The Effect of Pad-Asperity Curvature on Material Removal Rate in Chemical-Mechanical Polishing

Sanha Kim\textsuperscript{a,*}, Nannaji Saka\textsuperscript{a}, Jung-Hoon Chun\textsuperscript{b}

\textsuperscript{a}Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA

* Corresponding author. Tel.: +1-617-715-4783; fax: +1-617-253-1556. E-mail address: sanhkim@mit.edu

Abstract

In chemical-mechanical polishing (CMP), surface asperities of the polishing pad play a key role, for they transmit normal force and impart tangential motion to the hard, nano-scale abrasive particles in the slurry. It has been shown recently, however, that the soft pad asperities themselves often generate micro-scale scratches on the surfaces being polished. To mitigate scratching by pad asperities, therefore, topography control by flattening pad asperities has been proposed and experimentally validated. In this study, the effects of asperity-flattening on pad topography and the material removal rate are investigated. It is found both theoretically and experimentally that even at a relatively high pressures only the tallest of the asperities are flattened and the ratio of asperity radius-to-standard deviation of heights is increased, but the average roughness itself is little affected. Specifically, surface profiles of new and asperity-flattened pads indeed show that the average roughness of about $5 \mu m$ is changed less than ten percent. Concurrently, the material removal rate is increased by about 30 percent due in part to the increased real area of contact – the result of increased asperity radius of curvature and decreased standard deviation of asperity heights.

1. Introduction

Over the decades, the semiconductor industry has responded to the ever-increasing demand for high-performance ultra-large-scale integrated (ULSI) electronics by designing and fabricating submicron features of finer resolution, denser packing, and multi-level structures. In this relentless endeavor for meeting the ever-stringent specifications, the Chemical-Mechanical Planarization or Polishing (CMP) process has played a vital role due to its global and local planarization capabilities. The CMP process is now ubiquitous in many stages of the manufacture of ULSI circuits [1].

One of the primary focuses in CMP is the material removal rate (MRR). A phenomenological model for MRR in glass polishing was proposed by Preston as early as in 1927 [2]. It was empirically found that the rate of thickness change, $\frac{dh}{dt}$, is directly proportional to the applied nominal pressure, $p$, and the relative velocity, $v_r$, as

$$\frac{dh}{dt} = k_p v_r,$$  \hspace{1cm} (1)

where $k_p$ is a constant, now designated Preston constant. This equation represents the minimum number of variables required to describe the CMP process. Although the above relation has been experimentally validated adequately, the effects of several process variables are not explicit and are hidden in $k_p$. Several papers have been published recently to elucidate the role of the other process parameters on MRR [3-5].

As demands on metal interconnects and surface structures are becoming ever-stringent, however, micro- and nano-scale scratching has lately emerged as a critical problem in integrated circuit (IC) manufacturing due both to the shrinkage of the feature size and to the prevalence of low-$k$ dielectrics [6,7]. During the polishing process, scratches may
be generated either by abnormally large particles, due to agglomeration, or by pad-asperities, due to the height variation and high friction [8,9]. To optimize the CMP process, therefore, not only the improvement of MRR but also the mitigation of scratching must be addressed.

Recently, pad topography control by asperity-flattening has been introduced as a cost-effective method for mitigating pad scratching in CMP [10,11]. Contact mechanics models and scratching experiments have shown that pad scratching can be minimized by flattening the asperities, i.e., by increasing the radius of tall asperities and reducing their height variation. Topographical changes may also affect MRR, for topography plays a key role in material removal. The preliminary polishing experiments with topography-controlled pads have shown that MRR increases as the asperities are flattened [10].

The objective of this study, accordingly, is to investigate the effects of asperity-flattening on pad topography and on MRR. First, explicit equations for real contact area between a rough pad and a flat wafer are presented based on multi-asperity contact models. Then, the surface roughness, asperity radius and standard deviation of asperity heights of asperity-flattened pads are determined and used to predict the contact area ratio. Finally, the MRR results of polishing experiments employing the topography-controlled pads are compared with the theoretical predictions.

2. Contact area ratio and material removal rate

To elucidate the material removal and scratching mechanisms in CMP, interaction between the pad and the wafer should be examined at different scales: macro-, micro- and nano- scales, Fig. 1. The surface layer, oxidized by slurry chemicals, is polished primarily by the nano-sized particles entrapped at the pad asperity contacts, Fig. 1c. From the microscopic point of view, however, only a small fraction of the pad surface will be in contact with the wafer due to the relatively large roughness of polishing pads, Fig. 1b. The polishing rate, accordingly, strongly depends on the ratio of real contact area to the nominal contact area, which generally is less than a percent.

