Journal of The Electrochemical Society, **149** (1) G31-G40 (2002) G31
0013-4651/2001/149(1)/G31/10/$7.00 © The Electrochemical Society, Inc.

Evolution of Copper-Oxide Damascene Structures in Chemical Mechanical Polishing

I. Contact Mechanics Modeling

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

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

Nonplanarity arising from the chemical mechanical polishing of Cu-oxide damascene structures results in the exposure field (die-size) being partially out of focus in the subsequent lithography process. Thus the corresponding mechanisms of within-die polishing must be determined and the within-die nonplanarity due to polishing needs to be minimized to increase the process yield. In this paper, contact mechanics models were developed to explain the role of pattern geometry on the variation of material removal rate. The effects of Cu linewidth, area fraction, and the elastic properties of the polishing pad on pad displacement into low features were examined to focus on the mechanical aspects of the process. The pressure distribution on the high features was determined and the rate of pattern planarization was quantified. Experiments on patterned Cu wafers were conducted to verify the model. Based on these results, the planarization and polishing behavior and the within-die nonplanarity due to the variation of pattern geometry were discussed.
© 2001 The Electrochemical Society. [DOI: 10.1149/1.1420707] All rights reserved.

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

The continuing advances in ultralarge scale integration (ULSI) necessitate the design and fabrication of extremely small devices. The existing metallization schemes for interconnects are inadequate for new circuits. A new way to pattern Cu lines by a damascene scheme involving trench patterning and deposition followed by chemical mechanical polishing (CMP) has demonstrated a great potential for developments in interconnect technology. Figure 1 schematically shows the copper damascene structure. Copper and diffusion barrier layers (typically Ta, Ti, or TaN) are deposited on the etched interlevel dielectric (ILD) surface. Then a CMP process is employed to remove the excess Cu and form conductive paths in the ILD trenches. It has been reported in the literature that the excess Cu coating is planarized at a rate related to the local pattern geometry.[1-5] In order to remove all the Cu and barrier coating on the dielectric surface, so as to isolate the Cu interconnects, the pattern within a die is slightly overpolished and this results in surface nonplanarity. The within-die nonplanarity might result in the exposure field (die-scale) partially out of focus in the subsequent lithography process. Moreover, dishing of the soft Cu surface and erosion of oxide also reduce the thickness of Cu interconnects and increase the electrical resistance. Thus the corresponding mechanisms of planarization, dishing, and erosion must be determined and their impact on process yield must be addressed.

Much of the past research on the effects of pattern geometry on material removal has exclusively focused on the ILD layer (SiO_2) polishing.[6-12] Nevertheless, the results of ILD polishing can be adapted to the planarization stage in metal polishing since the pattern geometries are similar. A phenomenological model was proposed to relate the polishing rate of arrays of various features to the feature dimension and pattern density.[6] By experimentally determining the correlation between the polishing rate, feature dimensions, and the neighboring feature layout, the surface profile evolution was predicted. However, the correlation between the polishing rate and pattern geometry varies with different pattern design, and the tribological mechanisms of planarization were left unanswered in this model. Recently, the effects of pattern geometrical parameters, such as pattern density (*i.e.*, high feature area fraction), pitch, pattern area, and the ratio of perimeter to area, were extensively studied.[11] It was shown that the pattern density significantly affects the subdie-scale polishing. The influential range of a specific pattern on the neighboring area was characterized by a planarization length measured experimentally. A density-based numerical model was proposed to characterize the surface topography evolution for arbitrary layouts. Experiments were also conducted to verify this model.[11,13]

On the theoretical side, contact mechanics models have been employed to investigate the mechanisms of planarization. A planar elastic pad was assumed for predicting the pressure distribution on the die surface with various pattern layouts.[14-16] A generalized relation between pressure distribution and the pad displacement was proposed. An analytical solution for pattern evolution in steady-state regimes was also presented by Chekina *et al.* Based on the model, the nonunifom polishing rate across different pattern regions is attributed to the nonuniform pressure on the high features. The low features were assumed to stay intact without material being removed until the deformed pad contacts them. However, the steady-state analysis might be not applicable for some pattern layouts in Cu polishing. The pad might be in contact with the low features before

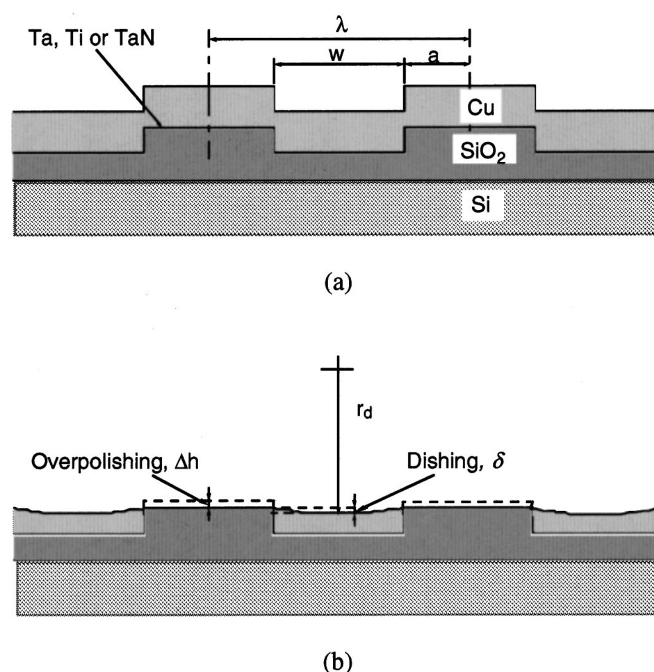


Figure 1. Schematics of Cu damascene structure: (a) before polishing and (b) after polishing.

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

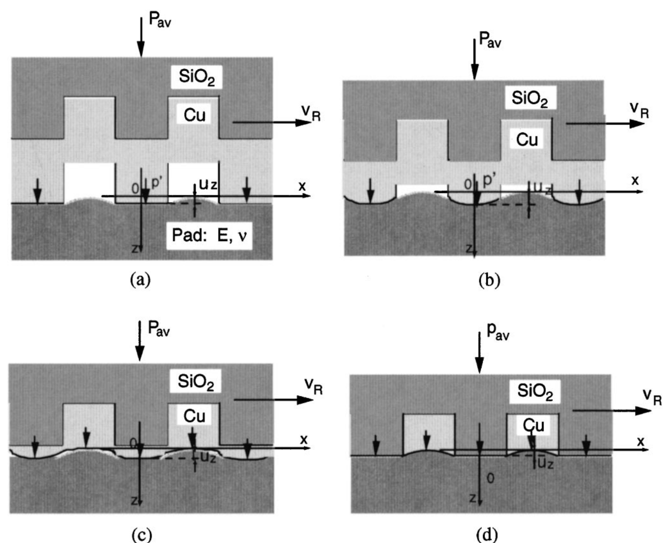


Figure 2. Schematics of the pattern/pad contact interface: (a) initial stage with uniform displacement specified on high feature, (b) planarization stage with pressure specified (either uniform pressure or elliptical distribution) on the contacting high features, (c) end of the planarization with pad in contact with low area, and (d) the onset of dishing and overpolishing.

