Journal of The Electrochemical Society, **149** (1) G41-G50 (2002)
0013-4651/2001/149(1)/G41/10/$7.00 © The Electrochemical Society, Inc.

Evolution of Copper-Oxide Damascene Structures in Chemical Mechanical Polishing

II. Copper Dishing and Oxide Erosion

Jiun-Yu Lai, Nannaji Saka,[z] and Jung-Hoon Chun

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

Test wafers comprising damascene structures were designed and fabricated to investigate Cu dishing and oxide erosion. The mask design covered a wide range of linewidths and pitches, from 0.5 to 100 μm, to represent such features as signal and power transmission lines, and probing or wire-bonding pads. Experiments were conducted to investigate the evolution of the pattern profile during polishing and to determine the onset and rates of dishing and erosion. The effects of Cu linewidth and area fraction on the rates of pattern planarization, Cu dishing, and oxide erosion have been quantified. The effect of hardness of the composite surface on dishing and erosion were examined. An optimization scheme, employing particle size, particle hardness, and pad stiffness, to enhance the selectivity between SiO_2, Ta, and Cu, and to reduce die-scale nonplanarity is proposed.
 [DOI: 10.1149/1.1420708]

Manuscript submitted November 14, 2000; revised manuscript received August 30, 2001. Available electronically November 29, 2001.

As shown in the authors' previous paper,[1] the local (die-scale) pattern geometry affects the local material removal rate significantly. The nonuniform pressure distribution resulting from the nonuniform area fraction and layout of the pattern introduces surface nonplanarity in the planarization stage of metal polishing. Consequently, in order to remove all the metal coating on the dielectric surface, so that the metal interconnects are isolated, the pattern is slightly overpolished. This results in dielectric thinning or so-called erosion. Concurrently, dishing occurs on the soft metal filled in the trenches and reduces the cross-sectional area of the interconnect. Both erosion and dishing result in surface nonplanarity and thickness variation of metal interconnects across a die area.[1-4]

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

$$\frac{dh}{dt} = k_p(x,\ y)p_{av}\phi(w,\ A_f,\ t^*,\ \ldots)v_R \qquad [1]$$

The Preston constant, k_p, is a function of position which relates to the physical layout of the oxide and Cu interconnects. It is assumed that 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,[5] and the slurry chemistry.[6-14] The pressure distribution is affected by the actual shape of the dished/overpolished surfaces, which in terms is 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 erosion is comparable to the surface roughness of the pad and the slurry particle size. Therefore, the contact mechanics model presented in the authors' previous paper, assuming a flat and homogeneous pad and neglecting the effects of particle, is not valid any more. However, with some simplified assumptions, a qualitative model for the rates of dishing and erosion at steady-state can be achieved.

Moreover, 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. An analytical model of this sort, however, is difficult to establish. 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.[15-21] Though a few semiquantitative models have been proposed,[22,23] the fundamentals of dishing and erosion 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 subquarter micrometer circuit design. The scaling issue must be addressed based on the similarity in fundamental material removal behaviors on different size features.

In this paper, the steady-state Cu dishing and oxide erosion are modeled based on the pressure distribution and the local-scale wear phenomenon. Dishing is also studied by considering the effects of pattern geometry, pad displacement, and particle size. Experiments are conducted to quantitatively establish the effects of Cu linewidth and area fraction on the rates of pattern planarization, Cu dishing, and oxide erosion. On this basis, the fundamentals of dishing and erosion phenomena and their mechanisms are examined and the important process parameters are identified. The results also are correlated with the contact mechanics model in the previous paper[1] to determine the effects of pressure distribution due to the remaining surface topography on dishing and erosion. Schemes for optimal Cu chemical mechanical polishing (CMP) to mitigate dishing and erosion, employing the effects of particle size, particle hardness, slurry pH, are discussed.

Theory

Referring to the wear equation, the Preston constant can be defined as the ratio of the wear coefficient k_w to the hardness H of material being polished.[5] Thus the intrinsic material removal rate at any point on the wafer surface can be determined by the Preston equation, which may be rewritten as

$$\frac{dh}{dt} = \frac{k_w}{H}pv_R \qquad [2]$$

where p is the local average pressure applied at the vicinity of the interested point on the wafer surface. As shown in the authors' earlier investigation, the wear coefficient depends on the polishing mechanism and is insensitive to the material polished.[5] It was found that k_w remains approximately a constant for various surface coatings, including Cu, tetraethylorthosilicate (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. Then the material removal rate on Cu and the oxide surface, as shown in Fig. 1, may be expressed as

[z] E-mail: nsaka@mit.edu

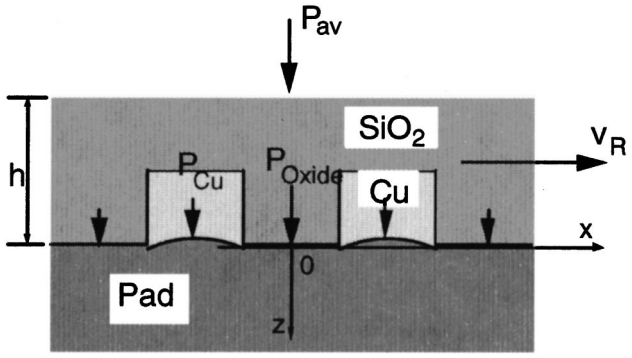


Figure 1. Schematic of the onsets of dishing and erosion.

$$\left(\frac{dh}{dt}\right)_{Cu} = \frac{k_w}{H_{Cu}} p_{Cu} v_R \quad [3]$$

$$\left(\frac{dh}{dt}\right)_{Oxide} = \frac{k_w}{H_{Oxide}} p_{Oxide} v_R \quad [4]$$

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

$$\left(\frac{dh}{dt}\right)_{Cu} = \left(\frac{dh}{dt}\right)_{Oxide} \quad [5]$$

As shown in the experimental results in a later section, the steady-state dishing profile usually occurs after clearing Ta and after a short period of oxide polishing. Therefore, only the rate balance between Cu and oxide is considered in Eq. 5. By equating Eq. 3 and 4, and 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

$$\frac{p_{Cu}}{p_{Oxide}} = \frac{H_{Cu}}{H_{Oxide}} \quad [6]$$

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

$$p_{Cu} w + p_{Oxide}(\lambda - w) = \overline{p}\lambda \quad [7]$$

where $\overline{p}$ is the average pressure on the specific area. In terms of area fraction A_f, $A_f = w/\lambda$, Eq. 7 can be rewritten as

$$p_{Cu} A_f + p_{Oxide}(1 - A_f) = \overline{p} \approx p_{av} \quad [8]$$

Since 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 $\overline{p}$ can be assumed approximately equal to the average pressure on the die p_{av}. Solving Eq. 8 by the relation given in Eq. 6, the pressure on the oxide surface at steady-state is given as

$$p_{Oxide} = \frac{\overline{p}}{[(H_{Cu}/H_{Oxide})A_f + (1 - A_f)]} \approx \frac{p_{av}}{[(H_{Cu}/H_{Oxide})A_f + (1 - A_f)]} \quad [9]$$

The pressure on the oxide surface is a function of the average pressure on the die, the hardness of materials present on the surface, and the area fraction of the pattern. If the pressure in Eq. 4 is replaced by Eq. 9, the Preston equation on the oxide surface can be rewritten as

$$\left(\frac{dh}{dt}\right)_{Oxide} = \frac{k_w}{H_{Oxide}} \frac{\overline{p}}{[(H_{Cu}/H_{Oxide})A_f + (1 - A_f)]} v_R = \frac{k_w}{H'} \overline{p} v_R \approx \frac{k_w}{H'} p_{av} v_R \quad [10]$$

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

$$H' \equiv H_{Cu} A_f + H_{Oxide}(1 - A_f) \quad [11]$$

Equation 10 can be interpreted as 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 the 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 designed to be as uniform as possible to reduce overpolishing. More details about process optimization are discussed in the later section. Additionally, the rate of oxide erosion is bound 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 right at the end point) and the steady-state rate of the surrounding oxide (which is very close to the blanket oxide polishing rate except the case with extreme high area fraction).