Assuming that particle spacing, $\lambda_p$, and removal rate by each particle, $(dV/dt)_p$, are uniform everywhere in the asperity contact, the wear rate of the surface layer, $dh/dt$, can be written as [4,5]:

$$\frac{dh}{dt} = \frac{N_{max}}{A_{p}} \left( \frac{dV}{dt} \right)_p = \frac{1}{\lambda_p^2} \left( \frac{dV}{dt} \right)_p A_{p}$$

Thus the material removal rate will be proportional to the ratio of the real and nominal contact areas, $A_r/A_n$ [12].

![Fig. 1. Schematic of the contact between a rough pad and a smooth wafer.](image)

When an asperity is pressed against a smooth flat surface, it experiences mainly three distinct deformation modes as the approach of distant points, $\delta$, increases: elastic ($0 \leq \delta \leq \delta_y$), elastic-plastic ($\delta_y \leq \delta \leq \delta_{f-p}$), and fully-plastic ($\delta_{f-p} \leq \delta$). For a homogeneous and isotropic material with Young’s modulus, $E_a$, yield strength, $\sigma_{y,a}$, and hardness, $H_a$, the contact area and pressure at an asperity with radius $R_a$ can be expressed as a function of $\delta$ as [13,14]:

### Nomenclature

- $A_a$: asperity contact area [m$^2$]
- $A_n$: nominal contact area [m$^2$]
- $A_r$: real contact area [m$^2$]
- $d$: separation distance [m]
- $k_p$: Preston constant [N$^{-1}$ m$^2$]
- $E_a$: Young’s modulus of pad asperity [N m$^{-2}$]
- $H_a$: hardness of pad asperity [N m$^{-2}$]
- $h$: surface layer thickness [m]
- $p$: applied pressure [N m$^{-2}$]
- $p_a$: applied pressure at an asperity [N m$^{-2}$]
- $R_a$: asperity radius [m]
- $t$: process time [s]
- $V$: volume of surface layer removed by particles [m$^3$]
- $v_r$: relative velocity [m s$^{-1}$]
- $z_a$: asperity height [m]
- $\delta$: approach of distant points [m]
- $\delta_y$: approach of distant points at the onset of asperity yielding [m]
- $\delta_{f-p}$: approach of distant points at the onset of fully-plastic deformation [m]
- $\lambda_p$: particle spacing [m]
- $\bar{z} = (z_a - d) / \sigma_z$
- $\sigma_{y,a}$: yield strength of asperity [N m$^{-2}$]
- $\sigma_z$: standard deviation of asperity heights [m]
- $\psi$: plasticity index
- $\phi(z_a)$: probability density of asperity height [m$^{-1}$]
where \( \delta_y \) can be estimated by [15]

\[
\delta_y = \frac{4}{3\pi} E_y \frac{d}{R_y} \quad \text{if } 0 < \delta < \delta_y
\]

and \( \delta_{f-p} \) is assumed to be \( \delta_{f-p} = 54 \delta_y \) [12,15].

Assuming further that the heights of pad asperities are exponentially distributed [16], the probability density function of asperity heights, \( \phi(z_a) \), can be written as

\[
\phi(z_a) = \frac{1}{\sigma_z} \exp \left( - \frac{z_a}{\sigma_z} \right)
\]

As \( n \) asperities per unit nominal contact area are pressed against a smooth, flat surface layer, only the asperities taller than the separation distance, \( d \), will be in contact. The ratio of real contact area to nominal contact area, \( A_r/A_n \), between the pad and layer surfaces is the sum (or the integral) of individual asperity contributions. Thus

\[
A_r = n \int_{\delta_y}^{\infty} A_n \phi(z_a) dz_a
\]

For a given separation distance \( d \), from Eqs. (3) and (6), and introducing \( \xi = (z_a - d)/\sigma_z \), the contact area ratio can be expressed as

\[
A_r = \frac{\pi R \delta}{2\pi R \delta} + 3 \left( \frac{d/\delta - 1}{53} \right) - 2 \left( \frac{d/\delta - 1}{53} \right)^2
\]

where \( f_4 \) is a function of \( \psi \):

\[
A_r = \frac{\pi R \sigma \exp \left( - \frac{d}{\sigma} \right) f_4(\psi)}{A_n}
\]

which indicates that the contact area ratio depends on the plasticity index, \( \psi \), in addition to the normalized nominal pressure, \( p/H_o \).