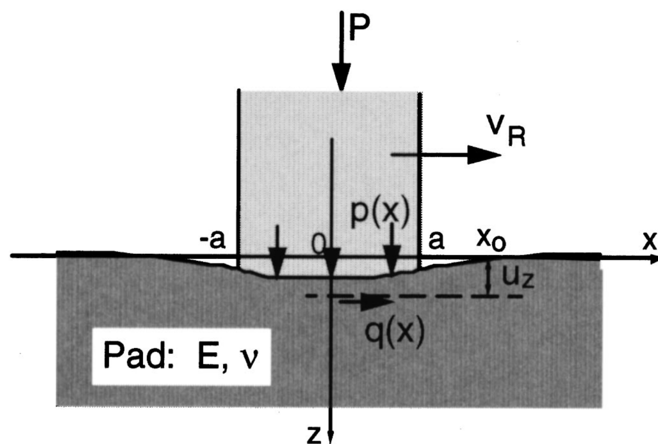


Figure 3. Schematic of a moving rigid line structure in contact with an elastic pad.

the high features reach the steady-state profile. Additionally, the pad might not be conformal with the high features as assumed in this model.

In this paper, models for different polishing regimes of Cu are proposed. Contact mechanics models, with conformal and nonconformal elastic pads, are employed for the planarization regime in this paper to focus only on the mechanical aspects of the polishing process. The effects of the pattern geometry (linewidth and area fraction), and the pad elastic properties on the pressure distribution and pad displacement, are investigated to explain the variation of material removal rate (MRR) on different patterns and the evolution of the surface profile. These contact mechanics models are examined by polishing patterned test wafers. Moreover, the understanding of nonplanar polishing due to pattern geometry will enlighten the study of Cu dishing and oxide erosion. The steady-state dishing and erosion will be modeled and examined experimentally in a companion paper. Accordingly, the fundamentals of Cu damascene polishing and the mechanisms of within-die nonuniformity of material removal are examined, and the important process parameters are identified.

Contact Mechanics Modeling

The local pattern geometry affects the pressure distribution $p(x,y)$ and thus results in a nonuniform MRR according to the Preston equation[17]

$$\frac{dh}{dt} = k_{\mathrm{p}} p(x, y) v_{\mathrm{R}} \quad [1]$$

where h is the thickness of the layer removed, t the polishing time, v_{R} the relative velocity, and k_{p} a constant known as the Preston constant. The Preston constant is affected by the hardness of the material being polished, abrasive size, slurry chemicals, and slurry transport phenomenon. At the die scale, these factors either remain unchanged for the same slurry, or vary insignificantly and can be assumed to be independent of the pressure and velocity. The objective of contact mechanics modeling, then, is to determine the pressure distribution on the feature surface. As shown schematically in Fig. 2, three regimes of polishing of patterned wafers may be postulated: planarization, polishing, and dishing/erosion. In the planarization regime in Fig. 2a and b, the step-height between the high and low features is much larger than the pad displacement u_z and hence the load is essentially supported by the high features only. The contact mechanics models are employed to determine the pressure variation on high features and to determine the pad displacement outside the high features to assure no wafer/pad contact on the low features. As the pad contacts the low area, however, both high and low features will be polished concurrently. The pressure distribution will become ever more uniform while the surface is being gradually smoothed down, Fig. 2c. In this regime, the thin, planar Cu layer is polished as in the blanket polishing situation. When the Cu layer is polished through, with the surface planarized, the pressure on both Cu and oxide would be about the same. As shown in Fig. 2d, the soft interconnect Cu now wears faster than the diffusion barrier layer (Ta, Ti, or TaN) and the ILD, and the surface of Cu interconnect is dished. Additionally, the oxide will be overpolished with the increase of overpolishing time, which may deteriorate surface planarity.

Geometry and boundary conditions.—Consider a rigid line structure on the wafer in contact with the elastic pad as shown in Fig. 3. The high features (shaded in Fig. 3) represent the Cu deposited on the underlying oxide and the low areas represent the Cu filled in the trenched oxide region. Since the length of the line is much larger compared with its lateral dimensions, the pattern/pad contact can be modeled as a two-dimensional (plane-strain) problem. The pad deformation is usually much smaller than the pad thickness, and thus the contact stresses are highly concentrated near the pad surface. With this approximation, the stresses can be calculated by assuming the pad as an elastic semi-infinite body. Now the boundary conditions along the pad surface must be specified to solve the stress and strain field in the elastic pad. On the boundary outside the loaded region, the surface is free of stress, *i.e.*

$$\sigma_z = \tau_{xz} = 0, \quad |x| > a \quad [2]$$

Within the loaded region, for simplification, the tangential traction is assumed to be negligible in the following analyses, *i.e.*

$$\tau_{xz} = -q(x) = 0, \quad -a < x < a \quad [3]$$

This assumption stands only when the friction coefficient at the wafer/pad contact interface is low. Prior experimental results support that the friction coefficient in Cu polishing is about 0.1.[18] Therefore, the effect of tangential traction on the stresses is negligible. Additionally, the normal stress within the loaded region can be specified to determine the pad tangential and normal displacements $u_x(x)$ and $u_z(x)$ on the entire surface, *i.e.*

$$\sigma_z = -p(x), \quad -a < x < a \quad [4]$$

In the contact mechanics literature, this problem has been solved and the formulation of the pad displacements within the contact region may be expressed in differential forms as[19,20]

$$\frac{\partial u_x(x)}{\partial x} = -\frac{(1-2v)(1+v)}{E}p(x) \quad [5a]$$

and

$$\frac{\partial u_z(x)}{\partial x} = -\frac{2(1-v^2)}{\pi E}\int_{-a}^{a}\frac{p(s)}{x-s}ds \quad [5b]$$

where v is the Poisson's ratio, E the Young's modulus of the pad, and s a dummy variable

Uniform pressure distribution.—The boundary condition of uniform pressure may be applied when a steady-state profile of high features is reached but without contact on the low features, Fig. 2b. The pressure distribution can be related to the load P on each high feature $|x - n\lambda| \leq a$ and the half width a of the feature as

$$p(x) = \frac{P}{2a} \quad [6]$$

where P is the force per unit length on each contact region, n the index of the high feature from the center of the specific area (from $-N$ to N, totally $2N + 1$ high features). In Eq. 6, the load P on the high features is assumed to be constant. This would be true when the features of interest are near the center of a specific pattern area with constant linewidth and pitch, and thus its pressure distribution will not be affected by the different pattern in the neighboring subdie area. Utilizing this boundary condition, Eq. 5b can be solved for the pad displacement over the specific area. The displacement of the pad can be written in normalized form as

$$\bar{u}_z(x) = -\frac{(1-v^2)}{\pi E}\frac{P}{2a}\sum_{n=-N}^{N}\left[\left(1+\frac{x-n\lambda}{a}\right)\times\ln\left(1+\frac{x-n\lambda}{a}\right)^2+\left(1-\frac{x-n\lambda}{a}\right)\times\ln\left(1-\frac{x-n\lambda}{a}\right)^2\right]+(2N+1)C_2 \quad [7]$$

where $\bar{u}_z(x)$ is the normalized displacement of the pad along the surface, defined as $\bar{u}_z(x) \equiv u_z(x)/a$, C_1 is an integration constant and determined relative to a datum x_o

$$C_1 = \frac{(1-v^2)}{\pi E}\frac{P}{2a}\left[\left(1+\frac{x_o}{a}\right)\ln\left(1+\frac{x_o}{a}\right)^2+\left(1-\frac{x_o}{a}\right)\ln\left(1-\frac{x_o}{a}\right)^2\right] \quad [8]$$

The datum x_o is arbitrarily chosen on the displaced surface referred to the initial surface plane, in which $u_z(x_o) = 0$. The choice of x_o usually needs to refer to the observation on the real deformed surface. This difficulty of determining C_1 is a general feature of the elastic half-space problems since the conditions at the ends far away from the contacting surface are undefined. To surmount this difficulty, the actual shape and dimension of the elastic body and the boundary conditions at the supporting sides must be considered. However, if our interest is on the relative shape of the pad surface and its displacement into low features instead of the movement of its surface level, the choice of x_o is not important and will not affect the profile of the pad surface.

