Analysis of Slurry Flow in Chemical-Mechanical Polishing

by

Krzysztof D. Kopanski

Bachelor of Science in Mechanical and Aerospace Engineering
Cornell University, 2003

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

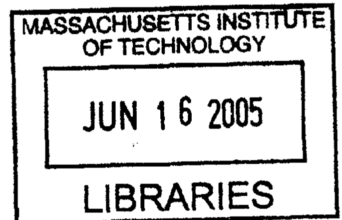
Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

© 2005 Massachusetts Institute of Technology
All rights reserved



Author ....
Department of Mechanical Engineering
May 6, 2005

Certified by ....
Jung-Hoon Chun
Professor of Mechanical Engineering
Thesis Supervisor

Certified by ....
Nannaji Saka
Research Affiliate
Thesis Co-supervisor

Accepted by ....
Lallit Anand
Chairman, Departmental Committee on Graduate Students

ARCHIVES

Analysis of Slurry Flow in Chemical-Mechanical Polishing

by

Krzysztof D. Kopanski

Submitted to the Department of Mechanical Engineering
on May 23, 2005 in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

Chemical-Mechanical Polishing (CMP) is one of the enabling processes used in the manufacture of semiconductor chips. In the relentless progress to make computer chips faster, smaller, and cheaper, the CMP process plays a prominent role. One of its limitations, however, is non-uniform polishing rate at the die and wafer scales. In this thesis, an innovative CMP machine configuration is proposed to minimize wafer-scale non-uniformity. The new *face-up* machine lowers wafer-scale non-uniformity by minimizing over-polishing of any particular area. The thesis discusses the kinematics and design considerations of this machine.

Additionally, this thesis develops an analytical model for slurry flow in CMP in two steps. First, a simple but useful method of estimating the effective gap between the wafer and the pad during polishing is developed. The method uses pressurized fluid flow and an analytical model to estimate the effective gap between the wafer and the pad. Second, this effective gap is used in the Couette model that describes the slurry behavior in CMP. The Couette model shows that rotational speeds of the wafer and pad, the effective gap, and the sizes of the wafer and pad dictate the slurry flow rate and flow pattern in both conventional CMP and the new face-up CMP. The Couette model can be used to estimate the slurry flow rate whenever the process parameters are changed.

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

Thesis Co-Supervisor: Nannaji Saka
Title: Research Affiliate

Acknowledgments

I would like to thank Prof. Chun for his support, advice, and for guiding me to attack the problem in scientific and practical ways.

I would like to thank Dr. Saka for his patience with me and for his wearisome attention to details. Throughout my stay at MIT it was apparent that we had many differences but nevertheless I give high regards for his consistency, knowledge, and competence.

I would like to thank Kyung-Yoon Noh for his camaraderie and for many shared moments while working in the lab. I wish you lots of luck and great success at Intel or wherever you might end up. I would like to congratulate you and your wife on your yet unborn baby.

I would like to give thanks to all the lab mates, Hady Joumaa, Sam Korb, and Munhee Sohn, for great discussions whether they were scientific or just entertaining. I wish all of you great luck on your future endeavors.

I want to acknowledge my Graduate Soccer team for giving me an outlet to relieve stress and enjoy the beautiful game. Thanks for pushing my physical limits and for the great time we had winning the BSSL Div 1 Championship in 2004.

I would like to thank my good friend Beto for good times whether they were discussions of our research, soccer games, or just partying Friday or Saturday night.

I want to give special thanks to my family who have been supporting me and giving me moral boosts whenever I needed one. My parents allowed me to pursue my own life by not pushing me in any particular direction. I want to give thanks to my brother who is there for me. I want to wish him luck and thank him for choosing the right path in life.

I want to give thanks to the LMP machine shop guys who helped me with almost infinite number of projects. I want to thank them for teaching me the tricks of the trade and not just doing it for me.

I want to thank all my friends whom I met at MIT and who are too numerous too list, but for sure they influenced my life one way or another. Good luck to all.

TABLE OF CONTENTS

Title Page..1
Abstract..2
Acknowledgments..3
Table of Contents..4
List of Figures..6
List of Tables..7

1. Introduction
   1.1 Background..8
   1.2 History..10
   1.3 Cu CMP Shortcomings..14
   1.4 Face-up Polisher's Promise..16
   1.5 Thesis Objective and Organization..17

2. Design and Analysis of the Face-up Polisher
   2.1 Design..21
   2.2 Wafer Carrier..22
   2.3 Pad and Slurry Distribution..26
   2.4 Kinematics..28
   2.5 Torques and Forces..33
   2.6 Material Removal Rate..35
   2.7 Polishing Condition..38

3. Pad Characterization for Fluid Flow
   3.1 Introduction..43
   3.2 Poiseuille Flow..44
   3.3 Flow Rate Analysis..49
   3.4 Experimental Setup..50
   3.5 Results..53
   3.6 Summary..57

4. Analysis of Slurry Flow
   4.1 Introduction..58
   4.2 The Couette Model..59
   4.3 Flow Rate Analysis..67
   4.4 Experimental Setup..70
   4.5 Results..72
   4.6 Discussion..74

5. Effect of Slurry Flow Rate on Material Removal Rate
   5.1 Introduction..80
   5.2 Material Removal Rate..81
   5.3 Results..84

5.4 Face-down Polishing....87
5.5 Summary....91

6. Conclusion
6.1 Concluding Remarks....93
6.2 Overview....94
6.3 Future Work....95

Appendix A Fluid Velocity Based on the Couette Model....97
Appendix B Experimental Data....103
Appendix C Photographs....107

Nomenclature....108
References....110

LIST OF FIGURES

Figure 1.1 Schematic of various CMP machines....9

Figure 1.2 Schematic of Al ILD process....12

Figure 1.3 Schematic of Cu damascene process....13

Figure 2.1 Schematic of face-up polisher....21

Figure 2.2 Photograph of the face-up polisher in action....24

Figure 2.3 Wafer carrier and slurry cup assembly....27

Figure 2.4 Kinematic analysis schematic for inverted CMP process....29

Figure 2.5 Average MRR as a function of radial position, pad outside the center....39

Figure 2.6 MRR as a function of radial position, pad overlapping the center....40

Figure 3.1 Schematic of axisymmetric Poiseuille flow problem....45

Figure 3.2 Experimental apparatus used to characterize the pad....51

Figure 3.3 Photograph of the actual experimental apparatus....52

Figure 3.4 Effective gap data....55

Figure 3.5 Pad profile compared to $\bar{h}$....56

Figure 4.1 Face-up polisher's pad coordinate system....61

Figure 4.2 Mean fluid velocity vector field, pad twice as fast as a wafer....66

Figure 4.3 Depiction of an arbitrary contour....67

Figure 4.4 Drawing of the face-up polisher....71

Figure 4.5 Experimental results for flow rates vs. linear velocity....73

Figure 4.6 Couette model for face-up polishing....75

Figure 4.7 Slurry behavior idealization a)Couette model b) More realistic view......79

Figure 5.1 Face-up polisher geometry..83

Figure 5.2 The effect of slurry flow rate on the Preston constant........................89

Figure A.1 Fluid velocity vector field for the given conditions..........................99

Figure C.1 Photograph of the prototype face-up polisher...............................107

LIST OF TABLES

Table 3.1 Effective gap as a function of pad pressure 55

Table 5.1 Slurry flow rate and MRR for two pads in face-up polishing...............85

Table 5.2 Face-down polishing conditions...87

Table B.1 Poiseuille flow experiments...104

Table B.2 Couette flow experiments...105

Table B.3 MRR as a function of slurry feed rate..106

Chapter 1

Introduction

1.1 Background

Chemical Mechanical Polishing (CMP) is a material removal process that was used to polish and planarize glass sheets and lenses for decades before the semiconductor industry applied it to the production of integrated circuit (IC) devices. The polishing process consists of using small abrasive particles to remove a thin layer of material which usually makes the surface both planar and smooth. All polishing processes fall into one of the two categories which are differentiated by the way the abrasive particles are introduced into the polishing area. Two-body polishing is done by fixing abrasive particles to a flexible surface and then rubbing that surface against the work piece. The second mode of polishing is called three-body abrasion and consists of a rough and compliant pad which rubs against the work piece while a fluid containing abrasive particles is introduced into the polishing area. Both types of polishing have their advantages but a vast majority of CMP is done by three-body abrasion which offers better planarization results and fewer scratches than two-body abrasion.

CMP has a few requirements: pressure between the work piece and the polishing pad, relative velocity between both, and the presence of abrasive particles. Typically, rotary CMP is done by pressing the wafer (work piece) off-center onto a large polishing pad. Both the wafer and the pad are rotated about their own axes. Abrasive slurry, a fluid containing the abrasive particles and other chemicals that enhance material removal rate, is introduced onto the pad, so that the motion of the pad will drag the slurry into the polishing region. Once in the polishing interface, the abrasives start removing material by plowing and/or cutting. Figure 1a is a schematic of the rotary type CMP machine. Of course, other ways of fulfilling the basic polishing requirements are possible as are depicted in Figure 1[1]. Typically, linear and orbital polishing machines are more complicated and expensive while at the same time do not offer any significant advantage

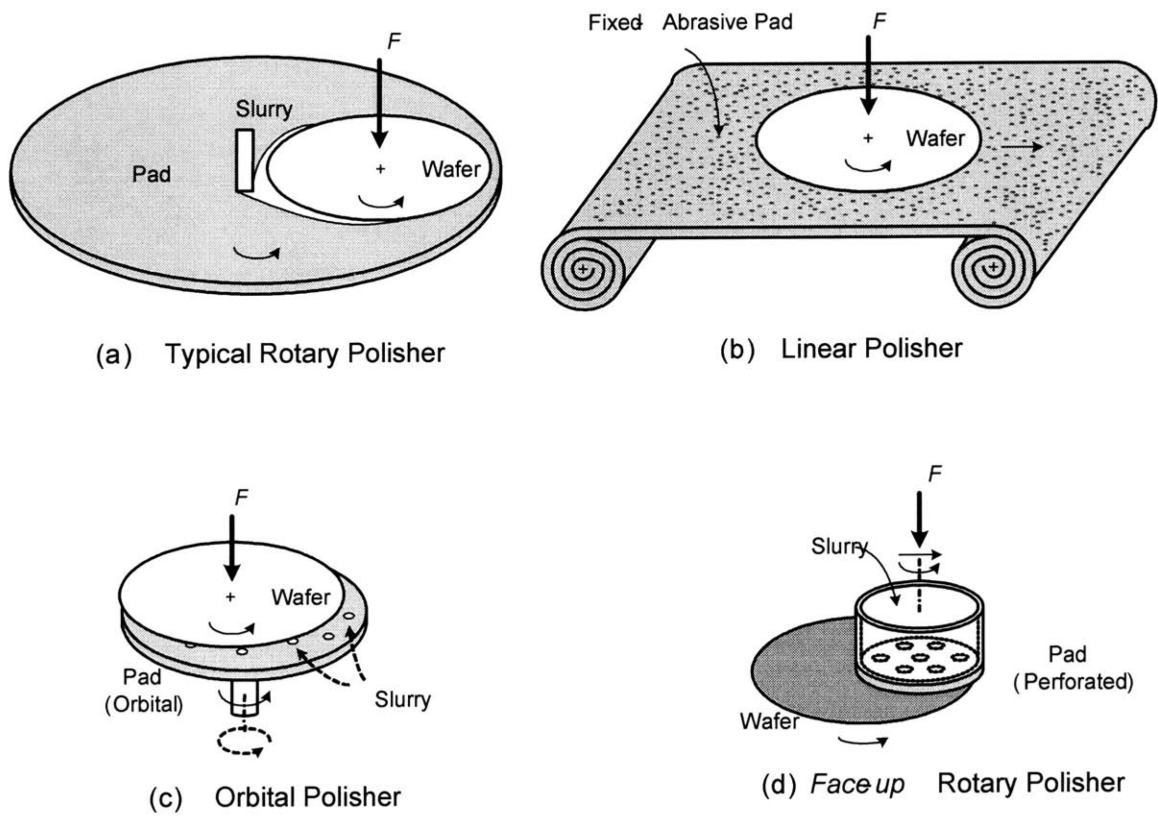


Figure 1.1. Schematic of various CMP machines

in polishing performance over the rotary type.

There is another possibility of fulfilling the polishing requirements while, at the same time, improving the polishing performance. This concept CMP machine was put together due to its great potential of improving the CMP results. It is depicted in Figure 1d and it is closely related to the rotary type CMP machine. This machine will be referred to as the *face-up* polishing machine, due to the fact that the wafer is held 'face' up while it is being polished. The advantage and disadvantage of the *face-up* polishing machine arise from the fact that the pad is smaller than the wafer and thus the wafer is only partially covered by the pad. This affects the throughput of the polisher since the pad only covers and removes material from a portion of the wafer at any given time as opposed to the entire wafer in the typical rotary polisher. This disadvantage will have to be counteracted by either increasing the rotational velocities or by using the slurry in a more effective fashion. The advantage of this concept is that a small pad can generally move away from a polished region and hence not over-polish it, which is a major problem in CMP. Since the wafer is larger than the pad, it can be easily scanned or imaged continuously to see whether or not the process is completed -- a process called end point detection which is cumbersome in the typical CMP machines[2].

1.2 History

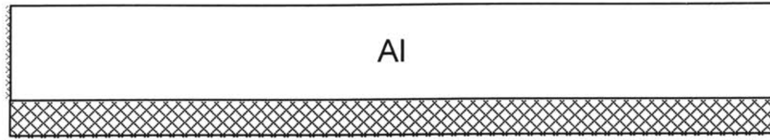
The earliest application of CMP in the semiconductor industry was to prepare silicon wafers for very-large-scale-integrated (VLSI) devices and circuits[3]. After the wafers are sawed from a single crystal silicon rod, the damaged surface is planarized and the scratches are removed by the CMP process. The term chemical mechanical polishing arose from the fact that chemicals are used to meet the specifications for both planarity and surface finish. The mechanical abrasive materials and sizes are also chosen to meet those specifications. It has been investigated that material removal in CMP is mostly due to mechanical abrasion[4-7] while the slurry chemistry softens the polishing material[8,9].

In 1980s, after years of using CMP to polish silicon wafers, IBM researchers discovered new applications for the CMP process, namely, inter-level-dielectrics (ILD) and shallow-trench-isolation (STI). In ILD, CMP was applied in conventional aluminum

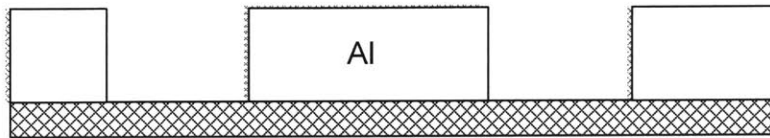
interconnect layers, where CMP produced a planar and defect free ILD surface, typically silicon dioxide, for the deposition of aluminum layer. The aluminum layer was then patterned by lithography which needs a very planar surface for focusing. After lithography, patterning was completed by etching of Al so that only desired interconnect wires remained. The remaining interconnect wires were then insulated by another ILD layer. The process was repeated as needed to form three dimensional electrical wiring called interconnects. An idealization of the ILD CMP is shown in Figure 1.2.

The second application of CMP to the semiconductor industry was STI which is used in the fabrication of transistors on the wafer. In STI, silicon nitride is deposited onto the wafer, patterned by lithography and etched to form trenches. Then, silicon oxide is deposited uniformly onto the surface and CMP is utilized to remove all of the oxide on top of silicon nitride leaving trenches filled with just silicon oxide. The discussion of STI is limited but the concept of STI is identical to Cu CMP, which came much later. Figure 1.3 shows a schematic of Cu CMP. STI is a similar process except on a slightly smaller scale and for different materials.

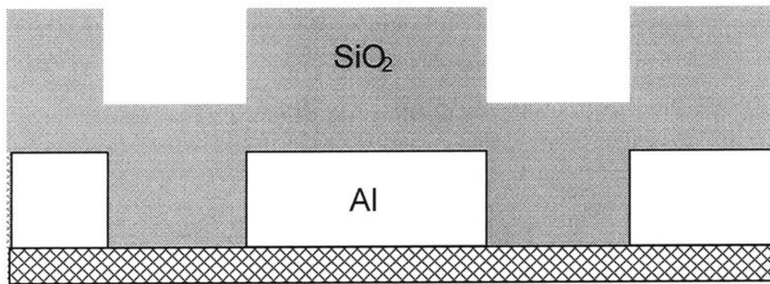
In the mid 1990s as the industry was trying to meet ever more stringent specifications required by the need to produce smaller and faster chips, Cu was looked at as a possible replacement for Al as the interconnect material. Copper was chosen for its low electrical resistivity and its resistance to electromigration. Copper had a promise of increasing the speed of chips by 20%, but the major problem with Cu was that, compared with Al, it was very difficult to etch. Hence, in order to fabricate multi-layered interconnect wiring, a process similar to STI CMP had to be employed as opposed to ILD used with aluminum. Cu CMP starts with a planar dielectric layer, typically silicon dioxide, which is patterned by lithography and then etched to form 'trenches'. A uniformly thick layer of copper is then deposited onto the surface by PVD, CVD, or electroplating which covers the entire surface in a uniform fashion. CMP is then employed to remove the excess copper on top of the dielectric layer. After CMP, some portion of the dielectric layer is exposed and the only copper remaining is in the trenches forming interconnects. This process is repeated to form as many as 8 levels of interconnect wiring on current Intel's Pentium 4 chips.



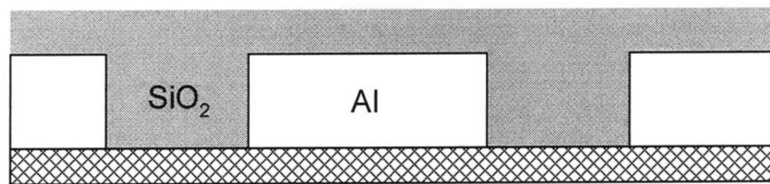
(a) Deposition



(b) Lithography and Etching



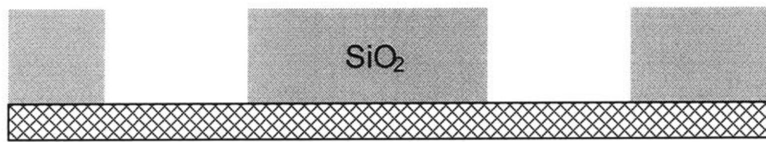
(c) Dielectric Deposition



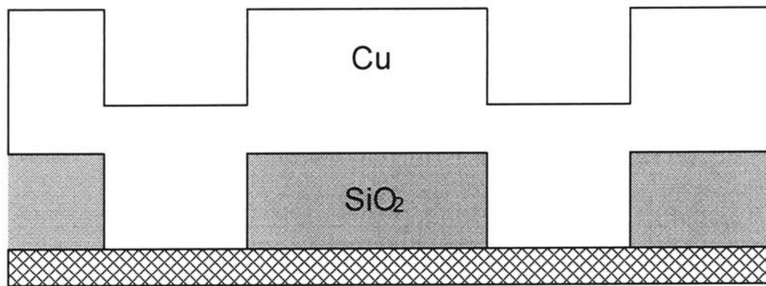
(d) ILD CMP

Figure 1.2. Schematic of Al ILD process

(a) Dielectric Deposition



(b) Lithography and Etching



(c) Copper Deposition

(d) Cu CMP

Figure 1.3. Schematic of Cu damascene process

1.3 Cu CMP shortcomings

The CMP process is definitely an enabling process in ultra-large-scaled-integrated (ULSI) devices and it began to be used in the fabrication of MEMS[10]. But like any other manufacturing process, it has limitations which will be discussed in light of Cu CMP. The most troublesome problem with Cu CMP is the polishing non-uniformity which occurs at three different scales: wafer, die, and feature scales. Feature-scale non-uniformities are mainly caused by local material variations and scratches which may cause the device to be defective. Local material variations may occur because of variation in deposition; hence CMP is not directly responsible for that defect. Scratches, however, are mainly due to CMP and can be eliminated by keeping the CMP environment clean and by controlling the abrasive size and particle agglomeration in the incoming slurry. On the other hand, CMP is directly responsible for die- and wafer-scale non-uniformities.

It is generally agreed that material removal in CMP is due to mechanical interaction between abrasive particles, the pad, and the chemically modified wafer surface. The material removal rate (MRR) can be modeled by the Preston equation[11]:

$$\frac{dh}{dt} = k_p p_{nom} \left|\vec{v}_R\right| \tag{1.1}$$

where h is the thickness of the layer removed, t the polishing time, p_{nom} the nominal pressure, $\left|\vec{v}_R\right|$ the magnitude of relative velocity vector, and k_p the proportionality constant known as the Preston constant. Many researchers demonstrated experimentally the validity of the Preston equation in CMP[5,12,13]. It must be understood that the Preston constant is not a constant on the wafer and die scales. The Preston equation describes material removal rate at some point on the wafer which can be characterized by local pressure and local relative velocity. Generally, another point on the same wafer will have different local pressure and relative velocity due to pattern geometry effects and/or spatial

pressure variations. Hence different points on the wafer will exhibit different material removal rates, and this difference is expressed as different Preston constants.

Spatial (radial) material removal variation occurs even if relative velocity and pattern geometry are the same across the wafer surface. This variation is called wafer scale non-uniformity and it is hypothesized to result from: (i) local pressure has spatial (radial) variation and (ii) abrasive slurry is not distributed uniformly throughout the wafer surface. Local pressure varies radially on the wafer due to the fact that a rigid wafer is pressed onto an elastic pad which cannot support infinite curvature near the edge; hence it causes radially varying pressure distribution[14,15]. The semiconductor industry resolves this problem by two means. First, they effectively enlarge the wafer by placing a retaining ring around the wafer which causes similar pressure distribution variation but within the wafer region it becomes more uniform. Second, they incorporate pressure chambers on the backside of the wafer which allows them to apply higher pressure in the center of the wafer, and thus counteracting this phenomenon.

The slurry distribution problem is mainly caused by the fact that the slurry is fed onto the pad outside the wafer. The slurry must enter the wafer/pad interface from the outside and thus the outer edge will be exposed to the freshest and most effective slurry which may cause the outer edge of the wafer to exhibit higher removal rate. This problem can only be solved by somehow feeding the slurry throughout the wafer-pad interface, probably through the pad. But on a rotary type CMP tool, this would cause a large slurry consumption which would increase an already high cost of CMP ownership (CoO).

Die-level non-uniformity is the variation of material removal rate within any die. Typically, a die would consist of a variety of geometries such as linewidth (interconnect width) and area fraction, which is a fraction of the projected area of interconnects. Within a die there may be some areas with large line widths and high area fraction, small linewidths with low area fraction or anything in between. These pattern geometries cause die-level non-uniformity which is a major problem for the CMP industry[18-22]. At present, CMP tools are helpless against die-level non-uniformity since all of adjustable parameters of any CMP tool such as nominal pressure, relative velocities, pad selection,

and slurry selection are global in nature and are expected to affect each portion of the die equally.

Wafer-level non-uniformity, however, is predominantly due to the design and geometry of the CMP tool. Pressure distribution non-uniformity is due to the fact that the entire wafer is pressed onto a larger and more compliant pad. Variations of slurry availability and effectiveness throughout the wafer-pad interface are also mostly due to the fact that the slurry is fed outside this interface and the pad is not providing an effective way of transporting the slurry. The inverted *face-up* CMP machine will eliminate both of these problems while introducing another cause of wafer-level non-uniformity which can be easily controlled.

1.4 *Face-up* polisher's promise

The proposed *face-up* CMP tool is a rotary type CMP tool that is completely inverted compared to a traditional rotary tool. The idea is to eliminate or at least significantly lower wafer-level non-uniformity by geometric and kinematic means. By holding the wafer *face-up* and polishing its surface with a smaller pad, the CMP tool will have much more control over polishing uniformity provided the kinematics are well understood and adequate end point detection is implemented. The *face-up* polisher can minimize wafer-level non-uniformity caused by pressure variation and slurry distribution.

The pressure variation can be significantly reduced by having a small, compliant polishing pad that is fully within the wafer. The only time the pressure variation will come into play is when some portion of the pad is outside the wafer. In order to effectively polish the edge, the pad must be moved outside the wafer so that a larger fraction of the wafer edge is covered by the pad. This will cause similar pressure variation as seen in conventional rotary type CMP machines. But a similar retaining ring concept can be used to counteract it. The wafer surface can be effectively enlarged by a retaining ring while the pad is polishing the edge area. This would eliminate any pressure variation as long as the wafer and the retaining ring are in the same plane to within a few micrometers.

But the greatest improvement over a conventional rotary type CMP tool comes from the way abrasive slurry can be distributed in the wafer-pad interface. In the *face-up* configuration, the pad is smaller than the wafer so the entire pad area is constantly within the wafer bounds and is always removing material. Hence it is possible and desirable to distribute the slurry through perforations in the pad in such a way that slurry will be equally available throughout the pad area ensuring a uniform Preston constant.

Both of these are benefits over a conventional rotary type CMP tool, and they will enable the process to minimize over-polishing time on any point on the wafer. Since incoming wafers are generally planar, as long as it does not over-polish any areas and hence remove more material than necessary, which will cause deviations from nominal planarity, the wafer will leave CMP with good planarity. This planarity is basically the definition of wafer-level non-uniformity. So by using this new CMP tool, it is possible to minimize over-polishing on the wafer-scale and keep within-wafer-non-uniformity (WIWNU) to the desired 5%[23]. However, die-scale non-uniformity may not be affected at all, for it is a separate problem on a smaller scale that the new face-up polisher cannot solve. The problem must be addressed before the wafers come to CMP. One way to solve die-scale non-uniformity is to ensure that Cu is deposited over the pattern trench in such a way that it appears like a blanket wafer and thus eliminating pattern dependency on polishing rates[1].

1.5 Thesis Objective and Organization

It is desirable to distribute the slurry through perforations in the pad uniformly so that the material removal can be described purely by the kinematics of the system. The object is to make the Preston constant spatially constant. It should not vary from one point of the interface to the next. However, the material removal rate will generally vary spatially due to the velocity gradients and the fact that the pad is round and it does not cover the wafer uniformly. But as long as the material removal rate only depends on the geometry and the kinematics, both of them can be altered during polishing to remove equal amounts of material at any point on the wafer and hence minimizing wafer-scale non-uniformity. But the alteration of the kinematics and geometry during polishing will

only produce improvements in wafer-scale non-uniformity if the pad can be assumed to remove material equally within its area.

The objective of this thesis is to analyze slurry flow between the wafer and the pad for both the face-down and for the *face-up* polishing machines. Such analysis is important because it will yield a simple yet meaningful model for slurry flow in CMP. In the past, many researchers studied slurry flow in the context of CMP but failed to provide any useful insights of which parameters are important[27-30]. The analysis presented in this thesis will yield a simple analytical model for slurry flow rate and general slurry velocity during polishing. This thesis will show which parameters are important and how they affect scaling. The analysis will also give guidance for designing the pad for the face-up polishing machine such that the Preston constant will not vary spatially. It is also desired to minimize the slurry consumption during CMP since it is a large portion of CMP CoO, in addition to the environmental issue of discarding hazardous slurry. It is believed that if the wafer/pad interface is flooded with slurry through many holes and grooves, the slurry distribution will be uniform, and the Preston constant will be uniform throughout that interface. However, it is too expensive and not a good practice to do that. Thus the pad perforations must be optimized to keep the slurry uniformly distributed but at the same time minimize its consumption. Additionally, the thesis will also focus on defining and characterizing certain pad properties that are important in determining slurry distribution and flow rates. Basic fluid models will be incorporated in determining these parameters which will give insights into how an optimized pad should be designed and what CMP parameters dictate the slurry consumption.