The plasticity index was first introduced by Greenwood and Williamson combining the topographical and mechanical properties of surfaces [17]. This dimensionless parameter characterizes the relative proportion of plastically deformed asperities in contact: higher plasticity index indicates greater proportion of plastic asperity contacts. Only when the index is much less than unity, can all asperities be assumed to deform elastically. If \( \psi \ll 1 \), Eq. (13) simplifies to

\[
A_r = \pi R \sigma \exp \left( - \frac{d}{\sigma} \right) f_4(\psi)
\]

which is the Greenwood-Williamson model. For such elasto-plastic contact, Young’s modulus of asperities governs the mechanical behavior of asperity deformation. On the other hand, if \( \psi \gg 1 \), the equation simplifies to

\[
A_r = \frac{\pi R \sigma \exp \left( - \frac{d}{\sigma} \right) f_4(\psi)}{H_o}
\]

which is the case of fully-plastic deformation. For such plasticity-dominant contact, the real contact area is determined by the asperity hardness and by the applied pressure. If \( \psi \) is near unity, which is the case of typical CMP pads [10], however, the asperity contacts comprise both elastically and plastically deformed asperities. Fig. 2 shows the contact area ratio versus the normalized pressure, \( p/H_o \).
for different plasticity indices, $\psi$. Therefore, Eq. (2) is consistent with the Preston equation, Eq. (1), as the contact area ratio is proportional to the applied pressure and as the removal rate per particle is proportional to the relative velocity. Material removal rate, accordingly, will increase if the contact area ratio between the pad and wafer increases.

3. Pad topography control by asperity-flattening

The area ratio of rough pads in contact with flat wafer surfaces can be increased by applying higher pressure, or by using softer pads. In addition, the real contact area can also be increased by decreasing $\psi$, Eq. (10): i.e., either by increasing $H_a/E_a$ or by increasing $R_a/\sigma_z$. The Young’s modulus and the hardness of pad asperities, $E_a$ and $H_a$, are difficult to change independently, whereas the topographical parameters, $R_a$ and $\sigma_z$, are relatively easy to control. Recently, a simple process was introduced to increase $R_a/\sigma_z$, i.e., by flattening the asperities using a smooth, flat plate or a smooth roller [10,11]. By pressing a flat metal plate at high pressure or by rolling/sliding a smooth metal roller, the radius of the tall asperities can be increased and the height variance reduced, thus increasing the value of $R_a/\sigma_z$. Although flattening requires much higher pressure (0.1 - 5 MPa) than the polishing pressure (10 - 50 kPa) typical pads, which have asperity hardness of about 100 MPa, will still contact but a few percent or less than the nominal contact area during flattening, Fig. 2. Therefore, only a few tall asperities will be flattened. It may be noted, however, that the compressed asperities may not be flat since flattened asperities may spring back partially by elastic recovery [18].

Asperities of circular disks, 20 mm in diameter, of the IC1000 pad were flattened by sliding a stainless steel cylinder, 4.8 mm in diameter, over the specimens at an average pressure of 2.3 MPa. The topographical properties, $R_a$ and $\sigma_z$, and their ratio of new and flattened IC1000 pads, were determined from surface profiles obtained by a stylus profilometer, Fig. 3, and listed in Table 1. Then, by Eq. (13), the contact area ratios for the pads with different $R_a/\sigma_z$ were estimated. A modulus-to-hardness ratio of 7.6 was used to estimate the plasticity indices. For selected applied pressures and constant mechanical properties of the pad, the real contact area increases as $R_a/\sigma_z$ increases, Fig. 4, and therefore MRR is expected to be greater.

Table 1. Asperity radius, standard deviation of asperity heights, their ratio, and plasticity index of new and asperity-flattened IC1000 pads. All values are in $\mu$m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>New (as-received)</th>
<th>Asperity-flattened</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_a$ ($\mu$m)</td>
<td>23.5</td>
<td>72.7</td>
</tr>
<tr>
<td>$\sigma_z$ ($\mu$m)</td>
<td>4.4</td>
<td>3.3</td>
</tr>
<tr>
<td>$R_a/\sigma_z$</td>
<td>5.3</td>
<td>22.0</td>
</tr>
<tr>
<td>$\psi$</td>
<td>4.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