Elliptical pressure distribution.—Another possible boundary condition on the high feature is an elliptical pressure distribution

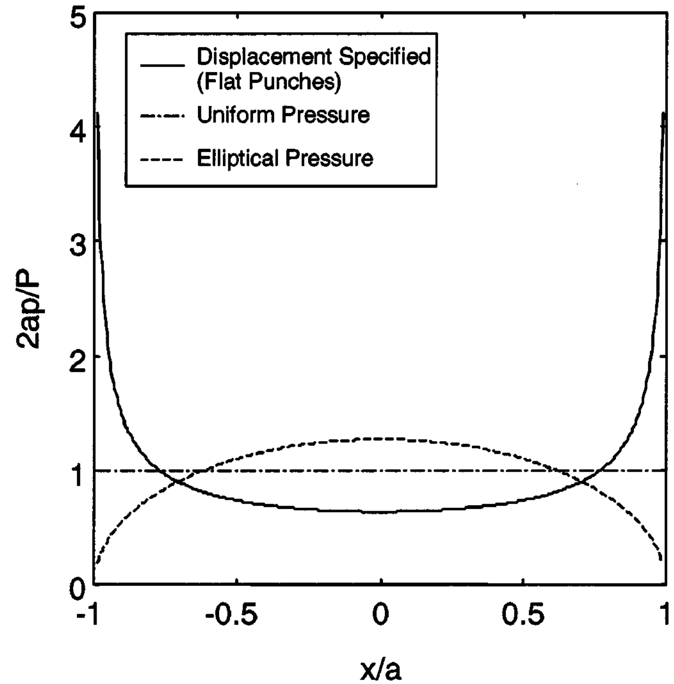


Figure 4. Pressure distribution in the contact region of the high feature for various boundary conditions.

given by the Hertz theory. In this case, both the wafer and the pad are modeled as nonconforming elastic bodies. The pressure distribution in the contact region, $|x - n\lambda| \leq a$, can be expressed as[20]

$$p(x) = p_o\left[1-\left(\frac{x}{a}\right)^2\right]^{1/2} \quad [9]$$

where p_o is maximum pressure on the feature and can be found by

$$p_o = \frac{2P}{\pi a} \quad [10]$$

As shown in Fig. 4, the pressure distribution can be rewritten as the normalized form, $\bar{p}$ (where $\bar{p}(x) \equiv 2ap(x)/P$), *vs.* the dimensionless distance, x/a

$$\bar{p}(x) = \frac{4}{\pi}\left[1-\left(\frac{x}{a}\right)^2\right]^{1/2} \quad [11]$$

The elliptical pressure distribution should result in a higher rate of material removal near the center of high features. However, this will change the profile of the high features quickly and promote the pressure distribution toward a more uniform fashion. Applying this elliptical pressure distribution on the boundary, the normalized displacement is given as

$$\bar{u}_z = -\frac{4(1-v^2)}{\pi E}\frac{P}{2a}\sum_{n=-N}^{N}\left\{\left(\frac{x-n\lambda}{a}\right)^2-\left[\frac{|x-n\lambda|}{a}\sqrt{\left(\frac{x-n\lambda}{a}\right)^2-1}-\ln\left(\frac{|x-n\lambda|}{a}+\sqrt{\left(\frac{x-n\lambda}{a}\right)^2-1}\right)\right]\right\}+(2N+1)C_2 \quad [12]$$

where C_2 is another dimensionless integration constant

$$C_2 = \frac{4(1-\upsilon^2)}{\pi E}\frac{P}{2a}\left\{\left(\frac{x_o}{a}\right)^2 - \left[\frac{|x_o|}{a}\sqrt{\left(\frac{x_o}{a}\right)^2 - 1} - \ln\left(\frac{|x_o|}{a} + \sqrt{\left(\frac{x_o}{a}\right)^2 - 1}\right)\right]\right\} \quad [13]$$

Rigid flat punches with specified displacement.—In some cases, it may be simpler to specify the normal displacement $u_z(x)$ instead of the pressure distribution within the contact region on the assumption that the pad conforms to the profile of the high feature. The solution for this "mixed boundary-value problem" is also given in the contact mechanics literature.[19] The general solution of the singular integral equation for $p(x)$, Eq. 5b, in the contact region given by

$$p(x) = -\frac{E}{2(1-\upsilon^2)(a^2-x^2)^{1/2}}\int_{-a}^{a}\frac{(a^2-s^2)^{1/2}}{(x-s)}\frac{\partial u_z(s)}{\partial s}ds + \frac{P}{\pi(a^2-x^2)^{1/2}} \quad [14]$$

For example, in the initial stage of Cu planarization the pattern profile is known from the prior deposition process, as shown in Fig. 2a. By assuming uniform indentation ($\partial u_z/\partial x = 0$) across the flat loaded region, the pressure distribution on an array of high features $|x - n\lambda| \leq a$ can be obtained from Eq. 14 as

$$p(x) = \frac{P}{\pi[a^2 - (x-n\lambda)^2]^{1/2}} \quad [15]$$

Figure 4 also shows the normalized pressure $\bar{p}$ *vs.* the normalized distance x/a across a high feature for the flat-punch contact. The pressure reaches a theoretical infinite at the edges of the high features due to the discontinuity in $\partial u/\partial x$. In reality, a finite radius retains the continuity of $\partial u/\partial x$ across the edge and the pressure would increase to a high but finite value near the edge. Moreover, since the pad material cannot sustain such a high stress, it will yield plastically. Thus the pressure would not be higher than the yield/fracture strength of the pad. Despite the pressure concentration near the edges, the pressure is distributed more or less uniformly over the feature and close to about 0.7 times the average pressure $P/2a$ in the contact region. Increase of the load on each high feature proportionally increases the pressure at each point but does not change its distribution. Additionally, Eq. 15 shows that the general shape of the curve is not affected by the elastic properties of the pad.

The pad displacement outside the high features can be found by substituting the pressure within loaded regions obtained from Eq. 15 into Eq. 5b and performing the integration. The pad displacement can be expressed in the normalized form as

$$\bar{u}_z(x) = -\frac{4(1-\upsilon^2)}{\pi E}\frac{P}{2a}\sum_{n=-N}^{N}\ln\left[\frac{x-n\lambda}{a} + \left(\left(\frac{x-n\lambda}{a}\right)^2 - 1\right)^{1/2}\right] + (2N+1)C_3 \quad [16]$$

where C_3 is the normalized integration constant, defined in the case of single flat punch indentation as

$$C_3 = \frac{2(1-\upsilon^2)}{\pi E}\frac{P}{2a}\ln\left[\frac{x_o}{a} + \left(\frac{x_o^2}{a^2} - 1\right)^{1/2}\right] \quad [17]$$

Results

Pad displacement.—The normalized pad displacements for the various boundary conditions are plotted in Fig. 5 based on Eq. 7, 12, and 16. The zero on the ordinate is set to refer to the bottom of the high features. The displacement is calculated based on the geometry close to current CMP practice that area fraction is 0.5 ($A_f = w/\lambda = 1 - 2a/\lambda = 0.5$), and the applied pressure on the wafer is 50 kPa (7 psi). As shown in Table I, the elastic modulus and Poisson's ratio of the pad are assumed to be $E = 500$ MPa and $\nu = 0.3$ (close to those of the bulk commercial pads). However, for features with dimensions smaller than the sizes of pad asperities (about 20-50 μm), the intrinsic elastic properties of the pad material (polyurethane) may be employed to calculate the pad displacement. Since the intrinsic Young's modulus of the pad is higher than that of the porous pad (*e.g.*, about 1.18 times higher for the pad with 15% of porosity in volume), the use of bulk properties will result in overestimation of the pad displacement for the small lines (about 18%). The properties of the materials being polished are also shown in the Table I. The assumption of rigid punch indentation in prior analyses is valid since the Young's moduli of all materials involved in Cu damascene pattern are much high than that of the pad.

