CHAPTER 1

INTRODUCTION

1.1 General Background

The electronics industry has grown rapidly in the past three decades. Ultra-large-scale integrated (ULSI) circuits, with 10^8 or more devices on a chip, can now be fabricated on semiconductor substrates, or wafers, to reduce cost and to increase the performance of electronic products. Figure 1.1 shows the growth of the number of components on a metal-oxide-semiconductor (MOS) memory chip. That number has approximately doubled every two years over the past two decades, matching the rate Moore forecast (Moore, 1965; Chang and Sze, 1996). Concurrently, the minimum dimension of the device-feature continues to shrink by about 13% per year, or by a factor of two every six years, due to the advances in fabrication technology. The decrease of feature length reduces the overall device size and increases the packing density, and thus reduces the cost of function. Moreover, device speed, which varies inversely to feature length, has been improved and power consumption, which approximately varies as the square of feature length, has been reduced. On the other hand, the complexity of microchip design and fabrication has increased continuously with integration and miniaturization. Extremely high degrees of repeatability and uniformity are required in wafer fabrication for a high production yield. One challenge is the demand for global surface planarity for the stringent depth of focus of the lithography in the back-end process integration. The surface topography of the inter-level dielectric (ILD) builds up after multiple levels of metal wiring. The Chemical Mechanical Polishing (CMP) process has emerged as a critical technique to smooth surface topography. Additionally, new materials such as Cu and W, introduced in ULSI fabrication, also require extensive use of the CMP process to form inlaid interconnect structures. A fundamental understanding of the CMP process is essential to improve process optimization and control, and to increase the process yield and throughput in the continuous integration and miniaturization in the semiconductor industry.

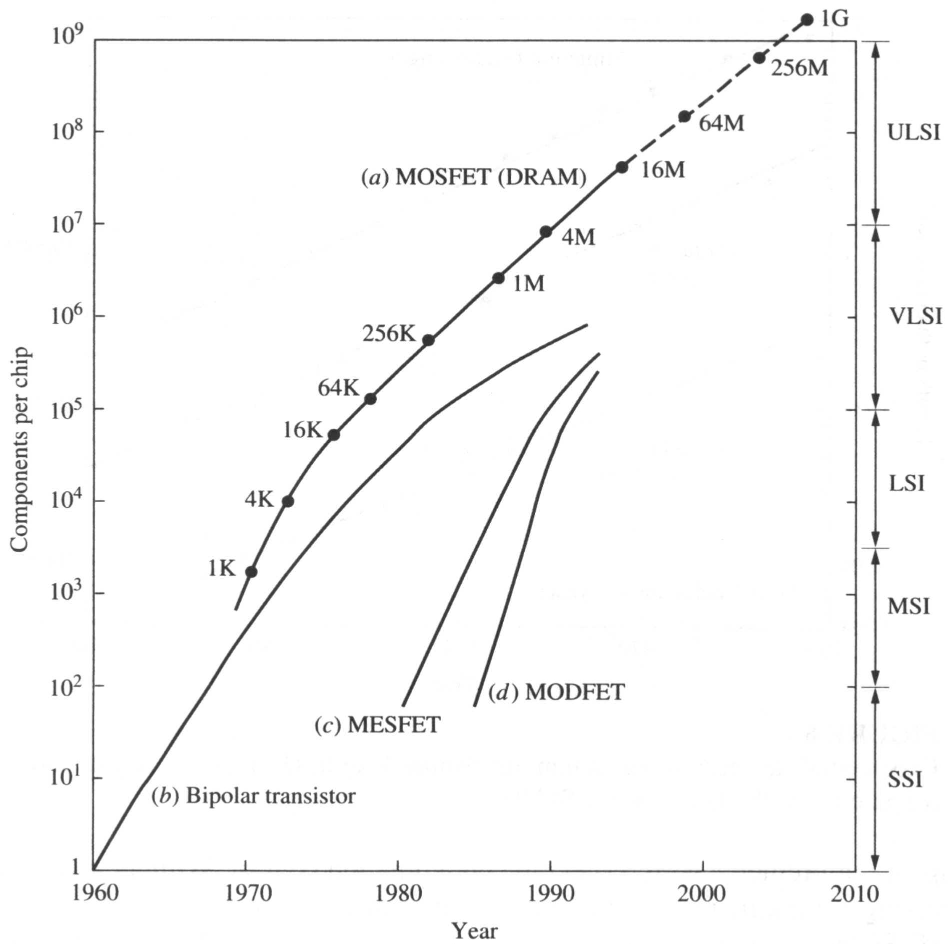

Figure 1.1 Experimental growth of the number of components per MOS IC chip (from Chang and Sze, 1996).

