CHAPTER 5

EVOLUTION OF COPPER OXIDE DAMASCENE STRUCTURES IN CMP: II. DISHING AND OVERPOLISHING

Test wafers comprising damascene structures are designed and fabricated to investigate Cu dishing and oxide overpolishing. The mask design coveres a wide range of linewidths and pitches to represent such features as signal and power transmission lines and probing or wire-bonding pads. Experiments are conducted to investigate the evolution of pattern profile during polishing and to determine the onset and rates of dishing and overpolishing. The effects of Cu linewidth and area fraction on the rates of pattern planarization, Cu dishing and oxide overpolishing are quantified. The effect of hardness of the composite surface on dishing and overpolishing are examined. An optimization scheme, employing particle size, particle hardness, and pad stiffness to increase MRR, to enhance the selectivity between SiO_2 and Cu, and to reduce surface nonplanarity is proposed.

5.1 Introduction

As shown in Chapter 4, the local (die-scale) pattern geometry affects the local MRR significantly. The nonuniform pressure distribution resulting from the nonuniform area fraction and layout of the pattern introduces surface non-planarity in the planarization stage of metal polishing. Consequently, the pattern is slightly overpolished to remove all the metal coating on the dielectric surface so that the metal interconnects are isolated. Concurrently, dishing occurs on the soft metal in the trenches and reduces the cross-sectional area of the interconnect. Both overpolishing and dishing result in surface nonplanarity and thickness variation of metal interconnects across a die area.

Dishing and overpolishing rates may be estimated by an extended version of the Preston equation:

$$\frac{dh}{dt} = k_p(x,y)\, p_{av}\phi(w, A_f, t^*, ...)v_R \tag{5.1}$$

The Preston constant k_p is a function of position which relates to the physical layout of the oxide and Cu interconnects. The Preston constants for different materials remain the same as those on blanket polishing. The Preston constant on blanket coating is a function of the coating hardness, abrasive size and hardness (as shown in Chapter 3), and the slurry chemistry (Kaufman et al., 1991; Steigerwald et al., 1995; Carpio et al., 1995; Zeidler et al., 1997; Fayolle and Romagna, 1997; Luo et al., 1998; van Kranenburg and Woerlee, 1998; Hariharaputhiran et al., 2000; Kondo et al., 2000). The pressure distribution is affected by the actual shape of the dished/overpolished surfaces, a function of Cu linewidth w, area fraction A_f and overpolishing time t*. The pressure distribution can be decoupled as a product of the average pressure on the die area and a geometrical function ϕ which includes the effects of pattern geometry. In practice, the geometrical function ϕ is not easy to find even when the surface topography is known. In this case, surface variation due to dishing and overpolishing is comparable to the surface roughness of the pad and the slurry particle size. Therefore, the contact mechanics model presented in Chapter 4, assuming a flat and homogeneous pad and neglecting the effects of particle, is no longer valid. However, a qualitative model for the rates of overpolishing and dishing at steady-state can be achieved with some simplified assumptions.

When the size of the planarized feature is close to or smaller than the abrasive particle diameter (0.2-0.3 μm) and pad surface roughness, the calculation of local pressure must take into account the particle distribution and the pad local topography. It is difficult, however, to develop an analytical model of this sort. Therefore, research on dishing and overpolishing has been confined to experimental characterizations and parametric studies on such pattern parameters as area fraction, linewidth and pitch (Murarka et al., 1993; Steigerwald et al., 1995; Gutmann et al., 1995; Stavreva et al., 1995, 1997; Park et al., 1999). Though a few semi-quantitative models have been proposed (Elbel et al., 1998; Tugbawa et al., 1999), the

fundamentals of dishing and overpolishing and their relation to pattern geometry and material properties are still not fully understood. Most of the experiments were conducted on larger size features. The results and associated problems (such as severe dishing on 100 μm features) may not be applicable to the current sub-quarter micron circuit design. The scaling issue must be addressed based on the similarity in fundamental material removal behaviors on different size features.

In this chapter, steady-state dishing and overpolishing are modeled based on the pressure distribution and the local-scale wear phenomena. Dishing is also studied by considering the effects of pattern geometry, pad displacement, and particle size. Experiments quantitatively establish the effects of Cu linewidth and area fraction on the rates of pattern planarization, Cu dishing, and oxide overpolishing. On this basis, the fundamentals of dishing and overpolishing phenomena and their mechanisms are examined and the important process parameters are identified. The results are correlated to the contact mechanics model in the previous chapter to determine the effects of pressure distribution due to surface remaining topography on the dishing overpolishing. Schemes for optimal Cu CMP to mitigate dishing and overpolishing, employing the effects of particle size, particle hardness, slurry pH, are discussed.

5.2 Theory

The Preston constant can be defined as the ratio of the wear coefficient k_w to the hardness H of material being polished. Thus the intrinsic MRR at any point on the wafer surface can be determined by the Preston equation, which may be rewritten as

$$\frac{dh}{dt} = \frac{k_w}{H} p v_R . \tag{5.2}$$

where p is the local average pressure applied at the vicinity of the points of interest on the wafer surface. As shown in Chapter 3, the wear coefficient depends on the polishing mechanism and is insensitive to the material polished. k_w roughly remains a constant for

various surface coatings, including Cu, TEOS, on blanket wafers in the CMP conditions. If we assume that the same k_w can be applied on both die-scale and feature-scale. The MRR on both Cu and oxide surface, as shown in Fig. 5.1, may be expressed as:

$$\left(\frac{dh}{dt}\right)_{Cu} = \frac{k_w}{H_{Cu}} p_{Cu} v_R \tag{5.3}$$

$$\left(\frac{dh}{dt}\right)_{Oxide} = \frac{k_w}{H_{Oxide}} p_{Oxide} v_R \tag{5.4}$$

where H_{Cu} and H_{Oxide} are the material hardnesses of Cu and oxide respectively. If the steady-state regime is assumed, i.e. the amount of Cu dishing remains constant with overpolishing time, the material removal on both the Cu and oxide surface should be uniform and at the same rate:

$$\left(\frac{dh}{dt}\right)_{Cu} = \left(\frac{dh}{dt}\right)_{Oxide} \tag{5.5}$$

By equating Eqs. (5.3) and (5.4), noticing that the relative velocity on the adjacent Cu and oxide regions are virtually the same, the relation between pressure distributed on Cu and oxide and the hardness of these materials can be expressed as

$$\frac{p_{Cu}}{p_{Oxide}} = \frac{H_{Cu}}{H_{Oxide}} \tag{5.6}$$

To solve p_{Cu} and p_{Oxide} with pattern geometry, the force equilibrium condition on the area of interest across a interconnect and the surrounding oxide spacing can be employed:

$$p_{Cu} w + p_{Oxide}(\lambda - w) = \bar{p}\lambda \tag{5.7}$$

where $\bar{p}$ is the average pressure on the specific area. In terms of area fraction A_f , $A_f = w/\lambda$, Eq. (5.7) can be rewritten as:

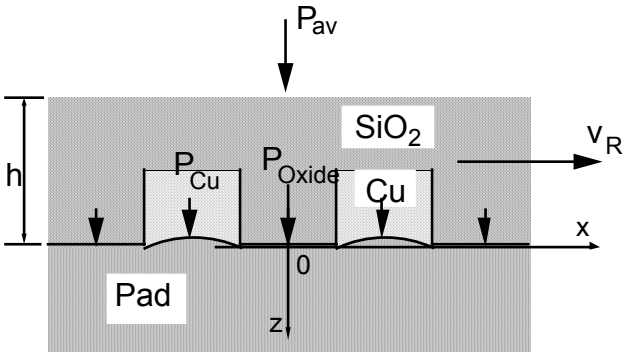


Figure 5.1 Schematics of the onsets of dishing and overpolishing.

$$p_{Cu} A_f + p_{Oxide}(1 - A_f) = \overline{p} \approx p_{av} \tag{5.8}$$

Because the surface variation (nonuniformity) will not be large, usually less than 100 - 200 nm after a short period of overpolishing, the local average pressure on $\overline{p}$ is approximately equal to the average pressure on the die p_{av}. Solving Eq. (5.8) by the relation given in Eq. (5.6), the pressure on the oxide surface in steady-state is given as:

$$p_{Oxide} = \frac{\overline{p}}{\left[(H_{Cu} / H_{Oxide}) A_f + (1 - A_f)\right]} \approx \frac{p_{av}}{\left[(H_{Cu} / H_{Oxide}) A_f + (1 - A_f)\right]} \tag{5.9}$$

The pressure on the oxide surface is a function of average pressure on the die, the hardnesses of materials presented on the surface, and the area fraction of pattern. If the pressure in Eq. (5.4) is replaced by Eq. (5.9), the Preston equation on the oxide surface can be rewritten as:

$$\left(\frac{dh}{dt}\right)_{Oxide} = \frac{k_w}{H_{Oxide}} \frac{\overline{p}}{\left[(H_{Cu} / H_{Oxide}) A_f + (1 - A_f)\right]} v_R = \frac{k_w}{H'} \overline{p} v_R \approx \frac{k_w}{H'} p_{av} v_R \tag{5.10}$$

where *H'* is defined as the "apparent hardness" and can be written as:

$$H' \equiv H_{Cu} A_f + H_{Oxide}(1 - A_f) \tag{5.11}$$

Equation (5.10) states that, in steady state, the polishing rate on a specific patterned area is equivalent to the rate in a field area with material hardness *H'* and the same average pressure $\overline{p}$ *(*$\approx p_{av}$*)*. If there is a variation of *H'* across a die area due to the variation of pattern area fraction, the deviation of the oxide and Cu thickness from the mean thickness will increase with overpolishing time. Hence the apparent hardness across the die should be as uniform as possible to reduce overpolishing. More details about process optimization will be discussed in the later section. Additionally, the rate of oxide overpolishing is bounded by the steady-state rate and the blanket oxide polishing rate. Based on force equilibrium, the pressure on the oxide will increase with the increase of dishing (less pressure will be applied on the Cu lines)

until it reaches a steady-state value. Similarly, the Cu polishing rate is bounded by the blanket Cu rate (as on the planar surface at the end-point) and the steady-state rate of the surrounding oxide (which is very close to blanket oxide polishing rate except in the case with extreme high area fraction).