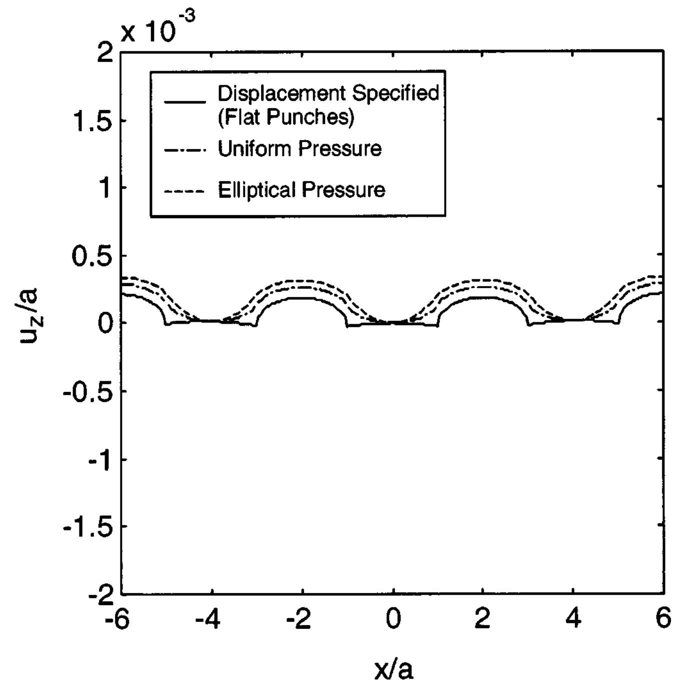


Figure 5. Surface profiles of deformed pad for various boundary conditions.

It is shown that the maximum displacements of the pad for the three boundary conditions are approximately of the same order of magnitude. For current circuit design, in which the width of small features is about $w = 0.18$-0.5 μm (or $a = 0.09$-0.25 μm for features with an area fraction of 0.5), the displacement of the pad outside the high features is about 0.03 to 0.08 nm, almost negligible compared with the surface roughness of the pad. Therefore, if the step-height is much larger than the pad displacement, as in the planarization stage of Cu polishing (for an initial step-height of about 0.5-1 μm), the pad will not be in contact with the low surfaces. Thus

Table I. Elastic properties of materials.

Material	Young's modulus (GPa)	Poisson's ratio
Cu	128[a]	0.30
Ta	186[a]	0.30
SiO_2 (TEOS)	74[b]	0.20
Rodel pad	0.5	0.30

[a] ASM Metals Handbook, ASM International.
[b] Handbook of Materials Science, CRC Press, Inc.

the material removal rate on low features will be relatively low because no abrasion would occur until the end of the planarization process. Additionally, the within-die nonplanarity will result in a variation of the average pressure on different patterns and affect the pad displacement. The pressure distribution on different high features at the nonplanar surface can be determined numerically based on Eq. 5 with a displacement specified boundary condition in the contact region corresponding to the actual surface topography. As shown later in the Experimental results, however, the effect of mean pressure variation due to within-die nonplanarity is not as significant as the local pattern geometry, linewidth, and area fraction on the pad displacement.

Effects of pattern linewidth.—The effects of linewidth on pad displacement are also illustrated in Fig. 5. For a pattern with constant area fraction, for example, 0.5, the increase of linewidth will proportionally scale-up the displacement of the pad. Consequently, the pad might start to contact the low area before the topography is planarized. For instance, the pad displacement is about 20 nm for a 100 μm wide line. In reality, when the size of pad asperities, about 100-200 nm, and particle size, about 200 nm, are considered, the low area of a wide feature (interconnect line, contacting pad) will be in contact with the pad in the planarization regime (about half of the initial step-height in this case). By contrast, for submicrometer size features, the pad asperities cannot reach the low features freely with the constraint of the surrounding high features. Hence, the pad does not contact the low area until the end of the planarization regime.

Another effect of increased linewidth for the same A_f is to decrease the (average) material removal rate on the high feature. Since part of the load at some point in the planarization regime is supported by the low features, the average pressure on the high features decreases and so does the material removal rate. The variation of the material removal rate for different area fraction regions results in a variation of process duration for clearing Cu on the same die. This results in part of the die being overpolished and introduces oxide thickness variation and Cu loss problems. Additionally, the earlier the low features are in contact with the pad, the more likely the surface topography may be partially retained until the end point of the process. This may increase the surface nonuniformity and start dishing the Cu in the trenches before the Cu layer is polished through.

Effect of pattern area fraction.—The effects of the area fraction are shown in Fig. 6. The normalized pad displacement, u_z/a, increases with the A_f. The three boundary conditions show the same trends and similar displacement values. Among the three, the elliptical pressure distribution yields the largest pad displacement for all A_f. For lower A_f, the uniform pressure boundary condition results in a slightly larger pad displacement in the low area than that of the (constant) displacement specified condition. When A_f is greater than approximately 0.7, the pad displacement is more in the low area for the boundary condition of specified displacement in the loaded region.

Except for the steep increase of u_z/a with A_f at higher area fraction region, *i.e.*, greater than 0.7, the displacement increases almost linearly with A_f. Between the area fractions of 0.2 to 0.6, the slope of the curve is about 1×10^{-3}. Thus the pad displacement will be about the same order of magnitude even with a slight variation of area fraction for the present circuit designs. Therefore, for the first few metal layers with fine interconnects, the surface will be planarized before the pad contacts the low area. Moreover, if the Cu linewidth is small and the effect of pad displacement is essentially negligible, the material removal rate and the rate of planarization will both increase proportionally with A_f because the average pressure on the high features is inversely proportional to A_f. This will result in a buildup of the surface nonplanarity within the die across different area fraction regions in the planarization stage of process.

Effect of pad elastic properties.—The effect of pad elastic modulus is shown in Fig. 7 with a dimensionless parameter p_{av}/E, which is close to 10^{-4} for the present pad and the nominal pressure employed in CMP. However, different pads might be tried to vary the degree of surface planarity. For example, some engineering plastics, with E about an order of magnitude greater than the polyurethane pad (about 1-5 GPa), or even some soft metals, with a two-order of magnitude greater E (about 10-20 GPa), may be tried to improve the

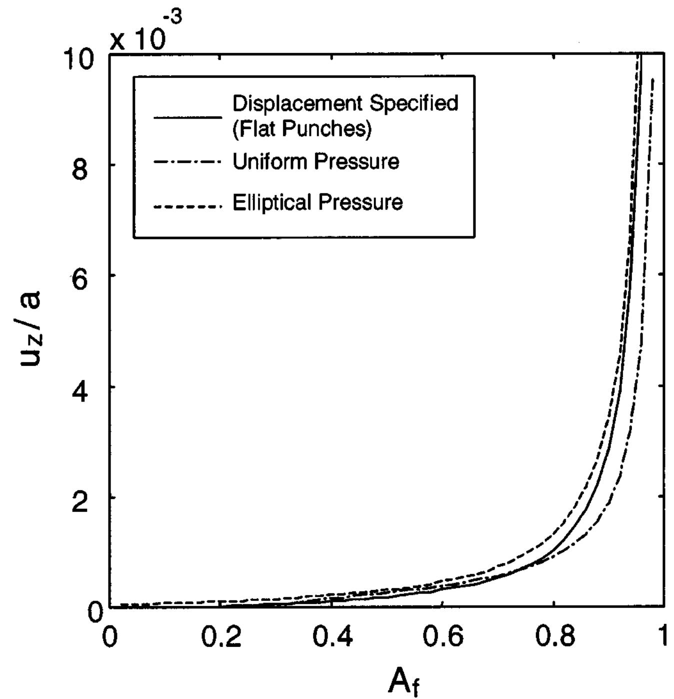


Figure 6. Pad displacement *vs.* pattern area fraction for various boundary conditions.