4. Polishing experiments and results

Cu-coated wafers were polished on a face-up polisher [10]. In experimental set 1, a slurry comprising 5 vol. % of Al$_2$O$_3$ abrasives of average size 300 nm was used. The polishing pressure was 13 kPa (2 psi) and the velocity was 0.87 m/s. In experimental set 2, a commercial slurry (HS-BT815, Hitachi Chemical Co.) was used. The pressure was 7 kPa (1 psi) and the velocity was 0.66 m/s. Table 2 lists the estimated contact area ratios from Eq. (13) and the results of the polishing experiments using new and asperity-flattened pads for the two different experimental conditions. The multi-asperity contact model predicts that the contact area ratio increases by about 37 and 36 percent, respectively. The polishing rates with the asperity-flattened pads indeed show 33 and 32 percent increase, respectively, compared with those using the new pads.
to reduce scratching in CMP, furthermore, the CMP practice adopts a method called “breaking-in”. Before polishing wafers a new pad is “broken-in” by polishing about 50 Cu-coated wafers while continuously roughening the surface using a diamond conditioner. About one to four hours are required to break in a new pad. It was experimentally determined that as more Cu wafers are used, and thus longer the “break-in” is, MRR would be stable and scratching decreased [19,20]. Therefore, both time and expensive consumables, such as wafers and slurry, are wasted in the “breaking-in” process. The major reason why the industry accepts this inefficient process is the lack of clear understanding of the phenomenon of pad scratching. Table 3 shows the \( R_z / \sigma_z \) values of pad asperities after typical “breaking-in” process used in the semiconductor industry compared to those of new pads as received from pad manufacturer. The process indeed increases the \( R_z / \sigma_z \), so that the “broken-in” pad can reduce scratching. However, since the process is not optimized for increasing \( R_z / \sigma_z \), it is inefficient and expensive.

Flattening the asperities, accordingly, by pressing a smooth, flat plate against, or by rolling/sliding a smooth roller over, the rough pad, accordingly, can be suggested as an alternative, inexpensive process to control the pad topography. Higher \( R_z / \sigma_z \) value can be achieved in much less process time, about 1 minute or less, compared with the current “break-in” process which in general takes one to four hours.

### Table 3. Asperity radius, standard deviation of asperity heights, their ratio, and the estimated plasticity indices of new and “broken-in” IC1000 pads.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>New pad</th>
<th>“Broken-in” pad</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_z (\mu m) )</td>
<td>23.5</td>
<td>53.8</td>
</tr>
<tr>
<td>( \sigma_z (\mu m) )</td>
<td>4.4</td>
<td>3.2</td>
</tr>
<tr>
<td>( R_z / \sigma_z )</td>
<td>5.3</td>
<td>16.8</td>
</tr>
<tr>
<td>( \Psi )</td>
<td>4.0</td>
<td>2.4</td>
</tr>
</tbody>
</table>

A greater challenge in pad topography control, however, is the maintenance of the enhanced \( R_z / \sigma_z \) value through the pad life. During the polishing of wafers, the pad asperities are worn out by the hard abrasives in the slurry and thus the \( R_z / \sigma_z \) will also continuously change. Locating a loaded roller between the conditioner and the wafer, as in Fig. 5, can be introduced for in-situ flattening while polishing, although more investigation is required to be conceived and tested for practical implementation.

Concerns may arise on the “asperity-flattening” process, as the process can affect the average roughness of the pad surface. Reduction in average roughness is known to reduce MRR due to the constriction of slurry flow from the edge to the center of the wafer or due to the hydroplaning [21-24]. It may be noted, however, that although lower surface roughness of the pad might result in smaller standard deviation of asperity heights, large asperity radius does not necessarily require low surface roughness. A pad surface can have a high \( R_z / \sigma_z \) and high average roughness simultaneously by having large average radius and large average height but with small variance. Flattening the tall asperities by compression or by rolling/sliding, can achieve higher polishing rate by decreasing the plasticity index without much decreasing in average roughness.
roughness appreciably. Shown in Table 4 are the roughness parameters of new and asperity-flattened pads. Because only the tall asperities are flattened, the decrease in average roughness of the pad surface is less than ten percent, whereas the decrease in the maximum peak height is over 30 percent.

Table 4. Surface roughness of new and asperity-flattened IC1000 pads. All values are in µm.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>New (as-received)</th>
<th>Asperity-flattened</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ra (arithmetic average)</td>
<td>6.18</td>
<td>5.52</td>
</tr>
<tr>
<td>Rq (root mean square)</td>
<td>8.24</td>
<td>7.24</td>
</tr>
<tr>
<td>Rp (maximum peak height)</td>
<td>22.72</td>
<td>15.50</td>
</tr>
<tr>
<td>Rv (maximum valley depth)</td>
<td>31.05</td>
<td>30.81</td>
</tr>
<tr>
<td>Rt (=Rp – Rv)</td>
<td>53.78</td>
<td>46.30</td>
</tr>
</tbody>
</table>

6. Conclusion

In this study, the effects of pad topography on the material removal rate in CMP have been investigated. Contact mechanics models predict that the material removal rate can be improved by increasing the ratio of asperity radius and the standard deviation of asperity heights, $R_r/\sigma_r$, which results in an increase in real area of contact. Asperity-flattening process significantly enhances the $R_r/\sigma_r$ while not reducing the average roughness appreciably. Cu polishing experiments using asperity-flattening process indeed showed an increase in polishing rate by about 30 percent, which agrees with theoretical predictions. Further improvements in MRR can be achieved by optimizing the asperity-flattening process.

Acknowledgments

The present study was funded by the Samsung Electronics Company, Ltd.