5.3 Experimental

5.3.1 Mask Design. A Cu damascene structure was designed to study the effects of geometry on metal dishing and oxide overpolishing. As shown in Fig. 5.2, the pattern on each die (10 mm x 10 mm) consists of a matrix of 2 mm x 2 mm blocks (sub-die area). These blocks in turn consist of line-space features with a minimum linewidth of 0.5 μm. Table 5.1 lists the design features of the pattern and Fig. 5.3 shows the physical layout of the pattern on the mask. The first type of features are compose of fine Cu lines of constant linewidth 0.5 μm and of various pitch ranging from 1 μm to 200 μm. These represent the metal interconnects with critical dimension and different packing density. The second type of features study the effect of linewidth on dishing. Various Cu lines, from 0.5 μm to 100 μm, with large pitch, 200 μm, provide wide spacing between adjacent Cu interconnects. For small Cu lines, the wide spacing reduces the effect of SiO_2 overpolishing on dishing. Two constant Cu area fractions, 0.01 and 0.5, with various linewidths and pitches are the third type of features to study the effects of scaling on dishing and overpolishing. The 0.5 area fraction is close to the present design rules of metal layer layout in ULSI circuits. By contrast, the features with 0.01 area fraction represent single, isolated lines.

Lithography transferred the pattern onto the 1.5 μm thick SiO_2 (TEOS) coating on a 100 mm, (100) orientation silicon wafer. After oxide trenches were etched to a depth of 1 μm, a 20 nm thick Ta barrier layer was deposited, followed by a 1.5 μm thick PVD Cu film. Figure 5.4 shows the scanning electron micrograph (SEM) of the cross-section of the patterned wafer.

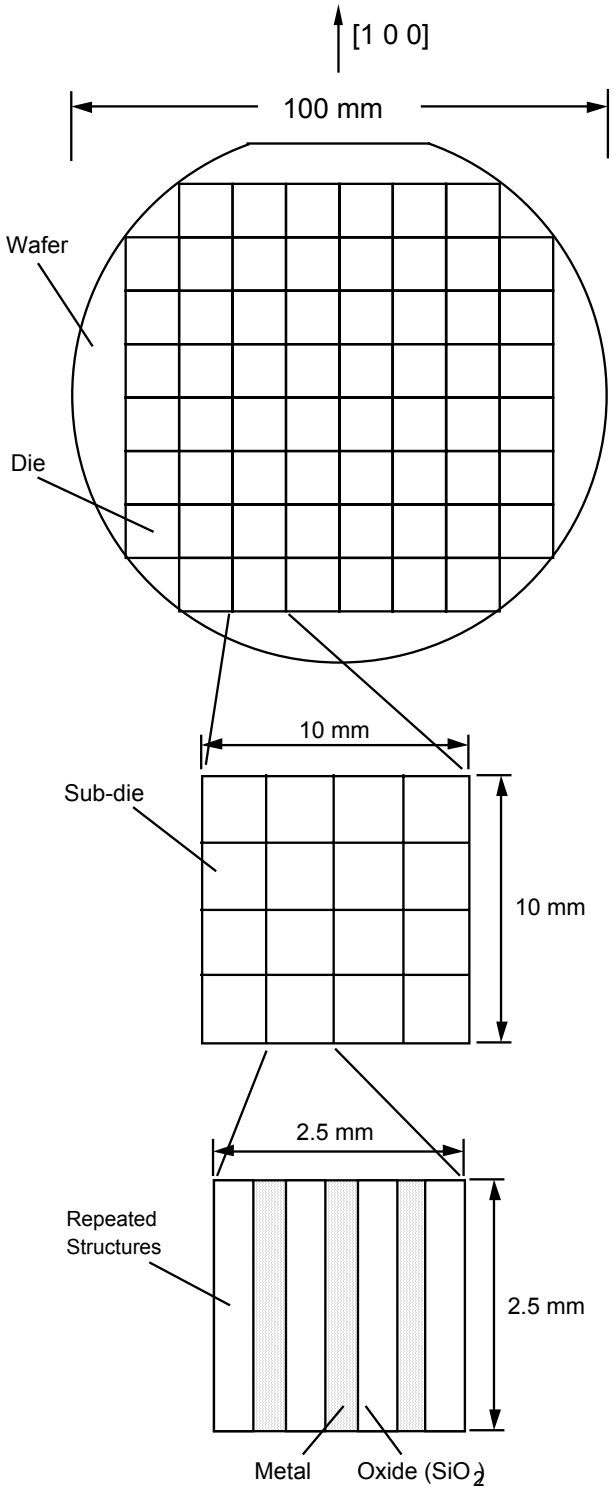
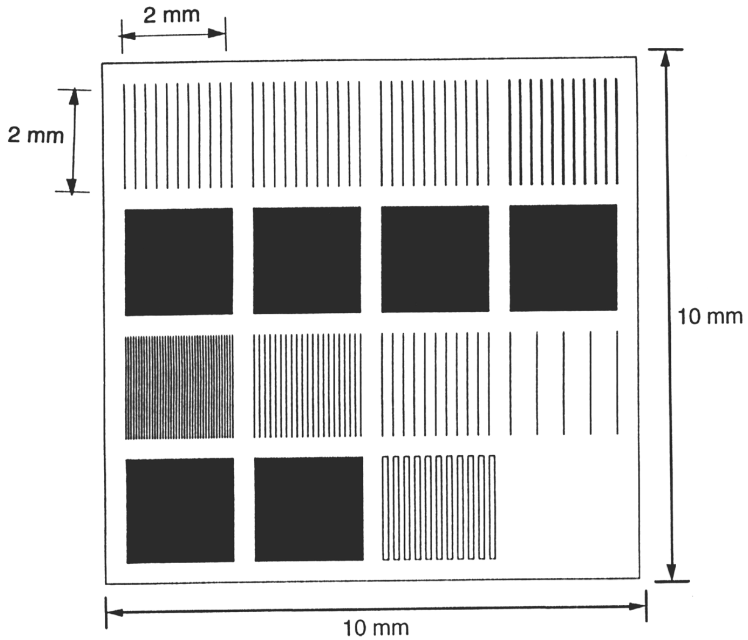


Figure 5.2 Schematic of the pattern layout on the test wafer.

Table 5.1: Linewidth (w), pitch (λ) and area fraction (A_f) of patterns on the test mask.

w (μm) \ λ (μm)	1	2	4	10	50	100	200	500
0.5	0.50	0.25	0.125	0.05	0.01		0.0025	
0.7							0.0035	
1.0						0.01		
2.0			0.50				0.01	
5.0							0.025	0.01
25.0					0.50		0.125	
100.0							0.50	

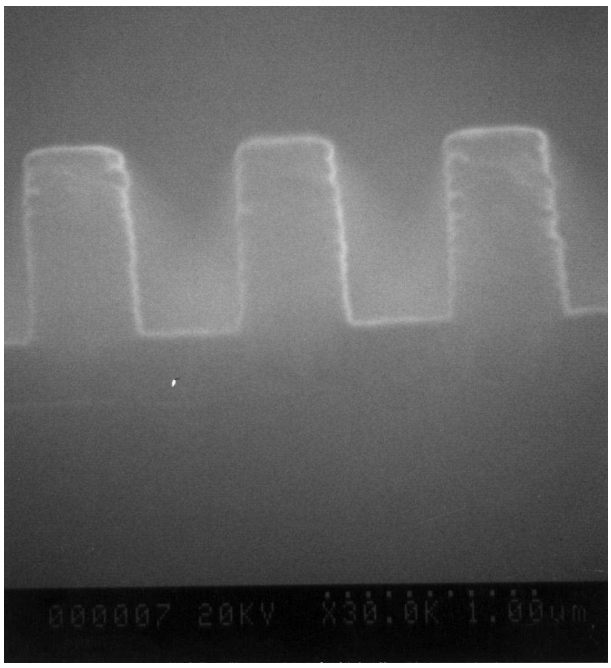


0.5/200 0.0025	0.7/200 0.0035	5/200 0.025	25/200 0.125
0.5/1 0.5	0.5/2 0.25	0.5/4 0.125	0.5/10 0.05
0.5/50 0.001	1/100 0.01	2/200 0.01	5/500 0.01
2/4 0.5	25/50 0.5	100/200 0.5	Field (No feature)

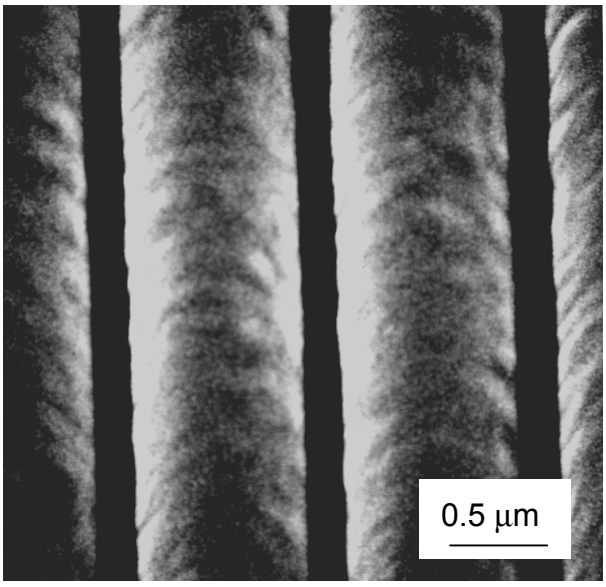
Linewidth (μm) / Pitch (μm)
Area Fraction

(a) (b)

Figure 5.3 Schematics of the CMP mask: (a) mask layout, and (b) pattern geometry layout.



