An Analysis of Potential 450 mm Chemical-Mechanical Planarization Tool Scaling Questions

L. Borucki, A. Philipossian, Araca Incorporated
M. Goldstein, Intel Corporation
Outline

• In idealized CMP tool and some easy scaling rules.
• Slurry flow rate scaling and application to 450 mm.
• Slurry flow simulation results for 200, 300 and 450 mm.
• Thermal simulation results for 200, 300 and 450 mm.
• Summary and conclusions.
理想的旋转CMP工具

Ground rules for process transfer to a larger size wafer on a larger tool:

- Identical polishing pressures.
- Identical pad/wafer relative sliding speed.

Since the distance between the wafer center and platen center must increase with wafer size, the platen rotation rate must decrease in order to maintain constant sliding speed.

What is an appropriate scaling law for the slurry flow rate?
Consumables

**Pad:**

<table>
<thead>
<tr>
<th>Platform</th>
<th>Pad Diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 mm</td>
<td>20&quot; (508 mm)</td>
</tr>
<tr>
<td>300 mm</td>
<td>30&quot; (762 mm)</td>
</tr>
<tr>
<td>450 mm</td>
<td>43&quot; (1092 mm), the current proposed diameter</td>
</tr>
</tbody>
</table>

IC plain pad
Thickness 0.082" (2.08 mm).
Mechanical properties: $\nu=0.25$, $E=272$ MPa, extracted from topo and contact data.
Summit height distribution extracted from conditioned pad surface topography data.

![Pearson fit](image)

Summit density 52.1/mm$^2$
Mean summit curvature 0.564 $\mu$m$^{-1}$
Greenwood and Williamson pressure-displacement coefficient $G=2.68\times10^{13}$ Pa/m$^{3/2}$

*Note:* Pad mechanical properties and summit data determine the mean pad/wafer gap and the wafer pitch and bank. For this pad, the mean gap is 14.06 $\mu$m, the pitch is $2.350\times10^{-6}$ (1.3x10$^{-4}$ degrees) and the bank is nearly 0.
Operating Conditions

**Pressures:**
- Wafer: 2 PSI
- Ring: 4 PSI

**Sliding Speed:** The sliding speed is held constant at $V=1.00$ m/sec.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Platen and Head Rotation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 mm</td>
<td>68.7 RPM</td>
</tr>
<tr>
<td>300 mm</td>
<td>47.0 RPM</td>
</tr>
<tr>
<td>450 mm</td>
<td>33.4 RPM</td>
</tr>
</tbody>
</table>

**Polishing Time:** 60 sec

**COF:** 0.35 for the wafer and 0.55 for the PEEK retaining ring.
The Thin Film Equation

Local slurry film thickness
\[
\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{v}) = 0
\]

- Local mean slurry velocity.
- Includes:
  - Gravity-drive flow
  - Centripetal acceleration
  - Surface tension
  - Direct transport by pores
  - Surface obstruction of pressure-driven flow

- Mass conservation law.
- Derived from the incompressible Navier-Stokes equations.
- Rigorously valid when there are no grooves.
- Applies outside of the retaining ring and wafer.
Thin Film Equation Boundary Conditions

\[-h < \vec{v} > \cdot \vec{n} = \frac{f_s}{2\pi r_s}\]

- $h$ is the flow rate
- $r_s$ is the stream radius

If there are $N$ injection points with identical flows and the total flow rate is $F$, then

\[f_s = \frac{F}{N}\]
Scaling the Thin Film Equation

General features common to all tool sizes can be understood by scaling the tool model

- Scale horizontal lengths by the platen radius $R$.
- Scale the time using the platen rotation rate $\Omega$ in radians per second.
- Scale the mean flow velocity by $R\Omega$. The maximum scaled platen speed is then 1.
- Scale the fluid film thickness by the pad surface height standard deviation $\sigma$.