The thesis is organized in chapters that address the problem in ever-greater detail until the conclusion. The second chapter defines and fully describes the design, geometry and kinematics of the *face-up* polishing machine. It is imperative to understand this concept tool since the thesis focuses on the slurry flow in the context of this tool. The general conclusions of this thesis can be applied to any CMP tool, but the analysis was done with the *face-up* polisher in mind.

The third chapter describes the methodology of characterizing the pad for fluid flow. Numerous papers have been published trying to determine the fluid film thickness between the pad and the wafer[24,25] and the effect of slurry delivery in CMP[26]. The new

methodology will enable researchers to obtain an effective gap between the wafer and the pad. The described experiment is an application of Poiseuille flow solution to a more complicated system used in CMP. However, even such a simple model can be used to characterize the pad in terms of fluid flow capability without using a sophisticated pad surface profilometric technique. The utility of this method is that it can be used to test different pads under loads experienced in actual CMP process, something profilometers cannot do. Profilometry is an excellent means of measuring roughness of the pad but during polishing such roughness is reduced due to loading. The Poiseuille method is used to find the effective gap between the wafer and pad even under loads experienced in CMP.

In Chapter 4, the obtained gap from Chapter 3 is used in another fluid flow model describing slurry flow in the face-up polisher. Since the face-up polisher uses a perforated pad, all of the slurry flows through the interface and it is limited by the effective gap between the wafer and the pad and their rotational speeds. The described model is basically a Couette flow model which is boundary induced flow. The slurry is driven from the pad perforation and out of the pad by the movement of the boundaries. The model is quite simple and can be readily used to estimate the slurry consumption of the polisher. The model is important to understand because in the face-up geometry, the CMP process draws as much slurry as it needs as long as there is some in a reservoir. It does not depend on the operator's decision of how much to feed; the given CMP conditions will restrict the flow. The operator's only choice is whether or not to starve the system. The desirability of such an action depends on the polishing pad used, its effective gap between the wafer and the pad and the effect of slurry starvation on material removal rate.

The effect of slurry flow rate on material removal rate will be discussed in Chapter 5. Some experiments performed on both face-up machine and conventional rotary machine will be discussed in light of slurry consumption and its effective use. In the face-down CMP, the slurry flow rates are typically established empirically for given CMP conditions. It is not clear how the slurry flow rates depend on CMP process conditions and wafer size. The Couette flow model describing flow in the face-up polisher gives some insights on the amount of slurry that needs to be fed. Its implications

can be easily applied to the conventional rotary geometry. Material removal rates or the Preston constants will be determined as a function of slurry feed rate for the conventional CMP machine. That data will then be compared to data from the face-up polisher.

The conclusion will elaborate on the performed work in the context of scaling and other CMP issues and suggest possible future research.

Chapter 2

Design and Analysis of the face-up polisher

2.1 Design

The novelty of the face-up polisher is that the pad is smaller than the wafer. This fact creates possibilities for an improvement in controlling wafer-scale non-uniformity. The conceptual idea for the face-up polisher was born during an investigation of erosion and dishing problems in CMP. The initial conclusion was that these problems were mainly caused by over-polishing. Over-polishing, in turn, is caused by non-uniform polishing rates across the wafer. For example, the outer area of the wafer shows faster material removal rates than the center, so by the time CMP is completed at the center; the edge will be noticeably over-polished. Even if a perfect end point detection is employed, the entire wafer continues to be polished until the slowest areas are ready, and in the meantime the fastest areas will be over-polished. This over-polishing makes the interconnect lines thinner by removing extra oxide material, erosion, and removing Cu below the diminished oxide level, dishing. The face-up polisher was conceived to minimize over-polishing on the wafer scale by progressively moving the pad away from the polished regions.

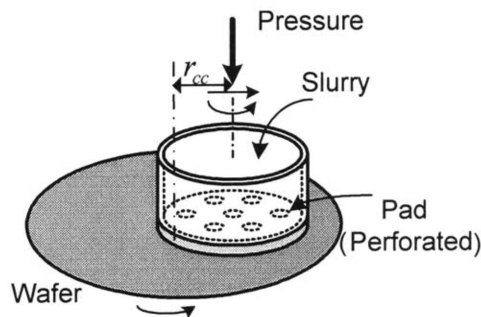


Figure 2.1. Schematic of face-up polisher

Figure 2.1 shows a schematic of the face-up polisher. It consists of a wafer which is held to-be-polished-side up and it is rotated at some angular velocity about its center. The pad, which typically is smaller than the wafer, is pressed onto the wafer off-center; the distance between the center of pad and the center of wafer being a variable, r_{cc}, which generally will change during the polishing process. The pad is also rotated about its center and at some angular velocity which generally will not be the same as that of the wafer. The slurry is distributed into the wafer pad interface by a number of perforations through the pad, thus enhancing the polishing rate and uniformity. These are the only functional requirements for the face-up CMP machine.

Due to the proposed configuration, there are a few potential concerns unique to the face-up machine. The wafer, while rotating rather fast (200-800 rpm) carries a normal load due to the pad and, since it is a tribological system, a tangential force which causes a moment in the normal direction on the wafer. Hence, an important design consideration is the way the wafer will be held during polishing. Another important consideration is how the alignment of the pad will be controlled so that pad can lay flat on the wafer surface and have a uniform pressure distribution and still be driven linearly and rotationally. The way the slurry is fed and distributed through the pad must also be considered. So, in order to design and fabricate the face-up polishing machine, these issues must be resolved. Those issues are not necessarily present on the typical face-down rotary CMP machine.

2.2 Wafer Carrier

The wafer carrier incorporated into the face-up polisher must be quite different from the typical rotary type. The basic differences arise from the fact that the normal load coming onto the wafer is usually off-centered. This will not only cause a large torque which the wafer must carry, but also a moment that will tend to lift the unloaded side of the wafer. On a conventional rotary CMP machine, the wafer is firmly pressed onto the pad with the entire wafer area supporting the load. The wafer is held by capillary forces between the wafer and the backside of the carrier, and it is constrained by a retaining ring. This will not allow the wafer to slip out of the wafer carrier since it is

firmly pressed against the pad and surrounded by the retaining ring. The face-up polisher is a little different since the off-center moment of the pad will tend to lift one side of the wafer upward. If the wafer is not constrained on that side it may slip out of the carrier all together. The ideal solution to this problem is to hold the wafer by a vacuum chuck that will provide enough normal force on the wafer that overcomes the off-centered loading. The vacuum normal load will also provide adequate normal force so a large friction force between the wafer and the chuck can be generated, which will easily resist the rotational torques associated with polishing. Figure 2.2 shows a photograph of the existing face-up polisher in operation. A cross-sectional view of the wafer carrier and the pad system is shown in Fig. 2.3.

The design of such vacuum chuck can be complicated since it must provide large forces to counteract the polishing forces. An important parameter in the vacuum chuck design is making sure that the wafer will not deflect considerably beyond the vacuum regions. It was observed that this bending is large enough during polishing so that low material removal rates were present in the deflected regions. The pad was basically skimming above the depressions in the wafer created by the vacuum. Hence, the vacuum chuck must be optimized in terms of a large vacuum area for generating a large clamping force and small vacuum areas for negligible deflections. It is believed that the optimum vacuum chuck design will consist of a large number of small holes, so that the wafer deflection overhanging the vacuum holes will be small, but there will be enough holes to generate a large clamping force.

It is best to provide continuous vacuum during polishing so that it can be a reliable clamping mechanism through out the duration of polishing. But since the wafer rotates during polishing, the vacuum chuck/wafer carrier must be equipped with a rotary seal that will provide a good seal and a way to continuously pump the air out. The rotary seal must be isolated from the slurry waste or be able to adequately deal with harsh chemicals. At present, the vacuum chuck has not been equipped with a rotary seal, instead

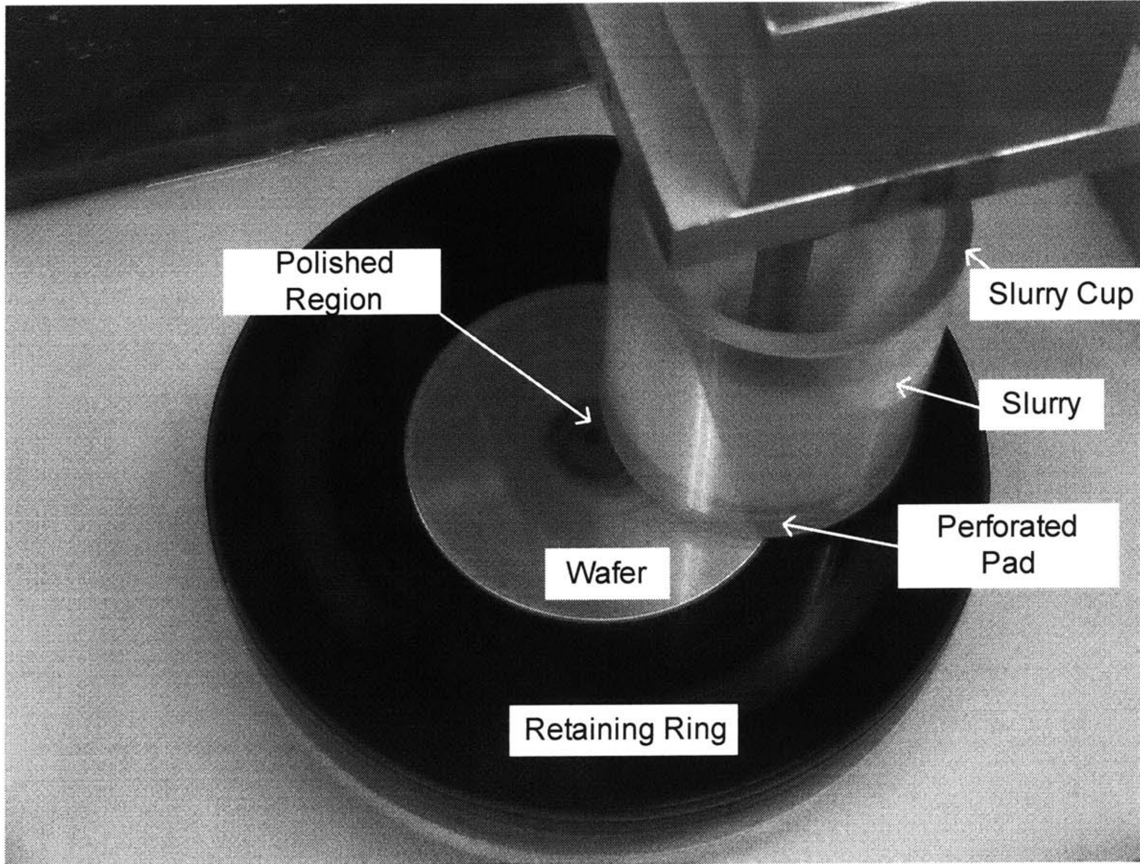


Figure 2.2. Photograph of the face-up polisher in action.

the vacuum chamber is evacuated before the polishing and it is sealed via gaskets and a valve so that it remains evacuated throughout the duration of polishing.

Another important design parameter concerning the wafer carrier has to do with the edge of the wafer. In the face-up configuration, the pad is smaller than the wafer so the pad must be moved, at least partially, out of the wafer area to effectively polish the wafer edge. Since some slurry supplying perforations would be outside the wafer at any given time, considerable volume of slurry would be wasted. It would be beneficial in terms of slurry consumption, therefore, to provide additional surface beyond the wafer to minimize slurry consumption. This surface too would be polished, so it is important that it be made of materials that do not interact with devices on the wafer or cause defects. This additional surface should be similar to the retaining ring used in the conventional polishers but this time it will be used for minimizing slurry consumption and helping with the pressure variation. The retaining ring must be as co-planar with the wafer surface as possible. Otherwise the pad will not polish the edge regions equally around the circumference. It is important for the retaining ring to be as planar as the wafer itself and that it should be adjustable to compensate for a range of wafer thicknesses and its wear. This is a critical issue with the face-up polishers since the entire reduction of wafer-scale non-uniformity will be eradicated if not made worse than the conventional rotary type CMP machine. Ideally, the retaining ring should be within a couple of microns below the wafer surface. In practice this may be hard to achieve so good care must be taken to minimize the deviation from the ideal.

Another important consideration with the wafer carrier has to do with the run-out and tilt of the wafer. Since the wafer carrier rotates and it carries off-center loading, it must be designed such that it does not provide large run-out and tilt. Run-out is caused by the rotation axis not located on the center of the wafer and tilt is the surface of the wafer not being perpendicular to the axis of rotation. Both of these problems can cause excessive vibration and non-uniform polishing. Care must be taken in designing the support bearings for the wafer carrier so that both of these problems are eliminated. Preferably, the wafer carrier should be independently supported by large radial and thrust bearings instead of being mounted on a shaft that is supported by smaller and less robust bearings.

2.3 Pad and Slurry Distribution

For uniform material removal rate throughout the contact region, the pad must be able to supply uniform pressure and slurry distribution within its bounds. For uniform pressure distribution, the polishing pad needs to include a self aligning feature to compensate for any tilt on the wafer carrier. This aligning feature must be able to transfer the required torque for polishing at any given speed and load, transfer the normal load, and be self-aligning by allowing certain freedom of rotation about axes that are co-planar with the pad. Currently, the alignment feature on the pad consists of a rotating load-bearing shaft pressing unto the pad carrier with a cross pin engaging two vertical pins fixed in the pad support. This way, the rotational torque is transferred through the pins and the alignment is carried by spherical ending of the rotating load shaft. The pad support is allowed to pivot as it desires since the spherical ending of the rotating shaft allows easy tilt. The torque carrying pins do not interfere with the alignment because they only transfer torque and not constrain any other motion. The cross-sectional view of the pad/slurry cup system is depicted in Fig. 2.3.

The slurry distribution system must also be integrated into the pad. Since the pad is perforated and it rotates, it is simplest to position a slurry reservoir above the pad that rotates with it. This slurry reservoir would tilt and rotate with the pad while supplying slurry for distribution through its perforations. The slurry reservoir can either be open to atmosphere so the slurry can be fed to it from a stationary nozzle, or it can be closed and pressurized in which case the slurry must be fed through a rotary seal. If the later option is chosen, the slurry reservoir walls must not leak the slurry while allowing some flexibility due to the movements of the aligning pad. It is obvious that the first option is much simpler and can provide adequate solution to the problem. The second option is more elegant and cleaner while providing an opportunity to save the slurry in the reservoir for later use.

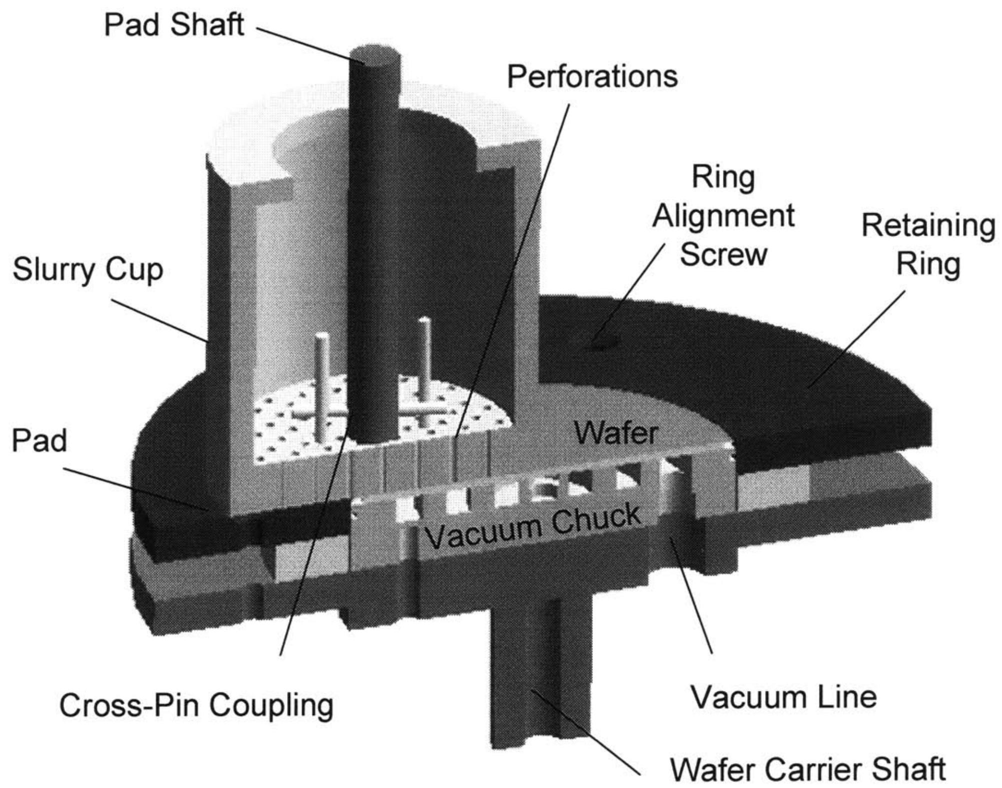


Figure 2.3 Wafer carrier and slurry cup assembly

2.4 Kinematics

The geometry of the face-up polisher is the inverse of the typical CMP machine. Hence, new kinematic analysis must be done so that one can predict material removal rates based on the operating conditions and geometry. As a first approximation, the Preston equation can be used to obtain material removal rates during polishing. However, the Preston equation expresses only the rate of material removal from relative velocity and applied pressure and integration must be performed to obtain the actual material removed in any given time. To use the Preston equation, the relative velocity vector (only its magnitude) has to be known, and generally it will vary from point to point within the pad.

The Preston equation predicts that the polishing rate is proportional to the pressure and relative velocity between the pad and the surface to be polished as given in Eq. 1.1. Therefore in order to analyze the uniformity of the polishing process, one must analyze how relative velocity changes spatially underneath the pad. The following is an analysis determining the relative velocity at any point P between the pad and the wafer. At any point P, the wafer velocity in wafer coordinate system can be obtained by differentiating position. Figure 2.4 shows the geometry and the coordinate systems used.

$$\vec{r}_{wafer} = r\cos\theta\hat{i} + r\sin\theta\hat{j} \tag{2.1}$$

where $\vec{r}_{wafer}$ is the position vector of point P in wafer coordinate system. Differentiating Eq. 2.1 to obtain local velocity of point P, $\vec{v}_{P_{wafer}}$:

$$\vec{v}_{P_{wafer}} = \frac{d\vec{r}_{wafer}}{dt} = -r\sin\theta\dot{\theta}\hat{i} + r\cos\theta\dot{\theta}\hat{j} \tag{2.2}$$

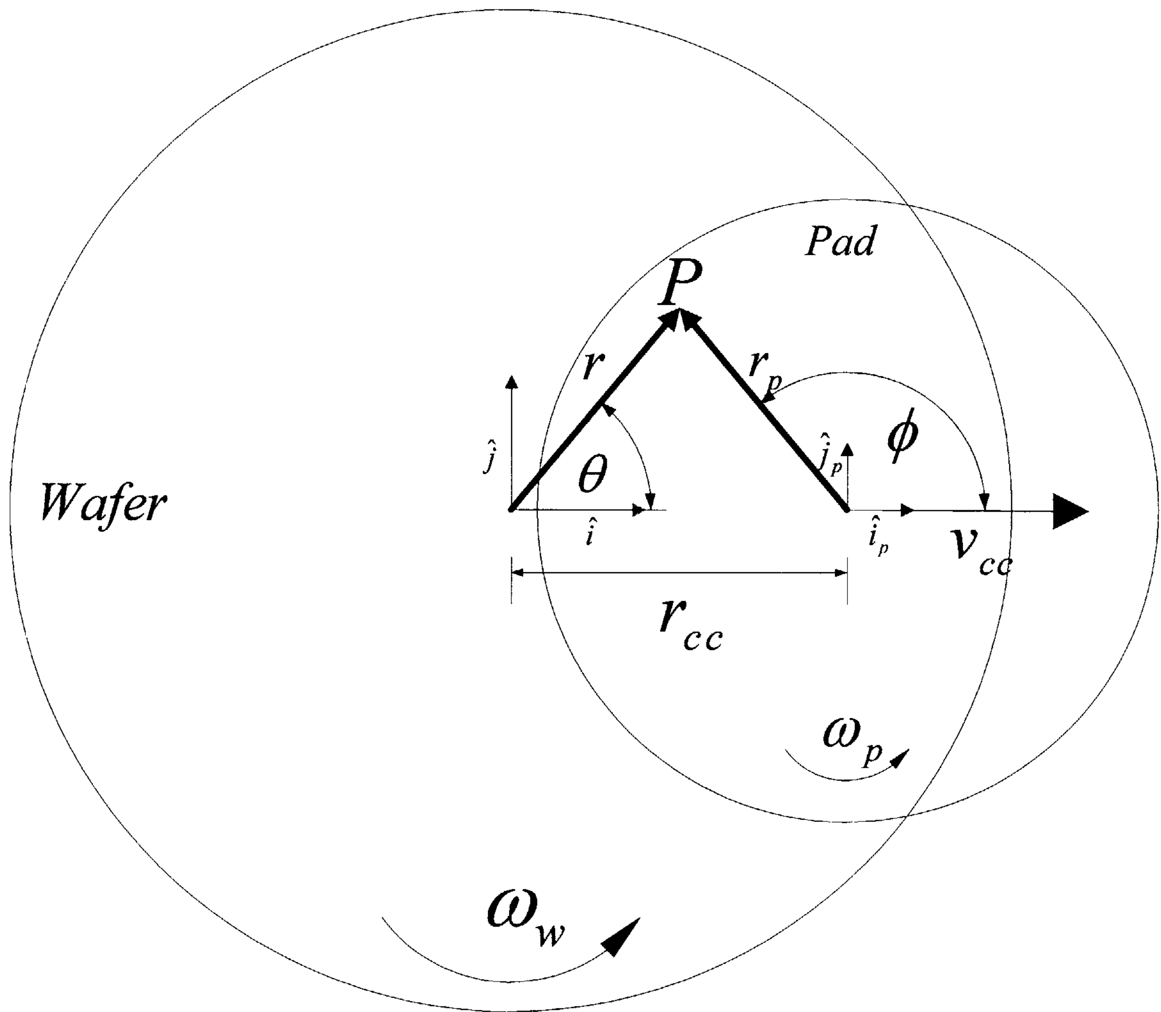


Figure 2.4. Kinematic analysis schematic for inverted CMP process

Considering constant angular speeds,

$$\vec{v}_{P_{wafer}} = -r\omega_w \sin\theta\hat{i} + r\omega_w \cos\theta\hat{j} \tag{2.3}$$

The same point P has pad's velocity in its coordinate system that can be obtained in a similar manner. The local position of point P is given by:

$$\vec{r}_{P_{pad}} = r_p \cos\phi\hat{i}_p + r_p \sin\phi\hat{j}_p \tag{2.4}$$

This can be differentiated to obtain the local velocity due to the pad.

$$\vec{v}_{P_{pad}} = \frac{d\vec{r}_{P_{pad}}}{dt} = -r_p \sin\phi\dot{\phi}\hat{i} + r_p \cos\phi\dot{\phi}\hat{j} \tag{2.5}$$

where $\vec{r}_{P_{wpad}}$ is the position vector of point P in pad coordinate system, $\vec{v}_{P_{pad}}$ the pad velocity at point P. Recognizing that the angular speeds are constant, $\vec{v}_{P_{pad}}$ becomes:

$$\vec{v}_{P_{pad}} = \left(v_{cc} - r_p\omega_p \sin\phi\right)\hat{i} + r_p\omega_p \cos\phi\hat{j} \tag{2.6}$$

$$v_{cc} = \frac{dr_{cc}}{dt} \tag{2.7}$$

The pad velocity can be converted to wafer coordinate system using the following relations:

$$r\cos\theta = r_{cc} + r_p \cos\phi \tag{2.8}$$

$$r\sin\theta = r_p \sin\phi \tag{2.9}$$

Converting pad velocity at point P in Eq. 2.6 to the wafer coordinate system:

$$\vec{v}_{P_{pad}} = \left(v_{cc} - r\omega_p \sin\theta\right)\hat{i} + \left(r\omega_p \cos\theta - r_{cc}\omega_p\right)\hat{j} \tag{2.10}$$

The relative velocity at point P is the difference between the wafer velocity and the pad velocity:

$$\overrightarrow{v_{P\,rel}} = \vec{v}_{P_{wafer}} - \vec{v}_{P_{pad}} \tag{2.11}$$

Substituting the expressions for both velocities in the wafer coordinate system.

$$\overrightarrow{v_{P\,rel}} = \left(-r\omega_w \sin\theta + r\omega_p \sin\theta - v_{cc}\right)\hat{i} + \left(r\omega_w \cos\theta - r\omega_p \cos\theta + r_{cc}\omega_p\right)\hat{j} \tag{2.12}$$

Combining like terms, the relative velocity vector is obtained as:

$$\overrightarrow{v_{P\,rel}} = \left\{-\left(\omega_w - \omega_p\right) r \sin\theta - v_{cc}\right\}\hat{i} + \left\{\left(\omega_w - \omega_p\right) r \cos\theta + r_{cc}\omega_p\right\}\hat{j} \tag{2.13}$$

The magnitude of relative velocity vector at point P is:

$$\left|\vec{v}\right|_{P_{rel}} = \sqrt{\begin{array}{l}\left(\omega_w - \omega_p\right)^2 r^2 \sin^2\theta + v_{cc}^2 + 2\left(\omega_w - \omega_p\right) v_{cc} r \sin\theta + \left(\omega_w - \omega_p\right)^2 r^2 \cos^2\theta + r_{cc}^2\omega_p^2 \\ +2\left(\omega_w - \omega_p\right)\omega_p r r_{cc} \cos\theta\end{array}}$$

$$\left|\vec{v}\right|_{P_{rel}} = \sqrt{\left(\omega_w - \omega_p\right)^2 r^2 + 2\left(\omega_w - \omega_p\right) r \left[v_{cc} \sin\theta + r_{cc}\omega_p \cos\theta\right] + v_{cc}^2 + r_{cc}^2\omega_p^2} \tag{2.14}$$

Typically, $v_{cc} << r_{cc}\omega_p$, so it can be omitted from consideration because it does not contribute much to the magnitude of relative velocity. Since Eq. 2.14 is valid for any point P in the interface, $\left|\vec{v}\right|_{P_{pad}}$ will be referred to as $\left|\vec{v}_R\right|$.

$$\left|\vec{v}_R\right| = \sqrt{\left(\omega_w - \omega_p\right)^2 r^2 + 2\left(\omega_w - \omega_p\right)\omega_p r r_{cc} \cos\theta + r_{cc}^2\omega_p^2} \tag{2.15}$$

Equation 2.15 can be expressed in a non-dimensionalized form by taking out $r_{cc}\omega_w$ out of the square-root term.

$$\left|\vec{v}_R\right| = r_{cc}\omega_w \sqrt{\left(1-\frac{\omega_p}{\omega_w}\right)^2\left(\frac{r}{r_{cc}}\right)^2 + 2\left(1-\frac{\omega_p}{\omega_w}\right)\left(\frac{\omega_p}{\omega_w}\right)\left(\frac{r}{r_{cc}}\right)\cos\theta + \left(\frac{\omega_p}{\omega_w}\right)^2} \tag{2.16}$$

Equation 2.16 shows that the magnitude of the relative velocity vector generally depends on the local position r and θ, and that it also depends on the ratio of pad to wafer rotational speeds ω_p/ω_w and on the linear speed of the center of pad, $r_{cc}\omega_w$. The non-dimensionalized square-root term basically just scales the most general linear velocity of the center of the pad. For completeness, here are the relative velocity vectors in both polar and cartesian coordinate systems and for both the wafer and the pad centered coordinate systems. Relative velocity vector expressed in pad's coordinate system is given by:

$$\vec{v}_R = \left\{r_{cc}\omega_w \sin\phi\right\}\hat{r} + \left\{\left(\omega_w - \omega_p\right)r + r_{cc}\omega_w \cos\phi\right\}\hat{\phi} \tag{2.17}$$

$$\vec{v}_R = \left\{-\left(\omega_w - \omega_p\right)r\sin\phi\right\}\hat{i} + \left\{\left(\omega_w - \omega_p\right)r\cos\phi + r_{cc}\omega_w\right\}\hat{j} \tag{2.18}$$

Relative velocity vector expressed in wafer's coordinate system is given by:

$$\vec{v}_R = \left\{r_{cc}\omega_p \sin\theta\right\}\hat{r} + \left\{\left(\omega_w - \omega_p\right)r + r_{cc}\omega_p \cos\theta\right\}\hat{\theta} \tag{2.19}$$

$$\vec{v}_R = \left\{-\left(\omega_w - \omega_p\right)r\sin\theta\right\}\hat{i} + \left\{\left(\omega_w - \omega_p\right)r\cos\theta + r_{cc}\omega_p\right\}\hat{j} \tag{2.20}$$

Once the relative velocity vector is known, it is easier to compute material removal rates and frictional forces and torques. It is important from the machine design perspective to know the forces and torques generated by the CMP process, so proper design of load bearing surfaces, motor size, and bearings can be selected.