(a)



(b)

Figure 5.4 SEM micrographs of the pattern (w = 0.5 µm and λ = 1 µm): (a) cross section of the patterned oxide ILD, and (b) surface topography after Cu deposition.

5.3.2 Experimental Conditions. Experiments were conducted on a rotary-type polisher. Table 5.2 lists the experimental conditions. The normal pressure and relative velocity were maintained at 48 kPa and 0.7 m/s respectively to ensure wafer/pad contact. The polishing duration was varied from one minute to six minutes to cover the under-polished, just-polished, and overpolished periods. The polishing slurry was composed of 4 vol.% of α-Al_2O_3 abrasives with average size 300 nm. In contrast to the acidic solutions used in commercial Cu CMP, the slurry pH was maintained at 7 to focus only on the mechanical aspects of polishing. The Rodel IC-1400 was used to polish the wafer. The pad was conditioned before polishing each wafer.

The profiles of the pattern surface at different polishing times were measured by stylus profilometry and by AFM for coarse and fine features, respectively. The Cu dishing was determined by measuring the amount of recess on the Cu lines relative to the oxide surface after the Cu coating on the oxide was cleared. The oxide overpolishing was determined by measuring the remaining oxide thickness. For coarse features, the oxide thickness was measured directly by ellipsometry. For fine features less than 20 μm wide a reference oxide thickness was measured by ellipsometry on the 400 μm wide oxide spacing between sub-die blocks. The thickness of the oxide features was determined by relating the surface profile inside the sub-die block to these reference spacings. All measurements were in the center of the sub-die block of the center die to minimize the effects of spatial variations due to wafer-scale polishing non-uniformity.

5.4 Results

5.4.1 Time Evolution of the Pattern. As shown in Fig. 5.5 under optical microscope, the patterned surface (w = 25 μm and λ = 50 μm) evolves with polishing time. Due to the high reflectance of Cu, the unpolished, scratch-free high features appear bright in the optical micrograph, Fig. 5.5 (a). The walls between the high and low surfaces appear dark in bright-field illumination because less normal incident light is reflected. In Fig. 5.5 (b), after one minute of polishing, the surface of high features was roughened. However, the surface of the low area did not change in microstructure, indicating that the pad did not contact the low

Table 5.2: Experimental conditions.

Experimental Parameters	Experimental Conditions
Diameter of Wafer (mm)	100
Normal Load (N)	391
Normal Pressure (kPa)	48
Rotational Speed (rpm)	75
Linear Velocity (m/s)	0.70
Duration (min)	1 – 6
Sliding Distance (m)	42 – 252
Slurry Flow Rate (ml/min)	150
Abrasive	α-Al_2O_3
Abrasive Size (nm)	300
pH	7

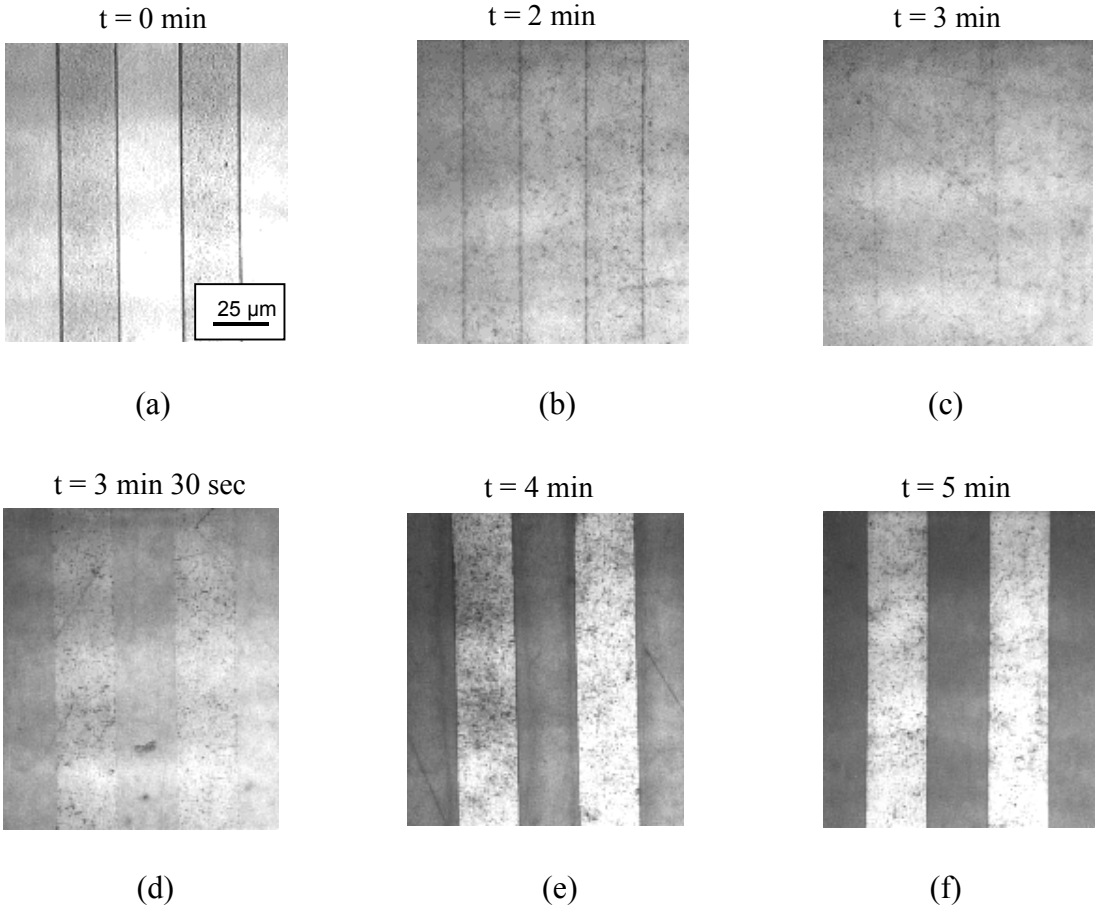


Figure 5.5 Optical micrographs of the evolution of pattern surfaces (w = 25 μm and λ = 25 μm).

area, as predicted by the contact mechanics models. After two minutes of polishing, as shown in Fig. 5.5 (c), the surfaces of both high and low features were roughened and the boundaries between the high and low features were less distinguishable. This is because the step-height decreased and the sharp edges of high and low features became rounded. Therefore, the pad was in contact with both the high and low features and both surfaces are polished. At three minutes, in Fig. 5.5 (d) the boundaries became indistinguishable, the step-height almost vanished, and the Cu surface was planarized. As shown in Fig. 5.5 (e), when the process almost hit the end-point at three and a half minutes, the less reflective barrier layer, Ta, started to appear. After thirty more seconds of polishing, the barrier layer was cleared and the underlying oxide exposed. The much darker oxide surfaces in Fig. 5.5 (f) indicate that the Ta layer has been polished off. The Cu lines are distinct because the reflectance of oxide is much less than that of Cu.

Figure 5.6 shows the evolution of the surface profile. At the beginning of polishing, the high features were removed faster than the low features, quickly smoothing the surface. The sharp corners were rounded in this period because of the pressure concentration at the edge. The high features may have reached a steady-state profile before the topography was planarized. The MRR in the planarization stage is about 500 nm/min for this feature. That is about twice of the blanket rate although the area fraction of the high features, 0.025, is very close to the blanket surface. One explanation is that the trenches on the surface improve the local slurry dispensing. The pressure on each subdie might be not uniform due to the die-level surface nonuniformity.

As the step-height between the high and low features decreased, the MRR on the high features approached that of the low features. This indicates that the pressure distributed more uniformly while the surface was smoothed out. Finally, both MRRs were close to the blanket Cu polishing rate, about 220 nm/min, and the surface was planarized. The Cu surface remained flat until the process end-point, which is consistent as shown in Fig. 5.5 (c). After passing the end-point, between three and four minutes, the Cu lines started dishing. The dishing increased with overpolishing time. The oxide was polished too, but at a rate much slower than for soft Cu. Therefore, the surface topography built up again.