1.2 The Role of CMP in Semiconductor Manufacturing

Figure 1.2 is a schematic of a conventional CMP process. It is an adaptation of the lapping technology used to polish plate glass. The wafer to be polished is mounted on a wafer carrier via back pressure or via surface tension by wetting its back surface. The wafer is pressed down against a rotating platen, which holds a compliant polishing pad. The wafer slides on the pad surface with a relative velocity generated by the rotation of the carrier and the platen. Concurrently, the abrasive slurry drips onto the platen surface and dispenses through the wafer/pad contact interface. The chemical slurry and abrasive particles retained on the porous pad surface remove the material on the wafer surface.

There are two major applications of CMP in ULSI manufacturing: to smooth surface topography of inter-level dielectrics (ILD, usually silicon dioxide), or to remove excess material to produce inlaid metal structure or isolation trenches. The inter-level dielectric CMP is applied in conventional aluminum metallization, where aluminum is deposited on the oxide ILD layer, patterned, and etched to form interconnects. Another layer of oxide is then deposited to insulate the aluminum interconnects. Thus three-dimensional electrical wiring is constructed. Device elements, such as resistors, capacitors and transistors, are connected to build up ICs. In this practice, CMP is employed on at least the top few layers of each ILD surface to provide a smooth surface for aluminum deposition and to provide a field flat enough that contact vias and metal wires can be patterned by lithography. The desired process end-point is determined based on the surface planarity and the thickness of the ILD layer required for electrical isolation of aluminum line. Figure 1.3 (a) shows the schematic for the ILD CMP. Wafers stacked with three or more layers of aluminum interconnects, such as are used in microprocessor applications, are usually subjected to ILD CMP to improve yield and reliability.

Unlike ILD CMP, copper CMP and other trench isolation processes such as shallow trench isolation (STI) are employed to remove the excess deposit covered on the trenches. As shown in Fig. 1.3 (b) for the copper damascene process, the underlying oxide (or other ILD material) layer is trenched by lithography and etching. A thin copper layer is deposited or electroplated to fill the trenches. Then a CMP process removes excess copper and forms

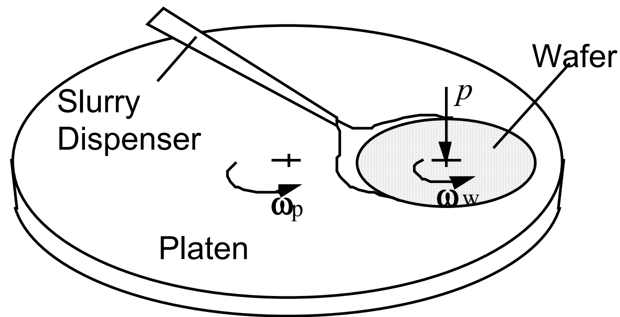

Figure 1.2 Schematic of Chemical Mechanical Polishing Process.