2.5 Torques and Forces

Based on the Coulomb friction law, the local frictional force between the pad and the wafer depends on the polishing pressure and it acts opposite to the local sliding direction. Hence, the frictional force acting on an infinitesimal area dA located in the wafer/pad interface can be represented by:

$$d\vec{f}(r,\theta) = -\mu_f p_{nom} dA \left(\frac{\vec{v}_R}{|\vec{v}_R|} \right) \tag{2.21}$$

where f is the frictional force, μ_f is the coulomb friction coefficient, $p_{nom} dA$ is the normal load acting on a small area dA, and $-\frac{\vec{v}_R}{|\vec{v}_R|}$ is the unit vector pointing in the direction opposite of sliding velocity. The torque generated by the frictional force can be determined in the following way. It must be noted that one should use a consistent coordinate system. Both the relative velocity vector and the position vector must be in the same coordinate system.

$$\vec{T} = \int \vec{r} \times -\mu_f p_{nom} \left(\frac{\vec{v}_R}{|\vec{v}_R|} \right) dA \tag{2.22}$$

The total torque on the pad can be calculated by:

$$T_{pad} = \int_0^{r_p} \int_0^{2\pi} \frac{\mu p_{nom} \left\{ \omega_w r_{cc} r \cos\phi + \left(\omega_w - \omega_p \right) r^2 \right\} r d\phi dr}{|\vec{v}_R|} \tag{2.23}$$

where r and ϕ are the polar coordinates located at the center of the pad. Keeping the same coordinate system, the total torque on the wafer can be calculated by:

$$T_{wafer} = \int_0^{r_p} \int_0^{2\pi} \frac{\mu p_{nom} \left\{ \left(2\omega_w - \omega_p \right) r_{cc} r \cos\phi + \omega_w r_{cc}^2 + \left(\omega_w - \omega_p \right) r^2 \right\} r d\phi dr}{|\vec{v}_R|} \tag{2.24}$$

From Eqn. 2.21, the frictional force can be obtained by:

$$\vec{F} = \int_0^{r_p}\int_0^{2\pi} -\mu p_{nom}\left(\frac{\vec{v}_R}{|\vec{v}_R|}\right) r\,d\phi\,dr \qquad (2.25)$$

The friction force can be resolved into x and y components as:

$$F_x = \int_0^{r_p}\int_0^{2\pi} \mu p_{nom}\frac{(\omega_w - \omega_p) r\sin\phi}{|\vec{v}_R|} r\,d\phi\,dr \qquad (2.26)$$

$$F_y = \int_0^{r_p}\int_0^{2\pi} -\mu p_{nom}\frac{\{(\omega_w - \omega_p) r\cos\phi + r_{cc}\omega_w\}}{|\vec{v}_R|} r\,d\phi\,dr \qquad (2.27)$$

Equations 2.23, 2.24, 2.26, and 2.27 can be used to find out the polishing torques and friction forces by substituting the magnitude of the relative velocity vector in pad coordinate system. Those equations have to be integrated numerically unless the kinematic condition of $\omega_w = \omega_p$ is used. In that case, the loads on the system become:

$$T_{pad} = 0 \qquad (2.28)$$

$$T_{wafer} = \pi\mu p r_{cc} r_p^2 \qquad (2.29)$$

$$F_x = 0 \qquad (2.30)$$

$$F_y = \pi\mu p r_p^2 \qquad (2.31)$$

where μ is the friction coefficient between the wafer and the pad, p is the nominal pressure, r_{cc} is the center-to-center distance between the wafer and the pad, and r_p is the radius of the pad.

2.6 Material Removal Rate

The Preston equation describes material removal rate based on applied nominal pressure and relative velocity between the pad and the wafer. However, the Preston equation has an implicit assumption built in. That is, if the relative velocity is higher a point has a higher MRR because more pad slid over that point per unit time. Basically, the Preston equation assumes that the process is cyclical; it assumes the pad will come back to the same point later in time hence the faster it moves the faster it will return, so more material removing cycles will be done per unit time. This notion seems obvious but it has larger implications for the face-up polisher. Typical face-down polisher have the wafer always contacting the pad so any given point on the wafer has some local MRR so it is relatively easy to integrate the Preston equation to get the total material removed for a given polishing time.

The face-up polisher, however, is much more complicated since at any point on the wafer material will be removed only when it is in the pad-wafer interface. The time it spends outside the pad, rotating around to get back to the pad, is basically lost since no material gets removed over that time period.

$$\frac{dh}{dt} = k_p p_{nom} \left| \vec{v}_R \right| \tag{1.1}$$

Integrating both sides in time during one wafer rotation will yield an expression for material removed, Δh.

$$\Delta h = \int_0^{2\pi/\omega_w} k_p p_{nom} \left| \vec{v}_R \right| dt \tag{2.32}$$

Changing the variables of integration will result in easier expression for material removed since time t is a dummy variable.

$$dt = \frac{d\theta}{\omega_w} \tag{2.33}$$

Equation 2.33 is valid because the angular rate is constant. Splitting Eq. 2.32 into parts:

$$\Delta h(r)=\int_{0}^{2\pi}k_p p_{nom}\left|\vec{v}_R\right|\frac{d\theta}{\omega_w}=\int_{0}^{\theta_p}\frac{k_p p_{nom}\left|\vec{v}_R\right|}{\omega_w}d\theta+\int_{\theta_p}^{2\pi-\theta_p}\frac{k_p p_{nom}\left|\vec{v}_R\right|}{\omega_w}d\theta+\int_{2\pi-\theta_p}^{2\pi}\frac{k_p p_{nom}\left|\vec{v}_R\right|}{\omega_w}d\theta$$

(2.34)

where $\theta_p=\cos^{-1}\left(\frac{r_p^2-r^2-r_{cc}^2}{-2rr_{cc}}\right)$ is the angular span of the pad on the wafer from the positive x-axis to the end of the pad at the wafer radial position r. During one wafer revolution, time spent outside the pad is wasted since the pressure is zero, so that term disappears. Combining the terms into single integral yields the following:

$$\Delta h(r)=\int_{-\theta_p}^{\theta_p}\frac{k_p p_{nom}\left|\vec{v}_R\right|}{\omega_w}d\theta=2\int_{0}^{\theta_p}\frac{k_p p_{nom}\left|\vec{v}_R\right|}{\omega_w}d\theta \tag{2.35}$$

It is clear that the amount of material removed is a strong function of radial position on the wafer, r, even if relative velocity is uniform when operating in $\omega_w=\omega_p$ mode. This is mainly due to the fact that the limit of integration θ_p varies strongly in r. Alternatively, one can define average MRR at any given radius by:

$$\overline{MRR}=\Delta h(r)\frac{\omega_w}{2\pi}=\frac{1}{\pi}\int_{0}^{\theta_p}k_p p_{nom}\left|\vec{v}_R\right|d\theta \tag{2.36}$$

The non-uniform material removal during polishing is the main drawback of the face-up polisher. It is clear that if the pad was rotating but with fixed r_{cc}, then wafer level non-uniformity would be quite large. On the one hand, it is a major problem since the goal is to lower wafer-scale non uniformity, but on the other hand it can be used to substantially lower non-uniformity. The idea is simple and comes from the MRR analysis. The angular velocities must be set in such a way so that material removed per

revolution is highest nearer the center of the wafer and decreases going outward. Once all of the Cu in the center region is removed as desired, the pad would move slightly outward increasing r_{cc} so that it does not over-polish the central region. The pad would continuously move outward very slowly once the next annulus area is completed. In this fashion, the entire wafer can be polished uniformly due to a good control of over-polishing. But to allow such algorithm to work, one must ensure that the average MRR is the highest in the pad area nearest the center of the wafer.

Material removal per wafer revolution is known from Eq. 2.36 which is clearly a function of wafer radial position r. It is also a function of wafer to pad distance r_{cc} which generally is a function of time t.

$$\Delta h_{tot} = \int_0^{t_f} \frac{\omega_w}{2\pi} \frac{\Delta h(r,t)}{rev} dt \tag{2.37}$$

where t_f is the total polishing time. The total material removal generally is a function of r but it is desirable to make it constant to minimize non-uniformity. Combining Eq. 2.37 with Eq. 2.35:

$$\Delta h_{tot}(r) = \int_0^{t_f} \frac{2}{2\pi} \left(\int_0^{\theta_p(t)} k_p p_{nom} \left| \vec{v}_R(r,t) \right| d\theta \right) dt \tag{2.38}$$

$$\text{for } r_{cc} - r_p \le r \le r_{cc} + r_p$$

The objective of this thesis is to design a pad so that the Preston constant can be assumed constant in the wafer-pad interface. If the Preston constant and the nominal pressure is uniform throughout the pad area, it can be taken out of the integral for simplicity.

$$\Delta h_{tot}(r) = \frac{k_p p_{nom}}{\pi} \int_0^{t_f} \int_0^{\theta_p(t)} \left| \vec{v}_R(r,t) \right| d\theta dt \tag{2.39}$$

$$r_{cc} - r_p \le r \le r_{cc} + r_p$$

The goal is to derive such an increasing function $r_{cc}(t)$ so that Eq. 2.39 is within 5%, the desired wafer scale non-uniformity as stated in the Semiconductor Industry Road Map. Analytically it is unsolvable, but numerically it can be done. Even if such $r_{cc}(t)$ is obtained, however, the polishing will still show some non-uniformity due to the fact that the Preston constant, k_p, may vary spatially and its spatial variation is generally unknown. The assumption that the Preston constant is uniform would be violated so it cannot be taken out of the integral and treated as a constant. Hence, the previously obtained solution for $r_{cc}(t)$ should be applied judiciously.

The optimal solution for this problem is not to follow some predetermined $r_{cc}(t)$ in open-loop control, but to have continuous feedback on the polishing performance and adjust r_{cc} based on that information. That way, the only requirement is that the kinematics of the polishers must ensure that the polishing starts at the center of the wafer and continuously expands outward. This can be easily obtained by adjusting the angular velocities of the pad and the wafer as described in the next section. As long as polishing starts at the center and keeps growing outward, an end point detection sensor can determine the size of the polished region and the pad location can be adjusted during the process to minimize over-polishing. This is the most important benefit of the face-up polisher. Even though it shows very non-uniform polishing, it can be controlled to obtain excellent uniformity.

2.7 Polishing Condition

The average material removal rate at any point on the wafer is given by:

$$\overline{MRR} = \Delta h(r)\frac{\omega_w}{2\pi} = \frac{1}{\pi}\int_0^{\theta_p} k_p p_{nom}\left|\bar{v}_R\right| d\theta \tag{2.36}$$

Plotting Eq. 2.36 in a normalized form can be beneficial in understanding the material removal rate and how to set the kinematics of the polisher to get higher removal rates

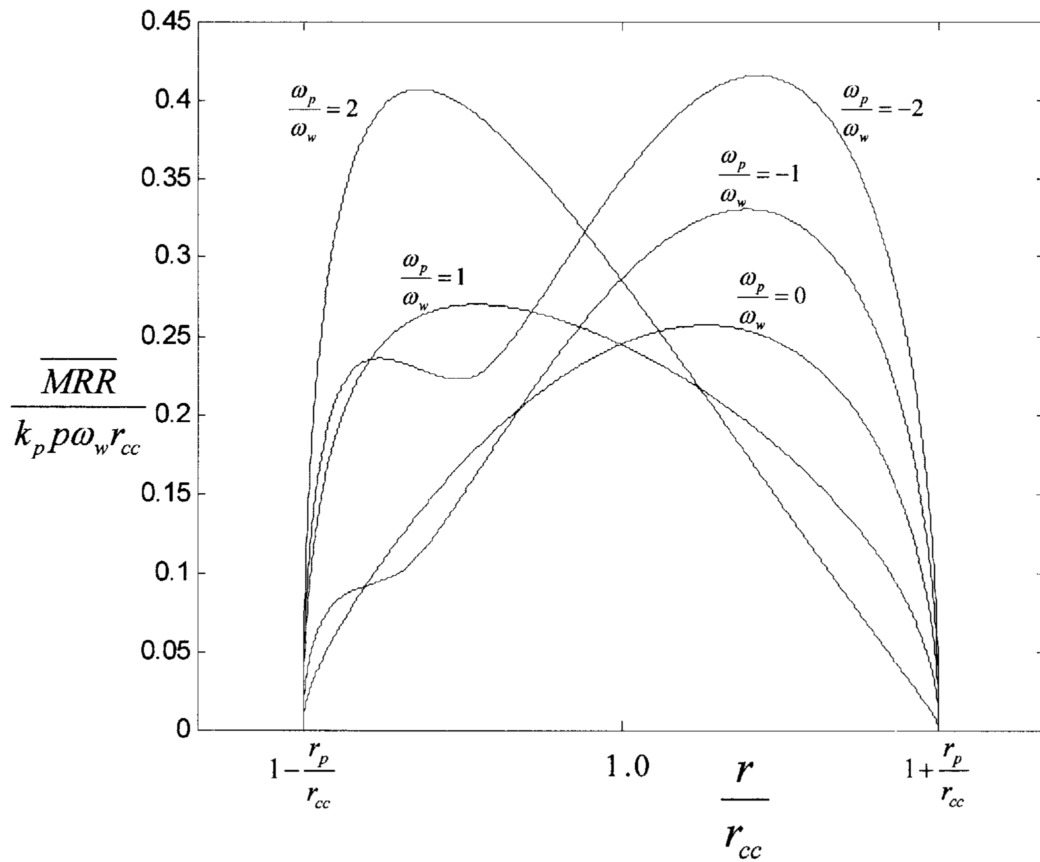


Figure 2.5. Average MRR as a function of radial position, pad outside the center.

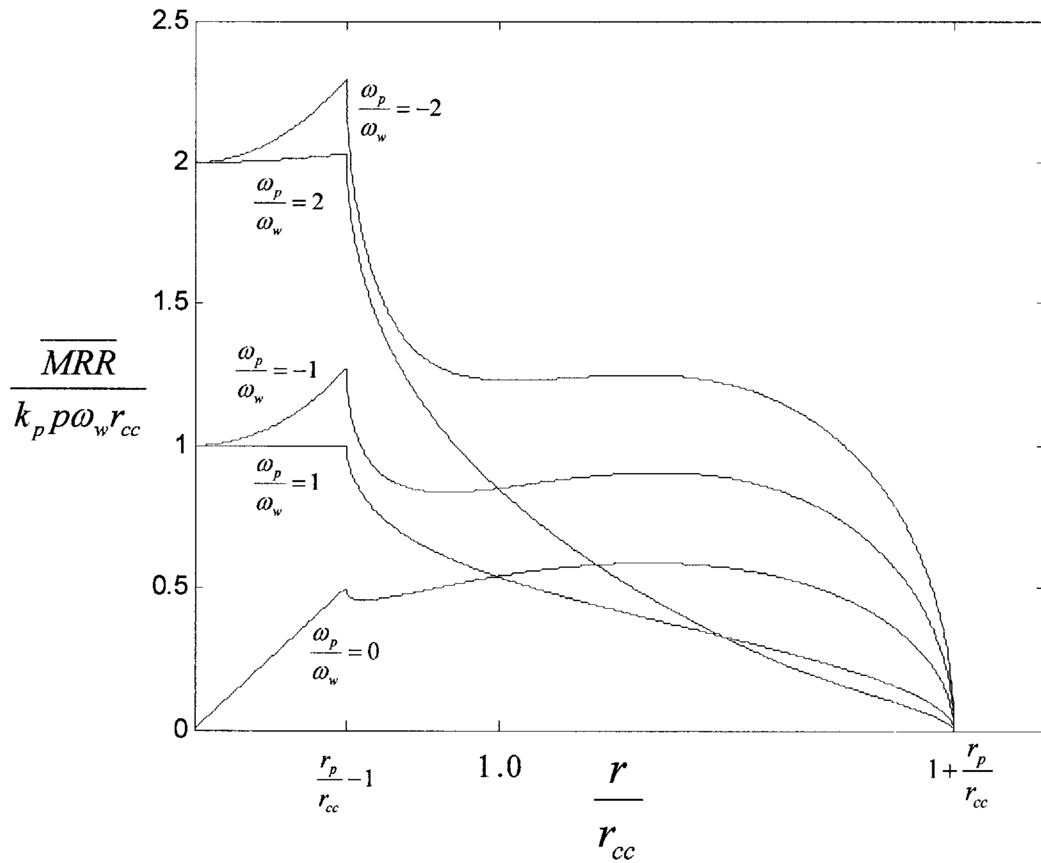


Figure 2.6. Average MRR as a function of radial position, pad overlapping the center.

nearer the center of the wafer. To obtain a condition of highest material removal rate at the side of the pad that is closest to the center of the wafer, $\omega_p/\omega_w > 1$ should be used. If this condition is met the material removal rate peaks near the inner region of the wafer.

Figs. 2.5 and 2.6 represent normalized MRR as a function of normalized radial position on the wafer for various pad-to-wafer speed ratios. Average MRR was normalized by $k_p p \omega_w r_{cc}$ because when the pad and wafer speeds are equal this term is constant and is taken out of the integral in Eq. 2.36. So, such normalization is only a function of the span of the pad and the pad-to-wafer speed ratio. The radial position was normalized by r_{cc} because it is the only radial distance that appears in $k_p p \omega_w r_{cc}$. This way, the only indirect variable, r_p, just scales the plots spatially (radially) and does not affect the material removal rates. In both figures the normalized average MRR was plotted as a function of normalized radial position on the wafer within the span of the pad for multiple pad-to-wafer speed ratios. Combining Eq. 2.16 and 2.36 to obtain Eq. 2.37 which is the normalized MRR:

$$\frac{\overline{MRR}}{k_p p \omega_w r_{cc}} = \frac{1}{\pi} \int_0^{\theta_p} \sqrt{\left(1 - \frac{\omega_p}{\omega_w}\right)^2 \left(\frac{r}{r_{cc}}\right)^2 + 2\left(1 - \frac{\omega_p}{\omega_w}\right)\left(\frac{\omega_p}{\omega_w}\right)\left(\frac{r}{r_{cc}}\right)\cos\theta + \left(\frac{\omega_p}{\omega_w}\right)^2}\, d\theta \tag{2.40}$$

$$\theta_p = \cos^{-1}\left\{\frac{\left(\frac{r_p}{r_{cc}}\right)^2 - \left(\frac{r}{r_{cc}}\right)^2 - 1}{-2\left(\frac{r}{r_{cc}}\right)}\right\} \tag{2.41}$$

It is apparent from Eqs. 2.40 and 2.41 that such a normalization seems natural. Equation 2.40, illustrates what fraction of maximum MRR is being accomplished with the maximum MRR occurring at the total pad coverage of the wafer at $\omega_p/\omega_w = 1$. Figures 2.5 and 2.6 are just graphical representation of Eq. 2.40 for various pad-to-wafer speed ratios and for two pad radius conditions; $r_p/r_{cc} < 1$ in Fig. 2.5 and $r_p/r_{cc} > 1$ in Fig. 2.6.

As long as $\omega_p/\omega_w \geq 1$, the inner part of the wafer will be polished first, allowing the pad to move out ever so slightly to continue to polish the un-finished outer region. Ideally, the material removal rate should be a decreasing function of wafer radial position so that the central region can be polished and one can continue polishing the outer region without over-polishing. However, the actual MRR is increasing before it peaks, so it is important to move the peak to the inner most position. Whatever material removal rate occurs between the beginning of the pad (MRR is zero) and the maximum MRR, will be virtually impossible to compensate for and will show up as wafer-scale non-uniformity. Ideally the wafer should exhibit highest MRR at the inner most point, but of course it is impossible because that inner-most point has zero area coverage of the wafer since at that point, $\theta_p = 0$ if $r_{cc} > r_p$. In practice the higher the ω_p/ω_w ratio the closer the highest MRR is to the left hand side. Figures 2.5 and 2.6 depict average MRR for various ω_p/ω_w ratios. Some of these ratios probably should not be used in polishing since they would cause larger non-uniformity but they are included in the figures for completeness.

Chapter 3

Pad Characterization for Fluid Flow

3.1 Introduction

In any polishing process that uses liquid slurry containing abrasive particles there are a few issues that must be addressed to ensure that the slurry is evenly distributed for uniform polishing. Moreover, the abrasive particles must be trapped at the interface for actual material removal. If the pad is modified by grooves to aid slurry distribution, it also allows high flow rates which are costly in CMP. In polishing there are three variables: slurry distribution, slurry consumption, and MRR and they are usually coupled. The semiconductor industry, of course, is striving toward uniform polishing with low slurry consumption and high material removal rates. It is essential to understand transport mechanisms of the slurry so that optimum pad topography can be designed to fulfill these requirements.

In order to analyze slurry transport mechanisms between a rough and a smooth surface, the pad must be characterized for fluid flow for a given set of polishing conditions. It is hypothesized that the fluid (slurry) flows between the asperities of the rough pad surface. The main point of the following experiments is to obtain effective gap between the wafer and the pad. The effective gap concept facilitates a fluid flow analysis, and the gap obtained in this set of experiments can be applied in CMP to estimate the flow behavior of the slurry.

The results of the following experiments can be compared with surface roughness measurements for validation. For instance, if the obtained effective gap between the pad and the wafer is much greater than the measured pad roughness, then it is clear that there was hydrostatic or hydrodynamic lift generated, which separates the two surfaces. The surface profiles are one-dimensional and some portions of the 'valleys' cannot facilitate any fluid flow because they are dead ends. Also, the pad roughness measurements are done at no load on the pad, which of course is not the actual polishing condition. The

advantage of the following technique of estimating pad roughness for fluid flow is that it can be done under pad compression simulating actual polishing conditions. The effective gap for fluid flow is obtained and used to determine slurry flow during CMP.

The experiment for determining the effective pad gap is a two-dimensional Poiseuille flow problem. The idea is to apply relatively simple fluid flow framework to a more complicated problem to obtain an effective parameter. Basically, the experiment is Poiseuille flow which is due to pressure gradient within a viscous fluid. The fluid is bounded by two surfaces, the polishing pad and the wafer.

3.2 Poiseuille Flow

In CMP, a rough polishing pad is pressed against a relatively smooth surface (wafer) and the fluid can flow between the high peaks on the pad. In order to determine this effective gap that characterizes the flow, several assumptions are made which describe Poiseuille flow. The main assumption is that the two bounding surfaces are smooth and flat and some small distance apart, $\bar{h}$. In reality, the pad is far from being smooth and flat and it definitely contacts the wafer. But an effective $\bar{h}$ can be defined which describes the pad's ability to transport fluid between the peaks contacting the wafer. The geometry of the problem is described in Fig. 3.1.