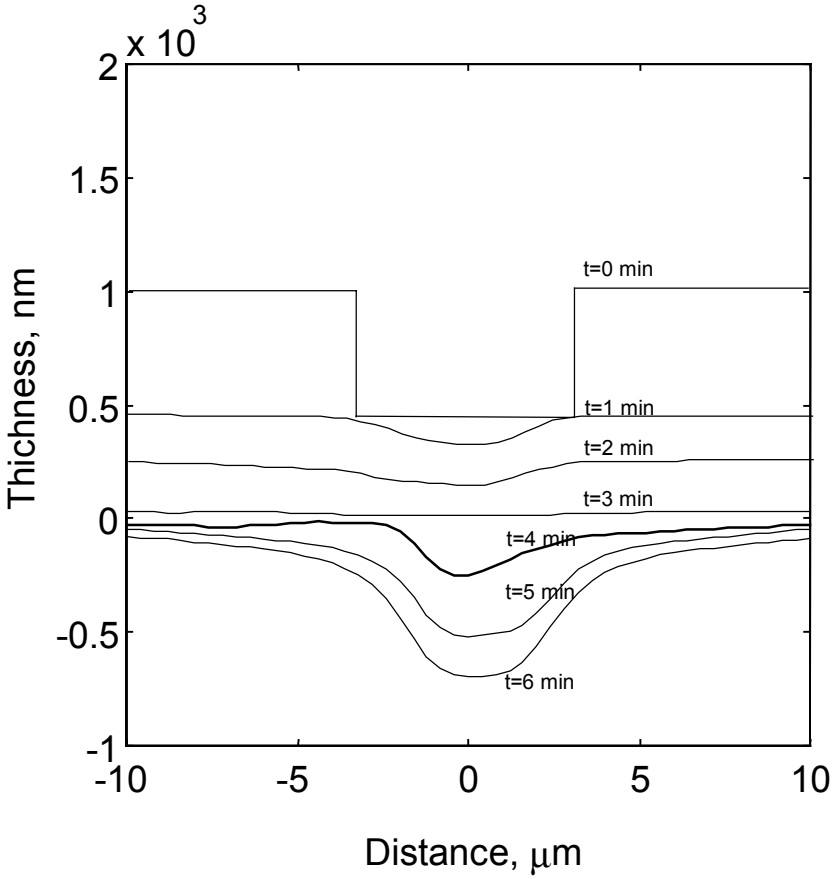
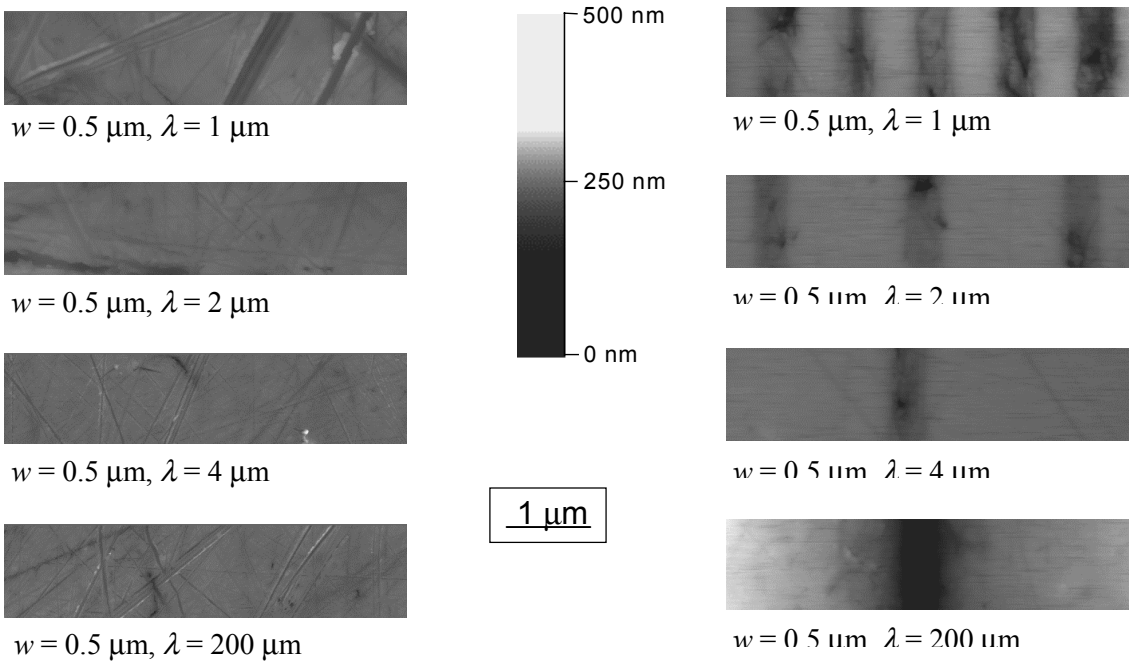


Figure 5.6 Cross-sectional profiles of the evolution of the pattern (w = 5 μm and λ = 200 μm).

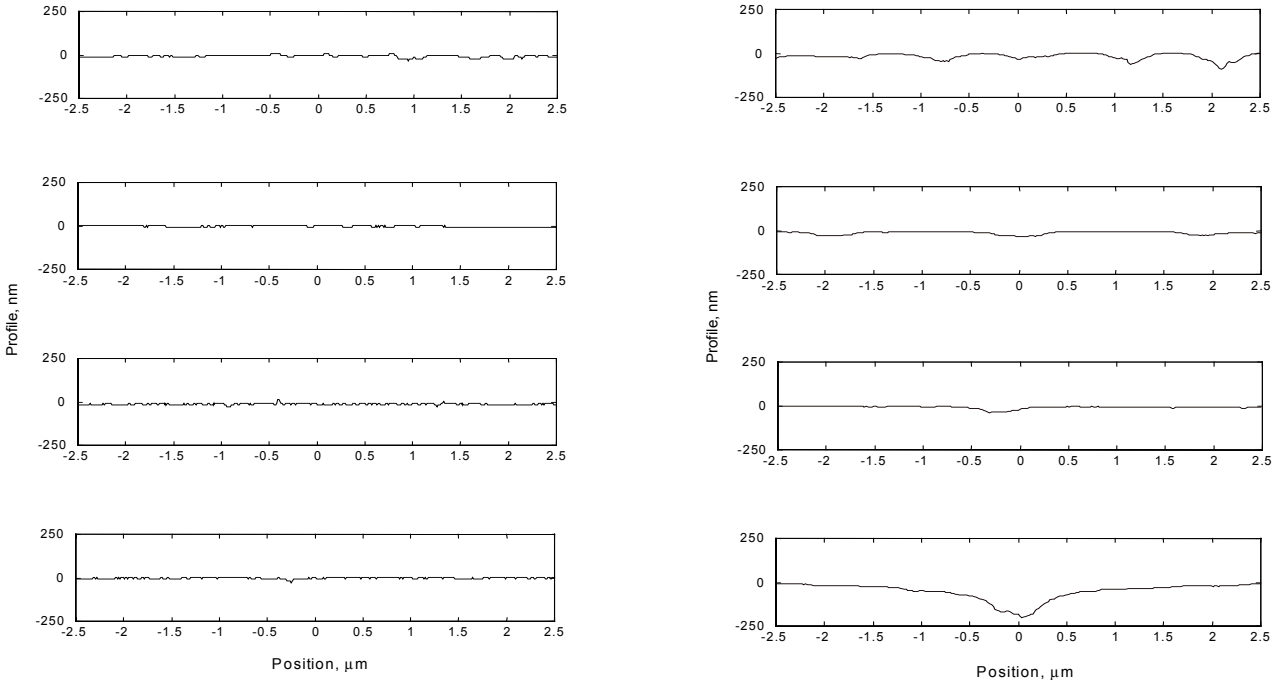
A similar trend of pattern evolution was observed on the smallest features, i.e. 0.5 μm Cu lines. Figure 5.7 shows the AFM micrographs and the cross-section plots of the features with 0.5 μm linewidth and different pitches (1, 2, 4 and 200 μm) at about process end-point (three minutes 30 seconds) and after overpolishing for one and a half minutes (total five minutes). All surfaces in the figure were planarized just before the end-point. Few shallow scratches due to particle abrasion of the soft Cu surfaces are evident. In the case of overpolishing, dishing occurs on the Cu lines, which appears dark in the AFM micrographs due to its low position relative to the surrounding oxide surface. For features with λ = 1 and 2 μm (or higher Cu area fraction, A_f= 0.5 and 0.25), the amount of dishing was less than 30 nm after overpolishing. In comparison, dishing is much significant, about 200 nm, for the isolated line-feature with 200 μm pitch. Significant rounding also occurs at the edges of oxide for the isolated line structures.

5.4.2 Copper Dishing. Table 5.3 lists the amount of dishing at different durations for structures on the center die of the test wafer. Before three min, when the surface was still covered with a thin layer of Cu, there was no dishing. The onset of dishing depends on the pattern geometry, characterized by the linewidth and the area fraction of Cu (or the pitch). From an earlier observation, dishing began when Cu was polished through. Because Cu was not cleared simultaneously for features with different linewidths, or area fractions, the onset of dishing varied with the same parameters. The time variation for the onset of dishing was about one minute for all patterns. In practice, this variation will require overpolishing on part of the wafer to clear all the Cu on the oxide surface. This creates surface nonuniformity. The table lists the rates of dishing resulted from the least square method for the data. The normalized rate of dishing, ranging from 0.04 to 1.39, is defined as the rate of dishing divided by the Cu blanket polishing rate, about 210 nm/min.

Figure 5.8 shows the effects of linewidth on dishing for 0.5 area fraction features, which is close to the present circuit design. For small-linewidth features, such as 0.5, 1, or even 25 μm lines, the amount of dishing levels off after a short period of overpolishing. The constant dishing level for 0.5 and 2 μm lines is 20 to 30 nm. Moreover, the rates of dishing are bounded by blanket Cu and oxide polishing rates as shown in Section 5.2. For 0.5 and 2 μm wide lines, the rate of dishing is close to the blanket oxide polishing rate, about 12 nm/min.



(a) (b)



(c) (d)

Figure 5.7: Time evolution of various patterns: AFM micrographs at (a) 3 minitues 30 seconds and (b) 5 minitues; surface profiles at (c) 3 minutes 30 seconds and (d) 5 minutes.