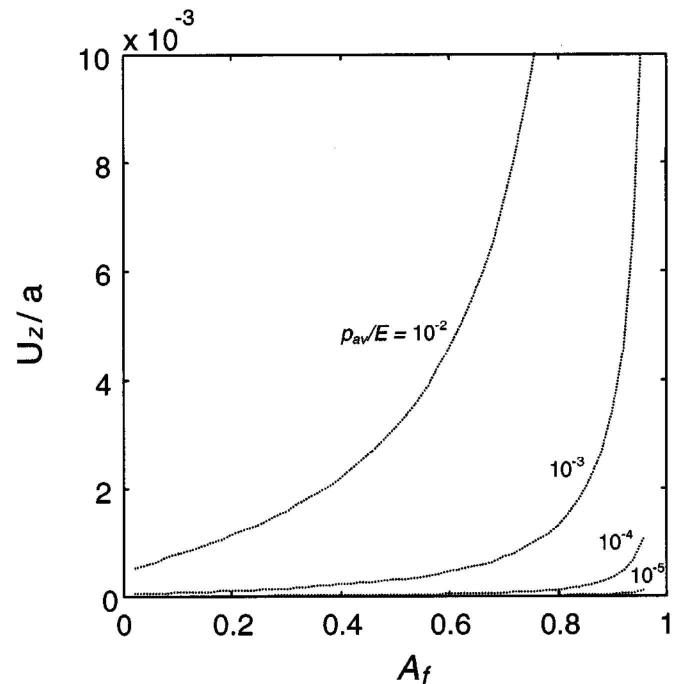
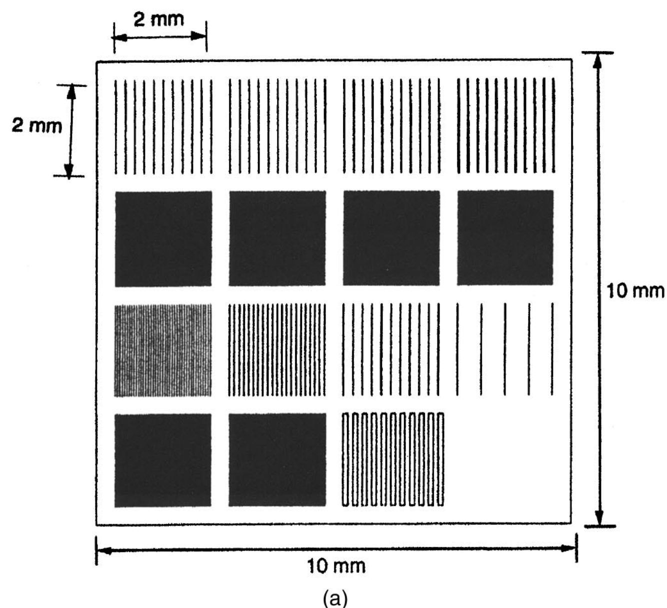


Figure 7. Effects of applied pressure p_{av} and Young's modulus of the pad E on the pad displacement (elliptical pressure distribution).



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

surface planarity. Figure 7 indicates that the pad displacement will decrease proportionally with the increase of E or the decrease of p_{av}/E value. The results can be applied to the pad displacement at both the low feature and the lower subdie region which has been polished down faster due to a higher area fraction than the surrounding low area fraction region. By using a stiffer pad, the surface level of two distinct area fraction regions may be maintained at a small difference and retain the surface planarity across the die. On the other hand, a compliant pad, such as some polymer foams with E ranging from that close to the present pad (about 500 MPa) down to one order magnitude less (about 10 MPa), may be employed in the final polishing regime to reduce the load on abrasive particles and prevent surface scratching. By employing the contact mechanics models, the desired range pad properties with respect to the range of applied pressure to satisfy the requirements at different process regimes can be determined.

Table II. Properties of slurry.

Abrasive	α-Al_2O_3
Particle size, μm	0.3
Particle hardness, MPa	20,500
Concentration, vol %	2-3
Viscosity, Pas	0.03
pH	7

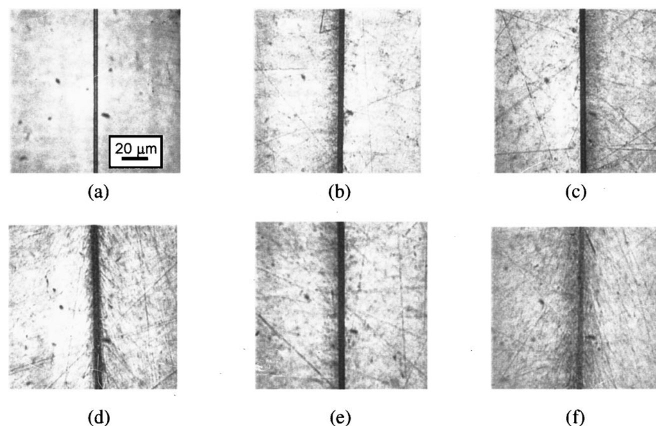


Figure 9. Optical micrographs of the evolution of pattern surfaces (w = 5 μm and λ = 500 μm): (a) initial surface, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6 min.

Experimental Validation and Discussion

Experimental conditions.—As shown in Fig 8, a Cu damascene structure has been designed to verify the results of contact mechanics modeling. The pattern for each die (10×10 mm) consists of a matrix of 2.5×2.5 mm blocks (subdie). These blocks in turn consist of line-space features, whose dimensions are close to the scale of the Cu interconnects on current chips with a minimum linewidth of 0.5 μm and a maximum linewidth of 100 μm. The area fraction of the experimental mask ranges from 0.01, representing an isolated line, to 0.5, representing a dense packing case. This pattern is transferred into a 2 μm thick SiO_2 coating by lithography on a 100 mm, (100) orientation silicon wafer. After oxide trenches are etched to a depth of 1 μm, a 20 nm thick Ta barrier layer is deposited, followed by a 1.5 μm thick Cu film, by physical vapor deposition (PVD). The step-height of the features is a function of the trench width and the

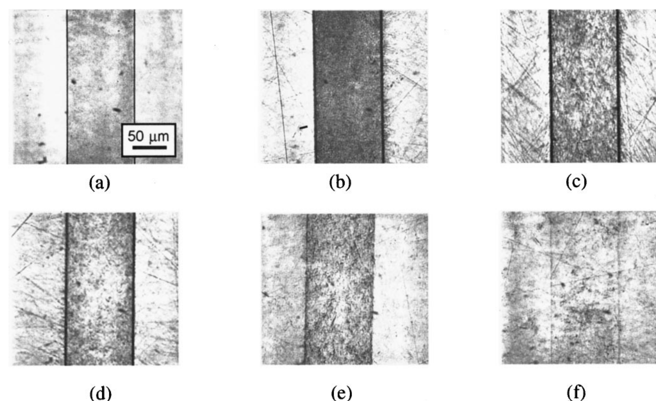


Figure 10. Optical micrographs of the evolution of pattern surfaces (w = 100 μm and λ = 200 μm): (a) initial surface, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6 min.

Table III. Experimental results of step-height evolution at planarization stage.