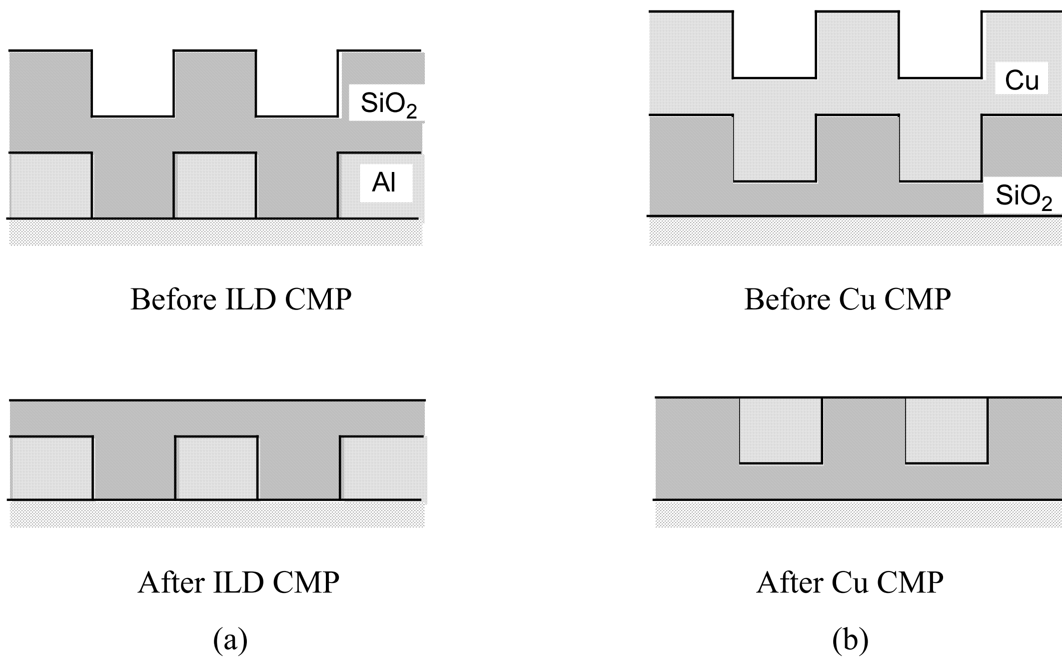

Figure 1.3 Schematics of (a) ideal oxide ILD CMP and (b) ideal Cu CMP processes.

isolated wirings of copper. The process is stopped while the copper layer and diffusion barrier layer (usually a thin Ta, TaN, Ti or TiN layer to prevent copper diffusion into the oxide and "poisoning" the underlying devices) are completely polished through and the oxide is exposed. Copper technology is expected to replace aluminum in new-generation chips with interconnect critical dimensions (CDs) below 0.25 µm. Since copper is difficult to pattern and etch, the damascene approach combined with extensive use of CMP seems to be the best solution for ULSI manufacturing.

1.3 Origins and Evolution of the CMP Process

1.3.1 Origins and Early Developments. The original commercial practice of CMP in the semiconductor industry was to prepare raw silicon wafers. After the wafer was sawed from a single-crystal silicon rod, the mechanically damaged surface layer was removed and the surface planarized to produce a flat, scratch-free surface for VLSI devices and circuits. Monsanto first developed this process and sold polished wafers in late 1962 (Walsh and Herzog, 1965; Hippel, 1988). Despite its recent use in VLSI fabrication, the polishing process has been employed for optical lens fabrication for centuries. In fact, the first machinery used by Monsanto was very similar to the commercial machine used in the optical industry. Monsanto's innovation was to redefine the polishing steps and choose proper polishing abrasives and slurry chemicals to achieve the specifications for both surface finish and planarity of silicon wafers. By using submicron silicon dioxide abrasives in an alkaline solution, the process time was reduced from one and a half hours to about five minutes.

A further improvement to the CMP process was made at IBM in the late seventies (Silvey et al., 1966). A cupric salt CMP slurry was employed to improve the throughput and surface finish on a conventional polishing machine. The new process was faster than the previous silica based polishing method and resulted in a mirror-like surface (Blake and Mendel, 1970). The slurry was tailored later to reduce defects and surface nonplanarity introduced by the etching and deposition processes. By the late 1980s, the CMP process was successfully applied in Japan for trench isolation (to insulate the implanted device region with planarized silicon dioxide walls for higher device packing density) and trench capacitors for DRAMs.

NEC, IBM, National Semiconductor and other chipmakers, were experimenting with different applications and products for both memory and logic ICs. In 1988, the first commercial polisher designed specifically for CMP was introduced by Cybeq in Japan. Later, SEMATECH identified CMP as a technology critical for future IC manufacturing. It launched a project with Westech to develop competitive, advanced CMP tools in the U.S.