Experimental

Mask design.—A Cu damascene structure has been designed to study the effects of geometry on metal dishing and oxide erosion. As shown in Fig. 2, the pattern on each die (10×10 mm) consists of a matrix of 2×2 mm blocks (subdie area). These blocks in turn consist of line-space features, with a minimum linewidth of 0.5 μm. Table I lists the design features of the pattern, and Fig. 3 shows the physical layout of the pattern on the mask. The first type of feature is composed of fine Cu lines of constant linewidth, 0.5 μm, and of various pitches ranging from 1 to 200 μm. These represent the metal interconnects with critical dimensions and different packing densities. The second type of feature was designed to study the effect of linewidth on dishing. Various Cu lines, from 0.5 to 100 μm, with large pitch, 200 μm, were chosen to provide wide spacing between adjacent Cu interconnects. For small Cu lines, the wide spacing helps study dishing mechanisms alone. Two constant Cu area fractions, 0.01 and 0.5, with various linewidths and pitches were designed on the third type of features to study the effects of scaling on dishing and erosion. The 0.5 area fraction is close to the present design rules of metal layer layout in ultralarge scale integrated (ULSI) circuits. By contrast, the features with 0.01 area fraction represent single, isolated lines.

The pattern was transferred onto the 1.5 μm thick SiO_2 (TEOS) coating by lithography 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 physical vapor deposited (PVD) Cu film. The scanning electron micrograph (SEM) of the cross section of the patterned wafer is shown in Fig. 4.

Experimental conditions.—Experiments were conducted on a rotary-type polisher and the experimental conditions are listed in Table II. The normal pressure and relative velocity were maintained at 48 kPa and 0.7 m/s, respectively, to assure wafer/pad interface in contact. The polishing duration was varied from 1 to 6 min to cover

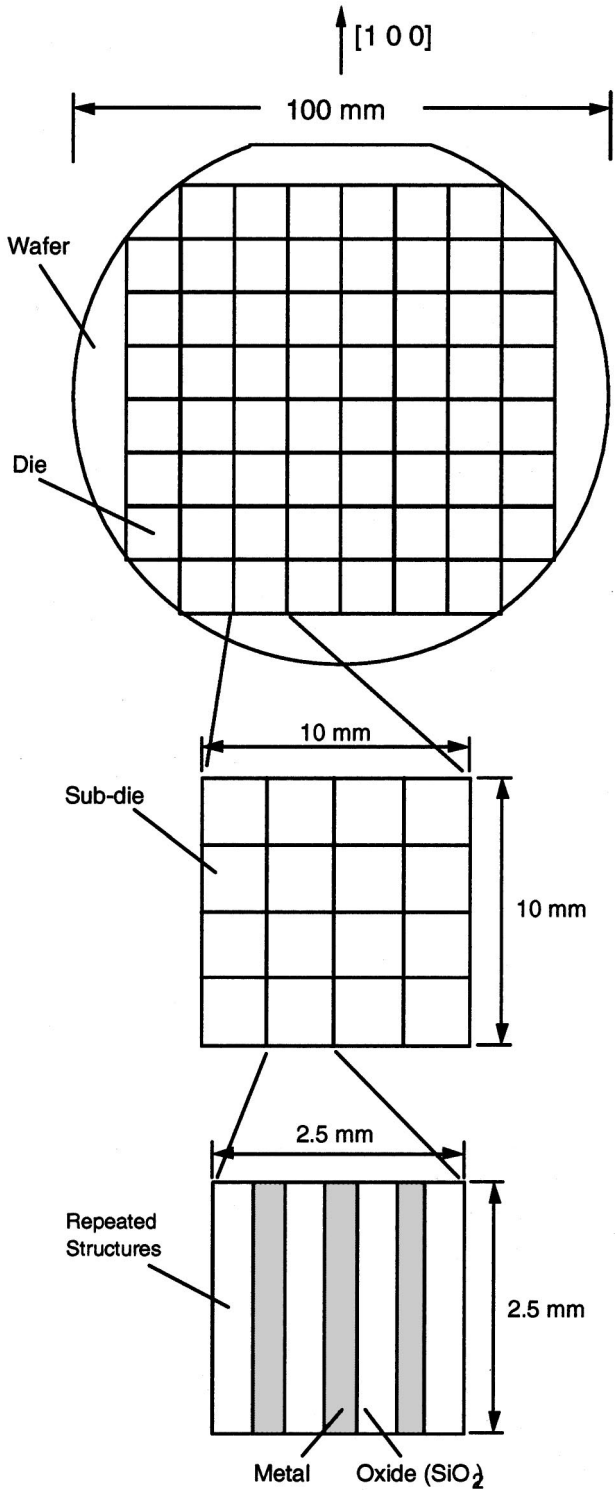


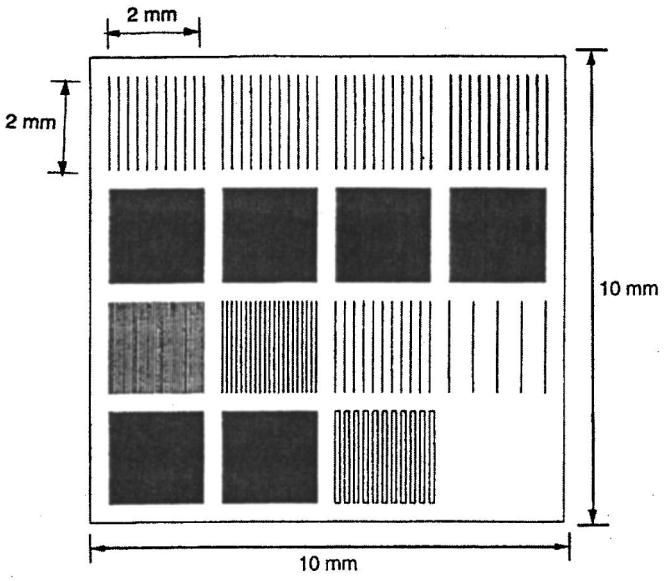
Figure 2. Schematic of the pattern layout on the test wafer.

the underpolished, just-polished, and overpolished periods. The polishing slurry was composed of 4 vol % of α-Al_2O_3 abrasives with average size 300 nm. Glycol and other additives were used in the slurry to prevent particle agglomeration. The slurry was mixed immediately before use and was stirred to prevent particle settling during polishing. 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 and the pad was conditioned before polishing each wafer.

