$$

Where, α is the bending factor and it is zero if there is no pad bending.

It should be noted that α has no co-relation to pad stiffness k,

Substituting (H.5) into (H.4) coresponding to the sixth assumption, we have an analytical solution as shown below.

$$\frac{dY_i}{dt} = C_1^i Y_1 + C_2^i Y_2 + \dots + C_{2^n-1}^i Y_{2^n-1} + C_{2^n}^i Y_{2^n} + C_{2^n+1}^i \qquad \text{[H.6]}$$

$$C_1^i = KVk\left(A_1^i - B^i - \frac{\alpha}{k}C_1^i\right)$$
$$C_2^i = KVk\left(A_2^i - B^i - \frac{\alpha}{k}C_2^i\right)$$
$$C_3^i = KVk\left(A_3^i - B^i - \frac{\alpha}{k}C_3^i\right)$$

Where, .

.

.

$$C_{2^n} = KVk\left(A_{2^n}^i - B^i - \frac{\alpha}{k}C_{2^n}^i\right)$$
$$C_{2^n+1} = KVk\left(-\frac{P}{k}\right)$$

Using the Laplace transform on (H.6), we have

$$s\widetilde{Y}_i - Y_i(0) = C_1^i\widetilde{Y}_1 + C_2^i\widetilde{Y}_2 + \ldots + C_{2^n-1}^i\widetilde{Y}_{2^n-1} + C_{2^n}^i\widetilde{Y}_{2^n} + \frac{C_{2^n+1}^i}{s} \qquad [H.7]$$

According to equations (H.7), the unknowns are $\widetilde{Y}_i s$. This equation can be rewritten in matrix form as:

$$\begin{bmatrix} s-C_1^1 & -C_2^1 & \ldots & -C_{2^n-1}^1 & -C_{2^n}^1 \\ -C_1^2 & s-C_2^2 & \ldots & -C_{2^n-1}^2 & -C_{2^n}^2 \\ . & . & \ldots & . & . \\ . & . & \ldots & . & . \\ -C_1^{2^n} & -C_2^{2^n} & \ldots & -C_{2^n-1}^{2^n} & s-C_{2^n}^{2^n} \end{bmatrix} \begin{bmatrix} \widetilde{Y}_1 \\ \widetilde{Y}_2 \\ . \\ . \\ \widetilde{Y}_{2^n} \end{bmatrix} = \begin{bmatrix} \frac{C_{2^n+1}^1}{s} + Y_1(0) \\ \frac{C_{2^n+1}^2}{s} + Y_2(0) \\ . \\ . \\ \frac{C_{2^n+1}^{2^n}}{s} + Y_{2^n}(0) \end{bmatrix} \qquad [H.8]$$

Solving the matrix in equation (H.8), we have

$$\begin{bmatrix} \widetilde{Y}_1 \\ \widetilde{Y}_2 \\ . \\ . \\ \widetilde{Y}_{2^n} \end{bmatrix} = \begin{bmatrix} s-C_1^1 & -C_2^1 & \ldots & -C_{2^n-1}^1 & -C_{2^n}^1 \\ -C_1^2 & s-C_2^2 & \ldots & -C_{2^n-1}^2 & -C_{2^n}^2 \\ . & . & \ldots & . & . \\ . & . & \ldots & . & . \\ -C_1^{2^n} & -C_2^{2^n} & \ldots & -C_{2^n-1}^{2^n} & s-C_{2^n}^{2^n} \end{bmatrix}^{-1} \begin{bmatrix} \frac{C_{2^n+1}^1}{s} + Y_1(0) \\ \frac{C_{2^n+1}^2}{s} + Y_2(0) \\ . \\ . \\ \frac{C_{2^n+1}^{2^n}}{s} + Y_{2^n}(0) \end{bmatrix} \qquad [H.9]$$

This inverse Laplace transform can be solved using any powerful mathematic software.

REFERENCES CITED

[1] Busetti, F. (2001), "Genetic algorithms overview," *http://www.geocities.com/francorbusetti/gaweb.pdf* (Date retrieved: 02 October 2005).

[2] Byrne, G., Mullany, B. and Young, P. (1999), "The effect of pad wear on the chemical mechanical polishing of silicon wafers," *Annals of the CIRP*, vol. 48, no. 1, pp. 143-146.

[3] Chen, D. Z., and Lee, B. S. (1999), "Pattern planarization model for chemical mechanical polishing," *J. Electrochemical Society*, vol. 146, no. 2, pp. 744 - 748.

[4] Cook, L. M. (2000), "CMP consumables II: pad," *Semiconductors and Semimetals*, vol. 63, pp. 155-182.

[5] Eamkajornsiri, S. (2002), "Yield improvement in chemical mechanical polishing process investigation of wafer scale," *Master thesis*, Iowa State University.

[6] Eamkajornsiri, S., Fu, G., Narayanaswami, R. and Chandra, A. (2001), "Simulation of wafer scale variations in CMP," *Transactions of NAMRI XXIX, SME*, pp. 221-228.

[7] Fu, G. (2002), "Modeling of chemical mechanical polishing at multiple scales," *Ph.D. Thesis*, Iowa State University.

[8] Fu, G. and Chandra, A. (2001), "A model for wafer scale variation of removal rate in chemical mechanical polishing based on elastic pad deformation," *J. Electronic Materials*, vol. 30, pp. 400 - 408.

[9] Fu, G. and Chandra, A. (2003), "An analytical dishing and step height reduction model for CMP," *IEEE Trans. on Semiconductor Manufacturing*, Vol 16, pp.477–485.