1.3.2 Oxide Planarization. Due to the fast shrinkage of the feature size and the increased number of devices, more metal-interconnect layers were required for each new generation of ICs, especially for logic chips and microprocessors. The quick build-up of surface topography with the increase of interconnect layers usually results in a poor step-coverage of the metal deposition. It thus requires a global planarization technique on the ILD surface. The smaller wavelength light employed in lithography to produce submicron features also sets a more stringent limit on surface planarity. The superior ability of CMP to produce a flat and smooth surface provides an ultimate solution for more complex chip design. Keeping the CMP performance robust and reliable became a key issue for integration of the CMP into the large-scale manufacturing operation. In 1993, Intel demonstrated intense use of CMP on the new three-level microprocessor, the Pentium chip. It was the first time that CMP was employed to produce high volumes of commercial devices.

The main impediment to oxide polishing is the slow rate of ILD oxide removal. Because the hardness of the abrasive is comparable to the surface coating, the process may rely on some chemicals to "soften" the surface and form a passive layer to be easily removed by particle abrasion (Cook, 1990). Moreover, the within-wafer and wafer-to-wafer uniformity of material removal varied significantly in early practice. Nonuniformity was partially due to the variation of process variables, such as pressure distribution, slurry transport, and degradation of pad surface during polishing. Due to the lack of knowledge about the fundamentals of the process, it was difficult to analyze these problems and thus improve process performance.

1.3.3 The New Era for Copper Polishing Technology. In the past, aluminum was exclusively employed as the main interconnect material. However, when the device dimensions below the sub-micron, the still high electrical resistivity of aluminum slows down the rate of digital signal transmission, the so-called RC delay (which is equal to the product of

the resistance and the capacitance of the line). The sub-micron aluminum interconnect cannot handle high power density applications because of its relatively low resistance to electro-migration. Electromigration is the phenomenon of metal atoms diffusing as currents carrying electrons "hit" them, which causes metal thinning. Because of these problems, the semiconductor industry focused on replacing aluminum with other metals when the SIA (Semiconductor Industry Association) noted that feature size would shrink down to 0.35 μm by 1995 and to 0.25 μm by 1997. Copper seemed to be the best candidate because of its lower electrical resistivity (about two thirds of aluminum) and high electromigration resistance. Cost reduction was another reason. Copper techniques require 20% to 30% fewer steps than conventional aluminum patterning due to the new damascene approach and the higher packing density of the smaller feature size (Singer, 1998).

For these reasons, the major chipmakers, such as IBM, Motorola and Texas Instruments, separately announced in 1997 aggressive plans to put copper into production in 1998. Table 1.1 presents the 1998 SIA roadmap for the interconnect technology. The main challenges for Cu CMP is to control the uniformity of the surface topography while the interconnect layers increase to meet the more stringent die-level planarity requirement. Moreover, the nonuniformity of the remaining Cu wires and the variation of local surface topography due to dishing and overpolishing must be reduced to prevent the retardation of signal transmission in interconnects.

Despite the many advantages of copper, copper patterning requires an entirely new manufacturing platform, damascene technique. As described earlier, the copper damascene technique circumvents metal etching, using the CMP process to form copper wiring inside the trenched oxide. The CMP process must be integrated horizontally and vertically to achieve high quality process performance. Horizontal integration, which includes reliability of the consumables, development of new slurries, wafer cleaning systems, and a new metrology for endpoint detection, focuses on an integral solution for a robust yield and throughput for the CMP. Vertical integration integrates upstream processes such as Cu/barrier deposition and etching and downstream processes such as ILD deposition and lithography. It opens a wide process window for more optimization options and increases the quality of the final products.

Table 1.1: SIA International technology roadmap for semiconductors (ITRS) for interconnect technology (1998 updated).