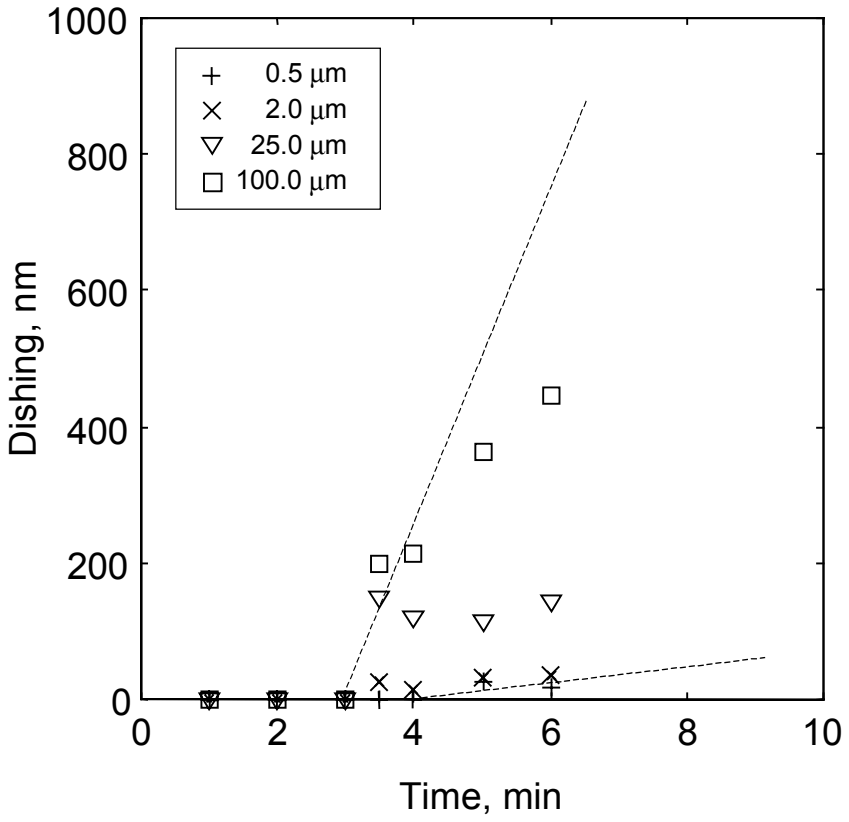


Figure 5.8 Time evolution of Cu dishing for patterns with constant area fraction 0.5 (w/λ = 0.5) and various linewidths.

For larger linewidths such as 100 μm, however, dishing increases with overpolishing time and did not reach a constant level (steady state) within the comparatively long overpolishing period in experiments. The dishing amount is about 450 nm after three minutes of overpolishing. Thus the dishing rate is about 150 nm/min, close to the polishing rate of blanket Cu, about 210 nm/min.

Dishing is as slow as the oxide removal rate for small features because the surrounding oxide constrains the polishing of fine Cu lines. As demonstrated in the contact mechanics analysis in Chapter 4, the pad cannot deform sufficiently into the small trenches. For instance, for a pattern with small linewidth and modest area fraction such as the 0.5 nm line with A_f 0.5, the pad displacement is about 0.08 nm and almost can be neglected compared to the pad roughness. This gives us a magnitude about the maximum depth that pad can indent into the dished Cu surface. The amount of dishing must be comparable to the sum of pad displacement and the indentation depth of the particle if no chemical reaction is involved. Based on experiment results, the indentation depth is about 10 to 20 nm for 300 nm Al_2O_3 abrasive at normal CMP conditions. Therefore, the maximum dishing is about 20 nm, which agrees with the measurements. After reaching steady state, the Cu will be polished at the same of oxide overpolishing predicted by Eq. (5.10).

In contrast, when the Cu line is wide enough, the pad easily conforms with the dished Cu surface. Pressure is uniformly applied on both the Cu and oxide surfaces as on the blanket wafer. For example 100 μm, the pad can deform into the dished area without the constraint of surrounding oxide. This will result in a large amount of Cu dishing, such as 300 - 400 nm, in these 100 μm lines (with the consideration of pad deformation, pad roughness and particle size). The ratio of dishing to linewidth is still very small, about 0.004. For such a small difference of strain between Cu and oxide contact regions, the normal pressure can almost be assumed uniform, i.e. $p_{Cu} \approx p_{Oxide} \approx \overline{p}$. Hence, the dishing rate will be close to the blanket Cu polishing rate, about 220 nm/min.

Figure 5.9 shows the effects of linewidth on the dishing behavior of isolated lines, with A_f = 0.01. The trend is similar to those on area fraction 0.5: dishing increases with overpolishing time. Its rate is bounded by blanket Cu and oxide polishing rates. The amount

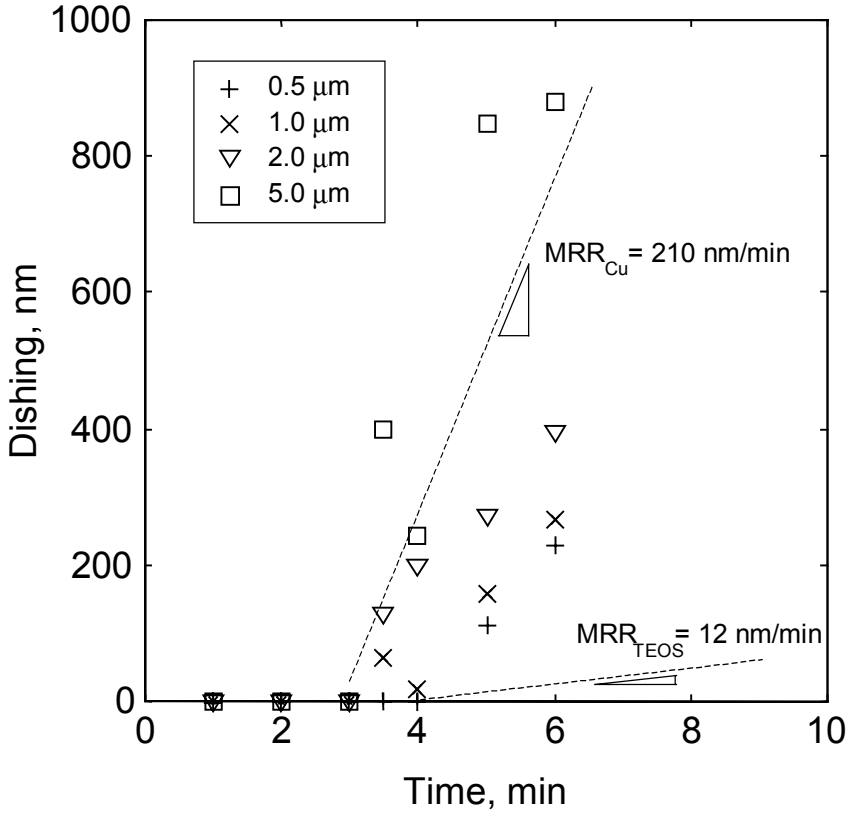


Figure 5.9 Time evolution of Cu dishing for patterns with constant area fraction ($w/\lambda = 0.01$) and various linewidths.

and the rate of dishing on isolated lines increase compared to those on 0.5 area fraction lines. The rate of dishing increases about 14 times for 0.5 and 2 μm features, less for features 5 μm or larger. This is relatively insignificant, considering the fifty-times decrease on the area fraction. Additionally, dishing on small Cu lines, such as 0.5 and 2 μm ones, does not reach a steady state with about two and a half minutes of overpolishing. Figure 5.10 shows the effects of area fraction on dishing for 0.5 μm lines. The results confirm that the area fraction does not significantly affect dishing. For area fraction ranging from 0.01 to 0.5, the rates of dishing are all close to the blanket oxide polishing rate. It is also shown that, except for the very low 0.01 area fraction, dishing will stay at low level, less than 35 nm, even with two-minute overpolishing.

Figure 5.11 compares the present work on features with 0.5 area fraction with the data from the literature (Park et al., 1999) with a commercial chemical slurry. The dishing behavior is not affected by the presence of chemicals in the slurry, for features less than 25 μm. In both experiments, dishing reaches the same steady-state level after overpolishing for about one minute. However, for a wide Cu area, for example 100 μm, the dishing rate is reduced by tailoring of the slurry pH and chemistry. The dishing amount is reduced by half, from 450 nm to 230 nm, after three minutes of overpolishing. Even for a one-minute period of overpolishing, dishing is reduced by a factor of 0.65 by using a chemical slurry. These results suggest that the effect of chemistry on dishing depends on the assistance of mechanical particle abrasion. The pure chemical etching is not very significant in the Cu polishing process. For small lines, the material removal due to particle abrasion decreases with the increase of dishing because the decrease of load on the particle. Thus the chemical effect of altering the hardness of surface material is not significant to the reduction of Cu dishing rate. On contrary, the pressure distribution is more uniform. It does not change much with the increase of dishing because the pad can conform with dished surfaces. Thus the change of surface properties by chemistry can change the rate of dishing, similar to the results observed on the blanket wafer. More discussion about process optimization by tailoring the slurry chemistry is given in Section 5.5.