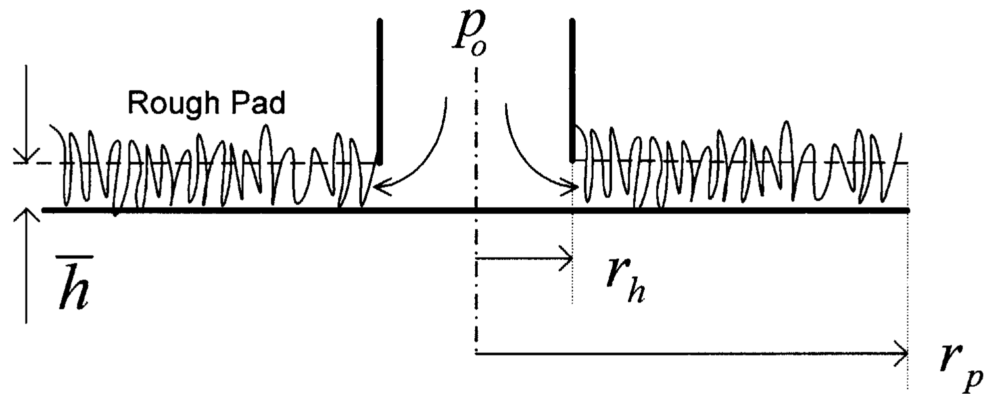
The complete Navier-Stokes (N-S) equations in polar coordinates are:

(r):

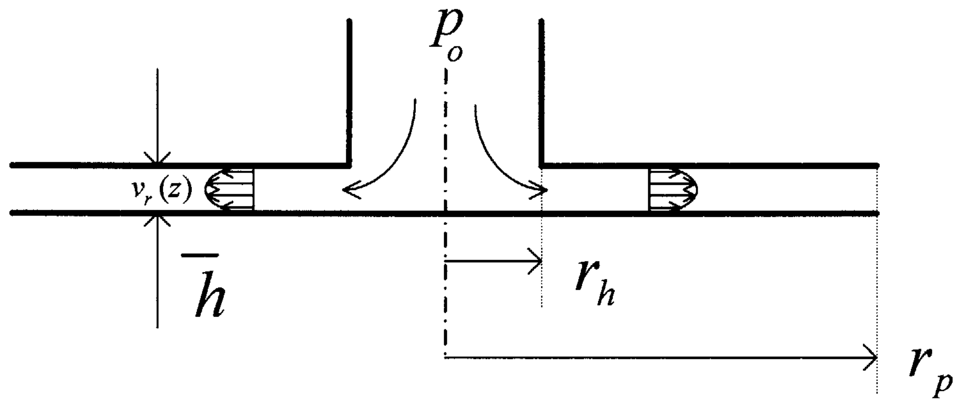
$$\rho\left(\frac{\partial v_r}{\partial t}+v_r\frac{\partial v_r}{\partial r}+\frac{v_\theta \partial v_r}{r\partial\theta}-\frac{v_\theta^2}{r}+v_z\frac{\partial v_r}{\partial z}\right)=-\frac{\partial p}{\partial r}+\mu\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r)\right)+\frac{1}{r^2}\frac{\partial^2 v_r}{\partial\theta^2}+\frac{\partial^2 v_r}{\partial z^2}-\frac{2}{r^2}\frac{\partial v_\theta}{\partial\theta}\right)+\rho g_r \tag{3.1}$$

(θ):

$$\rho\left(\frac{\partial v_\theta}{\partial t}+v_r\frac{\partial v_\theta}{\partial r}+\frac{v_\theta \partial v_\theta}{r\partial\theta}+\frac{v_r v_\theta}{r}+v_z\frac{\partial v_\theta}{\partial z}\right)=-\frac{1}{r}\frac{\partial p}{\partial\theta}+\mu\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_\theta)\right)+\frac{1}{r^2}\frac{\partial^2 v_\theta}{\partial\theta^2}+\frac{\partial^2 v_\theta}{\partial z^2}+\frac{2}{r^2}\frac{\partial v_r}{\partial\theta}\right)+\rho g_\theta \tag{3.2}$$



(a) Actual



(b) Idealization

Figure 3.1. Schematic of axisymmetric Poiseuille flow problem

(z):

$$\rho\left(\frac{\partial v_z}{\partial t}+v_r\frac{\partial v_z}{\partial r}+\frac{v_\theta \partial v_z}{r\partial \theta}+v_z\frac{\partial v_z}{\partial z}\right)=-\frac{\partial p}{\partial z}+\mu\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 v_z}{\partial \theta^2}+\frac{\partial^2 v_z}{\partial z^2}\right)+\rho g_z \tag{3.3}$$

where ρ is the fluid density, v_r the radial component of fluid velocity, v_θ the tangential component of fluid velocity, v_z the z-directional component of fluid velocity, t the time, p the pressure, μ the viscosity, g the acceleration due to gravity, and r, θ, and z are polar coordinates.

The boundary conditions relevant to the problem are:

1. No slip at stationary wafer

$$v_r(z=0)=0 \tag{3.4}$$

2. No slip at stationary pad

$$v_r(z=\bar{h})=0 \tag{3.5}$$

Since it is a steady state problem with axisymmetric flow, only the radial component of velocity remains due the following considerations:

1. There is no tangential velocity

$$v_\theta=0 \tag{3.6}$$

2. Radial component does not vary with θ

$$\frac{\partial v_r}{\partial \theta}=0 \tag{3.7}$$

3. There is no fluid moving in the z-direction (thin film assumption)

$$v_z=0 \tag{3.8}$$

Thus the N-S equations become:

(r):

$$\rho\left(v_r\frac{\partial v_r}{\partial r}\right)=-\frac{\partial p}{\partial r}+\mu\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r)\right)+\frac{\partial^2 v_r}{\partial z^2}\right) \tag{3.9}$$

(θ):

$$0 = \frac{1}{r}\frac{\partial p}{\partial \theta} \tag{3.10}$$

(z):

$$0 = -\frac{\partial p}{\partial z} + \rho g_z \tag{3.11}$$

Equation 3.10 shows that the pressure will not vary in the θ direction ($p = p(r)$), and Eq. 3.11 is hydrostatics. By introducing non-dimensional variables as:

$$v_r' = \frac{v_r}{U},\ r' = \frac{r}{r_p},\ z' = \frac{z}{\bar{h}},\ p' = \frac{p}{p_0} \tag{3.12}$$

where U is the maximum linear velocity of all the local velocities within the interface. Non dimensionalized N-S equations become:

(r):

$$\left(\frac{\rho U \bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right)\left(v_r' \frac{\partial v_r'}{\partial r'}\right) = -\frac{p_0 \bar{h}^2}{\mu U r_p}\frac{\partial p'}{\partial r'} + \left(\left(\frac{\bar{h}}{r_p}\right)^2 \frac{\partial}{\partial r'}\left(\frac{1}{r'}\frac{\partial}{\partial r'}(r' v_r')\right) + \frac{\partial^2 v_r'}{\partial z'^2}\right) \tag{3.13}$$

(z):

$$0 = \frac{\partial p'}{\partial z'} - \frac{\bar{h}}{p_0}\rho g_z \tag{3.14}$$

Since all of the non-dimensionalized terms, terms marked prime, are of the order of 1, one can easily see the criteria by which some terms can be neglected. The inertial term can be neglected if the following is satisfied:

$$\left(\frac{\rho U \bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right) << 1 \tag{3.15}$$

The remaining viscous term can be neglected if the following is satisfied:

$$\left(\frac{\bar{h}}{r_p}\right)^2 << 1 \tag{3.16}$$

Gravity can be neglected if the following is satisfied:

$$\left(\frac{\rho g \bar{h}}{p_0}\right) << 1 \tag{3.17}$$

Typically, the magnitudes of above parameters are as follows: $\bar{h} \cong 15\mu\text{m}$, $U \cong 0.01\,\text{m/s}$, $r_p \cong 0.0254\text{m}$, $\rho \cong 1000\,\text{kg/m}^3$, $\mu \cong 0.005\text{Pa}\cdot\text{s}$, $p_0 \cong 700\text{Pa}$. So the above dimensionless numbers can be estimated:

$$\left(\frac{\rho U \bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right) \cong 10^{-5} << 1 \tag{3.18}$$

$$\left(\frac{\bar{h}}{r_p}\right)^2 \cong 10^{-6} << 1 \tag{3.19}$$

$$\left(\frac{\rho g \bar{h}}{p_0}\right) \cong 10^{-4} << 1 \tag{3.20}$$

Hence, it is seen that all of these criteria are easily satisfied so that many of the remaining terms in the N-S equation can be neglected. The highly simplified N-S equation remains:

$$\frac{\partial p}{\partial r} = \mu\left(\frac{\partial^2 v_r}{\partial z^2}\right) \tag{3.21}$$

Equation 3.21 can be easily solved using the boundary conditions given as Eqs. 3.4 and 3.5. Integrating Eq. 3.21 in z and applying the boundary conditions, v_r can be obtained as:

$$v_r = \frac{\partial p}{\partial r}\left(\frac{z^2 - \bar{h}z}{2\mu}\right) \tag{3.22}$$

As one would expect with viscous flow bounded by non-moving surfaces, the velocity profile is parabolic and dependent on the viscosity of the fluid and the pressure gradient.

3.3 Flow Rate Analysis

Integration of Eq. 3.22 must be performed to obtain the flow rate coming out of the system at any radial position *r*. Assuming a steady state and hence the volumetric flow rate through any radial position *r* is equal due to continuity.

$$Q = \int_0^{\bar{h}} \int_0^{2\pi} v_r r \, d\theta dz \tag{3.23}$$

This yields a second differential equation:

$$\frac{\partial p}{\partial r} = \frac{-6Q\mu}{\pi \bar{h}^3 r} \tag{3.24}$$

It can be integrated using the pressure boundary conditions:

$$p(r = r_h) = p_0 \tag{3.25}$$

$$p(r = r_p) = 0 \tag{3.26}$$

yielding:

$$Q = \frac{\pi \cdot p_0 \bar{h}^3}{6\mu \ln\left(\frac{r_p}{r_h}\right)} \tag{3.27}$$

where μ is the fluid viscosity, r_p the outer radius of the pad, r_h is the inner radius of the pad (inlet of the fluid), p_o is the gauge pressure of the fluid, Q is the volumetric flow rate, and $\bar{h}$ is the effective gap between the two surfaces.

Equation 3.27 shows that for viscous flow between two disks $\bar{h}$ apart, the volumetric flow rate is proportional to the applied fluid pressure, inversely proportional to fluid viscosity and the logarithm of the ratio of the inlet and outlet radii. This result will be used to obtain the effective gap between the wafer and the pad, $\bar{h}$, by measuring flow rate coming out of such a system by knowing the fluid viscosity, internal pressure, and geometry.

3.4 Experimental Setup

A schematic of the experimental setup that was used in these experiments is shown in Fig. 3.2, while the actual apparatus is depicted in Fig. 3.3. The apparatus consists of a cylindrical structure with a single hole at the center of the bottom surface. The control surface (polished bearing steel disk) is attached to the bottom via a gasket seal. The holes at the bottom of the cylinder and the bearing steel piece are concentric so that fluid can flow from the cylinder through both holes and out radially through the pad roughness. The normal load is applied to the top of the cylinder with a pneumatic cylinder which is controlled with a pressure regulator and supplied from a gas tank. The top of the cylinder has a spherical hole made by a ball end mill so that a spherical rod end from the pneumatic cylinder could mate with it and align the cylinder. The cylinder itself is sealed and has two connections. The first is to allow water to be poured into it and once that is done it is closed. The second is an air connection that provides internal pressure for the fluid. After the cylinder is filled almost to the top with water, air at high pressure is introduced so that the water is forced out of the cylinder and through the pad roughness. The air pressure is controlled by a regulator and is equipped with a relief valve.

The experimental procedure is as follows. First, the wetted pad was placed underneath the control surface and it was cut to the same diameter to minimize pressure variation (normal pressure). The entire water container was thus loaded with the pneumatic cylinder to the appropriate normal pressure. The fluid (water) was introduced into the water container and the opening was sealed.

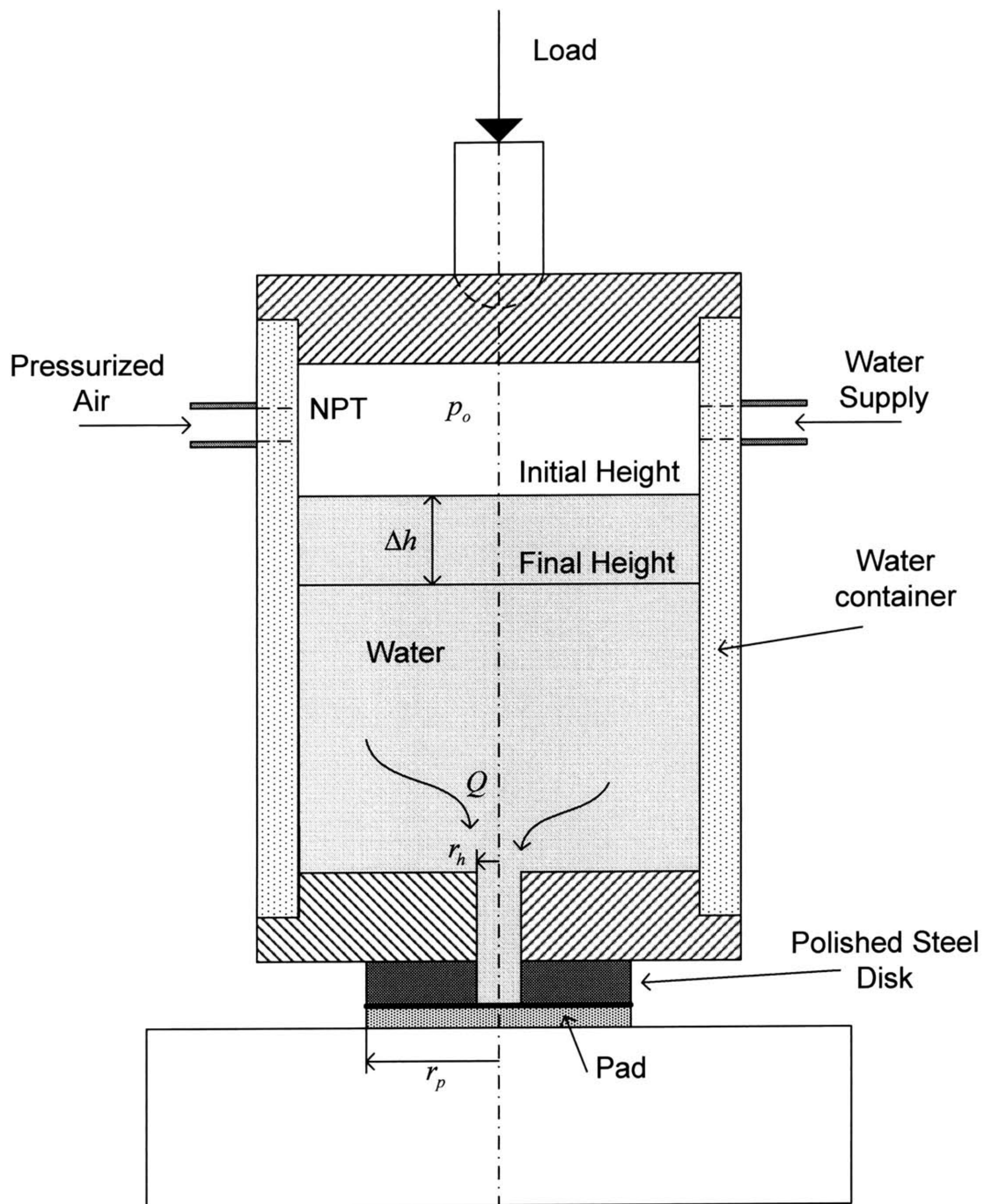


Figure 3.2. Experimental apparatus used to characterize the pad

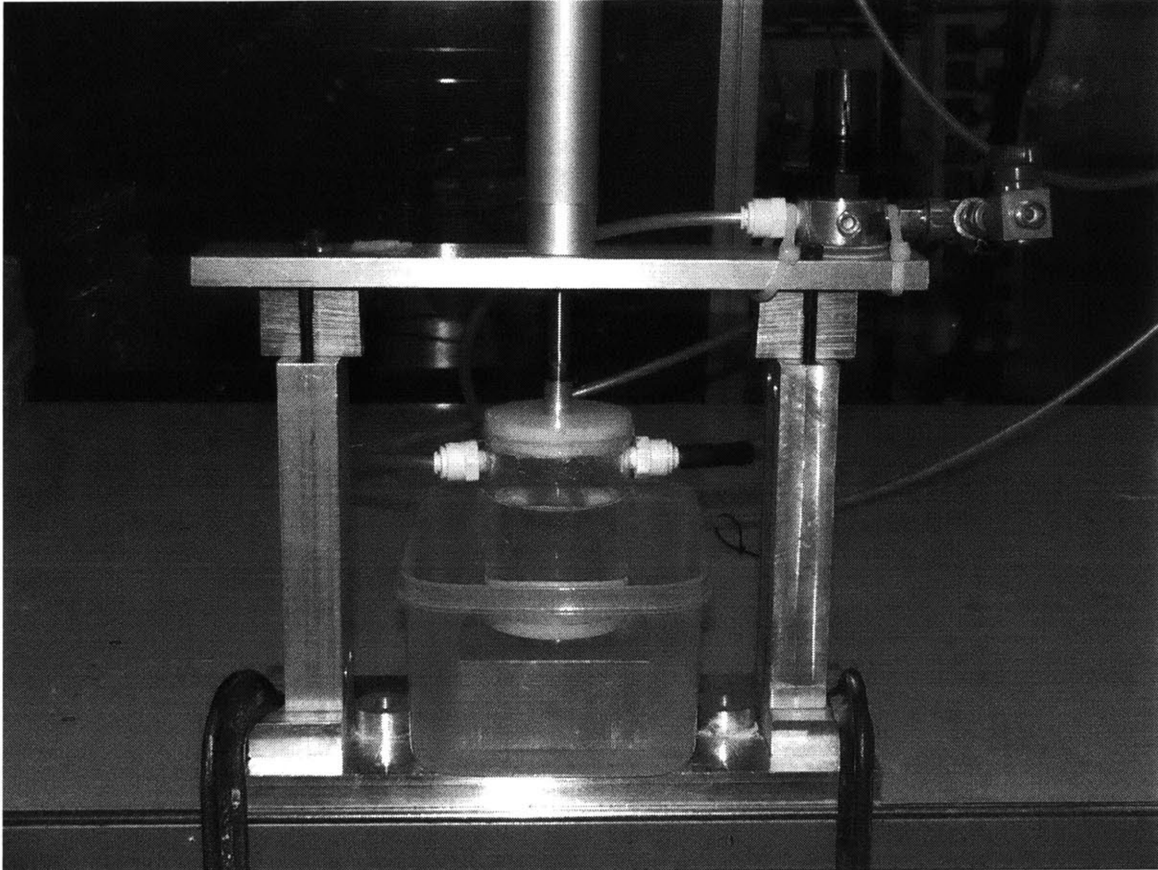


Figure 3.3. Photograph of the actual experimental apparatus.

Then the internal pressure was set to the desired value and after few seconds the initial height of the fluid was measured. The experiment was timed to start at the initial height reading. The entire experiment ran for no more than 20 minutes and the final fluid height was measured. Knowing the change in fluid height, the size of the water container, and the time, the flow rate was obtained which corresponds to the flow rate in Eq. 3.27. All the parameters but one in Eq. 3.27 are known; the missing parameter is flow rate and it was measured. Hence, Eq. 3.27 can be solved for the effective gap between the control surface and the pad, $\bar{h}$, which corresponds to the Poiseuille flow separation gap.

3.5 Results

Based on a set of experiments, several data points were obtained for different internal pressures and three different compressive pad pressures (loads). So far only one pad was measured due to the lack of pads without grooves. It turns out that grooves in the pad are much larger in cross-sectional area than any effective gap between the wafer and the pad so the grooves dominate the flow so one cannot obtain meaningful data describing the effective gap between the wafer and the pad. The only pad that was measured so far is Rodel IC 1400 pad without grooves (currently Rohm & Haas).

All the experiments were done at fairly high internal fluid pressures as compared with the head pressure that will be used in the face-up polisher. This was done so large enough flow rates were generated to get meaningful measurements. Of course by using internal fluid pressure, the fluid flow between the pad and the control surface is also pressurized which acts to reduce the compressive pad load. The fluid pressure wants to lift the pad and hence reduce the effective compressive load at the pad. So the quoted pad pressures of 14 kPa, 21 kPa, and 28 kPa are after compensation for this lift force. Calibration experiments were done to determine the lift force as a function of the internal fluid pressure. Hence, to obtain the quoted 14 kPa pad pressure, first, enough pressure was applied just to compensate the lift which depends on the internal fluid pressure and then additional load was applied to get to 14 kPa.

The results of the Rodel IC 1400 pad without grooves are shown in Fig. 3.4. One can see that there is some variation of obtained $\bar{h}$ but it is about 10μm and depends somewhat on the applied load. So the average $\bar{h}$ values for each pressure are given in the Table 3.1. It shows the pad is somewhat compressible even when considering the rather low down pressures. It shows that the effective gap is roughly 12 μm for 14 kPa pad load and decreases to 8 μm for 28 kPa pad load. The data show that the effective gap between the pad and the wafer is a function of pad pressure as one would expect based on contact mechanics. Similar experiments can be performed on different un-grooved pads yielding the amount of gap between the pad and the wafer during polishing which cannot really be done using any sort of profilometer measurements.

This set of experiments can estimate the effective gap between the wafer and the pad but the obtained flow velocity is not comparable to the flow velocity in actual polishing. For these experiments, the average flow velocity is given by:

$$\bar{v}_r = \frac{Q}{2\pi r_p \bar{h}} \tag{3.28}$$

The characteristic radial velocity in these experiments is velocity of the order of 0.01m/s, about two orders of magnitude below the polishing velocity. However, it was impractical and unsafe to try to reach fluid velocity of about 1 m/s in the interface since the internal fluid pressure have to be about 700 kPa (100 psi). At those pressures, the lift pressure would be roughly of the same magnitude thus the actual applied load would have to be about 700 kPa to compensate for lift plus 14 kPa to actually load the pad. Such a setup would have been impractical since the cylinder loading the pad would operate in the range of 1 MPa and the pressure control would have to be within 1% to make sure the pad would be loaded to 14 kPa or 21 kPa which the load during actual polishing.

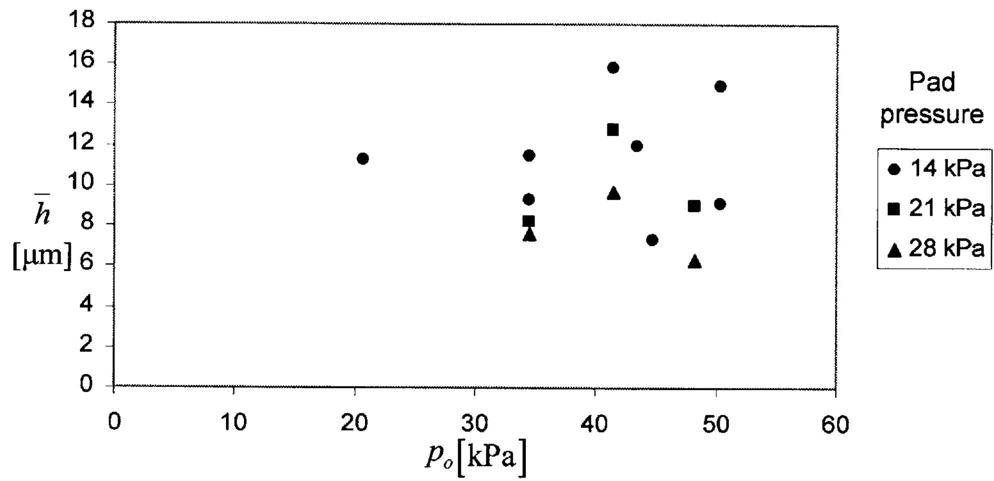


Figure 3.4. Effective gap data

Table 3.1 Effective gap as a function of pad pressure

Nominal Pad Pressure $[kPa]$	$\bar{h}_{avg}$ $[\mu m]$
14	11.4
21	10.0
28	7.9

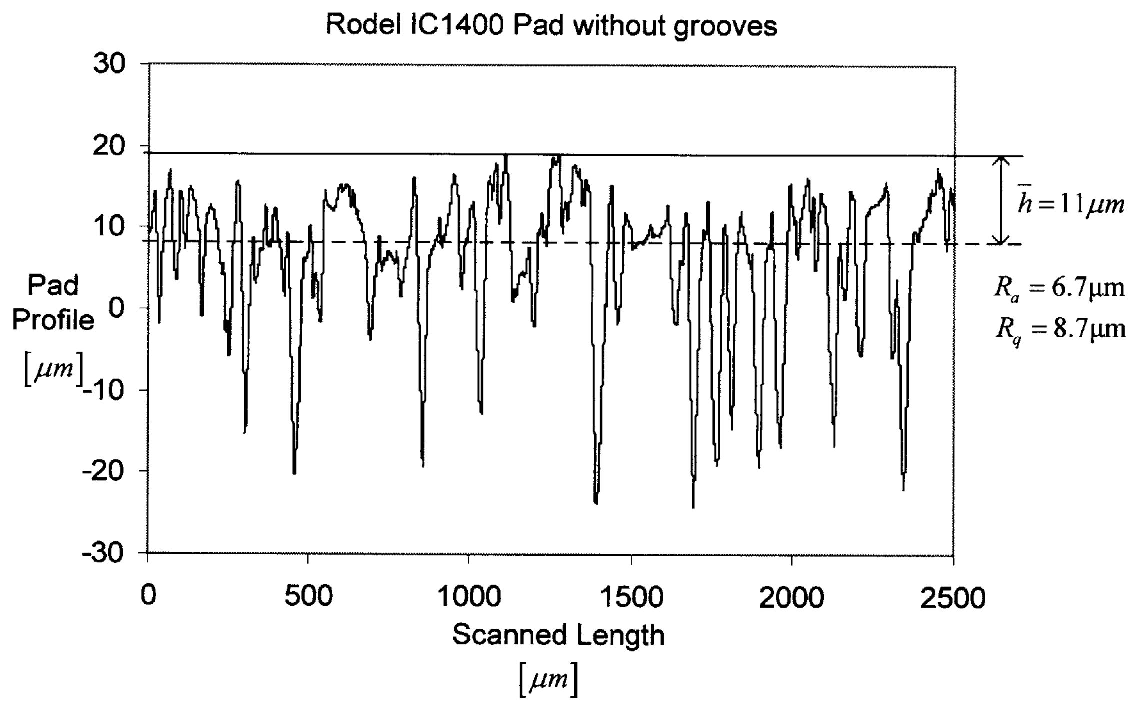


Figure 3.5 Pad profile compared to $\bar{h}$.

Figure 3.5 shows the actual profile of Rodel IC1400 pad without grooves and how it compares to the obtained $\bar{h}$ for the same pad. As may be seen, the obtained $\bar{h}$ of 11 μm seems to be validated by the profile. Obviously the actual pad is rough but one can see the obtained effective gap of 11 μm is of the same order as the RMS roughness. Of course, the profile was taken by a profilometer at no load while the obtained $\bar{h}$ is for 14 kPa loading. Nevertheless, the effective gap seems to be well within reason and can be used in the Couette model to estimate actual slurry flow rate during polishing.

3.6 Summary

The primary objective of the experiments described in this chapter is to idealize the interface between a rough pad and the relatively smooth wafer as a system of smooth, parallel plates separated by an effective gap, $\bar{h}$. The experiments clearly show that a loaded rough pad is capable of transporting fluid in the valleys of asperities. Poiseuille flow in such a system is compared to that in a system of two smooth surfaces $\bar{h}$ apart. It is a simple way of experimentally determining the effective gap between the wafer and the rough pad without the complex surface topography, contact mechanics, and hydrodynamic flow analyses.

Once the effective gap is determined by the Poiseuille flow experiments, it can be incorporated into the fluid flow model of rotating rough surfaces. For the known velocity boundary conditions and the effective gap, the fluid flow between rotating pad and wafer can now be modeled by recognizing that the flow is boundary-induced, as opposed to the pressure-induced Poiseuille flow. The expectation is that the effective gap, $\bar{h}$, would be the same for both flows for the same normal pressure on the pad and its surface topography.

Chapter 4

Analysis of Slurry Flow

4.1 Introduction

One of the most critical issues in CMP is the understanding of fluid transport at the wafer pad interface. There have been several studies investigating fluid transport within the CMP realm, but many of them do not offer any significant insight to the problem. Some researchers study the slurry flow assuming that the wafer is in hydrodynamic lubrication mode.[27,28] Others use extensive computer simulations which are difficult to interpret and do not present meaningful information in a clear and simple way[29,30]. It is widely believed that the slurry transport mechanism has a large effect on polishing performance and on slurry consumption. Hence it is imperative to have some basic understanding of the phenomenon and draw some conclusions relating to the kinematics and the pad design.