Table I. 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						0.50		0.125	
100								0.50	

The profiles of the pattern surface at different polishing times were measured by stylus profilometry and by atomic force micros-



(a)

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

(b)

Figure 3. Schematics of the CMP mask: (a, top) mask layout, and (b, bottom) pattern geometry layout.

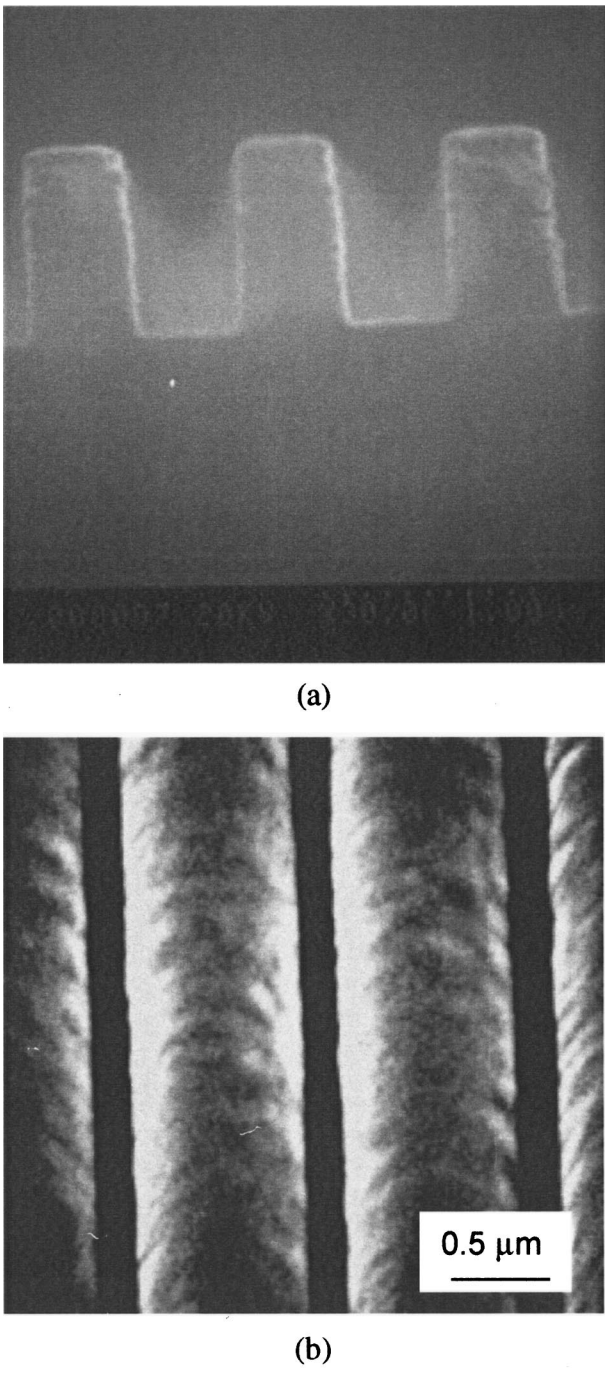


Figure 4. SEM micrographs of the pattern (w = 0.5 μm and λ = 1 μm): (a, top) cross section of the patterned oxide ILD, and (b, bottom) surface topography after Cu deposition.

copy (AFM) for coarse and fine features, respectively. By these data, 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 subdie blocks. The thickness of the oxide features was determined by relating the surface profile inside the subdie block to these reference spacing. All the measurements were carried out at the center of the subdie block of the center die to minimize the effects of spatial variations due to wafer-scale polishing nonuniformity.

Table II. 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	45-252
Slurry flow rate, mL/min	150
Abrasive	α-Al_2O_3
Abrasive size, nm	300
pH	7

Results

Time evolution of the pattern.—As shown in Fig. 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. 5a. The walls between the high and low surfaces appear dark in bright-field illumination since less normal incident light is reflected. In Fig. 5b, after 1 min of polishing, the surface of high features was roughened. However, the surface of the low area still remained the same microstructure from the Cu deposition, which indicates that the pad did not contact the low area as predicted by the contact mechanics models. After 2 min of polishing, as shown in Fig. 5c, the surfaces of both high and low features were roughened and the boundaries between the high and low features became less distinguishable. These occurred because the step-height decreased and the sharp edges of high and low features were rounded. Therefore, the pad was in contact with both the high and low features and both surfaces are polished. At 3 min, in Fig. 5d the boundaries became indistinguishable, the step-height almost vanished, and the Cu surface was planarized. As shown in Fig. 5e, when the process almost hit the end point at 3.5 min, the less reflective barrier layer, Ta, started to appear. After 30 s more of polishing, the barrier layer was cleared and the underlying oxide exposed. The

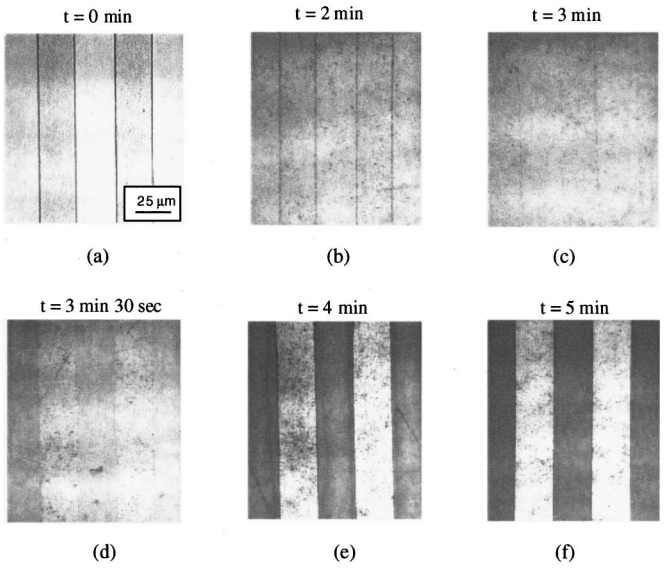


Figure 5. Optical micrographs of the evolution of pattern surface (w = 25 μm and λ = 25 μm).