Year of First Product Shipment	1997	1999	2002	2005	2008	2011
Technology Node	250 nm	180 nm	130 nm	100nm	70 nm	50 nm
Number of Metal Levels - DRAM	2 - 3	3	3	3 - 4		
Number of Metal Levels - Logic	6	6 - 7	7	7 - 8		
Maximum Interconnect Length - Logic (m/chip)	800	1,700	3,300	5,000	9,200	17,000
Planarity Requirements within Litho Field for Minimum Interconnect Critical Dimension (CD) (nm)	300	250	200	175	175	175
Minimum Contacted / Noncontacted Pitch - DRAM (nm)	550/500	400/360	280/260	220/200	160/140	110/100
Minimum Contacted / Noncontacted Pitch - Logic (nm)	640/590	460/420	340/300	260/240	190/170	140/130
Minimum Metal CD for Isolated Lines (nm)	250	180	130	100	70	50
Minimum Contact / Via CD (nm)	280/360	200/260	140/180	110/140	80/100	60/70
Metal Height / Width Aspect Ratio - Logic (Microprocessor)	1.8	1.8	2.1	2.4	2.7	3
Via Aspect Ratio - Logic	2.2	2.2	2.5	2.7	2.9	3.2
Minimum Metal Effective Resistivity ($\mu\Omega$/cm)	3.3	2.2	2.2	2.2	<1.8	<1.8
Barrier / Cladding Thickness (nm)	100	23	16	11	3	1
Minimum Interlevel Metal Insulator - Effective Dielectric Constant (k)	3.0-4.1	2.5-4.1	1.5-2.0	1.5-2.0	≤ 1.5	≤ 1.5

1.3.4 CMP Process in the Future. Beyond the adoption of copper interconnects, several technologies are necessary to continue the shrinkage of device dimension and the increase of packing density in ULSI manufacturing. The most promising one is low-k dielectric technology. Low-k materials can replace the present oxide ILD layer to reduce the capacitance loss and increase the signal transmission rate of the circuits. According to the SIA roadmap (Table 1.1), low-k dielectrics will be integrated with copper damascene technology in the next year or two on the 0.13μm-generation chips and beyond. However, the difference between the polishing rates of copper and the low-k materials available will significantly affect post-CMP surface planarity. New processes must be developed to address the problems associated with this nonuniform polishing phenomenon as well as the complexity of the materials structures. Additionally, it is necessary to explore the niche for the CMP process in shallow trench isolation and other applications such as back-side polishing and the fabrication of micro-electro-mechanical systems (MEMS).

1.4 Scope of Present Investigation

Because Cu damascene will emerge as the mainstream process to produce interconnects, this thesis focuses on developing an understanding of the Cu CMP process and designing the process to meet the necessary performance metrics. The target performance requirements include: material removal rate (MRR), within-wafer uniformity (WIWU) and wafer-to-wafer uniformity (WTWU, wafer-level), within-die planarity (die-level), dishing and overpolishing of an interconnect (device-level), and roughness and defects on the local surface (sub-device or line-level). These in turn define the yield and throughput of the Cu CMP. These requirements depend on a large number of process variables; even the fundamentals of the material removal in Cu CMP are still not clear. As shown in Fig. 1.4, process variables can be categorized as input variables or process parameters. The input parameters include both the geometrical shape of the initial wafer surface, such as the wafer curvature, the initial geometry of the coatings, and the local surface topography of the pattern, and the properties of the surface coatings, such as the mechanical and chemical properties of the coating materials. These variables are determined by the previous deposition processes, the initial chip design, or the initial wafer geometry. They generally will constrain CMP performance.