From the Poiseuille flow experiments, an effective gap between the wafer and the pad is established. This gap describes the effective spacing between the two surfaces which can be used in the next fluid flow model to estimate flow rate and general slurry behavior. During polishing, the wafer/pad system is not as simple as the Poiseuille setup since both are rotating. The wafer and the pad have their own local velocity which generally is different in magnitude and direction (two dimensionally). Contrary to the material removal rate calculation, the relative velocity vector alone is not sufficient to describe fluid flow. Both absolute velocity vectors determine the flow pattern in the polishing interface which determines slurry consumption and uniformity of slurry distribution.

The analysis is based on the recognition that slurry flow is driven by the moving boundaries, the wafer and pad surfaces being the two boundaries. Such flow is known as Couette flow. It is also assumed that the process is not in the hydrodynamic lubrication regime. The high peaks of the pad still support the load, but the slurry fills up the rest. It

is assumed that the slurry does not support any load and it is just entrapped by the two surfaces that rub against each other by the way of high peaks. The same high peaks ensure that there is some effective gap between the remaining areas of the two surfaces, $\bar{h}$. During polishing, the slurry fills up the effective gap between the wafer and the pad, and it is driven in various directions by the local velocity of the wafer from one side and local velocity of the pad from the other. Both velocities create an effective flow velocity which characterizes the local flow pattern.

In CMP, this relatively simple flow model can show some new insights to slurry distribution mechanism. It will identify which variables are important from the slurry consumption point of view. On a typical CMP machine, the slurry is fed from the outside and the user has complete control of how much to be fed. The user, however, does not really know what the minimum flow requirement for the process is and can either overfeed or starve the process. For the face-up polisher, the flow rate is determined by the velocities of the wafer and the pad. Also, the slurry flow rate is not controlled by the operator but the process draws as much as is allowed by the pad and thus it is always optimized in that sense. The following model will yield an estimate of the process needs based on the effective gap between the wafer and the pad for both face-up and face-down CMP.

4.2 The Couette Model

The fluid flow during polishing is similar to the Couette flow problem. The slurry is entrapped and dragged by two boundaries, the wafer on one side and the pad on the other. These boundaries have very different local velocities so the fluid near the pad is mostly following the pad and similarly near the wafer. One can consider an average fluid velocity which shows the direction and magnitude the fluid travels, on average, at a given local position. The following is an analysis of such flow neglecting many terms from the N-S equations which will be justified by making sure a few non-dimensional parameters are much less than one.

The boundary conditions for the N-S Eqs. in 3.1, 3.2, and 3.3 with coordinate system fixed in space and located at the center of the pad as shown in Fig. 4.1, are given by Eqs 4.1 through 4.4.

Radial component of wafer velocity:

$$v_r(z=0) = v_{r/wafer} = \omega_w r_{cc} \sin\phi \tag{4.1}$$

Radial component of pad velocity:

$$v_r(z=\bar{h}) = v_{r/pad} = 0 \tag{4.2}$$

Tangential component of wafer velocity:

$$v_\phi(z=0) = v_{\phi/wafer} = \omega_w r_{cc} \cos\phi + \omega_w r \tag{4.3}$$

Tangential component of pad velocity

$$v_\phi(z=\bar{h}) = v_{\phi/pad} = \omega_p r \tag{4.4}$$

From the fact that the gap between the wafer and the pad is small, the pad and wafer are impermeable, and the wafer is horizontal, one can recognize that

$$v_z(r,\phi,z) = 0 \tag{4.5}$$

$$g_r = g_\phi = 0 \tag{4.6}$$

The N-S equations get simplified to the following:

(r):

$$\rho\left(\frac{\partial v_r}{\partial t} + v_r\frac{\partial v_r}{\partial r} + \frac{v_\phi \partial v_r}{r\partial\phi} - \frac{v_\phi^2}{r}\right) = -\frac{\partial p}{\partial r} + \mu\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r)\right) + \frac{1}{r^2}\frac{\partial^2 v_r}{\partial\phi^2} + \frac{\partial^2 v_r}{\partial z^2} - \frac{2}{r^2}\frac{\partial v_\phi}{\partial\phi}\right) \tag{4.7}$$

(ϕ):

$$\rho\left(\frac{\partial v_\phi}{\partial t} + v_r\frac{\partial v_\phi}{\partial r} + \frac{v_\phi \partial v_\phi}{r\partial\phi} + \frac{v_r v_\phi}{r}\right) = -\frac{1}{r}\frac{\partial p}{\partial\phi} + \mu\left(\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_\phi)\right) + \frac{1}{r^2}\frac{\partial^2 v_\phi}{\partial\phi^2} + \frac{\partial^2 v_\phi}{\partial z^2} + \frac{2}{r^2}\frac{\partial v_r}{\partial\phi}\right) \tag{4.8}$$

(z):

$$0 = -\frac{\partial p}{\partial z} + \rho g_z \tag{4.9}$$

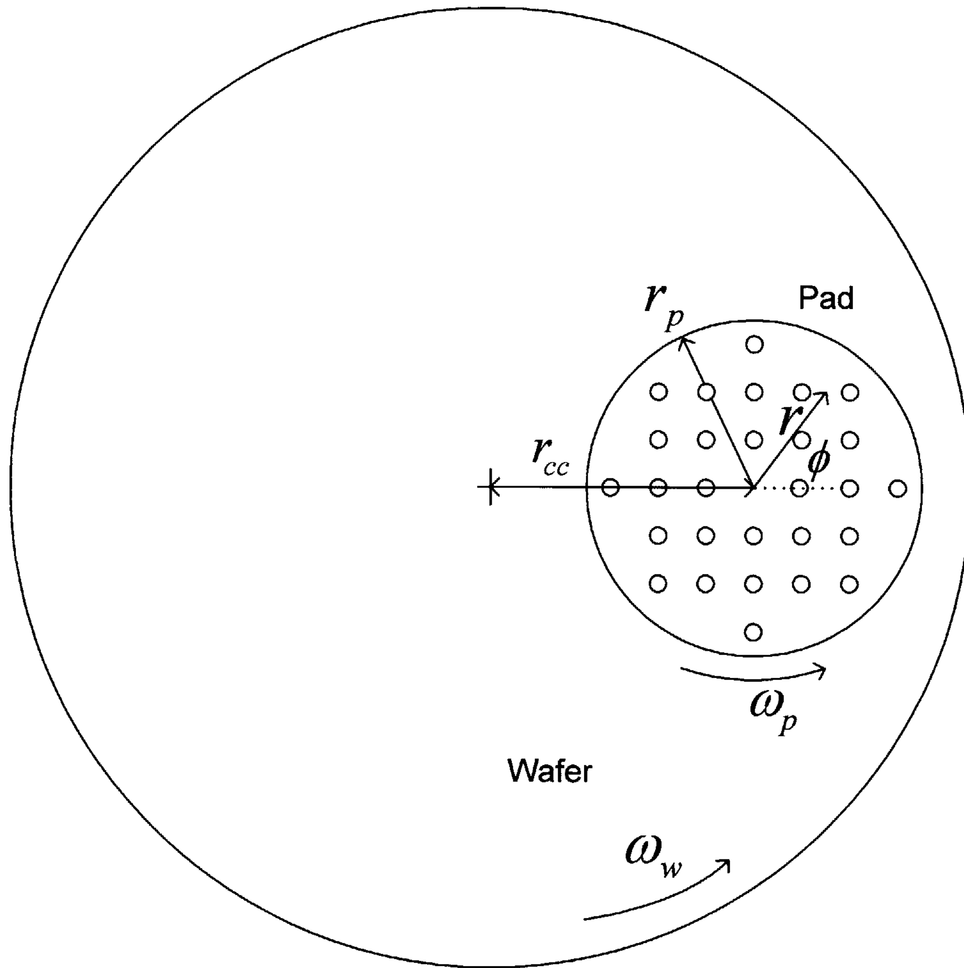


Figure 4.1 Face-up polisher's pad coordinate system

Equation 4.9 is just the hydrostatic pressure distribution.

By using a non-dimensional analysis one can quickly find out which remaining terms of N-S equations are significant and which can be neglected. The following are definitions of the non-dimensional variables:

$$v_r' = \frac{v_r}{U},\ v_\phi' = \frac{v_\phi}{U},\ r' = \frac{r}{r_p},\ p' = \frac{p}{p_0},\ z' = \frac{z}{\bar{h}},\ t' = \frac{t}{\tau} \qquad (4.10)$$

where U is the maximum local linear velocity in the polishing interface.

The Non-dimensionalized N-S equations become:

(r):

$$\left(\frac{\rho U\bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right)\left[\left(\frac{r_p}{U\tau}\right)\frac{\partial v_r'}{\partial t'}+v_r'\frac{\partial v_r'}{\partial r'}+\frac{v_\phi'\partial v_r'}{r'\partial\phi}-\frac{v_\phi'^2}{r'}\right]=-\left(\frac{p_0\bar{h}^2}{\mu r_p U}\right)\frac{\partial p'}{\partial r'}+\left[\left(\frac{\bar{h}}{r_p}\right)^2\left(\frac{\partial}{\partial r'}\left(\frac{1}{r'}\frac{\partial}{\partial r'}(r'v_r')\right)+\frac{1}{r'^2}\frac{\partial^2 v_r'}{\partial\phi^2}-\frac{2}{r'^2}\frac{\partial v_\phi'}{\partial\phi}\right)+\frac{\partial^2 v_r'}{\partial z'^2}\right] \qquad (4.11)$$

(ϕ):

$$\left(\frac{\rho U\bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right)\left[\left(\frac{r_p}{U\tau}\right)\frac{\partial v_\phi'}{\partial t'}+v_r'\frac{\partial v_\phi'}{\partial r'}+\frac{v_\phi'\partial v_\phi'}{r'\partial\phi}-\frac{v_r'v_\phi'}{r'}\right]=-\left(\frac{p_0\bar{h}^2}{\mu r_p U}\right)\frac{\partial p'}{r'\partial\theta}+ \qquad (4.12)$$
$$+\left[\left(\frac{\bar{h}}{r_p}\right)^2\left(\frac{\partial}{\partial r'}\left(\frac{1}{r'}\frac{\partial}{\partial r'}(r'v_\phi')\right)+\frac{1}{r'^2}\frac{\partial^2 v_\phi'}{\partial\phi^2}+\frac{2}{r'^2}\frac{\partial v_r'}{\partial\phi}\right)+\frac{\partial^2 v_\phi'}{\partial z'^2}\right]$$

(z):

$$0=-\frac{\partial p'}{\partial z'}+\left(\frac{\bar{h}}{p_0}\right)\rho g_z \qquad (4.13)$$

Since all of the non-dimensionalized terms, marked prime, are of the order of 1, given the following criteria, the N-S equations can be easily simplified.

The spatial inertial terms can be neglected if the following is satisfied:

$$\left(\frac{\rho U\bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right)\ll 1 \qquad (4.14)$$

The unsteady inertial term can be neglect if the following is satisfied:

$$\left(\frac{\rho \bar{h}^2}{\mu \tau}\right) \ll 1 \tag{4.15}$$

All other viscous terms can be neglected if the following is satisfied:

$$\left(\frac{\bar{h}}{r_p}\right)^2 \ll 1 \tag{4.16}$$

The gravity term can be neglected if the following is satisfied:

$$\left(\frac{\rho \bar{h} g}{p_0}\right) \ll 1 \tag{4.17}$$

Typically in CMP, the magnitudes of above parameters are as follows: $\bar{h} \cong 15\mu\text{m}$, $U \cong 1\,\text{m/s}$, $r_p \cong 0.0254\text{m}$, $\rho \cong 1000\,\text{kg/m}^3$, $\mu \cong 0.005\text{Pa}\cdot\text{s}$, $p_0 \cong 700\text{Pa}$, $\tau \cong 0.025\text{s}$. So the above dimensionless numbers can be estimated and all are less than 1 as desired.

$$\left(\frac{\rho U \bar{h}}{\mu}\right)\left(\frac{\bar{h}}{r_p}\right) \cong 0.001 << 1 \tag{4.18}$$

$$\left(\frac{\rho \bar{h}^2}{\mu \tau}\right) \cong 0.01 << 1 \tag{4.19}$$

$$\left(\frac{\bar{h}}{r_p}\right)^2 \cong 10^{-6} << 1 \tag{4.20}$$

$$\left(\frac{\rho g \bar{h}}{p_0}\right) \cong 10^{-4} << 1 \tag{4.21}$$

Since all of these parameters are less than unity, the N-S equations become much simpler and are of the form:

(r):

$$\frac{\partial p}{\partial r} = \mu \frac{\partial^2 v_r}{\partial z^2} \tag{4.22}$$

(ϕ):

$$\frac{\partial p}{r\partial\phi}=\mu\frac{\partial^2 v_\phi}{\partial z^2} \tag{4.23}$$

(z):

$$\frac{\partial p}{\partial z}=0 \tag{4.24}$$

The simplified N-S equations are quite easy to solve since they are uncoupled ordinary differential equations.

To solve Eq. 4.22, one needs the boundary conditions being the radial components of wafer and pad velocities:

$$v_r(z=0)=v_{r/wafer}=\omega_w r_{cc}\sin\phi \tag{4.1}$$

$$v_r(z=\bar{h})=v_{r/pad}=0 \tag{4.2}$$

Integrating both sides of Eq. 4.22 in z and applying the boundary conditions given in Eqs. 4.1 and 4.2:

$$v_r=\frac{1}{2\mu}\frac{\partial p}{\partial r}z^2-\left(\frac{\bar{h}}{2\mu}\frac{\partial p}{\partial r}+\frac{\omega_w r_{cc}\sin\phi}{\bar{h}}\right)z+\omega_w r_{cc}\sin\phi \tag{4.25}$$

To solve Eq. 4.23, one needs the boundary conditions being the tangential components of wafer and pad velocities:

$$v_\phi(z=0)=v_{\phi/wafer}=\omega_w r_{cc}\cos\phi+\omega_w r \tag{4.3}$$

$$v_\phi(z=\bar{h})=v_{\phi/pad}=\omega_p r \tag{4.4}$$

Integrating both sides of Eq. 4.23 in z and applying boundary conditions as given in Eqs. 4.3 and 4.4:

$$v_\phi = \frac{1}{2r\mu}\frac{\partial p}{\partial \phi}z^2 - \left(\frac{\bar{h}}{2r\mu}\frac{\partial p}{\partial \phi} + \frac{\omega_w r_{cc}\cos\phi + \left(\omega_w - \omega_p\right)r}{\bar{h}}\right)z + \omega_w r_{cc}\cos\phi + \omega_w r \quad (4.26)$$

Both components of the fluid velocity are functions of r, ϕ, and z. To eliminate the dependents on z, let the local mean fluid velocities be defined as:

$$\bar{v}_r = \frac{1}{\bar{h}}\int_0^{\bar{h}} v_r dz \quad (4.27)$$

$$\bar{v}_\phi = \frac{1}{\bar{h}}\int_0^{\bar{h}} v_\phi dz \quad (4.28)$$

Integrating both equations leads to both components of mean fluid velocity vector field under the pad.

$$\bar{v}_r = -\frac{\partial p}{\partial r}\frac{\bar{h}^2}{12\mu} + \frac{1}{2}\omega_w r_{cc}\sin\phi \quad (4.29)$$

$$\bar{v}_\phi = -\frac{\partial p}{r\partial \phi}\frac{\bar{h}^2}{12\mu} + \frac{1}{2}\left(\omega_w r_{cc}\cos\phi + \left(\omega_w + \omega_p\right)r\right) \quad (4.30)$$

Neglecting the flow due to the pressure gradient between the perforations in the pad and the outside atmospheric pressure, $\frac{\partial p}{\partial r} = \frac{\partial p}{r\partial \phi} = 0$. By substituting numbers from validation of N-S simplifications, the velocity due to pressure gradient is much lower than the velocity due to moving boundaries so an even simpler solution is obtained:

$$\bar{v}_r = \frac{1}{2}\omega_w r_{cc}\sin\phi \quad (4.31)$$

$$\bar{v}_\phi = \frac{1}{2}\left(\omega_w r_{cc}\cos\phi + \left(\omega_w + \omega_p\right)r\right) \quad (4.32)$$

where ω_w is the wafer angular speed, ω_p is the pad angular speed, r_{cc} is the distance between center of the pad and the center of the wafer, r and ϕ are the polar coordinate centered on the pad, $\bar{v}_r$ is the mean (z-directional) local radial fluid velocity, and $\bar{v}_\phi$ is the mean (z-directional) local tangential fluid velocity. Equations 4.31 and 4.32 show how the mean fluid velocity depends on the wafer and pad angular rates and the local position. Figure 4.2 is the plot of Eqs. 4.31 and 4.32 for the stated speed ratio. It shows the mean fluid velocity vectors whose radial component is Eq. 4.31 and tangential component is Eq. 4.32. Figure 4.2 illustrates how the local fluid velocities are distributed based on the Couette model. More of similar figures are shown in Appendix as Figs. A.1.a-d. The fluid flow velocity is basically a vector average of both the wafer and pad velocity. It shows how the slurry away from the center of the wafer has a much higher average velocity than the slurry nearer the center of the pad. The actual local slurry velocity may affect the distribution of fresh slurry coming from the perforations. Thus, it may affect polishing uniformity even if the relative velocity is uniform.

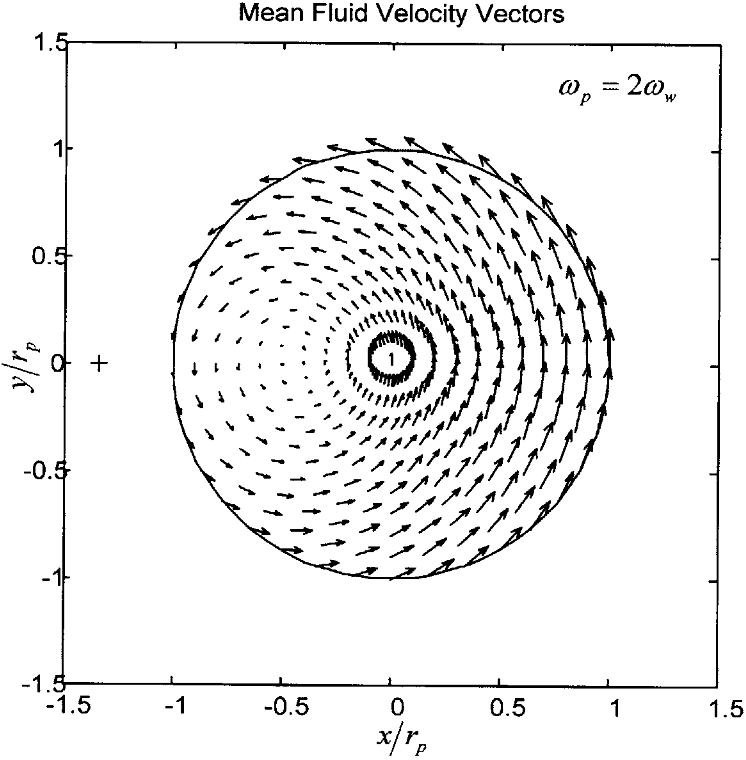


Figure 4.2 Mean fluid velocity vector field.

4.3 Flow Rate Analysis

The net flow across any arbitrarily drawn closed contour within the pad, including the pad circumference itself, is depicted in Fig 4.3 and given by:

$$Q = \oint_{CS} \bar{v} \cdot \hat{n} dA = 0 \tag{4.33}$$

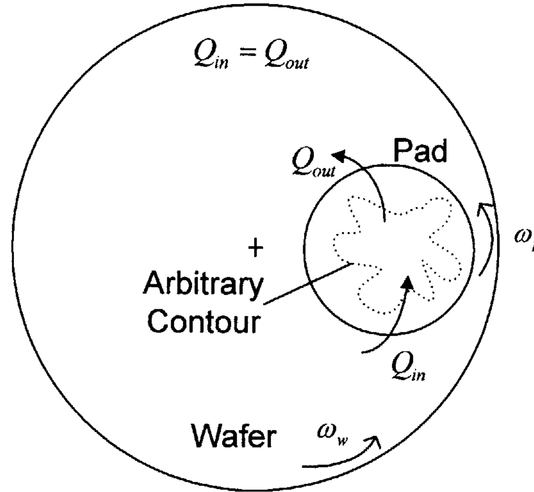


Figure 4.3. Depiction of an arbitrary contour

This is true since continuity was assumed in the derivation of the Eqs. 4.31 and 4.32.

$$\nabla \cdot \vec{v} = \frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{\partial v_\phi}{r \partial \phi} = 0 \tag{4.34}$$

Thus, the net flow rate around the circumference of the pad will be zero since inflow will equal outflow due to continuity. If we care to know the net flow through the pad or any area, only positive or negative part of $v \cdot \hat{n}$ should be considered. So, the net outflow is given by:

$$Q_o = \int_0^\pi v_r \bar{h} r_p d\phi = \int_0^\pi \frac{1}{2} \omega_w r_{cc} r_p \bar{h} \sin\phi d\phi = \omega_w r_{cc} r_p \bar{h} \tag{4.35}$$

Equation 4.35 shows that the net flow through the polishing interface depends on the wafer rotational speed but not on pad speed. Pad velocity does not contribute to the radial velocity of the fluid since the inertial term is so low. The flow rate also depends linearly on the gap between the wafer and the pad as long as the gap does not become too large and violate the assumptions. Grooves could violate assumptions of this model, so grooved pad cannot be used to verify the analysis.

The Couette model shows that there is an $\bar{h}$ thick fluid film leaving the wafer/pad interface. This fluid film can be either spun off the surface of the wafer or most of it can come back to re-enter the wafer-pad interface after one revolution of the wafer. If all of the fluid comes back around, there would be zero net flow rate due to continuity. No fluid would ever leave the wafer-pad system. In reality, some of the fluid gets spread out by inertial forces so that the fluid film coming back is smaller than $\bar{h}$. This fluid film thickness reduction would be the minimum flow rate experienced in CMP while the maximum would be the Couette model. To estimate the fluid film thickness reduction during one wafer revolution, one can use flow of viscous liquid on a rotating disk analysis derived by Emslie et al[31]

The analysis by Emslie et al. makes several assumptions which do not seem unreasonable for CMP application.

1. The rotating plane is infinite in extent.
2. The plane is horizontal so that there is no radial gravitational component.
3. The liquid layer is radially symmetric, and so thin that differences in gravitational potential normal to the surface of the disk have negligible effect in distribution of the liquid compared with the effect of centrifugal forces.
4. The viscosity is independent of the rate of shear; i.e., the liquid is Newtonian.
5. The liquid layer is everywhere so thin that shear resistance is appreciable only in horizontal planes.
6. The radial velocity is everywhere so small that Coriolis forces may be neglected.

Starting with initially uniform fluid thickness distribution, Emslie et al. solved for the fluid thickness as a function of time as:

$$h_f = \frac{h_0}{\sqrt{1+4\frac{\rho\omega_w^2 h_0^2 t}{3\mu}}} \tag{4.36}$$

where h_0 is the initial fluid thickness which is the $\bar{h}$ between the wafer and the pad, h_f the final fluid thickness after some time *t*, ρ the fluid density, μ viscosity, and ω_w is the angular velocity of the wafer. So in order to estimate how much fluid actually leaves the system one needs to figure out the change in fluid thickness during one wafer rotation. The time during one wafer rotation is $t = 2\pi/\omega_w$ and thus the reduction of fluid film thickness is given by the difference between the initial and final thicknesses.

$$\Delta h = h_0 - h_f = h_0\left(1-\frac{1}{\sqrt{1+8\frac{\rho\omega_w h_0^2\pi}{3\mu}}}\right) \tag{4.37}$$

Equation 4.37 estimates how much the fluid film thickness decreases during one wafer revolution. Fluid coming back around effectively reduces $\bar{h}$ but most of the fluid comes back around and reduces the space that new fluid can fill up, effectively reducing the flow rate coming out of the system. One can still use the Couette model to estimate what the reduced fluid flow is by substituting Eq. 4.37 into Eq. 4.35.

$$Q = \omega_w r_{cc} r_p \bar{h}\left(1-\frac{1}{\sqrt{1+8\frac{\rho\omega_w \bar{h}^2\pi}{3\mu}}}\right) \tag{4.38}$$

Equation 4.38 shows a significant reduction (about 95%) in flow rate as opposed to Eq. 4.35. In the following sections, it will be apparent that Eq. 4.35 is the upper bound and Eq. 4.38 is the lower bound on the actual polishing flow rates.

4.4 Experimental Setup

All the experiments for the Couette flow validation were done on the face-up polisher described in Chapter 2. The main goal of doing these experiments was to show whether or not the Couette model can be validated experimentally for conditions encountered in CMP. The polisher was equipped with a fairly small pad with diameter of 38 mm and 21 holes spaced 6.3 mm apart in a square matrix. The pad used for these experiments was Rodel IC 1400 without groves. It is the same pad that was characterized with an effective gap of about 11 μm by the Poiseuille experiment. The pad being smaller than the radius of the wafer (50 mm) was able to be fully within the wafer area but at the same time be off-center ($r_{cc} = 25.4\text{mm}$). The wafer was a Si wafer with a polished SiO_2 top surface. The wafer was held in the wafer carrier by a vacuum chuck. Fig 4.4 is a schematic of the face-up polisher used for these experiments.

The pad was then placed onto the wafer with the desired $r_{cc} = 25.4\text{mm}$ and the cup holding the pad was filled with water to about 63 mm in height. Before any motors were started, the entire wafer surface was wetted and the initial height of water was recorded. Then, the experiment proceeded to speed up the wafer and the pad simultaneously to the desired angular speeds. Once the speeds were stabilized, the timer was started to time the experiment. Typically, the experiment lasted 7-15min, and was marked by stopping the wafer and the pad and then reading off the remaining height of the water. For higher speeds the measurements were taken while the pad rotated. Because a rotating cup generates a parabolic shape in the fluid, the measurements were taken consistently at the top of this parabolic shape. These experiments were performed on the same face-up machine used for CMP and with similar parameters so that they effectively exhibit conditions encountered in CMP.

From the inner area of the cup ($A_{cup} = 19.6\,\text{cm}^2$), the height difference between the start and end of experiment, and the total experiment time, the total water flow rate was calculated and compared with the Couette model. The above experiment was repeated for various speeds of the wafer and the pad to make sure the general trends depending on velocity were observed.