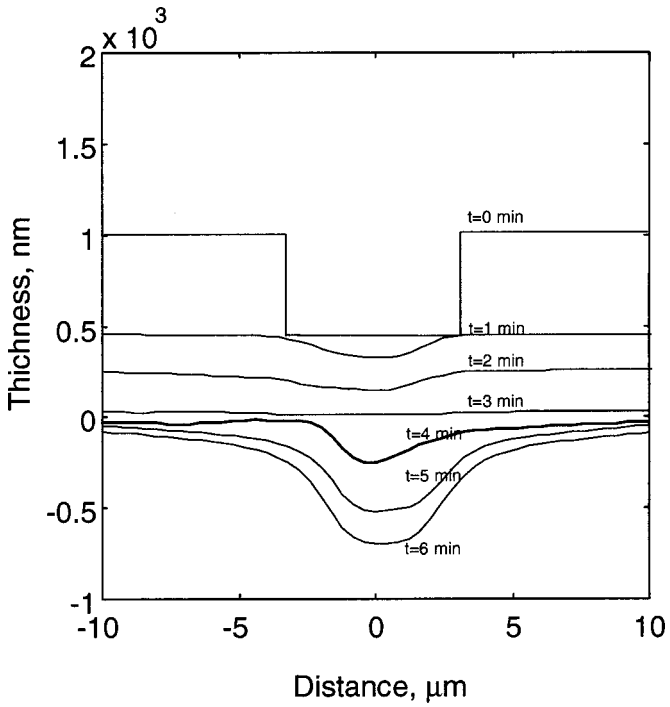


Figure 6. Cross-sectional profiles of the evolution of the pattern (w = 5 μm and λ = 200 μm).

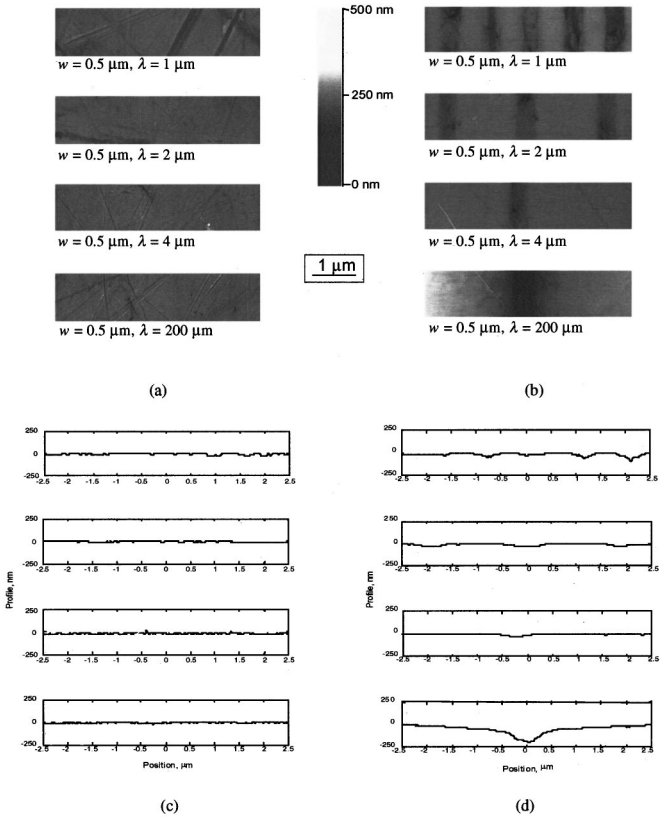


Figure 7. Time evolution of various patterns: AFM micrographs at (a) 3 min, 30 s, and (b) 5 min, surface profiles at (c) 3 min, 30 s, and (d) 5 min.

much darker oxide surfaces in Fig. 5f 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.

The evolution of the surface profile is shown in Fig. 6, in which the pattern with 5 μm linewidth and 200 μm pitch is employed as an example. At the beginning of polishing, the high features were removed faster than the low features, which quickly smoothed the surface. Also, the sharp corners were rounded in this period because of the pressure concentration at the edge. It is hard to say if the high features had reached a steady-state profile before the topography was planarized. The material removal rate in the planarization stage is about 500 nm/min for this feature and about twice that of the blanket rate although the area fraction of the high features, 0.025, is very close to blanket surface. One explanation for this is that the trenches on the surface improve the local slurry dispensing. Moreover, the pressure on each subdie might not be uniform due to the die level surface nonplanarity.

As the step-height between the high and low features decreased, the material removal rate 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 material removal rates were close to the blanket Cu polishing rate, about 220 nm/min, and the surface was planarized. Then the Cu surface remained flat until the process end point was reached, which is consistent as shown in Fig. 5c. After passing the end point, which was between 3 and 4 min, the Cu lines dished and the amount of dishing increased with polishing time. It is not clear if Cu started dishing during or after Ta clearing because the Ta is removed in a very short period in this case. The oxide was polished, too, but at a rate much slower than that of soft Cu. Therefore, the surface topography built up again.

A similar trend of pattern evolution was observed on the smallest features, *i.e.*, 0.5 μm Cu lines. Figure 7 shows the AFM micrographs and the cross-sectional plots of the features with 0.5 μm linewidth and different pitches (1, 2, 4, and 200 μm) at about the process end point (3 min 30 s) and after overpolishing for 1.5 min (total time, 5 min). 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 overpoishing, 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 a 200 μm pitch. Significant rounding also occurs at the edges of the oxide for the isolated line structures.

Copper dishing.—Table III lists the amount of dishing at different durations for structures on the center die of the test wafer. Before 3 min, when the surface was still covered with a thin layer of Cu, dishing did not initiate. 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 initiated at the time when Cu was polished through. Since Cu was not cleared at the same time 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 1 min for all patterns. In practice, this variation will require overpolishing part of the wafer in order to clear all the Cu on the oxide surface, and therefore creates surface nonplanarity. The rates of dishing resulted from the least square method for the data are listed in the table. 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.

The effects of linewidth on dishing for 0.5 area fraction features, which is close to the present circuit design, are shown in Fig. 8. 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 are about 20 to 30 nm. This verifies the assumption in the theory section that dishing on narrow lines reaches a steady state with continued further polishing. Moreover, the rates of dishing are bounded by blanket Cu and

Table III. Experimental results of dishing evolution.

Pitch, λ (μm)	Linewidth, w (μm)	w/λ	Amount of dishing at different duration (nm)					Rate of dishing (nm/min)	Normalized rate of dishing
			3 min	3.5 min	4 min	5 min	6 min		
1	0.5	0.50	0	0	0	26.6	18.5	9.3	0.04
2	0.5	0.25	0	0	0	27.5	71.3	35.7	0.17
4	0.5	0.125	0	16.4	27.6	36.4	92.1	27.7	0.13
	2.0	0.50	0	26.7	13.6	30.8	34.2	9.4	0.04
10	0.5	0.05	0	12.3	34.5	42.0	154.0	46.7	0.22
50	0.5	0.01	0	0	0	112.4	228.2	114.1	0.54
	25.0	0.50	0	150.0	120.0	113.0	143.5	28.8	0.14
100	1.0	0.01	0	65.3	16.5	158.0	267.1	87.5	0.42
200	0.5	0.0025	0	0	0	215.7	284.5	142.5	0.68
	0.7	0.0035	0	0	0	254.7	359.5	179.8	0.86
	2.0	0.01	0	127.6	200.0	273.4	395.7	121.0	0.58
	5.0	0.025	0	225.0	123.3	546.3	786.5	259.1	1.23
	25.0	0.125	0	183.5	400.0	628.3	705.0	236.5	1.13
	100.0	0.5	0	200.0	215.0	364.5	445.5	135.9	0.65
500	5.0	0.01	0	243.8	398.0	847.5	878.5	292.3	1.39

oxide polishing rates as shown in the same section. For 0.5 and 2 μm wide lines, the rate of dishing before reaching the steady state is close to the blanket oxide polishing rate, about 12 nm/min. 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 3 min of overpolishing. Thus dishing rate is about 150 nm/min and is close to the polishing rate of blanket Cu, about 210 nm/min.