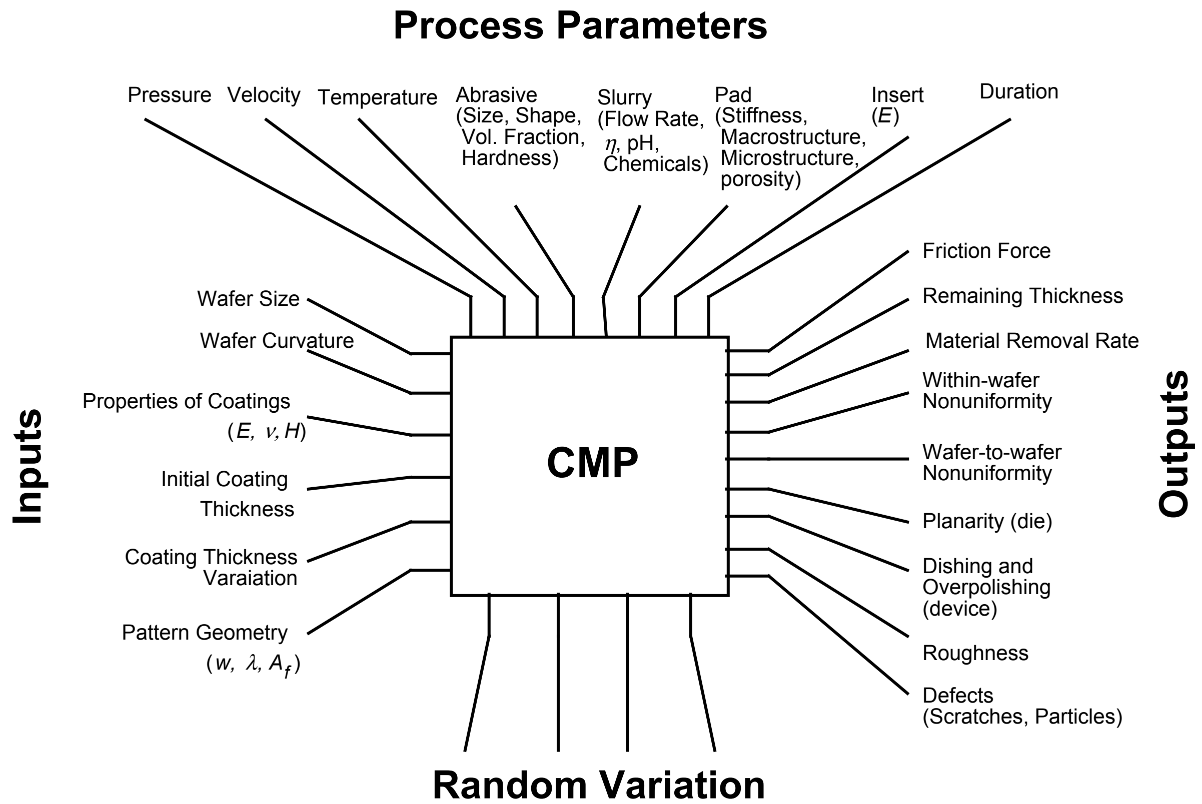

Figure 1.4 Complexity of the CMP process.

Process parameters can be manipulated to optimize process outputs and meet performance requirements. The process parameters, also shown in Fig. 1.4, may be grouped as the operation parameters (pressure, velocity, temperature, etc.), the slurry related parameters (abrasive and slurry fluid), and the pad characteristics.

The essence of process design is process decoupling. The wafer-level outputs of polishing, such as the average removal rate and within-wafer uniformity, generally are not affected by die layout and local pattern geometry. The local geometry after polishing (which defines the within-die planarity), oxide overpolishing, and Cu dishing, tend to be similar among different dies and will not be affected by the behavior of global material removal. This implies that process outputs at different length scales can be decoupled and that the mechanisms corresponding to each level of polishing behavior can be studied to identify the key process variables. As summarized in Table 1.2, this thesis characterizes the interfacial contact condition and its effects on the wafer-level material removal and investigates polishing uniformity are investigated. The possible mechanisms of material removal in CMP are modeled and examined for a wide range of coating materials and process conditions. Once the dominant mechanisms are determined, the key factors that affect material removal rate and surface quality can be identified and controlled to improve both the global (wafer-level) and the local level polishing performance. A neutral slurry is used in the present study to elucidate the mechanical aspects of the CMP process. The within-die planarity is studied based on the pressure distribution in the die field. A contact mechanics model is proposed to correlate the pressure distribution and pad displacement into low feature on the Cu patterned wafers to the pattern geometry and layout. Based on the model, the rates of planarization on various damascene structures are predicted and validated by experiments. For the nonuniform polishing behavior at the device-level, Cu dishing and ILD overpolishing are investigated through a systematic experimental approach. A semi-empirical model is established to study the effects of several device-scale factors such as the material selectivity, pattern linewidth, and pattern area fraction on the rates of dishing and overpolishing. Moreover, an *in situ* reflectance sensing and endpoint detection technique is developed to reduce overpolishing time and to account for the variation of process outputs. By integrating the endpoint control with the process optimization schemes, the variance of process outputs can be reduced to improve process yield.

Table 1.2: Scope of the process research.