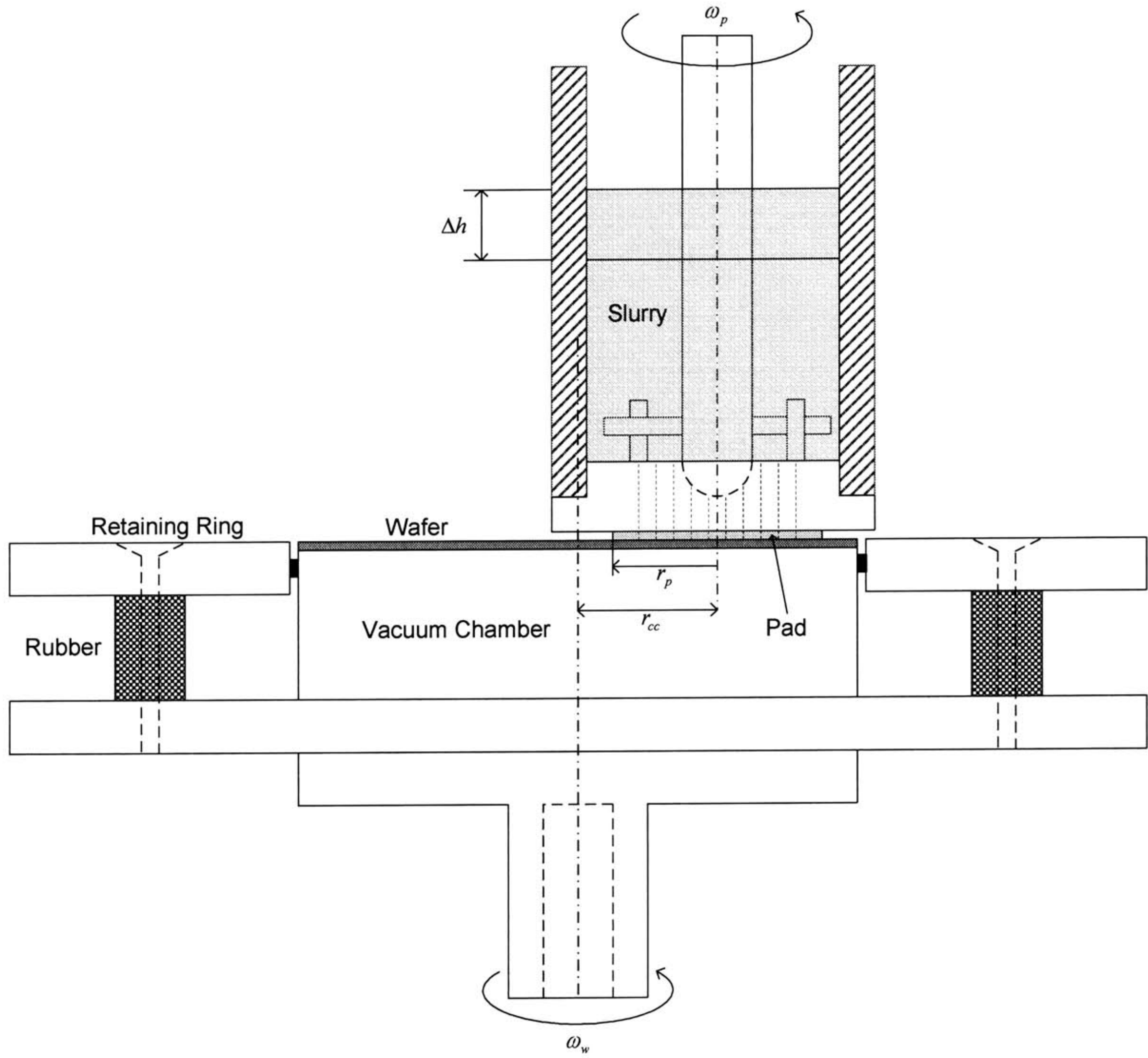


Figure 4.4. Drawing of the face-up polisher.

4.5 Results

Generally, the measured flow rates were between those predicted by the Emslie et al. Couette correction and the pure Couette model. The actual flow rates are lower than the pure Couette model because most of the fluid stays on the wafer and comes back around to re-enter the interface. This effectively reduces the available space for new slurry and hence it lowers the flow rate. The Emslie et al. Couette correction tries to estimate how much slurry stays on the wafer and is capable of re-entering the interface. It turns out over 95% of the slurry stays on the wafer because there is very little time for the thin layer of slurry to be spread out by the inertial forces of the rotation. However, the Emslie et al. Couette correction underestimates the actual slurry flow rates probably because not all of the returning 95% of the slurry is able to re-enter the interface. There is some efficiency at which the returning slurry can re-enter the interface at these relatively high speeds. Only a fraction of that 95% will actually be able to re-enter the interface, and the rest will be blocked and spun out by the pad.

The experimental results are summarized in the Fig. 4.5 and the data are available in Table B.2 in Appendix. These results are bounded by the Couette model from the above and the Emslie et al. Couette correction below. Both of these models do not completely describe the actual slurry flow rates because both portray worst cases. The Couette model assumes that no slurry will re-enter the interface which is not true. The Emslie et al. Couette correction assumes all of the returning slurry will re-enter the interface, which is also not true. Hence, the experimental results fall between these two extremes and are bounded by them.

The data, however, seem to validate at least one assumption made in the derivation of the Couette model. It was assumed that the inertial term of the N-S equation was negligible mostly because of very low fluid thickness, $\bar{h}$, so that rotation of the pad would not add to the total flow rate. The experiment with stationary wafer and the pad rotating as well as the stationary pad with wafer rotating ($r_{cc} = 0$) showed very little flow, as expected from the model since the model only depends on the product of wafer velocity and the center-to-center distance between the wafer and the pad.

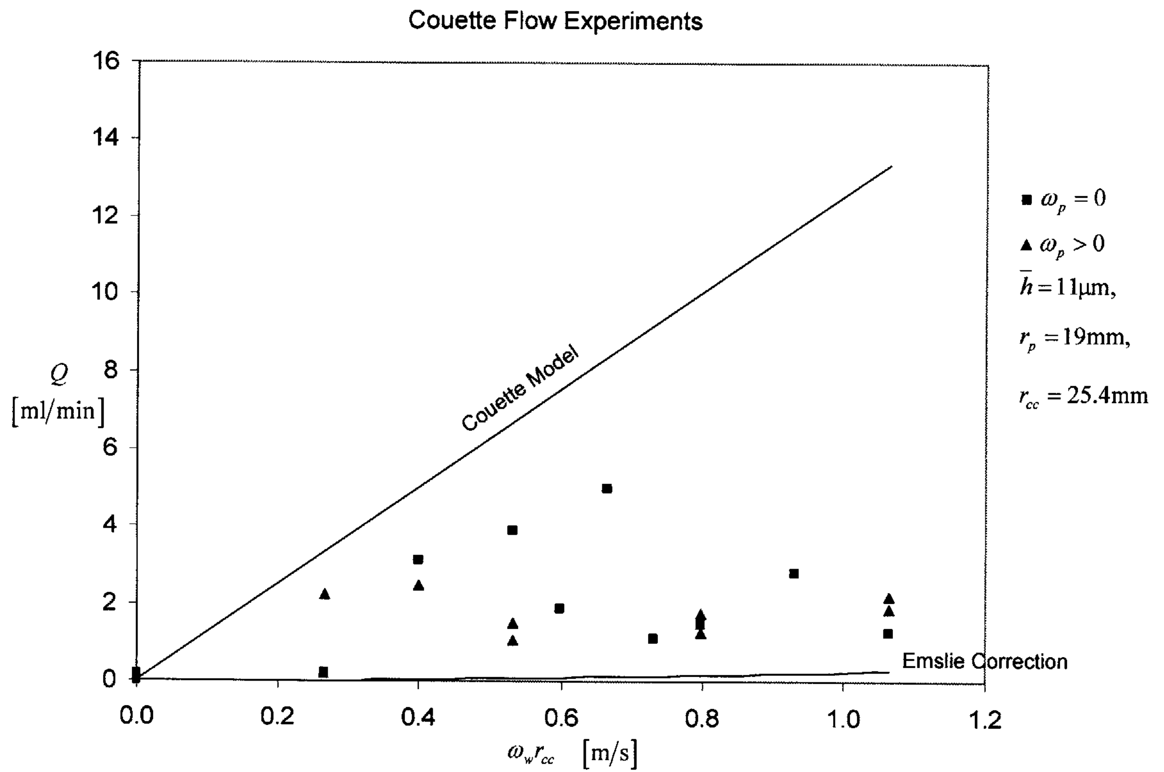


Figure 4.5. Experimental results for flow rates vs. linear velocity. Actual data is located in Table B.2.

The rotation of one of the surfaces does not cause any inertial forces on the fluid that would allow flow rate. So, the decision to neglect the inertial term seems valid. Otherwise, the inertial force due to the rotating surface would drive the flow outward and create non-zero flow rates.

4.6 Discussion

The two equations bounding the flow rate for the face-up polishing tool are:

$$Q_u = r_{cc}\omega_w r_p \bar{h} \tag{4.35}$$

and

$$Q_l = \omega_w r_{cc} r_p \bar{h}\left(1-\frac{1}{\sqrt{1+8\frac{\rho\omega_w \bar{h}^2 \pi}{3\mu}}}\right) \tag{4.38}$$

Equations 4.35 and 4.38, though simple, give some insight into how the slurry flow rate can be approximated in CMP. These equations have been derived specifically for the face-up polisher, but Eq. 4.35 can also be applied to the face-down polisher. Equation 4.38 however cannot be applied to the face-down configuration because the rough pad holds and carries the slurry that leaves the interface and thus violates the assumptions of the viscous flow model by Emslie et al., which assumes a smooth surface.

The Couette model is schematized in Fig 4.6. The flow rate equation for face-down polisher is:

$$Q = r_{cc}\omega_p r_w \bar{h} \tag{4.36}$$

The basic assumption of this model is that the polishing pad has a certain capacity to carry the slurry. The critical dimensions that characterize this capacity are the

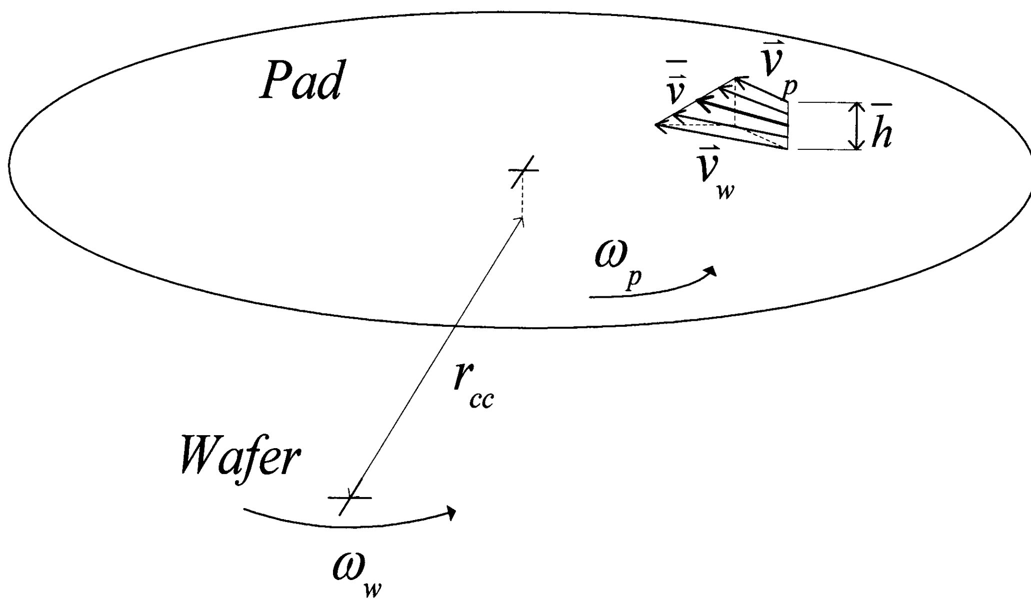


Figure 4.6. Couette model for face-up polishing.

effective gap between the wafer and the pad, $\bar{h}$, and the radius of the pad (wafer for face-down polishers). The product of the two gives a flow area through which slurry flows. The slurry velocity through the flow area is dictated by the linear velocity of the wafer for the face-up and linear velocity of the pad for the face-down polishers.

For the 100mm wafer polisher using Rodel IC 1400 pad without grooves, the relevant parameters are: $\omega_p r_{cc} = 1\,\text{m/s}$, $r_w = 0.05\text{m}$, $\bar{h} = 12\mu\text{m}$. Assuming no slurry reflux, the required flow rate is estimated to be $Q \cong 36\,\text{ml/min}$. In practice, however, one may feed less than this amount and still not adversely affect the CMP process if reflux is significant. For the 300mm wafers, the slurry feed rate is $Q \cong 108\,\text{ml/min}$, assuming the magnitude of relative velocity is the same. It may be noted that in industrial CMP practice, slurry is fed at a rate of $200\,\text{ml/min}$ in 300mm wafer polishing. Thus the Couette flow model, though simplistic, estimates the right order of magnitude of the slurry flow requirement. The difference may arise from several factors: the industry uses grooved pads that contain more slurry and/or allow slurry to leave the pad effectively through radial grooves. The relative velocity and the roughness of the pads may also be different from those in the present experiments.

The basic Couette model can be used in both face-down and face-up polishing, even with grooves so long as the grooves are closed within the pad, e.g., circumferential grooves as opposed to radial grooves. If the grooves are open to the outside of the pad, the Couette model fails to be the limiting case and large flow rates are expected. But if closed grooves are used, to obtain the effective $\bar{h}$, ungrooved version of the same pad should be used in the Poiseuille experiment. Grooves in closed form are beneficial because they help increase the distribution of the slurry while the total flow rate is still controlled by $\bar{h}$.

Due to large slurry reflux, however, the required slurry feed rate may indeed be quite low. Even so, it is desirable to feed fresh slurry frequently since the same slurry used repeatedly may result in lower MRR and/or defects. By feeding larger quantities of slurry than that predicted by the Couette model, the slurry will be replenished and such problems can be eliminated.

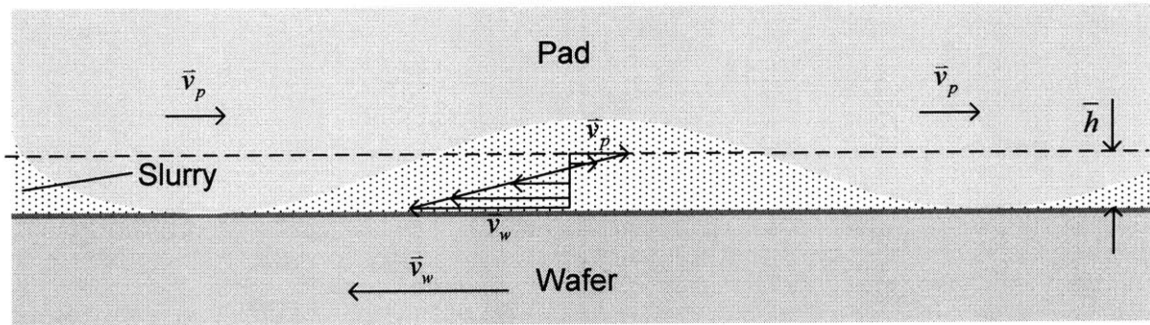
Based on the Couette flow model, increasing the size of the wafer is likely to result in higher wafer-scale non-uniformity because the slurry flow rate depends on the wafer diameter while the total volume of Cu removed depends on the area (diameter squared). For example, if the 100mm wafer is replaced with the 300mm wafer, assuming linear velocity is kept the same, the Couette model predicts the slurry flow rate would be tripled while the area to be polished is nine times. So, three times the slurry flow rate needs to remove nine times the volume of Cu. This would probably increase wafer scale non-uniformity. Since the pad is capable of drawing only three times that of the 100mm case, feeding more than that may not result in better uniformity. The pad is capable of holding only a certain volume of slurry and overflowing it does not add to increased MRR or uniformity. If more than the capacity is fed, only the volume dictated by this capacity would actually be used and the rest would be spun off. This scaling disparity between the slurry feed rate and the MRR is probably responsible for the wafer-scale non-uniformity. The larger the wafer, the longer the distance the slurry needs travel to be effective and keep non-uniformity low. By the time the slurry travels from the edge to the center of the wafer, the abrasive particles may be worn out and cause spatial variation in MRR.

In the face-up polisher, by contrast, the perforated pad allows efficient slurry replenishment at the wafer/pad interface. The slurry has to travel only from one perforation to the next where it gets mixed with the incoming slurry, thus slurry freshness and effectiveness are maximized. For scaling purposes, then, the face-up polisher is much better since it always has the same wafer-scale non-uniformity determined by the distance between perforations. Thus, polishing non-uniformity will be independent of the size of the wafer for the face-up polisher so long as the velocities are well controlled.

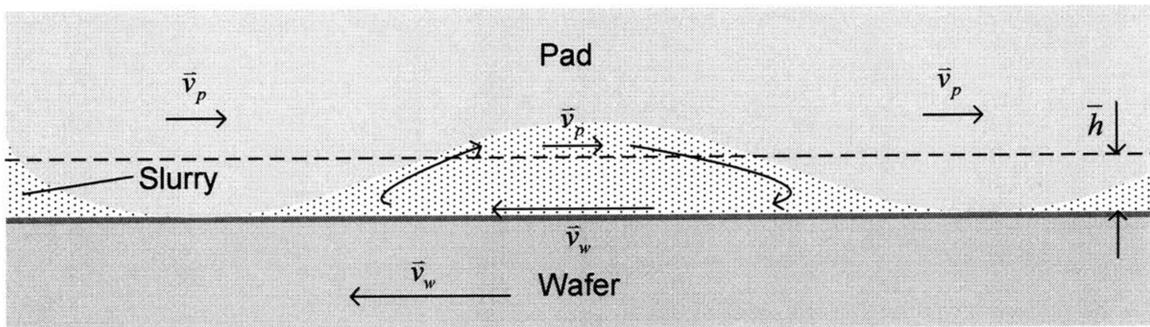
Another consequence of the Couette model is that for material removal rate, the magnitude of the relative velocity vector, the difference between the wafer velocity and the pad velocity vectors, is important while slurry flow pattern is determined by the sum of the wafer and pad velocities. This may be another cause of non-uniformity in polishing since optimization for MRR may adversely affect the slurry flow pattern and cause non-uniformity. For example, if both angular velocities are equal in magnitude and are in the same direction, the relative velocity vector field, too, is uniform in magnitude

and direction throughout the interface, but the slurry flow pattern is rather complex. However, if the angular velocities are in opposite directions with the same magnitude, the slurry flow pattern is uniform and unidirectional while the relative velocity vector field is complex. Hence optimization based solely on kinematics may not necessarily lead to best polishing performance or uniformity. Uniform relative velocity vector field is desired for uniform MRR, but even if the relative velocities are uniform the slurry velocity field will not be. Non-uniform slurry velocity field will affect the distribution of fresh slurry which in turn may affect local MRR. Thus, in the end MRR may not be uniform even if it is expected based on the relative velocity field. An investigation of both relative velocity and fluid velocity fields must be carried out to elucidate the contribution of each toward MRR.

The Couette flow model does not account for all the phenomena that occur at the polishing interface, which may be the reason for the discrepancy between the model and the experimental results. Perhaps the assumption that the slurry is transported by two smooth, moving boundaries $\bar{h}$ apart is not quite correct. In reality, the pad is a rough surface that actually contacts the wafer at many points, and the fluid does not have the freedom to flow as driven by the boundaries because of the many obstacles (the asperities) along its path. Figure 4.7 shows how the slurry behavior may differ from the idealized Couette model. The Couette model is more valid on the face-down polishers since the pad is probably more influential over the slurry and at the same time the pad speed is the dominant term in the Couette model. On the face-up polishers, however, while the wafer speed is the dominant term in the Couette model, the details of pad topography may be more influential in slurry flow. The profile and spacing of the asperities may substantially affect the flow rate. Because the asperities are randomly distributed, the streamlines of flow are disturbed, as shown in Fig. 4.7. This may be the reason for the discrepancy between the Couette flow model and the experimental data.



(a)



(b)

Figure 4.7. Slurry behavior idealization: (a) Couette Model, (b) More realistic view

Chapter 5

Effect of Slurry Flow Rate on Material Removal Rate

5.1 Introduction

In the face-down polisher the operator has complete control over the slurry feed rate. The operator, however, does not know the optimum slurry feed rate without an analytical model or extensive empirical studies. At the very least, the Couette model is able to give an upper bound on the required slurry feed rate. It is also important to study its effects on material removal. It is not clear how MRR depends on slurry feed rate but it seems intuitive to expect that as the slurry feed rate is reduced, the MRR should also be reduced. It is desirable to find out what the MRR is as a function of slurry feed rate.

The face-up polisher is fundamentally different from the face-down polisher because in the face-up configuration the operator does not have control over the slurry flow rate. As long as the slurry reservoir is full, the process will draw as much slurry as the pad allows. The only decision the operator makes is whether or not the slurry reservoir should be full. Therefore, it is important to find the face-up MRR at the natural slurry flow rate and how it compares to the face-down configuration. Such a study will be able to show whether or not face-up is able to obtain comparable MRR at least at the local scale.

The Preston constant will be used to compare local MRR on both face-up and face-down machines. The Preston constant is an ideal parameter since it is independent of relative velocity, pressure, and geometry, so it reflects the overall polishing effectiveness. The local Preston constant is an indicator of slurry distribution in the face-up configuration. As long as the Preston constant of the face-up configuration is similar to the Preston constant of the face-down configuration, the slurry is adequately distributed in the interface by the given perforation geometry. A starved polishing interface is likely to have a low Preston constant (MRR), so the Preston constant can indicate whether the slurry is adequately supplied or starved in the polishing interface.

If the face-up Preston constant is lower in magnitude than the one obtained in the face-down configuration, the pad perforation system does not allow the slurry to be adequately drawn into the polishing interface. Thus, the perforation geometry needs to be re-designed by adding more holes and/or closed grooves. However, if the Preston constant is larger than in the face-down configuration, then the slurry supply system is sufficient and one can probably conclude that the face-down machines needs a re-design of the slurry distribution system.

5.2 Material Removal Rate

It is important from a manufacturing process point of view to determine the Preston constant. In the conventional CMP process, the average Preston constant can be found by polishing a blanket Cu wafer half way or so and knowing the initial Cu height and the time taken. Thus, in regular CMP process if the angular speeds are the same ($\omega_w = \omega_p$) the Preston constant is given by:

$$k_p = \frac{\Delta h}{p|\vec{v}_R|\Delta t} \tag{5.1}$$

In the face-up polisher the Preston constant cannot be so easily determined because the polishing is non-uniform even if the Preston constant is uniform across the contact. Thus it is necessary to develop a methodology to measure Preston constant for some substantial area as spatial average. Of course. the direct way to measure the thickness of Cu removed is to measure it by electrical resistivity at some specific radial point.

A convenient and simple way to determine the Preston constant is by measuring the total volume of material removed in a given polishing time. This can be easily done by measuring the change in mass of the wafer before and after polishing, and knowing the density of the material. Once the total volume removed is measured it can be used to solve for Preston constant as follows:

$$\frac{dh}{dt} = k_p p |\vec{v}_R| \qquad (1.1)$$

The change in Cu height is obtained by simple integration in time:

$$\Delta h = k_p p \int_0^{t_f} |\vec{v}_R| dt \qquad (5.2)$$

Figure 5.1 shows the geometry of the system along with the variables used in the analysis. By changing variables to eliminate time we obtain:

$$\Delta h' = k_p p \int_{-\theta_p}^{\theta_p} |\vec{v}_R| \frac{d\theta}{\omega_w} \qquad (5.3)$$

Total height removed is the integral of Eq. 5.3 over total number of revolutions:

$$\Delta h = \frac{\omega_w}{2\pi} t_{\exp} k_p p \int_{-\theta_p}^{\theta_p} |\vec{v}_R| \frac{d\theta}{\omega_w} = \frac{1}{2\pi} t_{\exp} k_p p \int_{-\theta_p}^{\theta_p} |\vec{v}_R| d\theta \qquad (5.4)$$

Total volume of material removed is the integration of Eq. 5.4 over affected wafer area:

$$\Delta V = \int_0^{2\pi} \int_{r_1}^{r_2} \Delta h(r) r dr d\theta = \frac{1}{2\pi} t_{\exp} k_p p \int_0^{2\pi} \int_{r_1}^{r_2} \int_{-\theta_p}^{\theta_p} |\vec{v}_R| d\theta r dr d\theta = t_{\exp} k_p p \int_{r_1}^{r_2} \int_{-\theta_p}^{\theta_p} |\vec{v}_R| d\theta r dr \qquad (5.5)$$

If the entire pad area is within the wafer, the integrals in Eq. 5.5 are nothing more than integrals over the pad area hence if $|\vec{v}_R|$ is expressed in the pad coordinate system, it reduces to the following.

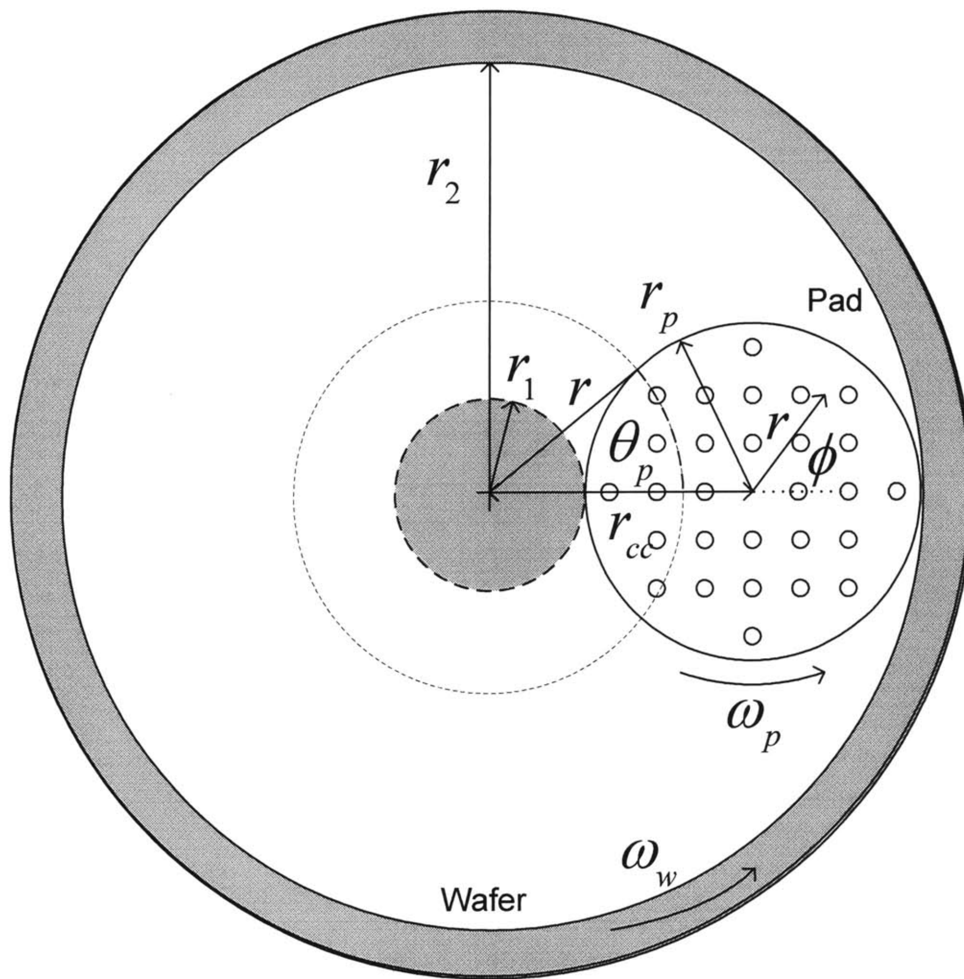


Figure 5.1 Face-up polisher geometry

$$\Delta V = t_{\exp} k_p p \int_0^{r_p} \int_{0_p}^{2\pi} \left| \bar{v}_R(\phi, r) \right| d\phi r dr \tag{5.6}$$

Equation 5.6 is a general equation valid for any speeds, but only if the entire pad is within the wafer. If the goal is to determine the Preston constant, then the simplest way is to run an experiment with $\omega_w = \omega_p$. Then $\left|\bar{v}_R\right| = \omega\ r_{cc}$ which is constant over the contact interface so that change in volume is given by:

$$\Delta V = t_{\exp} k_p p \int_0^{r_p} \int_{0_p}^{2\pi} \omega \cdot r_{cc} d\phi r dr = t_{\exp} k_p p \omega r_{cc} \int_0^{r_p} \int_{0_p}^{2\pi} d\phi r dr \tag{5.7}$$

Where the right-hand side integral is just the area of the pad:

$$\Delta V = t_{\exp} k_p p \omega r_{cc} A_{pad} \tag{5.8}$$

Thus:

$$k_p = \frac{\Delta m}{p \omega r_{cc} A_{pad} t_{\exp} \rho_{Cu}} \tag{5.9}$$

where k_p is the Preston constant, Δm the measured change in mass, p the nominal pressure, ω the angular rate of both the pad and the wafer, A_{pad} the area of the pad (or the area of the pad within the wafer), $t_{\exp}$ the total polishing time, and ρ_{Cu} is the density of Cu.