Linewidth w (μm)	Pitch, λ (μm)	w/λ	Step-height at different durations (nm) 2 min	3 min	4 min	5 min	6 min	Rate of planarization (nm/min)
2.0	4.0	0.50	857	582	270	114	28	213
2.0	200.0	0.01	800	636	489	355	201	148
5.0	200.0	0.025	851	700	565	315	222	164
5.0	500.0	0.01	849	711	550	435	320	133
25.0	50.0	0.5	897	601	253	47	10	233
25.0	200.0	0.125	795	654	503	321	212	150
100.0	200.0	0.5	945	695	345	78	49	241

pattern density. The mask layout and the corresponding area fractions of metal interconnect lines with respect to the specific w and λ are shown in Fig. 8.

In this set of experiments, low nominal pressure of 28 kPa (4 psi) and linear velocity of 0.46 m/s were employed to decrease the material removal rate for a longer period of planarization than the normal CMP conditions. A neutral slurry with 4 vol % of 0.3 μm Al_2O_3 abrasive particles was used to focus on the mechanical aspect of the process and verify the contact mechanics models. The composition and the properties of the slurry are listed in Table II. In addition, the polishing experiments were conducted on the grooved commercial polyurethane pad, Rodel IC-1400, and conditioning was performed between each polishing to retain a similar surface condition of the pad. The initial and polished profiles of the patterns were recorded by a stylus profilometer and the atomic force microscope (AFM), for coarse and fine structures, respectively.

The evolution of patterns.—The evolution of the patterned surface ($w = 5$ μm and $\lambda = 500$ μm) is shown in Fig. 9. Due to the high reflectance of Cu, the unpolished, scratch-free high features appear bright in the optical micrograph, Fig. 9a. The walls between the high and low surfaces appear dark in bright-field illumination since less normal incident light is reflected. The low features are less reflective because of the coarse microstructure from the Cu deposition. After 2 min of polishing as shown in Fig. 9b, the surface of high features was slightly roughened, but is still highly reflective compared with the low area. The same coarse microstructure of the low area indicates that the pad does not contact the low area, as predicted by the contact mechanics models. Moreover, the boundaries between the high and low features become less distinguishable and wider due to the rounding at the edges. This has been explained in the earlier theory section by the high stress concentration around the sharp edges in the early stage of planarization. Similar surface morphology was found in Fig. 9c, d, and e for polishing 3, 4, and 5 min. Nevertheless, after 6 min of polishing, as shown in Fig. 9f, the surfaces of both high and low features are roughened. Therefore, the pad is now in contact with both the high and low features and both surfaces are polished.

Figure 10 shows the evolution of a different pattern with much larger linewidth, 100 μm, and 0.5 area fraction with 200 μm pitch. The trend of surface evolution is similar to the prior case with small line and area fraction. Few scratches were observed on the low area because large contaminant or agglomerate particles might be transported into the trenches and contact with pad asperities to abrade the surfaces of low features. Additionally, the low area starts to be roughened at about 4 min of polishing, earlier than that of the prior case. This implies that the pad contacts with the low features in the planarization stage for wider lines, which suggests that pad displacement increases with the linewidth. A qualitative comparison between the theoretical prediction and the experimental results of the pad displacement is given in the next section.

Effect of pattern geometry.—The remaining step-heights of various patterns at different polishing durations are listed in Table III. For the features of interest, the process is still in the planarization stage since the step-height after 6 min of polishing is generally larger than twice the rms surface roughness, about 10 nm. The rate of planarization is defined by the reduction rate of step-height, which indicates how fast the surface topography can be removed in the planarization process. It is found that the rate of planarization is strongly affected by the area fraction of the pattern but less by the linewidth. For $A_f = 0.5$, the surface was planarized at a rate about 220 nm/min, twice the rate of blanket Cu polishing, about 100 nm/min. For small A_f of 0.01, the planarization rate is close to the blanket rate, about 140 nm/min.

The time evolution of step-heights for patterns with 0.5 area fraction is shown in Fig. 11. Again, the high features of 2, 25, and 200 μm line structures were planarized at a similar rate. Since the low area is not polished in the earlier stage of planarization, up to 4 min according to the experimental observation, the rate of material removal at the high feature can be estimated based on the rate of step-height reduction. The dashed line in the figure represents the least square fit of the data before the low area is in contact with pad, which indicates that the MRR on the high features is about 305 nm/min. After about 5 min of polishing, the pad starts contacting the low area when the step-height reaches about 50 to 200 nm, depending on the linewidth of the patterns. The rate of planarization reduces to about 130 nm/min before the surface is planarized, about half that before the pad contacts with low features. This suggests

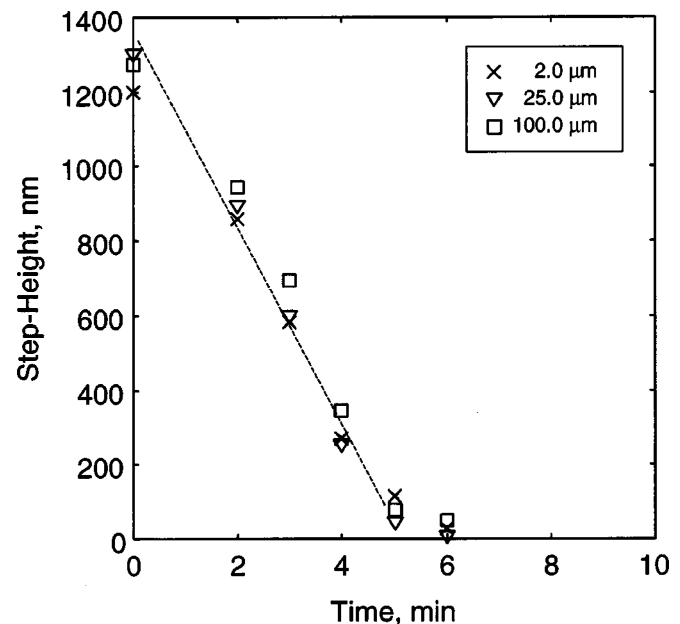


Figure 11. Time evolution of step-heights for patterns with constant area fraction 0.5 ($w/\lambda = 0.5$) and various linewidths.

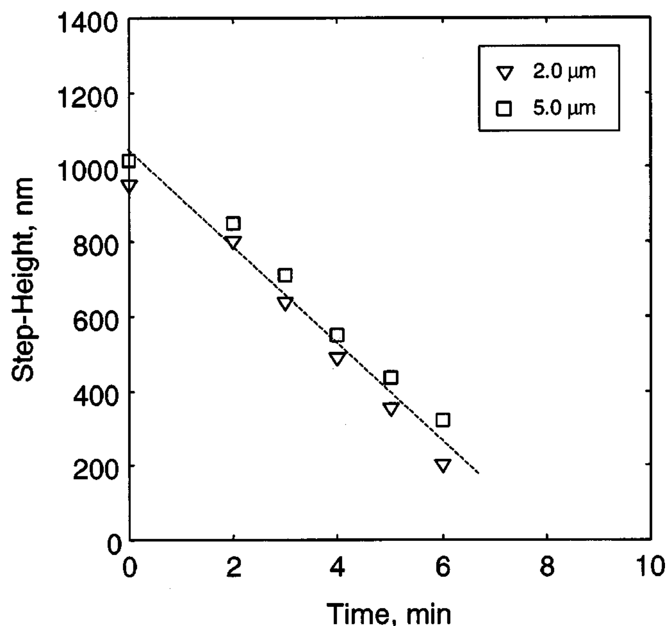


Figure 12. Time evolution of step-heights for patterns with constant area fraction 0.01 ($w/\lambda = 0.01$) and 2 and 5 μm linewidths.

that now the load is uniformly supported by both low and high areas rather than just 50% of the high area in the earlier stage.