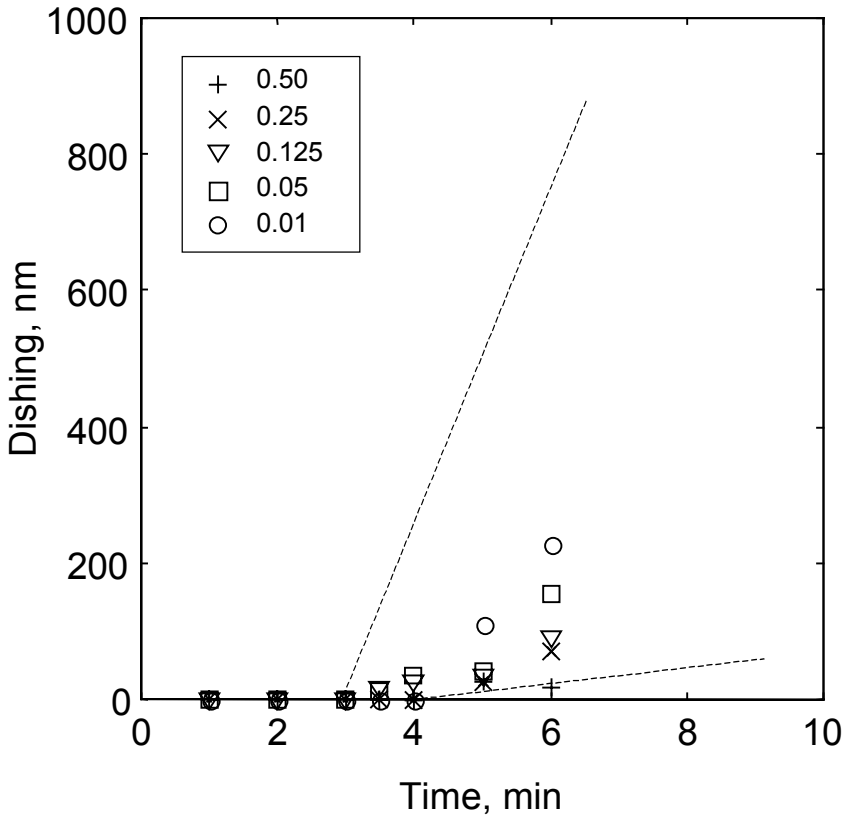


Figure 5.10 Time evolution of Cu dishing for patterns with constant linewidth (w = 0.5 μm) and various area fraction (w/λ).

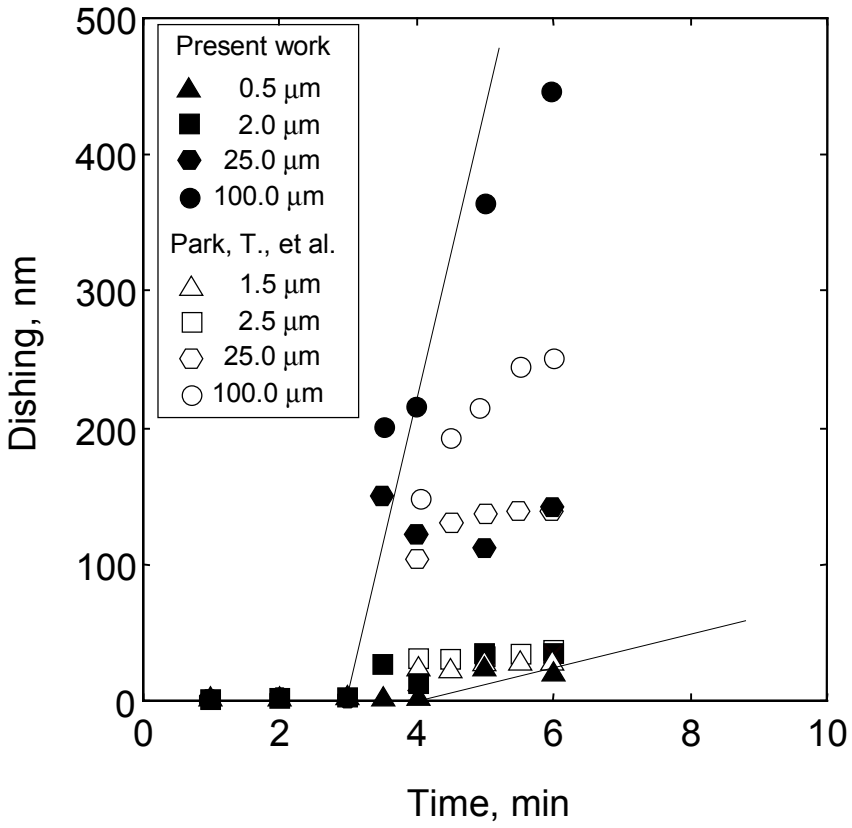


Figure 5.11 Comparison between the present work on dishing with neutral slurry and the results from literature (Park, et al., 1999) with chemical slurry.

5.4.3 Oxide Overpolishing. Oxide overpolishing starts at the onset when the Cu layer is polished through. Different patterns on the die take from three to four minutes of polishing. Figure 5.12 shows the amount of oxide overpolishing versus the polishing time for constant linewidth of 0.5 μm. The amount of overpolishing increases with overpolishing time. The overpolishing rate increases with pattern area fraction. For small area fraction, such as 0.01, 0.05, and 0.125, the rate of overpolishing is similar to the rate of blanket oxide polishing. For area with larger area fraction, such as 0.25 and 0.5, the rate of overpolishing increases with area fraction. Figure 5.13 shows the comparisons between analytical and experimental results for the effect of area fraction on the rate of overpolishing. The solid line represents the analytical results of Eq. (5.10) with blanket polishing of Cu and oxide at 270 and 26 nm/min, respectively. The experimental results agree with the model well, especially when the area fraction is less than 0.25. For higher area fraction like 0.5, the rate is higher than that predicted by the model. A possible explanation for this discrepancy is that the slurry transfers more efficiently at the interface on a dense pattern area than that on a blanket area or a less dense area. The surface features may help dispense slurry locally. The dished Cu recesses will improve the intrinsic rates of material removal of Cu and oxide and thus increase the rate of overpolishing.

Figures 5.14 and 5.15 show oxide overpolishing of various patterns with different linewidths and constant area fractions of 0.5 and 0.01, respectively. Overpolishing does not strongly depend on the linewidth for either small or large area fraction. In the case of area fraction 0.5, the rate of overpolishing is about 100 nm/min, for linewidths ranging from 0.5 to 100 nm. For small area fraction of 0.01, which mimics the area with isolated interconnects on the surface, the oxide overpolishing rate is very close to the blanket rate of oxide polishing for linewidths ranging from 0.5 to 5 nm. This implies that scaling does not significantly change the pressure distribution on both Cu and oxide. The pressure distribution on the surface during the overpolishing stage essentially is affected by the area fraction only. The average material removal rate across a sub-die area is constrained by the oxide overpolishing rate, which depends on the area fraction of pattern. The pressure distribution on each sub-die area will be similar and close to the average pressure applied on the wafer, which verifies the assumption employed in Section 5.2.

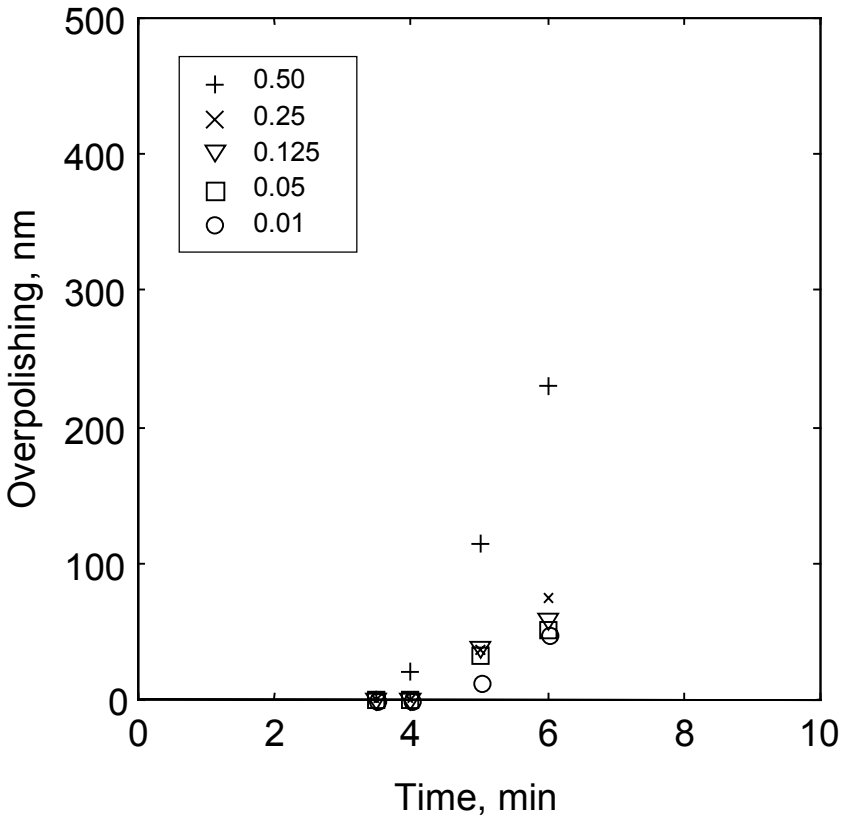


Figure 5.12 Oxide overpolishing for patterns with constant linewidth (w = 0.5 μm) and various area fraction (w/λ).