Dishing is as slow as the oxide removal rate for small features since the surrounding oxide constrains the polishing of fine Cu lines. As demonstrated in the contact mechanics analysis in the authors' earlier paper,[1] the pad cannot deform sufficiently into small trenches. For instance, for a pattern with a small linewidth and modest area fraction such as the 0.5 nm line with A_f, 0.5, the pad displacement into low is about 0.08 nm and can almost be neglected compared to the pad roughness. This gives us a magnitude of the maximum depth that the pad can indent into the dished Cu surface.

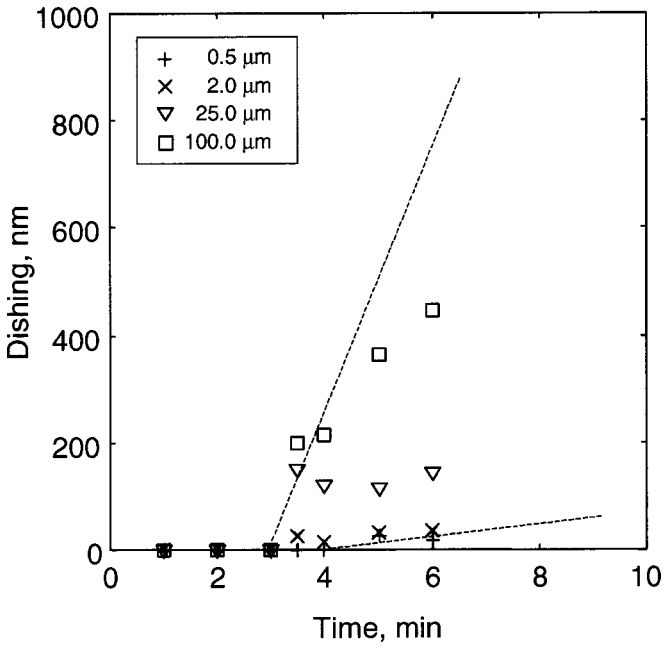


Figure 8. Time evolution of Cu dishing for patterns with constant area fraction 0.5 ($w/\lambda = 0.5$) and various linewidths.

It is expected that 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 experimental 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 expected to be about 20 nm, which agrees with the measurements. After the steady state, the Cu will be polished at the same rate of oxide erosion predicted by Eq. 10. More details about oxide erosion are given later.

In contrast, when the Cu line is wide enough the pad easily conforms with the dished Cu surface and applies uniform pressure on both 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 a surrounding oxide. Consider when an even larger amount of dishing occurs, such as 300-400 nm in 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 to be 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. Moreover, for a design with a Cu depth of 1 μm, more than 40% of the Cu is lost due to dishing in the worst case.

The effects of linewidth on the dishing behavior of isolated lines, with $A_f = 0.01$ is shown in Fig. 9. The trend is similar to those on an area fraction of 0.5: dishing increases with overpolishing time and its rate is bound by blanket Cu and oxide polishing rates. The amount and the rate of dishing on isolated lines increased compared to those on 0.5 area fraction lines. The rate of dishing increases about 14 times for both 0.5 and 2 μm features, and smaller for 5 μm or larger features. However, it is not so significant when the fifty-times decrease on the area fraction is considered. Additionally, dishing on small Cu lines, such as 0.5 and 2 μm ones, does not reach a steady state with about 2.5 min of overpolishing. Figure 10 shows the effects of area fraction on dishing for 0.5 μm lines. It is confirmed again that the area fraction does not affect dishing significantly. For an 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 one with the very low 0.01 area fraction, dishing will stay at a low level, less than 35 nm, even with 2 min of overpolishing.

In Fig. 11, the present work on features with 0.5 area fraction is compared with the data from the literature with a commercial chemical slurry.[20] Similar Cu dishing levels for the same sizes of lines at the end point and after a short period of overpolishing were reported by other researchers.[9,10,18,19,21] It is apparent that the dishing behavior is not affected by the presence of chemicals in the

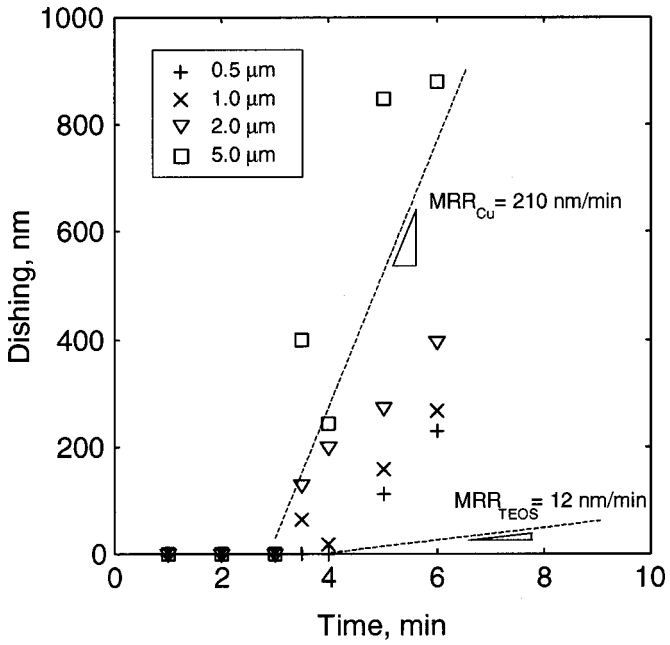


Figure 9. Time evolution of Cu dishing for patterns with constant area fraction ($w/\lambda = 0.01$) and various linewidths.

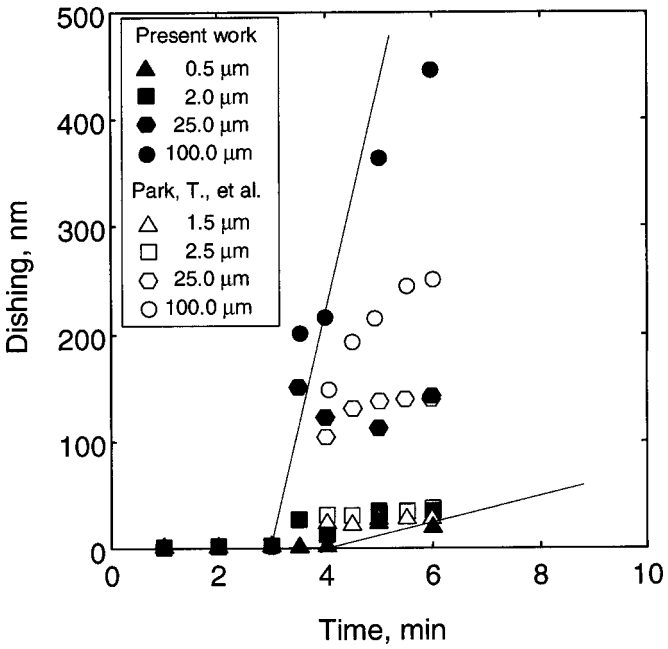


Figure 11. Comparison between the present work on dishing with neutral slurry and the results from the literature[20] with chemical slurry.

slurry, less than 25 μm. In both experiments, dishing reaches the same steady-state level after overpolishing for about 1 min. However, for a wide Cu area, for example, 100 μm, the dishing rate is reduced by proper tailoring the slurry pH and chemistry. The dishing amount is reduced by half, from 450 to 230 nm, after 3 min of overpolishing. Even for a short period of overpolishing, 1 min, 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 of the decrease of load on the particle. Thus the chemical effect of altering the hardness of the surface material is not significant to the reduction of the Cu dishing rate. On the contrary, the pressure distribution is a lot more uniform and 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. Further discussion about process optimization by tailoring the slurry chemistry is given in a later section.