For the wide features, it is shown that the pad starts to contact the low area much earlier before the topography is planarized. For instance, the low area for the 100 μm line has been polished when the step-height was about 350 nm. This step-height is larger than the pad displacement predicted by the contact mechanics model, which is about 20 nm. In reality, when the size of pad asperities, about 20-50 μm, and particle size, about 300 nm, is considered, the low area of a wide feature (interconnect line, contacting pad) will be in contact with the pad at the earlier planarization regime. By contrast, for micrometer or submicrometer size features, the pad asperities cannot reach the low feature freely with the constraint of the surrounding high features. Hence, the pad does not contact the low area until the end of the planarization regime. Therefore the high feature will be polished further down before the low feature has been polished, as in the case of 2 μm lines.

On the other hand, as shown in Fig. 12 for the patterns with 0.01 area fraction and linewidth 2 and 5 μm, the high features are removed at a lower rate of about 140 nm/min for the first 6 min. The high area is polished slowly as on a blanket area because of the small area fraction. Since the removal rate on the high features is low and the linewidth is small, the low area remained unpolished during the polishing period. The MRR on the high area rate is about half that for 0.5 A_f in the earlier planarization stage. These imply that the average pressure on each subdie is similar and the nonuniformity of material removal rate across a die is essentially due to the nonuniform pressure distribution on the high features affected mainly by the pattern area fraction.

Moreover, the MRR of high features for a 0.01 area fraction is much higher than the blanket rate, and is slightly higher but close to the blanket rate after the pad is in contact with the low area with uniform pressure distribution. This can be explained by the effects of surface nonplanarity across subdie areas of different area fractions. The high area fraction area will be planarized faster than the low area fraction. Based on the contact mechanics model, the area of low area fraction will on average support a higher normal pressure. However, compared with a single line, the pressure difference between the high and low area fraction regions usually will not be large because the subdie area is wide so that the pad can displace into the low subdie region and even out the pressure. Additionally, the trenches on the wafer surface might help transport the slurry and thus increase the MRR slightly.

Conclusions

The mechanical aspects of planarization behavior in the CMP of Cu patterns in the early stages was studied based on the contact mechanics models. Experimental results to support the models are also presented. Based on the model and the experimental results, the following conclusions can be drawn.

1. Three different boundary conditions, displacement, uniform pressure, and elliptical pressure distribution (Hertzian contact) on high features, were specified to represent the initial stage, steady-state stage, and pattern/pad nonconformal situation. For each condition, the pad penetration into low features of submicrometer lines is insignificant compared with the surface roughness of the pad. Therefore, the applied load is carried by the high features during planarization and the material removal rate increases with the area fraction of Cu interconnects (*i.e.*, the percentage of the low feature area).

2. The pad displacement increases with the increase of linewidth and area fraction and with the decrease of the pad elastic modulus. For the wider Cu features (about 100 μm) at the higher levels of interconnects, the pad may start to contact low areas and reduce the rate of planarization and retain the surface topography. For area fractions of about 0.5, the pad displacement does not vary significantly with the area fraction. Additionally, the models can be used to choose the desired elastic modulus of the pad to help improve the surface planarity for given conditions (pressure, pattern geometry, etc.).

3. The contact mechanics models cannot be applied when the surface uniformity is small, as at the end of the planarization stage and the dishing and erosion regimes, because pad roughness and the particle/wafer contact conditions are difficult to obtain. On the other hand, the nonuniform MRR in different density regimes across a die can be explained by the contact mechanics model.

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.

Appendix

Microcutting Model of Polishing

The wear equation is another way to represent the volume worn in sliding or abrasive wear[21,22]

$$V = k_w \frac{LS}{H} \qquad \text{[A-1]}$$

where V is the volume removed, L the load on the sample, S the relative sliding distance, H the hardness of the worn material concerned, and k_w is the wear coefficient. It may be noted that the wear equation is concerned only with the total worn volume, regardless of whether the surface is worn uniformly or the size of the contacting surface. It is also found that the wear coefficient, to a first approximation, can be used to categorize the wear processes, *i.e.*, the wear coefficient will be approximately the same for various materials if the wear mechanism is the same.

Equation A-1 can be expressed as the volume wear rate form

$$\frac{dV}{dt} = \frac{k_w L}{H} v_R \qquad \text{[A-2]}$$

where v_R ($= dS/dt$) is the relative velocity of the two surfaces in contact. If the removal across the surface in contact is assumed to be uniform, *i.e.*, thickness of the layer removed, h is constant across the surface, Eq. A-2 may be further written as

$$\frac{dh}{dt} = \frac{k_w}{H} p v_R \qquad \text{[A-3]}$$

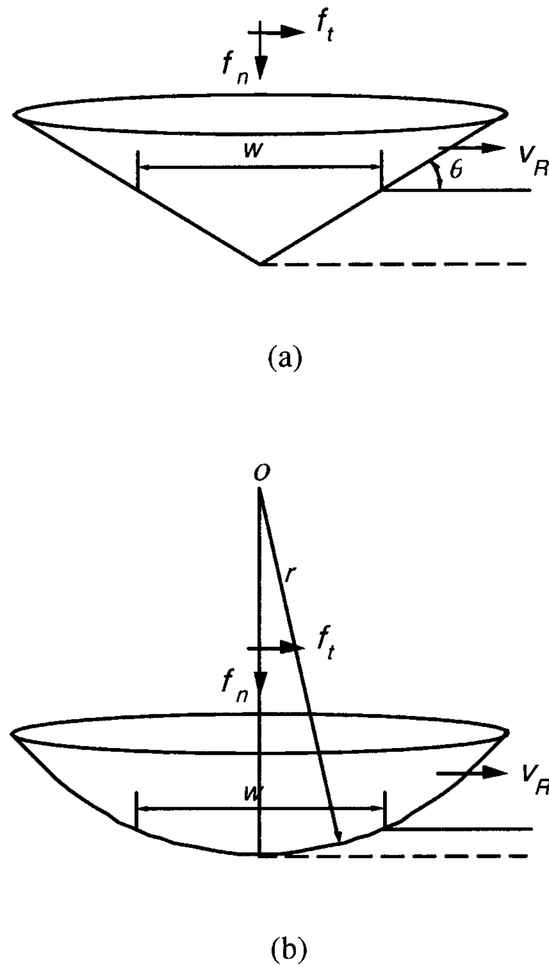


Figure A-1. Schematics of microcutting model in polishing: (a) with conical shape abrasive, (b) with spherical shape abrasive.

This is identical to the Preston equation in Eq. 1, and the Preston constant k_p is the ratio of the wear coefficient to the hardness of the worn surface. It may be noted that Eq. A-3 can be used on both global as well as local scales.