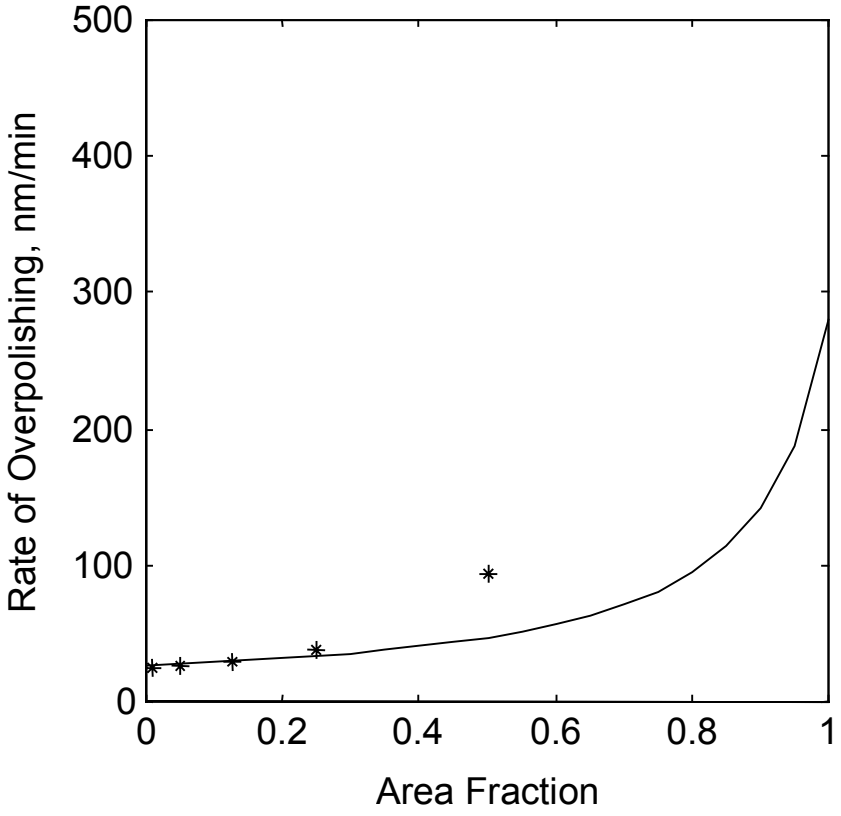


Figure 5.13 Comparison between the theoretical and experimental results for rate of oxide overpolishing for various pattern with constant linewidth 0.5 μm and various area fraction.

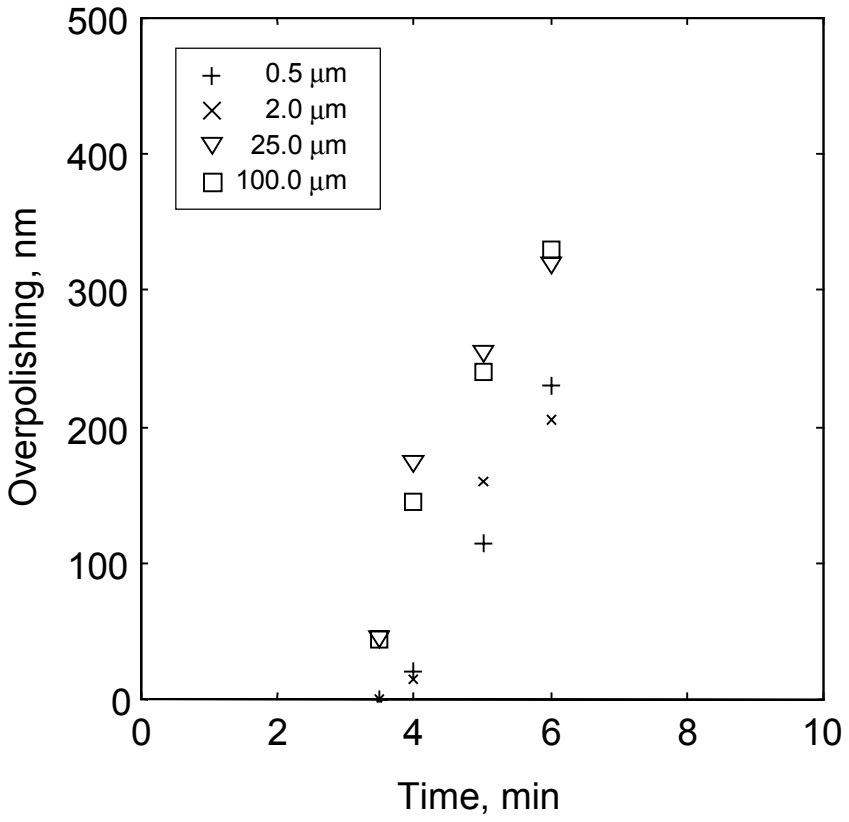


Figure 5.14 Time evolution of Cu dishing for patterns with constant area fraction 0.5 (w/λ = 0.5) and various linewidths.

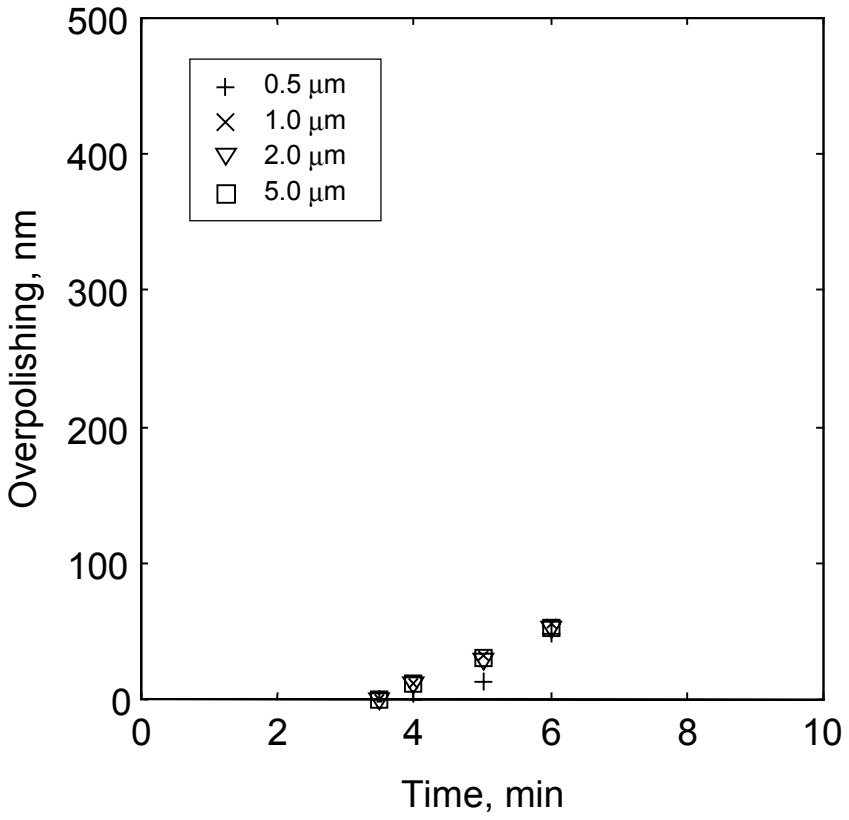


Figure 5.15 Oxide overpolishing for patterns with constant area fraction ($w/\lambda = 0.01$) and various linewidths.

Figures 5.14 and 5.15 show the oxide overpolishing rate to be smaller at the early stage of overpolishing before the steady-state regime. This may be explained by the variation of pressure distribution on both the Cu and oxide surface during overpolishing. Consider the case in which the surface is planar when the Cu is just cleared. The pressure is distributed uniformly on both Cu and oxide surfaces. Because the Cu dishes and the surface nonuniformity increases, the pressure on the Cu will decrease because the pad might slightly deformed into the dished area to relax stresses on the contacting surface. Based on force equilibrium, the load on the oxide might increase and thus the MRR increase until it reaches a steady value. The MRR of oxide is bounded by the steady-state rate given by Eq. (5.10) and by the blanket rate (at the stage at which the surface is planar).

5.5 Discussion: Process Optimization

The effects of Cu pattern geometry on the MRR in the planarization stage and on dishing and oxide overpolishing has been studied earlier. The prior analytical and experimental results not only help understand the fundamental mechanisms of patterned wafer polishing, but also provide an opportunity to improve the process outcomes. In Cu CMP, there are two important process requirements within a die area: remaining Cu interconnect thickness and its within-die uniformity (which also represents the variation of the surface topography). The remaining Cu thickness at any point (at a randomly chosen point k in the j-th subdie region of the i-th die of the wafer) can be expressed as:

$$h_{ijk} = h_o - (\mu_i + \xi_{j(i)} + \delta_{j(i)} + r_{k(ij)}) \tag{5.12}$$

where h_o is the initial designed thickness of the Cu interconnect which is the same as the depth of the oxide trench, μ_i the mean of oxide overpolishing on a specific die i, $\xi_{j(i)}$ the deviation of the amount of oxide overpolishing from μ_i on the subdie area j (with the same pattern geometry) on the die i. Therefore, the amount of Cu loss due to overpolishing is the sum of μ_i and $\xi_{j(i)}$. Also in Eq. (5.12), $\delta_{j(i)}$ is the amount of dishing on the subdie area j on the die i, and $r_{\mathrm{k(ij)}}$ the random error at a specific point k in the subdie area j on the die i. The random error for each observation in the subdie area is estimated by randomly choosing n

replicants of Cu interconnect thickness. If a specific subdie with repeat features is large enough, i.e., the different pattern of neighboring sub-die will not affect the pressure distribution and slurry flow in most of the subdie area, the random error represents the error form measurement and other random factors. The intention of employing Eq. (5.12) is to just help identify the effects of each geometry or process parameters for process optimization.

Each variable on the right-hand-side of Eq. (5.12) must be minimized, both mean and variance, to minimize the Cu loss. The mean of oxide overpolishing, μ_i, is affected by the average Cu area fraction and increases with overpolishing time. Its variance across a wafer increases with the increase of within-wafer polishing non-uniformity, which is determined by the global (wafer-scale) factors such as wafer/pad contact conditions, slurry dispensing, and pad stiffness (as discussed in Chapter 4). In practice, the average area fraction is limited to 0.3 to 0.5 and does not vary too much for similar IC products. Thus minimization of μ_i relies mostly on the reduction of within-wafer polishing non-uniformity so that the overpolishing time required to remove the excess Cu at different dies can be minimized. Detailed schemes of reducing within-wafer nonuniformity can be found in Section 2.4.4.