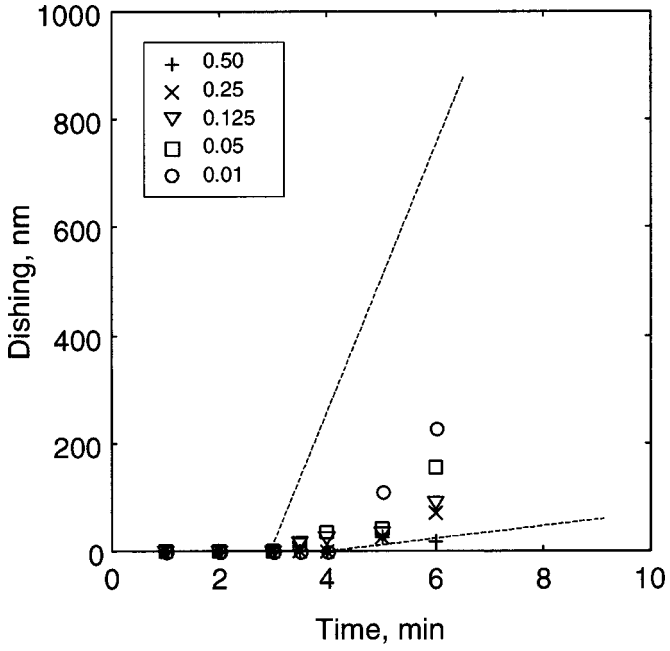


Figure 10. Time evolution of Cu dishing for patterns with constant linewidth ($w = 0.5$ μm) and various area fractions (w/λ).

Oxide erosion.—Oxide erosion starts at the onset when the Cu layer is polished through, which depends on the geometry of pattern. In this case, it varies from 3 to 4 min of polishing for different patterns on the die. Figure 12 shows the amount of oxide erosion *vs.* the polishing time for constant linewidth of 0.5 μm. The amount of erosion increases with overpolishing time. It shows that the rate of oxide erosion increases with pattern area fraction. For a small area fraction, such as 0.01, 0.05, and 0.125, the rates of erosion are similar and close to the rate of blanket oxide polishing. For an area with a larger area fraction, such as 0.25 and 0.5, the rates of erosion increase with area fraction. Figure 13 shows the comparisons between analytical and experimental results for the effect of area fraction on the rate of erosion. The solid line represents the analytical results of Eq. 10 with blanket polishing of Cu and oxide at 270 and 26 nm/min, respectively. It is shown that the experimental results agree with the model well, especially when the area fraction is less than 0.25. For a 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 on a blanket area or a less dense area. The dished Cu recesses will improve the intrinsic rates of material removal of Cu and oxide and thus increase the rate of erosion.

Figures 14 and 15 show oxide erosion of various patterns with different linewidths and constant area fractions of 0.5 and 0.01, respectively. It is shown that the linewidth has less effect than the area fraction on the rate of oxide erosion. In the case of an area fraction of 0.5, the rate of erosion is about 100 nm/min for linewidths ranging from 0.5 to 100 nm. For a small area fraction of 0.01, which mimics the area with isolated interconnects on the surface, the rate of oxide erosion is very close to the blanket rate of oxide polishing,

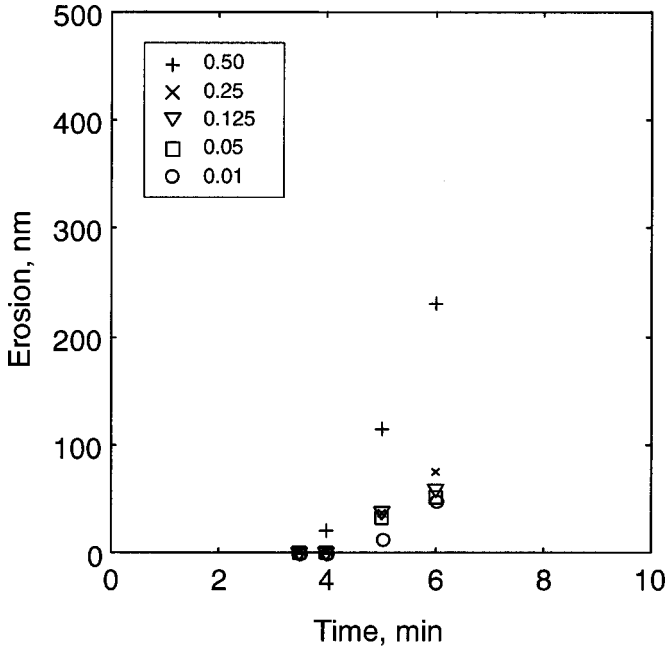


Figure 12. Oxide erosion for patterns with constant linewidth (w = 0.5 μm) and various area fractions (w/λ).

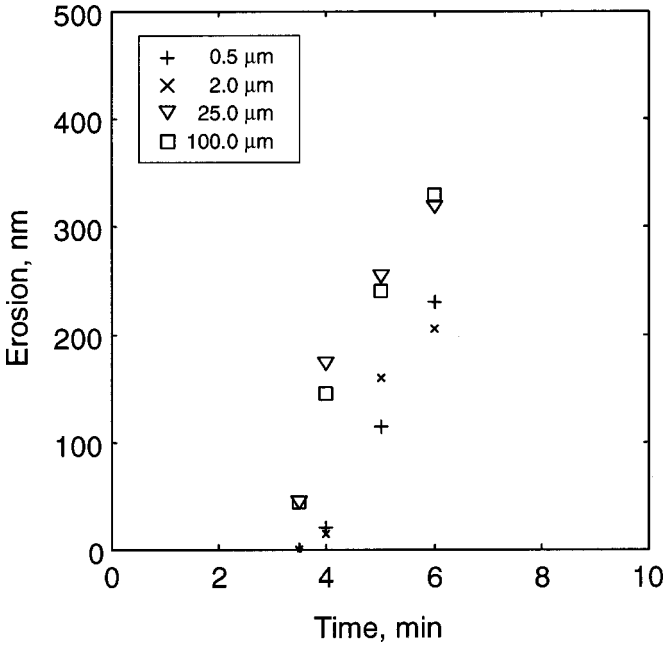


Figure 14. Time evolution of oxide erosion for patterns with constant area fraction 0.5 (w/λ = 0.5) and various linewidths.

for linewidths ranging from 0.5 to 5 nm. This implies that scaling does not change the pressure distribution on both Cu and oxide very significantly since only mechanical polishing is involved in this case. The pressure distribution on the surface during the overpolishing stage essentially is affected by the area fraction only. This is because the average material removal rate across a subdie area is constrained by the rate of oxide erosion, which in turn depends on the area fraction of pattern.