If only the mechanical aspect of the polishing is concerned, it was shown by the authors that plastic deformation of the surface material by particle abrasion is the dominant mode of material removal.[5] In this case, the microcutting model can be employed to estimate k_w and the material removal rate. In this model, abrasive particles act as single-point cutting tools and produce shallower and narrower grooves under the smaller load than those in grinding or other abrasion processes. Hardness determines the depth of particle penetration and, therefore, the material removal rate and surface finish. The abrasive is normally harder than the surface being polished to maintain a high rate of material removal. The frictional force is mostly due to the resistance of the soft surface being polished to plastic flow. Upper bound estimates for the material removal rate in the microcutting mode have been made in the past by idealizing the shape of the abrasive tip as a cone or a sphere.[23-25]

For a conical tip of abrasive particle, as shown in Fig. A-1a, the depth and width of cut are related by

$$d = \frac{w}{2} \tan \theta \qquad \text{[A-4]}$$

where d is the depth of the grooves, w the width of the grooves, and θ the angle between the abrasive and the surface. The normal force, f_n, on the particle is support by the pressure, p, at the contacting interface, which is assumed to be the pressure on the soft surface at yielding. Since only the front half of the cone is in contact while sliding, the normal force can be calculated simply by multiplying the projected area of the contacting surface with the pressure at yielding of the worn surface. Thus

$$f_n = \frac{\pi w^2}{8} p \qquad \text{[A-5]}$$

It may be noted that in the above equation the yield pressure is approximated as the hardness, H, of the worn surface. The material removed can be estimated by the product of the cross section of the groove, $dw/2$, and the sliding distance of the particle, S. Hence, the wear coefficient thus can be determined by Eq. A-1 as

$$k_w = \frac{VH}{LS} = \frac{(dwS/2)H}{f_n S} = \frac{\tan \theta}{2\pi} \qquad \text{[A-6]}$$

The same treatment can be applied to an abrasive with a spherical tip, as shown in Fig. A-1b, which might be closer to the actual shape of the abrasive particles. The geometry of the groove can be related to the radius of the particles, r, by

$$r^2 = (r - d)^2 + (w/2)^2 \qquad \text{[A-7]}$$

In polishing, the penetration is usually very shallow and therefore Eq. A-7 can be simplified and the depth of the groove can be approximated by

$$d = \frac{w^2}{8r} \qquad \text{[A-8]}$$

Thus, the normal forces acting on the particle to penetrate the surface and the wear coefficient can be determined as before

$$f_n = \frac{\pi w^2}{8} H \qquad \text{[A-9]}$$

$$k_w = \frac{2}{\pi}\left(\frac{2r}{w}\right)^2 \left\{ \sin^{-1}\left(\frac{w}{2r}\right) - \left(\frac{w}{2r}\right) + \frac{1}{2}\left(\frac{w}{2r}\right)^3 \right\} \qquad \text{[A-10]}$$

Prior experimental results by the authors showed that the wear coefficient in CMP conditions is less than these estimated by the microcutting model.[5] This implies that the material is not cleanly sheared off by a single pass of the abrasive particle. This point is also supported by the observation that no chip-like wear particles were found on the worn wafer or post-CMP pad surfaces. Instead, ridges were formed along with some deeper and wider grooves, which suggests particle plowing. In fact, early research in abrasive wear[24] found that when the attack angle θ between the abrasive and the coating surface is small (*i.e.*, the rake angle is a large negative value), large plastic deformation occurs at the surface and in the subsurface region below the contact. Since the penetration depth in the CMP condition is usually very shallow, the attack angle for any shape of particles will be very small. Therefore, it is more likely that plowing will prevail rather than cutting.

On the other hand, if the chemical aspect is important or dominant in material removal, the wear coefficient and the Preston constant are also determined by the chemical reaction, slurry composition, and the slurry rate and dispensing phenomenon at the interface. This is beyond the focus of this paper and is not discussed here.

List of Symbols

A_f	area fraction of metal pattern
a	half-linewidth of the metal pattern
C_1, C_2, C_3	integration constants
d	depth of the surface scratch
f_n	normal force applied on the abrasive particle
h	thickness of the material removed on wafer surface
k_p	Preston constant
k_w	wear coefficient
L	normal load on the wafer
P	load per unit length on a high feature
p	normal traction on the wafer surface
p_o	maximum normal traction on a high feature
q	tangential traction on the wafer surface
S	relative sliding distance
t	experiment duration
u_x, u_z	normal and tangential displacements of the pad
$\bar{u}_z$	normalized normal displacement of the pad
V	volume loss due to wear
v_R	relative linear velocity of wafer
w	pattern linewidth
x, y, z	Cartesian coordinates
x_o	position of a datum on the displaced surface of pad

Greek

Δh	oxide overpolishing
δ	Cu dishing
λ	pattern pitch
θ	angle between the abrasive and the sworn surface
ν	Poisson's ratio
σ_z	normal stress on the pad surface
τ_{xz}	shear stress on the pad surface

References

1. J. M. Steigerwald, R. Zirpoli, S. P. Murarka, D. Price, and R. J. Gutmann, *J. Electrochem. Soc.,* **141**, 2842 (1994).
2. S. P. Murarka, J. Steigerwald, and R. J. Gutmann, *MRS Bull.,* **18**, 46 (1993).

3. T. Park, T. Tugbawa, D. Boning, J. Chung, S. Hymes, R. Muralidhar, B. Wilks, K. Smekalin, and G. Bersuker, in *Proceedings of 4th International Chemical-Mechanical for ULSI Multilevel Interconnection Conference: CMP-MIC*, p. 184, Santa Clara, CA (1999).
4. T. Park, T. Tugbawa, D. Boning, S. Hymes, T. Brown, K. Smekalin, and G. Schwartz, 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. 94, The Electrochemical Society Proceedings Series, Pennington, NJ (1999).
5. J. T. Pan, P. Li, W. Kapila, S. Tsai, F. Redeker, T. Park, T. Tugbawa, and D. Boning, in *Proceedings of 4th International Chemical-Mechanical for ULSI Multilevel Interconnection Conference: CMP-MIC*, p. 423, Santa Clara, CA (1999).
6. J. Warnock, *J. Electrochem. Soc.,* **138**, 2398 (1991).
7. P. A. Burke, in *Proceedings of the 8th International VLSI Multilevel Interconnection Conference: VMIC*, p. 379 (1991).
8. S. R. Runnels, *J. Electrochem. Soc.,* **141**, 1900 (1994).
9. S. R. Runnels, *J. Electron. Mater.,* **25**, 1574 (1996).
10. J. M. Boyd and J. P. Ellul, *J. Electrochem. Soc.,* **143**, 3718 (1996).
11. B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung, D. Hetherington, I. Ali, G. Shinn, J. Clark, O. S. Nakagawa, and S.-Y. Oh, in *Proceedings of 2nd International Conference on Chemical-Mechanical Polishing for ULSI Multilevel Interconnection CMP-MIC*, p. 266, Santa Clara, CA (1997).
12. 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); Society of Photo-optical Instrument Engineering.
13. T. H. Smith, S. J. Fang, D. Boning, G. B. Shinn, and J. A. Stetani, in *Proceedings of 4th International Conference on Chemical-Mechanical Polishing for ULSI Multilevel Interconnection CMP-MIC*, p. 93, Santa Clara, CA (1999).
14. O. G. Chekina, L. M. Keer, and H. Liang, *J. Electrochem. Soc.,* **145**, 2100 (1998).
15. D.-Z. Chen and B.-S. Lee, *J. Electrochem. Soc.,* **146**, 744 (1999).
16. D.-Z. Chen and B.-S. Lee, *J. Electrochem. Soc.,* **146**, 3420 (1999).
17. F. W. Preston, *J. Soc. Glass Technol.,* **11**, 214 (1927).
18. J. Y. Lai, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA (Feb 2001).
19. L. A. Galin, in *Contact Problems in the Theory of Elasticity*, translated by H. Moss and I. N. Sneddon, Editor, North Carolina State College, Raleigh, NC (1961).
20. K. L. Johnson, *Contact Mechanics*, Cambridge University Press., Cambridge (1985).
21. R. Holm, *Electric Contacts*, Almqvist and Wiksells, Stockholm (1946).
22. A. J. G. Archard, *J. Appl. Phys.,* **24**, 981 (1958).
23. J. Goddard and H. Wilman, *Wear,* **5**, 114 (1962).
24. H.-C. Sin, N. Saka, and N. P. Suh, *Wear,* **55**, 163 (1979).
25. K. Komvopoulos, N. Saka, and N. P. Suh, *ASME J. Tribol.,* **107**, 452 (1985).