[10] Kaanta, C. W., Bombardier, S. G., Cote, W. J., Hill, W. R., Kerzkowski, G., Landis, H. S., Poindexter, D. J., Pollard, C. W., Ross, G. H., Ryan, J. G., Wolff, S. and Cronin, J. R. (1991), "Dual damascene: a ULSI wiring technology," *Proc. VMIC Conf.*, pp. 144-152.

[11] Kranenberg, H. and Woerlee, P. H. (1998), "Influence of overpolish time on the performance of W damascene technology," *J. Electrochemical Society*, vol. 145, no. 4, pp. 1285-1291.

[12] Luo, J. and Dornfeld D. A. (2004), "Integrated modeling of chemical mechanical planarization for sub-micron IC fabrication," *Springer*, Berlin Heidelberg, Germany.

[13] Obitko, M. (1998), "Introduction to genetic algorithms," *http://cs.felk.cvut.cz/~xobitko/ga/* (Date accessed: 02 October 2005).

[14] Ouma, D. O. (1998), "Modeling of chemical mechanical polishing for dielectric planarization," *Ph.D. Thesis*, MIT.

[15] Ouma, D., Boning, D., Chung, J., Shinn, G., Olsen, L. and Clark, J. (1998), "An integrated characterization and modeling methodology for CMP dielectric planarization," *Proc. Int. Interconnect Technology Conf.*, June, San Francisco, pp. 67-69.

[16] Ouma, D. O., Boning, D. S., Chung, J. E., Easter, W. G., Saxena, V., Misra, S. and Crevasse, A. (2002), "Characterization and modeling of oxide chemical-mechanical polishing using planarization length and pattern density concepts," *IEEE Trans. on Semiconductor Manufacturing*, vol. 15, pp. 232-244.

[17] Preston, F. W. (1927), "The theory and design of plate glass polishing machine," *J. Society Glass Technology*, vol. 11, no. 44, pp. 214-256.

[18] Runnels, S. R. (1994), "Feature-scale fluid-based erosion modeling for chemical-mechanical polishing," *J. Electrochemical Society*, vol. 141, no. 7, pp.1900-1904.

[19] Runnels, S. R. (1996), *J. Electronic Materials*, vol. 25, p.1574.

[20] Runnels, S. R. and Eyman, L.M. (1994), "Tribology analysis of chemical-mechanical polishing," *J. Electrochemical Society*, vol. 141, no. 6, pp.1698-1701.

[21] Runnels, S. R. and Renteln, P. (1993), "Modeling the effects of polish pad deformation on wafer surface stress distributions during chemical-mechanical polishing," *Dielectric Sciences Technology*, pp. 110-121.

[22] Sivaram, S., Bath, H., Lee, E., Leggett, R. and Tolles, R. (1992), "Measurement and modeling of pattern sensitivity during chemical mechanical polishing of interlevel dielectrics," *SEMATECH Technical Report*, Austin, TX.

[23] Sasaki, Y., Aoyama, H., Inasaki, I., Miyairi, H. and Shibaya, H. (1998), "Evaluation of effective CMP conditions by estimation of pressure distribution on semiconductor wafer," *Proc. 1998 Spring Topical Meeting on Silicon Machining*, vol. 17, pp. 92-95.

[24] Srinivasa-Murthy, C., Wang, D., Beaudoin, S. P., Bibby, T., Holland, K. and Cale, T. S. (1997), "Stress distribution in chemical mechanical polishing," *Thin Solid Flims*, vol. 308-309, pp. 533-537.

[25] Steigerwald, J. M., Murarka, S. P. and Gutmann, R. J. (1997), "Chemical mechanical planarization of microelectronic materials," *John Wiley & Sons Pub.*, New York.

[26] Stine, B., Mehrotra V., Boning D., Chung J. and Ciplickas, D. (1997), "A simulation methodology for assessing the impact of spatial/pattern dependent interconnect parameter variation on circuit performance," *IEDM Tech. Digest*, pp. 133-136.

[27] Tugbawa, T., Park, T., Lee, B., Boning, D., Lefevre, P. and Camilletti, L. (2001), "Modeling of pattern dependencies for multi-level Copper chemical-mechanical polishing processes," *Mat. Research Soc.*, Spring meeting, April, San Francisco.

[28] Wang, D., Lee, K. Holland, Bibby, T., Beaudoin, S. and Cale, T. (1997), "Von Mises stress in chemical-mechanical polishing processes," *J. Electrochemical Society*, vol. 144, no. 3, pp. 1121-1127.

[29] Warnock, J. (1991), "A two-dimensional process model for chemical mechanical polish planarization," *Journal of the Electrochemical Society*, vol. 138, no. 8, pp. 2398-2402.

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude for my advisors, Dr. Frank Peters and Dr. Abhijit Chandra, whose excellent guidance, and inspiring encouragement helped make this thesis possible.

My gratitude also goes to members of my committee, Dr. John Jackman, Dr. Sigurdur Olafsson and Dr. Wallapak Tavanapong, for their assistance and encouragement. Dr. Guanghui Fu also deserves a lot of appreciation for the time he took to make me clearly understand his CMP process models and the underlying equations.

I would especially like to thank Tanita Methiyapun, who not only proofread my dissertation, but also encouraged me with my graduate studies. I would also like to thank my sisters for their love, support and encouragement.

Finally, I want to thank the Carver Trust Fund Grant of Iowa State University and the NSF Grant No. DMI 0084736 and DMI 0323069 for their financial supports. Any opinions, findings, and conclusions or recommendations expressed in this material are those of me and do not necessarily reflect the views of the Carver Trust Fund or the National Science Foundation.