Moreover, in Fig. 14 and 15, the oxide erosion rate is 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 that the surface is planar when the Cu is just cleared. The pressure is distributed uniformly on both Cu and oxide surface. Since the Cu dishes and the surface nonplanarity increases, the pressure on the Cu will decrease because the pad might be slightly deformed into the dished area and relax stresses on the contacting surface. Based on force equilibrium, the load on the oxide might increase and thus the material removal rate (MRR) increase until it reaches a steady value. The MRR of oxide is bound by the steady-state rate given by Eq. 10 and by the blanket rate (at the stage when the surface is planar).

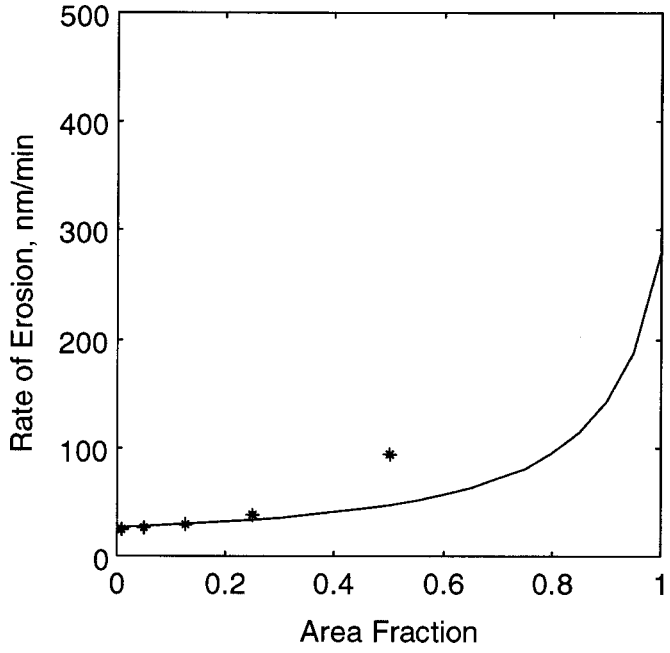


Figure 13. Comparison between the theoretical and experimental results for rate of oxide erosion for various patterns with constant linewidth 0.5 μm and various area fractions.

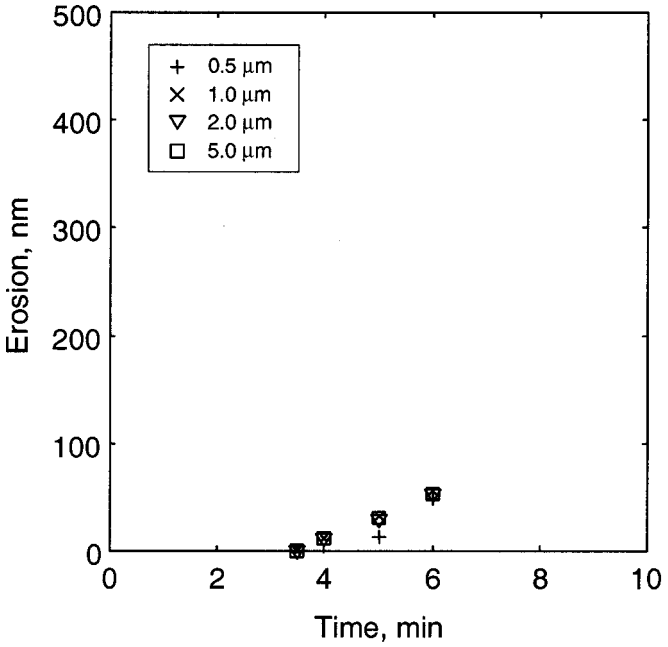


Figure 15. Oxide erosion for patterns with constant area fraction (w/λ = 0.01) and various linewidths.

Discussion: Process Optimization

The effects of Cu pattern geometry on the material removal rate in the planarization stage and on dishing and oxide erosion has been studied earlier. The prior analytical and experimental results not only help in understanding the fundamental mechanisms of patterned wafer polishing, but also provide an opportunity for improving the process outcomes. In Cu CMP, there are two important process requirements within a die area: remaining Cu interconnect thickness and its within-die nonplanarity (which also represents the variation of the surface topography). The remaining Cu thickness at any point (at a randomly chosen point k in the jth subdie region of the ith die of the wafer) can be expressed as

$$h_{ijk} = h_o - (\mu_i + \xi_{j(i)} + \delta_{j(i)} + r_{k(ij)}) \quad [12]$$

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 erosion (usually referred to the global oxide loss and can be measured on a wide oxide area) on a specific die i, $\xi_{j(i)}$ the deviation of the amount of oxide erosion 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 oxide erosion is the sum of μ_i and $\xi_{j(i)}$. Also in Eq. 12, $\delta_{j(i)}$ is the amount of dishing on the subdie area j on the die i, and $r_{k(ij)}$ 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 larger enough, *i.e.*, the different pattern of neighboring subdie 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. 12 is to just help identify the effects of each geometry or process parameters for process optimization.

In order to minimize the Cu loss, each variable on the right side of Eq. 12 must be minimized, both mean and variance. The mean of oxide overpolishing, μ_i, is affected by the average Cu area fraction and increases with overpolishing time. Also its variance across a wafer increases with the increase of within-wafer polishing nonuniformity, which is determined by the global (wafer-scale) factors such as wafer/pad contact conditions, slurry dispensing, and pad stiffness (as discussed in a previous paper). 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 nonuniformity 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 the authors' previous investigation.[5]

As suggested by Eq. 12, the rate of $\partial\xi/\partial t$, due to the local pattern layout and the variation of pattern density 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. In order to minimize the effects of pattern local layout on erosion, $\partial\xi/\partial t$ 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 dishing and erosion). The rate of oxide polishing decreases with the wear coefficient. It might be possible to optimize the process by employing soft abrasive particles, in which the hardness of the abrasive is close to interlevel dielectric (ILD) oxide but still higher than those of Ta and Cu. Hence less erosion will occur even when the overpolishing time for cleaning up all the excess Cu still remains the same. A slight overpolishing is necessary in practice to remove a thin layer of oxide to ensure metal clearing. However, an endpoint detection scheme must be adopted to control overpolishing.

It is shown that dishing is strongly related to the Cu linewidth. For submicrometer lines, the rate of dishing is very low (close to the oxide blanket rate and is not sensitive to the slurry chemistry) and steady-state dishing is very small. The effects of dishing on the Cu loss and surface nonuniformity might be negligible for current and future circuit designs. However, for some designs with a large metal pad or wide power transmission lines, 50-100 μm wide, it is shown that dishing rate is close to that in blanket polishing. In these cases, dishing results in Cu loss and surface nonplanarity. This suggests that Cu dishing might be reduced by increasing the Young's modulus of the pad to reduce the pad indentation, or by employing a slightly basic slurry to retard the Cu polishing rate but without increasing the rate of oxide erosion.