5.3 Results

Experiments were performed on the face-up polisher to determine the Preston constant for two different pads. The first pad was Rodel IC 1400 without grooves. The

second was similar Rodel IC 1400 pad with grooves. Both pads were cut to a diameter of 38 mm so that the entire pad would lay within the wafer area. This condition is not a necessity for actual CMP but a good way to simplify Eq. 5.6 in order to determine the Preston constant. Also, both pads were perforated in the same manner. The perforations were done by a punch and their diameter was 2.4 mm. The perforations were spaced in a square matrix 6.3 mm apart. Each pad contained 21 perforations.

Both experiments were performed on Cu blanket wafers 100 mm in diameter. The center-to-center distance between the wafer and the pad was 25.4 mm and was held constant throughout the experiment. From the above analysis it is clear that the angular velocities of the pad and the wafer should be equal, for simplicity. For both experiments the angular velocities were set at 250 rpm. The nominal polishing pressure was 14 kPa. Both experiments were performed using Cabot ICUE 5001 slurry with 10% by volume addition of 30% Hydrogen Peroxide in accordance with manufacturer specifications. The change in mass of the wafer was obtained by weighing before and after polishing. The results of the two experiments are listed in Table 5.1.

Table 5.1 Slurry flow rate and MRR for two pads in face-up polishing.

Pad	t_{exp} [min]	*Initial mass* [g]	*Final mass* [g]	Δm [mg]	ΔV [mm^3]	Q [ml/min]	k_p [1/Pa]
Rodel IC1400 without groves	12	11.2601	11.2500	10.1	1.13	0.83	$1.5\text{x}10^{-13}$
Rodel IC1400 with grooves	15	9.2657	9.2503	15.4	1.73	5.50	$1.8\text{x}10^{-13}$

p_{nom} =14 kPa, $\omega_w = \omega_p = 26.18$ rad/s, $r_{cc} = 25.4$ mm, $r_p = 19$ mm, $\rho_{Cu} = 8900$ kg/m^3

The experiments were performed to investigate the effect of slurry flow rate on material removal rate, too. The Rodel IC 1400 pad without grooves allowed average

slurry flow rate of about 0.8 ml/min and obtained the average Preston constant of $1.5\times10^{-13}\,1/\text{Pa}$. The same pad but with grooves allowed slurry to flow at about 5.5 ml/min and obtained the average Preston constant of $1.8\times10^{-13}\,1/\text{Pa}$. The grooves were roughly linear as the pad was cut from the outer region of a much larger pad with concentric grooves. Both pads had the same topography except for the grooves. It seems the grooves themselves allow higher slurry flow rates because they act like deep and wide channels compared with the effective gap, $\bar{h}$. As a result, slurry was drained from the interface without actually allowing it to be effectively used. The experiment seems to show that by providing grooves that allow slurry to flow outside from the interface, the Preston constant is only marginally increased and it is probably due to a slightly more uniform slurry distribution. The slurry consumption, however, is greatly increased. Hence, to minimize slurry consumption and keep the MRR constant, pads without grooves should be used in face-up polishing. Alternatively, pads with concentric and closed grooves may be used since they do not allow slurry to be drained from the interface. Such a pad would also provide an effective way of distributing the slurry more uniformly in addition to the perforations.

These two experiments are meant to show the effect of slurry flow rate on MRR in face-up polishing. The quoted slurry flow rates must be put in context, since the pad was smaller than necessary and the rotational speeds were also low. Consequently, the overall material removal rate (wafer-scale) was quite small even though the Preston constant was about what it is in regular CMP. In regular CMP, however, the entire wafer is being polished at once so that total polishing time is around 5 min. In the present experiment, the pad only covered a small portion of the wafer (roughly $1/7\text{th}$) so that overall material removal rate was low. In actual face-up polishing the pad would be almost twice the diameter and the speeds should be roughly doubled so the slurry consumption should be quadrupled (based on Couette flow model). Hence, to compare slurry consumption in the face-up polisher with that in regular CMP, it is necessary to consider the total polishing time required to finish the wafer and not just the Preston constant.

5.4 Face-down Polishing

The following experiment was performed to determine the effect of slurry flow rate on MRR in conventional rotary (face-down) CMP. The effect of slurry feed rate on MRR is not clearly understood. It is imperative to understand this phenomena from the slurry consumption point of view. The following is an empirical attempt to show the effect of slurry flow rate on the material removal rate.

The experiment attempted to measure the average Preston constant of Cu blanket wafers for various slurry flow rates. It was performed on a 100mm face-down CMP machine with applied pressure of 14 kPa and with equal angular rates for both the wafer and the pad. The pad was Rodel IC1400 with grooves (concentric K-grooves) that was thoroughly cleaned before each experiment with water and high pressure air to rid of any residual abrasive or wear particles.

Table 5.2 Face-down polishing conditions

Wafer Diameter [mm]	100
r_{cc} [mm]	89
ω_w [rad/s]	7.8
ω_p [rad/s]	7.8
Pressure [kPa]	14
Pad	Rodel IC1400 with grooves
Slurry	Cabot ICUE 5001 with H_2O_2

Cabot ICUE 5001 was the slurry with the addition of 10% by volume of 30% Hydrogen Peroxide, as specified by the manufacturer. The slurry was supplied by a low flow rate peristaltic pump to obtain such a low-range flow rates. The polishing conditions are given in Table 5.2.

Each experiment began with a new 100mm blanket wafer with a 1μm Cu coating by PVD. The overall spatial Preston constant can be determined from Eq 5.1 by determining average Cu height reduction from the total change in mass.

$$k_p = \frac{\Delta h_{avg}}{p|\vec{v}_R|\Delta t} \tag{5.1}$$

where Δt is the total polishing time ($t_{\exp}$) and Δh_{avg} is the average Cu height reduction which can be calculated using the total change in mass of the wafer by:

$$\Delta h_{avg} = \frac{\Delta V}{A_{wafer}} = \frac{\Delta m}{\rho_{Cu} A_{wafer}} \tag{5.10}$$

Where Δm is the measurable change in mass as a result of polishing, A_{wafer} is the wafer area, and ρ_{Cu} is the density of Cu. The basic assumption behind this analysis is that only Cu is being removed from the wafer by polishing. It is a reasonable assumption since the total polishing time was not long enough to completely remove Cu from the wafer. The entire wafer area is used to average the material removal rate to obtain the average Preston constant throughout the wafer area. Substituting in Eq 5.10 into Eq 5.1:

$$k_p = \frac{\Delta m}{p|\vec{v}_R| t_{\exp} A_{wafer} \rho_{Cu}} \tag{5.11}$$

The experiments were performed with equal wafer and pad's angular velocities so Eq. 5.11 becomes:

$$k_p = \frac{\Delta m}{p \omega r_{cc} t_{\exp} A_{wafer} \rho_{Cu}} \tag{5.12}$$

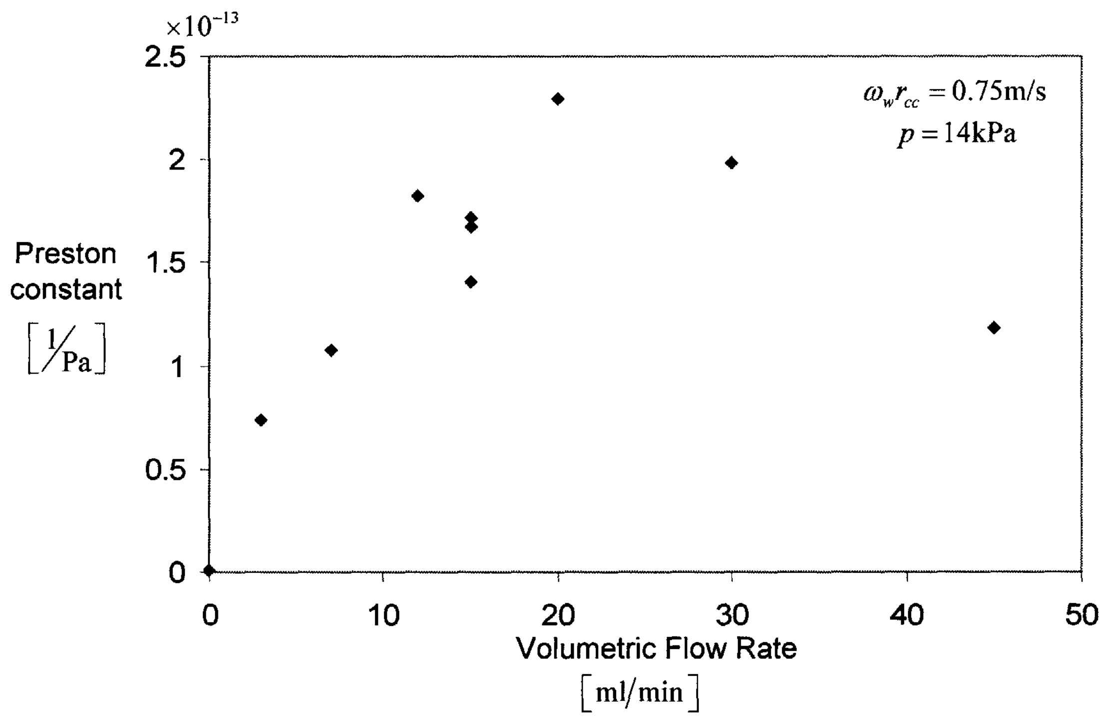


Figure 5.2 The effect of slurry flow rate on the Preston constant

Equation 5.12 was used to obtain the average Preston constant by measuring the total change in mass of the wafer. This analysis is only valid for an average Preston constant throughout the wafer area and cannot be used to obtain a point-wise Preston constant. The experiment consisted of polishing a new wafer for about 4 min at the conditions listed in Table 5.2 at various slurry flow rates. The wafer was weighed before and after the experiment to obtain the change in mass. The overall Preston constant was determined from equation 5.12 and the experiment was repeated with another new blanket wafer for another slurry flow rate condition.

The results of these experiments are summarized in Figure 5.2 and the data are listed in Table A3 in Appendix. The Preston constant is a useful way to describe MRR because it factors out speed and pressure, but depends on the slurry feed rate in a complicated way. First, MRR increases as the slurry feed rate is increased up to 20ml/min. This shows that a starved polishing interface will exhibit lower MRR and as the starvation is reduced, MRR is increased. At slurry feed rate over 20ml/min, MRR seems to drop as feed rate is increased. This is counterintuitive since it is expected that as the feed rate is increased past the pad capacity, MRR should level off. These experiments show that MRR actually drops as slurry feed rate is increased beyond the optimum feed rate. Based on the Couette flow model, this optimum which is the pad capacity to accommodate slurry flow rate is around 23ml/min from Eq. 4.35, where $\omega_p r_{cc} = 0.7$ m/s, $r_p = 0.05$ m, and $\bar{h} = 11$ μm. It appears that once the slurry feed rate is higher than this pad capacity then MRR drops off. Also, the point of zero flow rate was obtained by polishing with 50 ml/min flow rate of just water. That way the effect of residual abrasives or contaminants can be accounted for.

It was noticed that when the slurry feed rate was below 30ml/min then there was noticeable stick-slip motion of the wafer carrier during polishing. Hence, one can conclude that the friction coefficient was higher than the friction coefficient during polishing with higher slurry feed rates. Since the friction coefficient provides the shear stress for material removal, higher shear stress would create higher MRR as long as there are enough abrasives present. All the experiments below 23ml/min experienced some sort of starvation and hence high friction since the interface is not well lubricated. As more and more slurry is fed, MRR increases because there is a much larger volume of

abrasives flow through the interface even if the shear force decreases. At 23ml/min, shear force is low enough that adding more abrasives will not increase MRR; instead it would lubricate the system and hence reduce the shear force even more. As more slurry is fed, the system is better and better lubricated with lower and lower shear force while the total number of abrasives remains the same inside the interface because the pad can only carry so many. Thus, an optimum MRR condition is obtained at around 23ml/min. It must be noted full fledged hydrodynamic lubrication is not occurring in CMP. However, localized lubrication may be occurring consistent with mixed lubrication regime since there seems to be a decrease of friction if slurry feed rate is larger than 23 ml/min.

5.5 Summary

Based on experiments, both the face-up and the face-down polishers have the same Preston constant. It may be assumed therefore that the current perforation geometry is optimal enough to be used in the face-up polisher. But there is room for improvement, for the grooved pad showed a slightly higher MRR than the un-grooved pad. The slurry flow rate with the grooved pad, however, was much greater since the grooves drained the slurry away from the interface faster. Thus, by cutting closed grooves on the pad, in addition to the perforations, the slurry can be distributed more evenly and thus increased MRR and polishing uniformity are expected.

While the Preston constants are about equal and the local MRR is the same for both configurations, the face-up polisher takes longer to polish the entire wafer. This is because the pad covers only a fraction of the wafer in the face-up configuration, whereas in the face-down configuration the pad covers the entire wafer at once. This is one of the limitations of the face-up design, but can be easily overcome by using higher pad speeds, also required for achieving uniformity in polishing.

Indeed, this is an unexpected benefit, for the slurry flow rate is determined only by the wafer angular speed (by the Couette model) while MRR is a function of the

angular speeds of both the wafer and the pad. So, by increasing the pad speed as required to get high uniformity, overall MRR is also increased and slurry consumption is held low. On the face-down polisher, however, different angular speeds for the wafer and the pad will result in non-uniform polishing. So, the face-up polisher has an added advantage in that the increased pad speed results in higher uniformity and MRR, but not in slurry flow rate. Nevertheless, it is also necessary that fresh slurry enter the contact interface continuously. Otherwise, MRR in fact will decrease.

Chapter 6

Conclusion

6.1 Concluding Remarks

In this thesis, an innovative face-up polisher configuration is proposed to control over-polishing and hence dielectric erosion and Cu dishing in CMP. In the face-up configuration, over-polishing is controlled kinematically. A smaller pad initially removes Cu at the center of the wafer and then moves radially out, progressively removing Cu from center to edge. As a result, it will not over-polish the completed central regions. However, this method works only if the rotational speed of the pad is greater than that of the wafer, i.e. $\omega_p \geq \omega_w$. The kinematics and the frictional forces of the face-up CMP tool are described in Chapter 2.

A methodology of characterizing the effective gap between the wafer and the pad is developed by Poiseuille flow analysis and experiments in Chapter 3. Asperities of the pad contact the wafer and the fluid fills the space between them. The effective gap, $\bar{h}$, is of the same order of magnitude as the roughness; it is closer to the RMS roughness of the pad than to the average roughness. Thus, the Poiseuille method for characterizing the pad's capacity for fluid flow seems reasonable.

In Chapter 4, a model based on Couette flow is developed to study the slurry flow behavior in both face-up and face-down CMP tools. The Couette model is simple in that it does not account for slurry reflux. Thus, the Couette model overestimates the flow rate. A correction to the Couette model is also developed to estimate the remaining slurry film thickness in the face-up configuration, which would re-enter the pad/wafer interface and reduce slurry consumption. The Emslie et al. correction shows that over 95% of the slurry that leaves the interface will stay on the wafer and re-enter, but it grossly underestimates the flow rate possibly because not all the returning slurry re-enters the interface. Therefore, the actual slurry flow rate is bounded from above by the Couette model and from below by the Emslie et al. correction to the Couette model.

The Couette model provides an analytical solution for the average local fluid velocity at the polishing interface. The average fluid velocity vector is proportional to the vector sum of the local pad velocity and the local wafer velocity, not to the difference as in MRR. Average fluid velocity vectors for various polishing speeds are shown in the Fig. A.1 in Appendix A.

Polishing experiments show that the Preston constants of the face-up and face-down configurations are about equal. This suggests that in the face-up scheme, the slurry distribution generated by the current perforation geometry is adequate, or at least equivalent to that of the face-down configuration. The grooved pad yields a slightly higher Preston constant than the un-grooved pad, but also causes much larger slurry flow rates. Since the grooves were not closed, they allowed the slurry to be drained from the interface at high rates. This result suggests, however, that the Preston constant for the un-grooved pad can be increased substantially by providing a better slurry distribution network, in addition to the perforations. Possibly, the pad should have closed grooves that aid slurry distribution at the interface but not allow the slurry to be drained from it.

6.2 Overview

Several objectives have been accomplished in this work. First, an innovative method of characterizing the pad roughness for fluid flow is developed. The procedure outlined in Chapter 3 may be followed to determine the effective gap between the pad and the wafer. Basically, a fixture has to be made so that a small circular pad with one central hole can be pressure loaded and simultaneously allow pressurized fluid to flow from the central hole to the outside. To accomplish this task, a design similar to the one shown in Figure 3.2 may be employed. The pad should not have grooves since grooves by far dominate the flow and skewed measurements will be obtained. If for actual polishing purposes a grooved pad is used, to characterize that pad an un-grooved version of it must be used. It should have the same surface roughness excluding the grooves. In face-down polishing, the circular grooves do not contribute to slurry flow since they are closed and act as reservoirs—once filled, they remain filled. The important parameter is $\bar{h}$, the effective gap between the wafer and the pad due to the roughness of the pad.

Experiments similar to those described in Chapter 3 should be done and Eq. 3.27 used to calculate the effective $\bar{h}$.

The next contribution of this thesis is the analysis of slurry flow during polishing. The model discussed in detail in Chapter 4 may be used to estimate the slurry consumption in CMP or in similar processes. From the effective gap between the wafer and the pad, $\bar{h}$, the required slurry feed rate can be estimated for either the face-down or the face-up polishers. The model is not exact in predicting the flow rates obtained experimentally, but is able to bound the actual flow rates. The upper bound is the simpler Couette model, which overestimates the flow rate because it does not take into account slurry reflux. The lower bound is the Emslie et al. correction to the Couette model, which estimates the amount of slurry that remains on the wafer and assumes that it re-enters the interface. The Emslie et al. correction basically re-establishes the effective gap at the interface as the difference between $\bar{h}$ and the thickness of returning slurry. Only about 5% of the gap will be able to accommodate new slurry and hence allow some flow rate. The Emslie et al. correction, however, grossly underestimates the flow rate because substantial portion of the returning slurry will not be able to re-enter the interface. That difference will allow fresh slurry to fill the remainder of the gap and hence the actual flow rates will be higher than the correction due to Emslie et al. to Couette flow.

6.3 Future work

It is apparent that the Couette flow model overestimates the slurry flow rate. Future work may include research to refine this model and incorporate corrections for the discrepancies. For instance, some secondary phenomena might be taking place so that the $\bar{h}$ value obtained from Poiseuille experiments does not fully describe an effective gap between the rotating wafer and pad. Future work may include tying together $\bar{h}$ from Poiseuille flow to Couette flow by incorporating pad waviness into the Couette model.

Future work may also involve continuing improvements on the face-up polisher. The current machine has numerous problems and limitations. Stiffness of the pad carrier

needs to be increased to avoid chatter due to stick-slip motion. The wafer carrier needs to be modified so that the height differential between the wafer and the retaining ring is minimized. Also, endpoint detection sensors should be incorporated into the design and integrated with automated kinematic controls.

Once a fully operational and fully integrated with sensors and kinematic controls, the face-up polishing machine should be used to obtain numerous data on reducing wafer scale non-uniformity. Several algorithms may also be proposed to uniformly polish the wafer depending on incoming wafer non-uniformities.

Appendix A

Fluid Velocity Based on the Couette Model

In this work, the Couette flow is used for modeling the slurry transport in CMP. Equations 4.31 and 4.32 derived by that model can be used to plot the integrated mean fluid velocity vector as a function of position in the wafer pad interface for various speed ratios. Equation 4.31 and 4.32 respectively show the mean radial component and the mean tangential component of the fluid velocity. To obtain effective local fluid velocity both radial and tangential components were averaged in z. Equation 4.31 and 4.32 express that average fluid velocity.

$$\bar{v}_r = \frac{1}{2}\omega_w r_{cc} \sin\phi \tag{4.31}$$

$$\bar{v}_\phi = \frac{1}{2}\left[\omega_w r_{cc} \cos\phi + \left(\omega_w + \omega_p\right) r\right] \tag{4.32}$$

where ω_w is the wafer angular speed, ω_p the pad angular speed, r_{cc} the distance between center of the pad and the center of the wafer, r and ϕ are the polar coordinates centered on the pad, $\bar{v}_r$ is the mean (z-directional) local radial fluid velocity, and $\bar{v}_\phi$ the mean (z-directional) local tangential fluid velocity. Figure A.1 shows the mean fluid vector field for various angular speed combinations where each vector comprises radial and tangential components from Eqs. 4.31 and 4.32, respectively. The magnitude of the vectors is automatically scaled for each of the setting so they can easily fit into the graph. The magnitudes show, relatively, the speed of the flow within each figure. However, they should not be compared across figures since not only the speed ratio is important but also the absolute velocity. Each figure may have a different scaling normalization so the magnitudes should not be compared between the figures.

Figure A.1 shows the flow pattern of the fluid. For instance, if the angular speeds are set the same and the midpoint of the pad and wafer centers lies within the pad, there is

a point of zero mean fluid velocity. On the average, the net flow is zero at that point, but in reality the top surface drags the top portion of the fluid in one direction and the bottom drags the bottom portion of the fluid with equal velocity in the opposite direction.

Another interesting point is that when both angular speeds are set equal but in opposite directions the mean fluid velocity field is uniform. The net flow will be equal in magnitude and direction at every point within the pad. All of the Figs. A.1 are mere visualizations of Eq. 4.31 and 4.32 and just show how on average (in z-direction) the fluid is flowing at the wafer/pad interface. Such analysis of the fluid flow could be used in the future to possibly explain polishing non-uniformity.

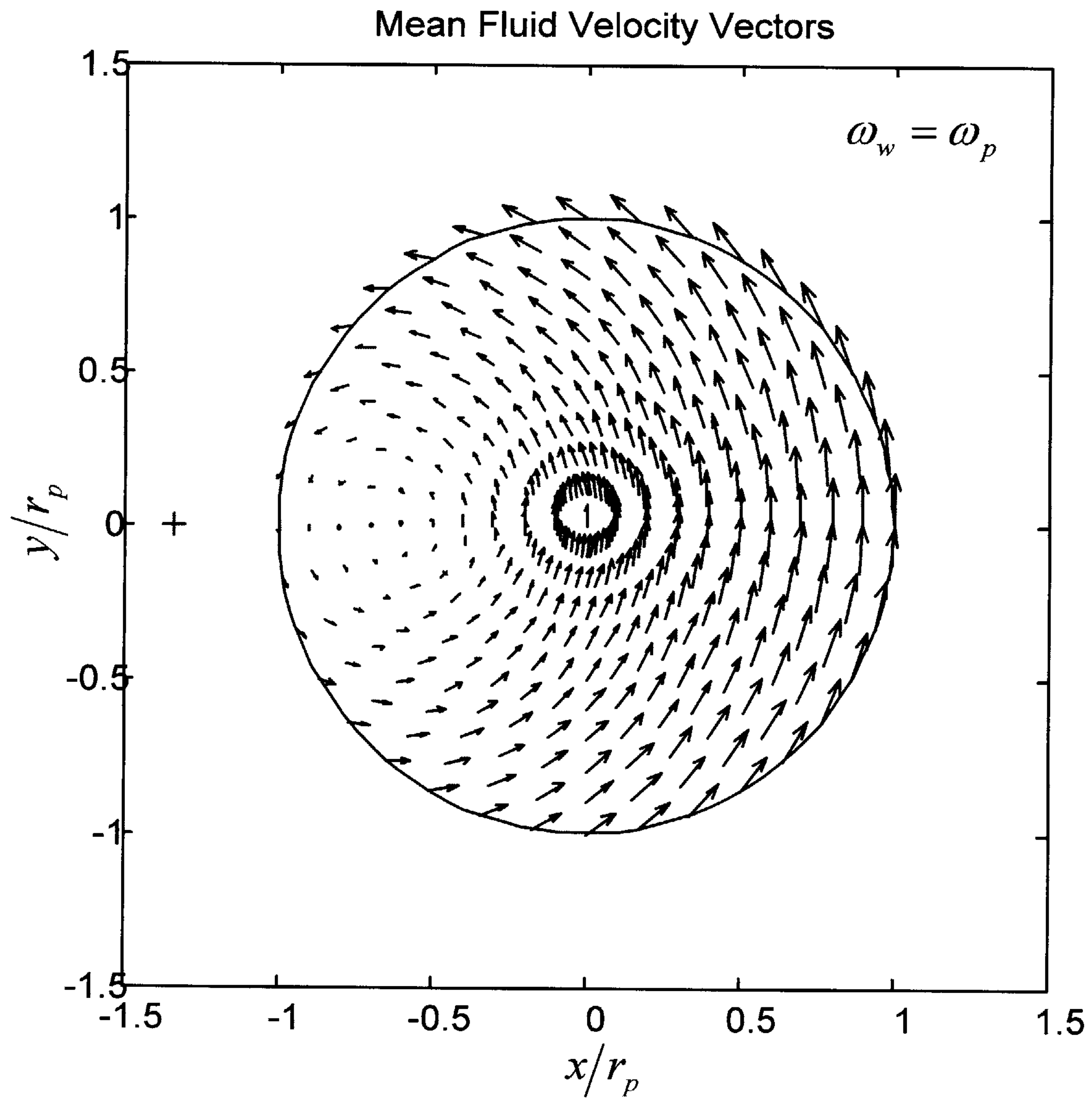


Figure A.1a. Fluid velocity vector field for a given condition.