Geometry Level	Problem and Issue	Mechanisms	Process parameters of Interest
Wafer	**Material removal rate**	• Interfacial contact condition	• All
		• Mode of polishing	
	Within-wafer uniformity	• Interfacial contact condition	• *p, v*, slurry viscosity
	Wafer-to-wafer uniformity	• Slurry transport	• Macrostructure of pad
		• Global pressure distribution	• Mechanisms of pressure application
Die	**Within-die planarity**	• Nonuniform pressure distribution due to pattern geometry	• Pattern area fraction
			• Pad elastic properties
		• Selectivity	• Coating/abrasive hardness
Subdie & Device	**Device level uniformity (Dishing and Overpolishing)**	• Nonuniform pressure due to device or interconnect geometry	• Pattern linewidth and area fraction
			• Pad properties
Sub-device	**Roughness**	• Mechanisms of removal	• Abrasive size and hardness
	Scratching	• Particle agglomeration and contamination	• Abrasive, pH, chemistry

1.5 Thesis Organization

The overall goal of this thesis is to study the mechanics and the mechanisms of the polishing process and to model polishing behavior at different length scales. This will provide a theoretical framework to understand the process fundamentals and provide optimization schemes for the Cu CMP process. The results not only can be applied to the current Cu CMP practice, but can also be employed to guide CMP research for the next-generation IC fabrication, involving new materials and new geometrical complexity.

The present chapter describes the background and scope of the thesis. Chapter 2 studies kinematics of the polishing tools, rotary and linear. A technique is developed to characterize the wafer/pad interfacial contact condition and is related to such global process parameters as pressure, relative velocity, and slurry viscosity. The effects of contact mode on the material removal rate and within-wafer nonuniformity are discussed. Optimization schemes to control the process in the contact mode and to determine the preferable pressure and velocity based on the consideration of frictional heat dissipation are proposed. Chapter 3 describes possible mechanisms of the fine-abrasive polishing process. The dominant regime in CMP is determined based on the experimental results and surface topography of the polished surface. Both abrasive size and hardness effects are employed for process optimization to improve the material removal rate and limit surface roughness. Chapter 4 provides a contact mechanics model to determine the pressure distribution on the patterned Cu wafer. Experiments verify the model in the Cu planarization regime. Chapter 5 provides the effects of pattern geometry on Cu dishing and oxide overpolishing. A general trend of the evolution of surface topography in Cu damascene polishing is also studied. A robust process design to increase the Cu/ILD polishing selectivity, based on the effects of abrasive hardness, size, and the elasticity of the pad, are proposed to improve the wafer- and die-level surface uniformity. Chapter 6 offers the modeling of the reflectance sensing technique and an *in situ* endpoint sensing algorithm for Cu CMP. Experiments are also conducted to verify the detection schemes. Chapter 7 presents conclusions and Chapter 8 offers a few suggestions for future research on CMP. The theories of oxide CMP are briefly reviewed in Appendix A. The pin-on-disk polishing results on various materials similar to the wafer coatings are presented in Appendix B.

References

Baliga, J., 1999, “Economic Forecast: Slowly Turning Upward,” *Semiconductor International*, Vol. 22, pp. 56-60.

Blake, L.H. and Mandel, E., 1970, “Chemical-Mechanical Polishing of Silicon,” *Solid State Technology*, Vol. 13, pp. 42-46.

Brown, N.J., Baker, P.C., and Maney, R.T., 1981, “Optical Polishing of Metals,” *Proc. SPIE*, Vol. 306, pp. 42-57.

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

Cook, L.M., 1990, “Chemical Processes in Glass Polishing,” *J. Non-Crystalline Solids*, Vol. 120, pp. 152-171.

Hipple, E., 1988, *The Source of Innovation*, Oxford University Press, Oxford, UK.

Moore, G.E., 1965, “Cramming More Components onto Integrated Circuits,” *Electronics*, p. 114.

Newton, Sir Isaac, 1695, *Opticks*, Dover Publication, Inc., New York, 1952, based on the 4th ed., published in 1730.

Silvey, G.A., Regh, J. and Gardiner, 1966, U.S. Patent No. 3,436,259, assigned to IBM 1966.

Singer, P., 1998, “Tantalum, Copper and Damascene: The Future of Interconnects," *Semiconductor International*, Vol. 21, pp. 91-98.

Walsh, R.J. and Herzog, A., 1965, U.S. Patent No. 3,170,273, issued 23 February 1965.