Conclusions

Both analytical and experimental studies on Cu dishing and oxide erosion, are presented in this paper and the following conclusions can be drawn.

1. The steady-state Cu dishing and oxide erosion are modeled and the material removal rate 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 nonplanarity and the variance of the 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 patterns (A_f ranging from 0.01 to 0.5) to reach planar surface is about 1 min. After the surface has been planarized, the remaining Cu is removed at a rate close to the blanket polishing rate. Also 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 nonplanarity increases because of dishing and erosion. Experiments show that linewidth is an important geometrical parameter for dishing. For thin lines, less than 1 μm, the dishing rate is close to the 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 in the literature with chemical slurry, it is shown that 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 erosion depends more on pattern area fraction but less on linewidth. It is shown that erosion reaches a steady-state rate after a short period. The steady-state rate of erosion is dependent on the apparent hardness and the intrinsic wear coefficients of Cu and oxide. Experiments also show that for a pattern with a large fraction, the rate of oxide erosion may increase than that predicted by the model due to the improvement of slurry transport. Moreover, since erosion does not depend on linewidth significantly, when the device scale shrinks down, the within-die nonplanarity will mainly be attributed to the erosion but not dishing if a large variation area fraction is shown on the pattern layout.

5. The objectives of process optimization are to maximize the Cu removal rate and to reduce surface nonplanarity due to dishing and erosion. The key is to reduce oxide erosion and to minimize the variance of dishing and erosion resulting from the effects of different area fractions and linewidths. Thus the surface topography will not be uneven even with a short period of overpolishing. It is proposed to employ SiO_2 abrasive or other particles with hardness close to the ILD silicon oxide to reduce the oxide removal rate and to increase the polishing selectivity between Cu and oxide. For patterns with wide Cu lines, the results suggested that the dishing rate might be reduced by a stiff pad and/or by a slight basic slurry.

Acknowledgments

This work was supported by the Silicon Valley Group, Inc., San Jose, CA. Thanks are due Papken Der Torossian, Dr. Larry Oh, and Sattar Al-Lami of SVG for encouragement. Many helpful discussions with Professor Nam P. Suh and Professor Duane Boning of MIT are gratefully acknowledged.

Massachusetts Institute of Technology assisted in meeting the publication costs of this article.

References

1. J.-Y. Lai, N. Saka, and J. H. Chun, *J. Electrochem. Soc.*, **149**, G31 (2002).
2. D. Ouma, B. Stine, R. Divecha, D. Boning, J. Chung, G. Shinn, I. Ali, and J. Clark, in *Proceedings of SPIE Microelectronics Manufacturing Conference*, p. 236 (1997); SPIE-Society of Photo-Optical Instrument Engineering.
3. T. H. Smith, S. J. Fang, D. Boning, G. B. Shinn, and J. A. Stefani, in *Proceedings of 4th International Conference on Chemical-Mechanical Polishing for ULSI Multilevel Interconnection CMP-MIC*, p. 93, Santa Clara, CA (1999).
4. J. T. Pan, P. Li, W. Kapila, S. Tsai, F. Redeker, T. Park, T. Tugbawa, D. Boning, in *Proceedings of 4th International Conference on Chemical-Mechanical Polishing for ULSI Multilevel Interconnection CMP-MIC*, p. 423, Santa Clara, CA (1999).
5. J. Y. Lai, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA (2001).
6. F. B. Kaufman, D. B. Thompson, R. E. Broadie, M. A. Jaso, W. L. Guthrie, D. J. Pearson, and M. B. Small, *J. Electrochem. Soc.*, **138**, 3460 (1991).
7. J. M. Steigerwald, S. P. Murarka, R. J. Gutmann, and D. J. Duquette, *Mater. Chem. Phys.*, **41**, 217 (1995).
8. R. Carpio, J. Farkas, and R. Jairath, *Thin Solid Films*, **266**, 238 (1995).
9. D. Zeidler, Z. Stavreva, M. Plötner, and K. Drescher, *Microelectron. Eng.*, **33**, 259 (1997).
10. M. Fayolle and F. Romagna, *Microelectron. Eng.*, **37/38**, 135 (1997).
11. Q. Luo, S. Ramarajan, and S. V. Babu, *Thin Solid Films*, **335**, 160 (1998).
12. H. van Kranenburg and P. H. Woerlee, *J. Electrochem. Soc.*, **145**, 1285 (1998).
13. M. Hariharaputhiran, J. Zhang, S. Ramarajan, J. J. Keleher, Y. Li, and S. V. Babu, *J. Electrochem. Soc.*, **147**, 3820 (2000).
14. S. Kondo, N. Sakuma, Y. Homma, Y. Goto, N. Ohashi, H. Yamaguchi, and N. Owada, *J. Electrochem. Soc.*, **147**, 3907 (2000).
15. S. P. Murarka, J. Steigerwald, and R. J. Gutmann, *MRS Bull.*, **1993**, 46.
16. J. M. Steigerwald, R. Zirpoli, S. P. Murarka, D. Price, and R. J. Gutmann, *J. Electrochem. Soc.*, **141**, 2842 (1994).
17. R. J. Gutmann, J. Steigerwald, L. You, D. T. Price, J. Neirynck, D. J. Duquette, and S. P. Muraka, *Thin Solid Films*, **270**, 596 (1995).
18. Z. Stavreva, D. Zeidler, M. Plötner, and K. Drescher, *Appl. Surf. Sci.*, **91**, 192 (1995).
19. Z. Stavreva, D. Zeidler, M. Plötner, G. Grasshoff, and K. Drescher, *Microelectron. Eng.*, **33**, 249 (1997).
20. T. Park, T. Tugbawa, D. Boning, J. Chung, S. Hymes, R. Muralidhar, B. Wilks, K. Smekalin, and G. Bersuker, in *Proceeding of 4th International Conferences of Chemical-Mechanical for ULSI Multilevel Interconnection (CMP-MIC)*, p. 184, Santa Clara, CA (1999).
21. J. M. Steigerwald, S. P. Murarka, and R. J. Gutmann, *Chemical Mechanical Planarization of Microelectronic Materials*, p. 258, John Wiley & Sons, Inc., New York (1995).
22. N. Elbel, B. Neureither, B. Ebersberger, and P. Lahnor, *J. Electrochem. Soc.*, **145**, 1659 (1998).
23. T. Tugbawa, T. Park, D. Boning, T. Pan, S. Hymes, T. Brown, and L. Camilletti, in *Chemical Mechanical Polishing IC Device Manufacturing III*, R. L. Opila, C. R. Simpson, K. B. Sundaram, I. Ali, Y. A. Arimoto, and Y. Homma, Editors, PV 99-37, p. 605, Electrochemical Society Proceedings Series, Pennington, NJ (1999).