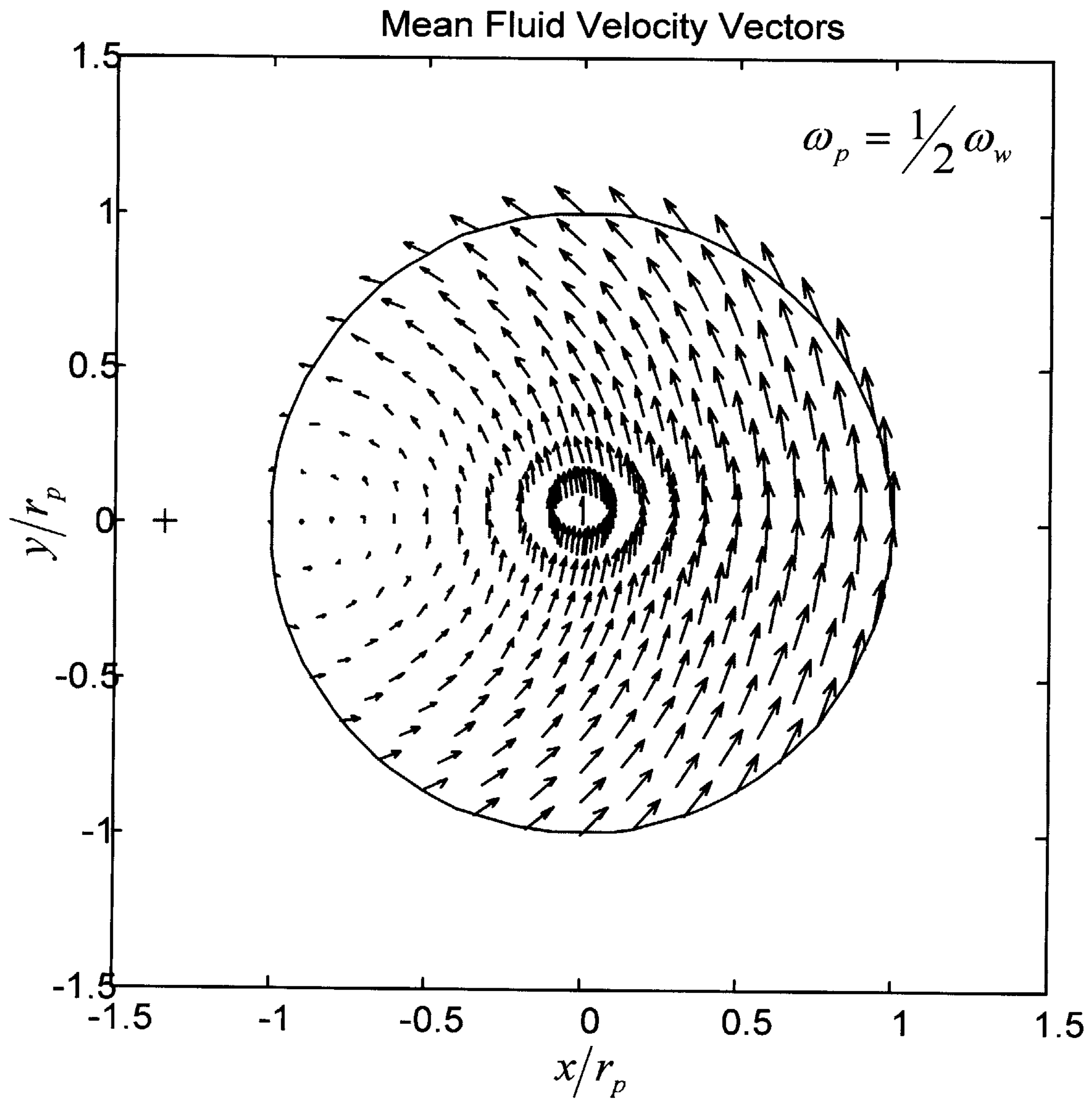


Figure A.1b. The average velocity vector field for a $\omega_p = \frac{1}{2}\omega_w$.

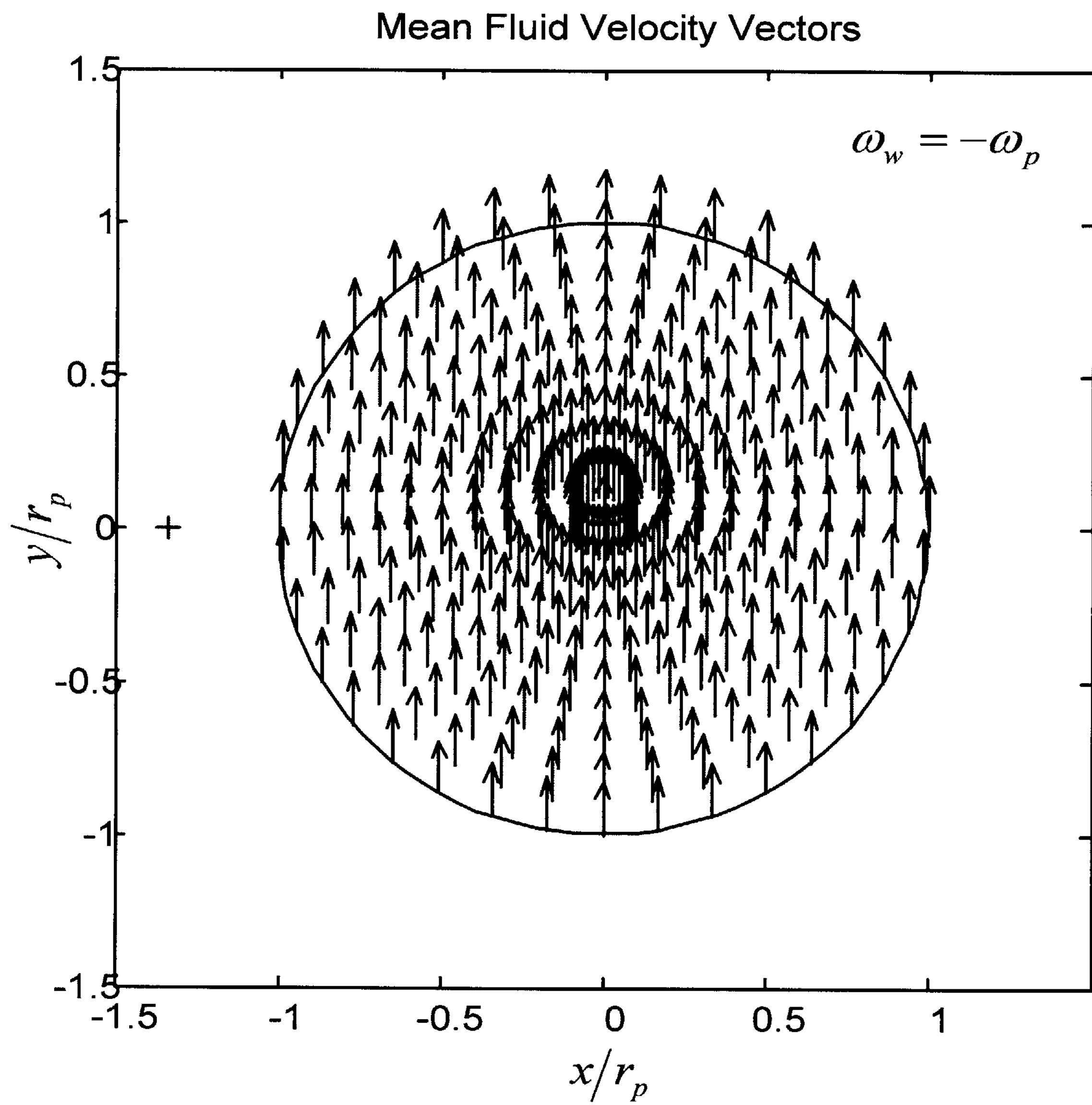


Figure A.1c. The average velocity vector field for $\omega_p = -\omega_w$.

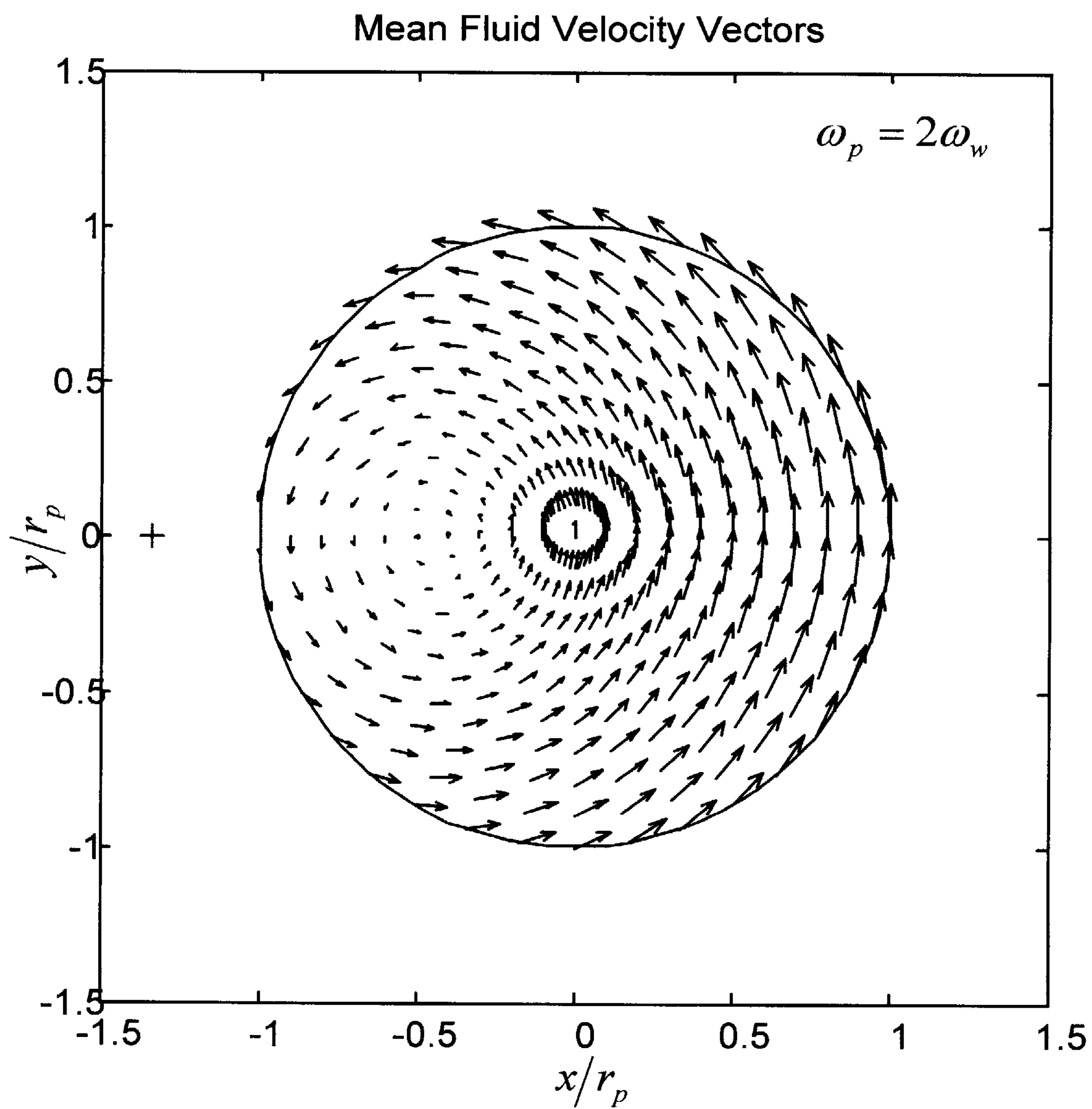


Figure A.1d. The average velocity vector field for $\omega_p = 2\omega_w$.

Appendix B

Experimental Data

In Chapter 3, the Poiseuille flow is used to characterize the pad for fluid flow. The derived model, though simple, is used to estimate the effective gap between the wafer and the pad that facilitates fluid flow. An experiment was performed to estimate the effective gap between the two surfaces. Table B.1 lists the data.

In Chapter 4, the Couette flow model is used to characterize the behavior of slurry in CMP. An experiment was performed for various operating speeds to verify some of the assumptions derived in the model and also to validate its applicability. Table B.2 lists the data from the experiment where the change in height of fluid in the cup was converted to the flow rate for given operating speeds.

In Chapter 5, material removal rate was mapped as a function of slurry feed rate for the face-down polishers. MRR was represented by the Preston constant and was obtained for Cu blanket wafer polishing. The average Preston constant was obtained from the change in mass of the wafer after a known time in polishing at known conditions. Table B.3 shows the data obtained from these experiments.

Figure C.1, shows the prototype face-up CMP machine. One can see the wafer carrier and the pad mounted on an arm that supplies the load and it is able to move sideways on a linear stage. Numerous accessories are required for the machine to work. For instance, the vacuum pump below evacuates the vacuum chamber in the wafer carrier to hold the wafer fixed during polishing. Two peristaltic pumps were used to pump slurry into the slurry cup and water for cleaning after polishing. Various controllers are also required for the operation of the wafer carrier motor, the pad motor, and the linear stage. In the future a data acquisition system can be used along with the endpoint detection to control wafer level non-uniformity during polishing.

Table B.1. Data of the Poiseuille Flow Experiments

Fluid Pressure [kPa]	Lift Force [N]	Applied Load [N]	Effective Pad Pressure [kPa]	Change in Fluid Height [mm]	Δt [min]	Q [ml/min]	$\bar{h}$ [μm]
17	14	15	1.4	2.5	20	0.3	10.4
28	24	31	5.9	2.5	20	0.3	8.9
24	21	23	1.4	5.1	25	0.5	10.8
50	46	61	13.8	5.1	20	0.7	9.1
50	46	61	13.8	10.2	9	2.9	15.0
43	38	53	13.8	12.7	25	1.3	12.1
34	31	46	13.8	3.8	20	0.5	9.4
21	19	34	13.8	5.1	25	0.5	11.4
41	36	51	13.8	19.1	17	2.9	16.0
45	40	55	13.8	3.8	32	0.3	7.4
34	31	46	13.8	8.9	25	0.9	11.6
41	36	59	20.7	12.7	22	1.5	12.8
34	31	53	20.7	2.5	20	0.3	8.2
48	44	66	20.7	7.6	32	0.6	9.1
41	36	66	27.6	5.1	20	0.7	9.8
34	31	61	27.6	2.5	25	0.3	7.6
48	44	74	27.6	2.5	30	0.2	6.4

Rodel IC1400 pad without grooves, $\mu_{\text{water}} = 0.001\ \text{Pa} \cdot \text{s}$, $r_{outer} = 19\,\text{mm}$, $r_{inner} = 3.1\,\text{mm}$, $A_{cup} = 25.6\ \text{cm}^2$

Table B.2. Data of the Couette Flow Experiments

Wafer Speed [rpm]	Pad Speed [rpm]	r_{cc} [mm]	$r_{cc}\omega_w$ [m/s]	Initial fluid height [mm]	Final fluid height [mm]	Δh [mm]	Δt [min]	Q [ml/min]
100	0	0	0.00	67.3	67.3	0.0	15	0.00
300	0	0	0.00	67.3	66.0	1.3	15	0.17
100	0	25.4	0.27	68.6	67.3	1.3	15	0.17
100	0	25.4	0.27	69.9	68.6	1.3	12	0.21
150	0	25.4	0.40	67.3	54.6	12.7	8	3.11
200	0	25.4	0.53	71.1	53.3	17.8	9	3.87
225	0	25.4	0.60	61.0	53.3	7.6	8	1.87
250	0	25.4	0.66	71.1	50.8	20.3	8	4.98
275	0	25.4	0.73	53.3	48.3	5.1	9	1.11
300	0	25.4	0.80	71.1	62.2	8.9	12	1.45
350	0	25.4	0.93	69.9	58.4	11.4	8	2.80
400	0	25.4	1.06	71.1	64.8	6.4	10	1.24
100	200	25.4	0.27	61.0	50.8	10.2	9	2.21
150	200	25.4	0.40	61.0	50.8	10.2	8	2.49
200	200	25.4	0.53	53.3	47.0	6.4	12	1.04
300	200	25.4	0.80	47.0	40.6	6.4	10	1.24
200	300	25.4	0.53	78.7	71.1	7.6	10	1.49
300	300	25.4	0.80	74.9	66.0	8.9	10	1.74
400	300	25.4	1.06	72.4	63.5	8.9	8	2.18
400	400	25.4	1.06	73.7	66.0	7.6	8	1.87

Rodel IC1400 pad without grooves, $\mu_{\text{water}} = 0.001\ \text{Pa}\cdot\text{s}$, $r_p = 19\ \text{mm}$, $A_{cup} = 19.6\ \text{cm}^2$,
Pad with 21 perforations $\phi_{hole} = 2.4$ mm each, arranged in 6.3 mm square grid.

Table B.3. Effect of slurry flow rate on the Preston constant in the face-down polisher

Q [ml/min]	Initial Mass [g]	Final Mass [g]	Polishing Time [min]	Δm [mg]	ΔV [mm^3]	k_p $\left[1/Pa \times 10^{13}\right]$
0	11.2651	11.2650	4	0.1	0.01	0.0
3	9.2418	9.2256	5	16.2	1.82	0.7
7	9.1330	9.1140	4	19.0	2.13	1.1
12	9.1140	9.0818	4	32.2	3.61	1.8
15	9.1197	9.0950	4	24.7	2.77	1.4
15	9.0945	9.0642	4	30.3	3.40	1.7
15	9.3926	9.3631	4	29.5	3.31	1.7
20	9.0795	9.0390	4	40.5	4.55	2.3
30	9.2256	9.1906	4	35.0	3.93	2.0
45	9.1459	9.1250	4	20.9	2.34	1.2

Rodel IC1400 Pad with concentric K-grooves, $\omega_w = \omega_p = 8.37\text{rad/s}$, $r_{cc} = 88.9\text{mm}$, $p = 14\text{kPa}$, 100mm wafer diameter, Cabot ICUE 5001 slurry with 10% by vol. of 30% H_2O_2.

Appendix C

Photographs

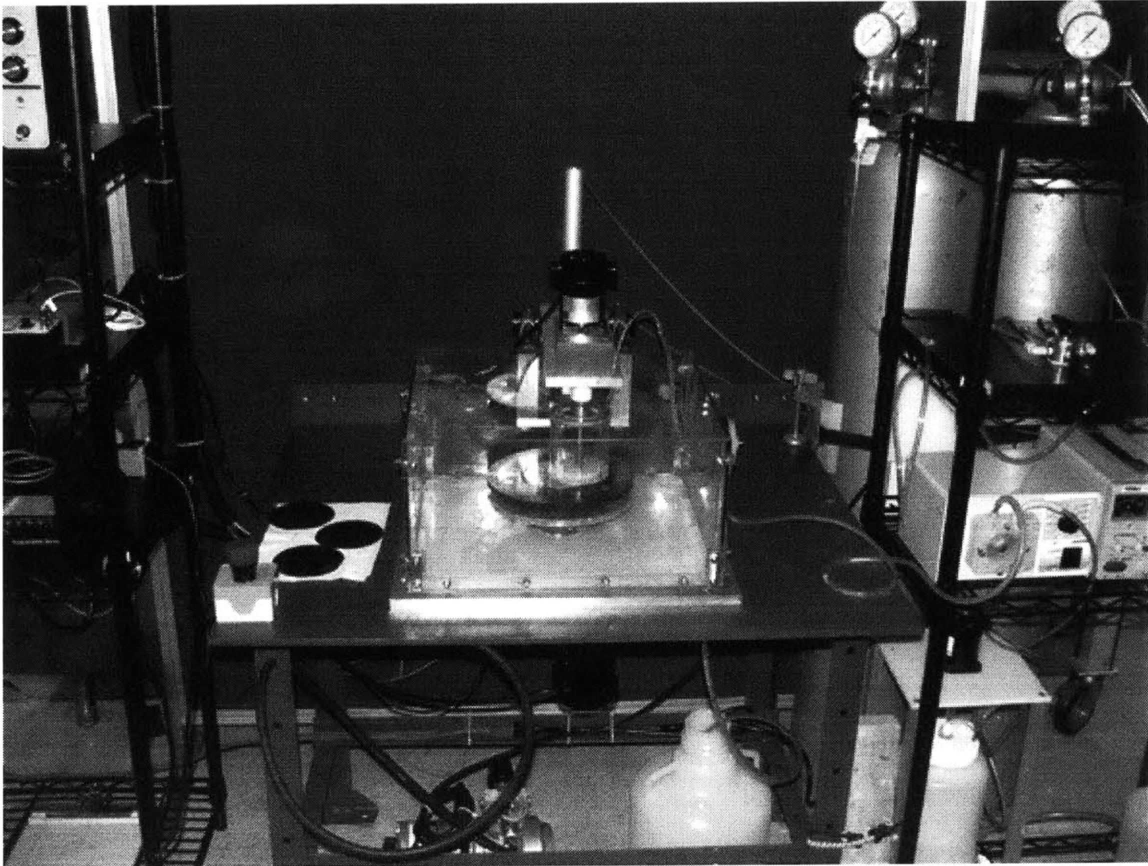


Figure C.1. Photograph of the prototype face-up polisher.

Nomenclature

g	Gravitational acceleration
h	Height
$\bar{h}$	Effective gap between the wafer and the pad
k_p	Preston constant
m	Mass
p	Pressure
p_0	Characteristic pressure
p_{nom}	Nominal applied pressure
p'	Normalized pressure
r	Radial position
r,θ	Wafer-centered polar coordinates
r,ϕ	Pad-centered polar coordinates
r_{cc}	Center-to-center distance between the pad and the wafer
rev	Revolution of the wafer
r_p	Pad radius
r'	Normalized radial position
t	Time
$t_{\exp}$	Total experiment time
t'	Normalized time
v	Velocity
$\bar{v}$	Average velocity
v_{cc}	Linear outward velocity of the pad
v_P	Velocity at point P
v_R	Relative velocity
v'	Normalized velocity
z	Vertical position
z'	Normalized vertical position
A	Area
F	Force
Q	Volumetric flow rate
T	Torque
U	Characteristic velocity
V	Volume
μ	Viscosity
μ_f	Coefficient of friction

θ_p	Angular span of pad on the wafer
ρ	Density
τ	Characteristic time
ω_p	Pad angular velocity
ω_w	Wafer angular velocity

References

[1] Noh, K., "Modeling of Dielectric Erosion and Copper Dishing in Chemical-Mechanical Polishing", PhD Thesis, Department of Mechanical Engineering, MIT, 2005.

[2] Adams, B.W., Swedek, B., Bajaj, R., Redeker, F., Birang M. and Amico, G., "Full-Wafer Endpoint Detection Improves Process Control in Copper CMP," *Semiconductor Fabtech*, vol. 12, pp. 283-289, 2001.

[3] Fury, M. A., "The early days of CMP," Solid State Technology, pp. 81-86, May 1997.

[4] Luo, J. and Dornfeld, D.A., "Optimization of CMP From the Viewpoint of Consumable Effects," *J. Electrochemical Soc.*, vol. 150, pp. G807-G815, 2003.

[5] Luo, J. and Dornfeld, D.A., "Material Removal Mechanism in Chemical Mechanical Polishing: Theory and Modeling," *IEEE Trans. on Semiconductor Manufacturing*, vol. 14, pp. 112-133, 2001.

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

[7] Guo, L. and Subramanian, R. S., "Mechanical Removal in CMP of Copper Using Alumina Abrasives," *J. Electrochemical Soc.*, vol. 151, pp. G104-G108, 2004.

[8] Jindal, A., Li, Y. and Babu, S.V., "Effect of pH on Chemical Mechanical Polishing of Copper and Tantalum," *Proc. Mat. Res. Soc. Symp.*, April 16-20, San Francisco, CA, vol. 671, pp. M.6.81-M.6.86, 2001.

[9] Steigerwald, J. M., Murarka, S.P., Gutmann, R.J. and Duquette, D.J., "Chemical Processes in the Chemical Mechanical Polishing of Copper," *Materials Chemistry and Physics*, vol. 41, pp. 217-228, 1995.

[10] Sniegowski, J.J., "Chemical-Mechanical Polishing: Enhancing the Manufacturability of MEMS," *SPIE Micromachining and Microfabrication '96 Symposium*, vol. 2879, pp. 104-115, 1996.

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

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

[13] Luo, Q., Ramarajan, S. and Babu, S.V., "Modification of the Preston Equation for the Chemical-Mechanical Polishing of Copper," *Thin Solid Films*, vol. 335, pp. 160-167, 1998.

[14] Johnson, K.L., "Contact Mechanics," Cambridge University Press, pp. 35-44, 1985.

[15] Churilov, V. A., "Action of an Elliptic Stamp Moving at a Constant Speed on an Elastic Half-Space," *Journal of Applied Mathematics and Mechanics*, vol. 42, no. 6, pp. 1074-1079, 1978.

[16] Braun, A.E., "Copper CMP Advances, Low-k Retreats," *Semiconductor International*, pp. 54-58, August 2003.

[17] Morrison, B., Joshi, S. and Tolles, R., "Copper and STI CMP Technology: the Challenges and the Cost," *Future Fab International*, vol. 11, pp. 269-273, 2000.

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

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

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

[21] Ouma, D.O., Boning, D.S., Chung, J.E., Easter, W.G., Saxena, V., Misra, S. and Crevasse, A., "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, 2002.

[22] Park, T., Tugbawa, T. and Boning, D.S., "Overview of Methods for Characterization of Pattern Dependencies in Copper CMP," *Proc. CMP-MIC Conf.*, February 28-March 3, Santa Clara, CA, pp. 196-205, 2000.

[23] "International Technology Roadmap for Semiconductors," 2003 Edition, http://public.itrs.net, 2003.

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

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

[26] Thakurta, D.G., Borst, C. L., Schwendeman, D. W., Gutmann, R. J. and Gill, W. N., "Pad porosity, compressibility and slurry delivery effects in chemical-mechanical planarization: modeling and experiments," *Thin Solid Films*, vol 366, pp. 181-190, 2000.

[27] Sundararajan, S., Thakurta, D.G., Schwendeman, D.W., Murarka, S.P. and Gill, W.N., "Two-Dimensional Wafer-Scale Chemical Mechanical Planarization Models Based on Lubrication Theory and Mass Transport," *J. Electrochemical Soc.*, vol. 146, pp. 761-766, 1999.

[28] Thakurta, D.G., Borst, C. L., Schwendeman, D. W., Gutmann, R. J. and Gill, W. N., "Three-Dimensional Chemical Mechanical Planarization Slurry Flow Model Based on Lubrication Theory," *J. Electrochemical Soc.*, vol. 148, pp. G207-G214, 2001.

[29] Muldowney, G. P., Hendron, J. J. and Crkvenac, T. T., "The Impact of Slurry Backmixing in Determining Optimal CMP Process Conditions," *Proc. CMP-MIC Conf.*, February 24-March 26, Santa Clara, CA, pp. 224-231, 2004.

[30] Muldowney, G. P., "A Computational Study of Slurry Flow in Grooved CMP Polishing Pads," *Proc. CMP-MIC Conf.*, February 24-March 26, Santa Clara, CA, pp. 147-156, 2004.

[31] Emslie, A. G., Bonner, F. T., Peck, L. G., "Flow of a Viscous Liquid on a Rotating Disk," *Journal of Applied Physics*, vol. 29, no. 5, pp. 858-862, 1958