The scaled thin film equation:

$$\sigma \Omega \frac{\partial \tilde{h}}{\partial \tilde{t}} + \frac{1}{R} \tilde{V} \cdot (\sigma \tilde{h} R \Omega \langle \tilde{v} \rangle) = 0$$

After cancellation:

$$\frac{\partial \tilde{h}}{\partial \tilde{t}} + \tilde{V} \cdot (\tilde{h} \langle \tilde{v} \rangle) = 0$$

The dimensionless conservation equation thus has the same form as the dimensional equation and is solved on a dimensionless domain of radius 1 with a rotation rate of 1 regardless of the dimensional pad radius and rotation rate.
Scaling the Injection Boundary Conditions

Apply the same scaling to the boundary condition. The stream radius is a horizontal length, so

\[-\sigma \tilde{h} R \Omega \mathbf{<} \tilde{v} \mathbf{>} \cdot \mathbf{n} = \frac{F}{N2\pi \tilde{s} R}\]

or

\[-\tilde{h} \mathbf{<} \tilde{v} \mathbf{>} \cdot \mathbf{n} = \frac{F}{N2\pi \tilde{s}(\sigma R^2 \Omega)}\]

The injection boundary conditions are the same for all platform sizes if the total slurry flow rate \(F\) is scaled by \(\sigma R^2 \Omega\)
**Flow Rate Scaling**

If the pad and wafer co-rotate, this implies that if the pad radius is increased from $R_0$ to $R$ and the wafer center is moved outward from $d_0$ to $d$, then the rotation rate $\Omega$ should be selected so that

$$d\Omega = d_0\Omega_0$$

where $\Omega_0$ is the rotation rate of the smaller tool. The scaling law for the flow rate is then

$$\frac{F}{F_0} = \frac{\sigma R^2 \Omega}{\sigma R_0^2 \Omega_0}$$

$$\frac{F}{F_0} = \frac{R^2}{R_0^2} \frac{d_0}{d}$$

Rather than just scaling the flow rate by the pad area ratio $A/A_0 = (R/R_0)^2$, this scaling law takes into account that the sliding speed is held constant. The following table compares the flow rate scaling without and with the sliding speed correction.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Area Scaling</th>
<th>Area + Sliding Speed Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 mm</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>300 mm</td>
<td>338</td>
<td>231</td>
</tr>
<tr>
<td>450 mm</td>
<td>693</td>
<td>339</td>
</tr>
</tbody>
</table>
Slurry Film – 200 mm
Slurry Film – 300 mm
Slurry Film – 450 mm
During the first platen rotation, the volume of slurry on the pad rapidly decreases due to loss from the bow wave. At steady state, the volume on the pad is roughly proportional to the pad area.

The total time-integrated volume lost increases with time. At longer times, the slope of each volume loss curve is the same as the slurry application rate.
Each simulation starts with the same slurry layer thickness. During the first one to two pad rotations, the excess slurry is removed from the pad by the squeegee action of the pressure ring and by centripetal acceleration. All three tools have about the same mean slurry thickness at steady state, but the time to steady state increases as the platform size increases due to the decrease in platen rotation rate.
The minimum slurry thickness is determined by the pad properties and by the load, which is the same for all three tools. The increase in slurry thickness at the edge of the pad is caused by backflow from the bow wave.
The next three slides show the age distribution under the wafer. In all three cases, the area to the left is the backflow region.
Slurry Age Under the Wafer

300 mm
Slurry Age Under the Wafer

450 mm
This graph compares the mean slurry age under the wafer for all three platforms. The filled black circles are the means and the red squares are the ages at individual grid points. The grid points are visible in the previous three slides.

There is no significant difference in mean slurry age under the wafer between the three platforms.
Steady State Pad Temperature Rise

The pad temperature increase is predicted to be highest at the inside trailing edge of the ring. This happens because the ring has a higher coefficient of friction than the wafer and because points on the pad close to the inner edge spend a higher fraction of each rotation in contact.
Steady State Pad Temperature Rise

300 mm
Steady State Pad Temperature Rise

450 mm

Temperature Increase (°C)
Temperature Transient Comparison

The graph illustrates the temperature increase over time at different positions on a 200 mm wafer. The lines represent:
- **Leading**
- **Center**
- **Trailing**