Equation (5.12) suggests that the rate of overpolishing, $\partial\xi/\partial t$, due to the local pattern layout in the subdie area is determined by the wear coefficient, Cu area fraction, and the hardness of both Cu and oxide. The arrangement of the subdie area fraction is usually prescribed by the circuit designers and cannot be changed. To minimize the effects of pattern local layout on overpolishing, $\partial\xi/\partial t$ must be adjusted to be as low as possible and/or less sensitive to the local geometry variation in the final polishing stage (or after the onset of overpolishing). The overpolishing rate decreases with the wear coefficient. One efficient way is to employ soft abrasive particles, in which the hardness of the abrasive is close to ILD oxide but still higher than that of Cu. Less overpolishing will occur even when the same overpolishing time needed to clean up excess Cu. Another method is to increase the hardness of Cu or reduce the oxide hardness (essentially reducing the ratio of Cu MRR to oxide MRR, or the so-called selectivity) by tailoring the slurry pH and chemistry. This will reduce the sensitivity to area fraction variation on the variation of overpolishing rate (or the variation of the "relative hardness" across different subdies). Reducing the oxide hardness by increasing slurry pH (but no too high to retard the Cu removal) is better than increasing the Cu hardness

because the overpolishing time will not increase. However, an end-point detection scheme must be adopted because the oxide is overpolished at a much faster rate.

Dishing is strongly related to the Cu linewidth. For sub-micron lines, the rate of dishing is very low (close to oxide blanket rate and insensitive to the slurry chemistry). The steady–state dishing is very small. The effects of dishing on the Cu loss and surface non-uniformity may be negligible for current and future circuit designs. However, for some designs with large metal pad or wide power transmission lines, 50–100 μm wide, the dishing rate is close to that in blanket polishing. In these cases, dishing results in Cu loss and surface non-uniformity. It might be necessary to increase the Young's modulus of the pad to reduce the pad indentation or to employ a slightly basic slurry to retard Cu polishing rate without increasing the oxide overpolishing rate.

5.6 Conclusions

Both analytical and experimental studies on Cu dishing and oxide overpolishing were presented in this chapter. The following conclusions can be drawn:

(1) The steady-state overpolishing and dishing were modeled. The MRR in a subdie area (with same pattern geometry) is related to the "apparent hardness" of that area. Both area fraction and the material hardnesses (Cu and oxide) will affect the polishing uniformity across different pattern regions in the die. The die-scale surface nonuniformity and the variance of remaining Cu thickness will increase with overpolishing time before reaching the steady state.

(2) Experiments were conducted on patterned Cu wafers. The pattern, with minimum dimension 0.5 μm, was designed to study the effects of linewidth, area fraction and scaling effect. The results agree with trends shown by contact mechanics modeling. The initial topography is planarized quickly and the time variation for different pattern (A_f ranging from 0.01 to 0.5) to reach planar surface is about 1 minute. After the surface has been planarized, the remaining Cu is removed at a rate close to blanket polishing rate. The surface variation will remain until part of the Cu is polished through in some subdie areas.

(3) After the Cu is cleared, the surface nonuniformity increases because of dishing and overpolishing. Experiments show that linewidth is an important geometrical parameter for dishing. For thin lines, less than 1 μm, the dishing rate is close to oxide blanket rate and might reach a steady-state profile after a short period of overpolishing. For wider lines, about 50 to 100 μm, the Cu is dished at a rate close to the blanket rate. Compared with the results with chemical slurry in the literature, the slurry pH and chemicals do not increase the amount or rate of dishing for small lines but might retard the dishing of wider lines. This implies that the load distribution due to the deformation of the pad and mechanical action of the particles play an important role in Cu dishing, especially for small lines.

(4) Compared to dishing, oxide overpolishing depends more on pattern area fraction than on linewidth. Overpolishing reaches a steady-state rate after a short period. The steady-state rate of overpolishing depends on the apparent hardness and the intrinsic wear coefficients of Cu and oxide. Experiments show that for a pattern with large fraction, the overpolishing rate may increase over that predicted by the model due to the improvement of slurry transport. Moreover, overpolishing does not depend on linewidth significantly. When the device scale shrinks down, the within-die nonuniformity will cause by the overpolishing but not dishing if a large variation area fraction is shown on pattern layout.

(5) The objectives of process optimization are to maximize Cu removal rate and to reduce surface nonunifomity due to dishing and overpolishing. The key is to reduce oxide overpolishing and to minimize the variance of dishing and overpolishing resulting from the effects of different area fraction and linewidth. The surface topography will not be uneven even with a short period of overpolishing. A SiO_2 abrasive or other particles with hardness close to the ILD silicon oxide should reduce the oxide polishing rate and increase the polishing selectivity between Cu and oxide. For patterns with wide Cu lines, dishing rate can be decreased by a stiff pad and by a slightly basic slurry.

Nomenclature

A_f = area fraction of metal pattern
H = hardness of coating material (N/m^2)
H' = apparent hardness of a composite surface (N/m^2)
h = thickness of the material removed on wafer surface (m)
h_o = initial coating thickness (m)
k_p = Preston constant (m^2/N)
k_w = wear coefficient
p_{av} = nominal pressure on wafer (N/m^2)
$\overline{p}$ = average pressure on a pattern (N/m^2)
r = random error in thickness measurement (m)
t = experiment duration (s)
t^* = overpolishing duration (s)
v_R = relative linear velocity of wafer (m/s)
w = pattern linewidth (m)
x, y, z = Cartesian coordinates (m)
Δh = oxide overpolishing (m)
δ = Cu dishing (m)
λ = pattern pitch (m)
μ = average overpolishing on a die
ϕ = dimensionless geometrical function
ν = Poisson's ratio
ξ = deviation of overpolishing on the specific pattern from average overpolishing on a die (m)

References

Carpio, R., Farkas, J., and Jairath, R., 1995, "Initial Study on Copper CMP Slurry Chemistries," *Thin Solid Films*, Vol. 266, pp. 238-244.

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

Fayolle, M. and Romagna, F., 1997, "Copper CMP Evaluation: Planarization Issues," *Microelectronic Eng*., Vol. 37/38, pp. 135-141.

Gutmann, R.J., Steigerwald, J., You, L., Price, D.T., Neirynck, J., Duquette, D.J. and Muraka, S.P., 1995, "Chemical-Mechanical Polishing of Copper with Oxide and Polymer Interlevel Dielectrics," *Thin Solid Films*, Vol. 270, pp. 596-600.

Hariharaputhiran, M., Zhang, J., Ramarajan, S., Keleher, J.J., Li, Y., and Babu, S.V., 2000, "Hydroxyl Radical Formation in H_2O_2-Amino Acid Mixtures and Chemical Mechanical Polishing of Copper," *J. Electrochem. Soc.*, Vol. 147, pp. 3820-3826.

Kaufman, F.B., Thompson, D.B., Broadie, R.E., Jaso, M.A., Guthrie, W.L., Pearson, D.J., and Small, M.B., 1991, "Chemical-Mechanical Polishing for Fabricating Patterned W Metal Features as Chip Interconnects," *J. Electrochem. Soc.*, Vol. 138, pp. 3460-3464.

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

Murarka, S.P., Steigerwald, J. and Gutmann, R.J., 1993, "Inlaid Copper Multilevel Interconnections Using Plnanarization by Chemical-Mechanical Polishing," *MRS Bulletin*, pp. 46-51.

Ouma, D., Stine, B., Divecha, R., Boning, D., Chung, J., Shinn, G., Ali, I. and Clark, J., 1997, "Wafer-Scale Modeling of Pattern Effect in Oxide Chemical Mechanical Polishing," *Proc*. *SPIE Microelectronics Mfg. Conf*., pp. 236-247.

Pan, J.T., Li, P., Kapila, W., Tsai, S., Redeker, F., Park, T., Tugbawa, T., Boning, D., 1999, "Copper CMP and Process Control," *Proc. 1999 CMP-MIC Conf.*, pp. 423-429.

Park, T., Tugbawa, T., Boning, D., Chung, J., Hymes, S., Muralidhar, R., Wilks, B., Smekalin, K., Bersuker, G., 1999, "Electrical Characterization of Copper Chemical Mechanical Polishing," *Proc. 1999 CMP-MIC Conf.*, pp. 184-191.

Kondo, S., Sakuma, N., Homma, Y., Goto, Y., Ohashi, N., Yamaguchi, H., and Owada, N., 2000, "Abrasive-Free Polishing for Copper Damascene Interconnection," *J. Electrochem. Soc.*, Vol. 147, pp. 3907-3913.

Smith, T.H., Fang, S.J., Boning, D., Shinn, G.B. and Stefani, J.A., 1999, "A CMP Model Combining Density and Time Dependencies," *Proc. 1999 CMP-MIC Conf.*, pp. 93-104.

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

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

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

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

Tugbawa, T., Park, T., Boning, D., Pan, T., Hymes, S., Brown, T., and Camilletti, L., 1999, "A Mathematical Model of Pattern Dependencies in Copper CMP Processes," *The 3rd Int. Symposium on Chemical Mechanical Polishing in IC Device Manufacturing, Electrochem. Soc. Meeting*, Honolulu, HA.

van Kranenburg, H. and Woerlee, P.H., 1998, "Influence of Overpolish Time on the Performance of W Damascene Technology," *J. Electrochem. Soc.*, Vol. 145, pp. 1285-1291.

Zeidler, D., Stavreva, Z., Plötner, M., Drescher, K, 1997, "Characterization of Cu Chemical Mechanical Polishing by Electrochemical Investigations," *Microelectronic Eng*., Vol. 33, pp. 259-265.