The y-axis represents temperature increase (°C) and the x-axis represents time (sec). The inset image shows the wafer center and the leading and trailing edges, with a 1 cm mark for scale.
Temperature Transient Comparison
The 450 mm tool is predicted to have the lowest pad temperature. This is due to the lower platen rotation rate and higher slurry injection rate. This conclusion depends critically on polishing at the same speed on all three platforms.
The next three slides show the temperature increase on both the wafer and the pressure ring. The temperature distributions are viewed from above with the pad center to the left. Streaks in this graphic are caused by cool slurry from the injection points.
The ring has constant width and therefore is smaller relative to the wafer diameter as the platform size increases.
Wafer Temperature Comparison (60 sec)

The temperature distribution for 450 mm is remarkably similar to 300 mm.
Radial temperature increment averages (solid lines) and grid point temperatures are shown at the right. The ring on average is hotter than the wafer. Temperature scatter also increases toward the edge of the wafer due to the proximity of the injector to the leading edge.

The radial temperature distribution for 450 mm is a smooth extension of the distribution for 300 mm.

In all cases, the wafer is hotter in the center than at the edge.
The next slides compare the flash temperature increment at contacting pad summits. The flash increment is calculated with a transient finite element thermal model. Heat partitioning includes loss to the wafer and the slurry. We think that the flash increment is the main contribution to the reaction temperature, the other two components being the ambient temperature and the wafer body temperature. The previous wafer thermal slides show the body temperature.
The maximum flash temperature is a little higher for 300 mm than for 200 mm, but the maximum occurs in the ring.
The trend in the maximum continues for 450 mm.
One limitation of the flash estimates is that they depend on pad properties, so the absolute numbers may be different from pad to pad. By forming a ratio, this dependence is removed. This graph shows the ratio of the 450 mm flash increment to the 300 mm increment after scaling the wafers to the same size. The 450 mm flash increment is predicted to be 10-20 percent higher than the 300 mm increment.
Similarly, the model indicates that 200 mm flash temperatures are 10-20 percent lower than for 300 mm.
Summary and Conclusions

Three CMP tool sizes have been simulated under very similar conditions. The same IC pad thermal and mechanical properties and the same pad topography properties are used throughout. The applied wafer and ring pressures are all the same, as are the wafer/pad and ring/pad friction coefficients. It is assumed that the relative sliding speed between the pad and wafer is held constant, implying that the platen rotation rate must decrease with increasing separation between the wafer center and pad center.

Slurry is applied in all cases using a bar injector with a 1” injection point spacing, so that a 450 mm tool has more injection points than a 300 mm tool. The total slurry flow is equally divided between the injection points. We analyze the total slurry flow requirements using the thin film equation. This equation indicates that for constant pad surface roughness, the slurry flow rate should be scaled in proportion to the square of the pad radius divided by the distance from the wafer center to the pad center. This scaling is a rational method of approaching slurry use and provides significant savings over scaling by the pad area. The validity of the scaling depends on the assumption of constant sliding speed.

Under these conditions, we find that the slurry thickness behind the trailing edge of the wafer at steady state is similar for all three tools, being determined by the pad surface properties and by wafer and ring load and moment balance. The three tool sizes also have similar steady state mean slurry thicknesses. However, the transient time to reach steady state slurry flow increases as the platform size increases because of the decreasing platen rotation rate. The transient time in all cases is approximately the time for 1-2 pad rotations. While the distribution of slurry ages is complex due to multipoint injection, there is no significant difference in mean slurry age under the wafer with platform size.

Thermal simulations indicate that the wafer is cooler on average than the ring and that the maximum temperature occurs on the ring at the trailing edge close to the platen center. A comparison of transient temperatures at selected points at the pad leading edge, the pad trailing edge and the wafer center suggests that the 450 mm platform pad should have a marginally lower pad temperature. Radial comparisons of the wafer body temperature increment, however, indicate that 450 mm is almost identical to 300 mm at constant sliding speed and that in all cases the wafer is slightly hotter in the center than at the edge. The flash temperature increment at pad summits that contact the wafer and ring is found to increase by 1-2 C (or 10-20%) for each increase in platform size.