Study on Stiffness and Conditioning Effects of CMP Pad Based on Physical Die-Level CMP Model

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1149/1.3369963">http://dx.doi.org/10.1149/1.3369963</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Electrochemical Society</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Mon Feb 11 23:29:34 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/71165">http://hdl.handle.net/1721.1/71165</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Study on Stiffness and Conditioning Effects of CMP Pad Based on Physical Die-Level CMP Model

Wei Fan,a,z Duane Boning,a,* Leslie Charns,b Hiroyuki Miyachi,c Hiroyuki Tano,c and Shoel Tsujic

aDepartment of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
bJSR Micro Incorporated, Sunnyvale, California 94089, USA
cJSR Corporation, Yokkaichi Research Center, Yokkaichi 510-8552, Japan

Chemical mechanical planarization (CMP) pad stiffness and conditioning effects were evaluated based on a physical die-level CMP model, with pad effective modulus and asperity height as parameters. In one study, patterned dielectric wafer were polished using polymeric pads of different stiffnesses. In a second study, wafers were polished by standard pads using different conditioning disks. Polishing experimental data (dielectric thickness and step height) were fitted by the physical model, enabling the extraction of the pad effective modulus and asperity height model parameters. A higher pad stiffness gives better within-die uniformity, and the conditioning disk with blocky diamonds achieves up-area only polishing for longer times. Polishing simulations using the physical model reflect a clear pattern density dependence.

© 2010 The Electrochemical Society. [DOI: 10.1149/1.3369963] All rights reserved.

Manuscript submitted December 3, 2009; revised manuscript received March 1, 2010. Published April 1, 2010.

Physical Die-Level CMP Model

Our physical die-level CMP model describes the dependence of pressures on the local pattern density and surface step height of a single die. It assumes that the elastomeric polishing pad can be decomposed into bulk and asperity regions, as shown in Fig. 1. The bulk material can be treated as an elastic body, deforming in response to long-range wafer height differences. The surface asperities come in contact with the wafer surface, and the compression of the asperities depends on both the wafer surface profile and the pad bulk bending. Figure 2 illustrates the model framework and the coordinate system. The wafer is assumed to press into the polishing pad in the positive Z direction. For convenience, the surface normal of the wafer is taken as the positive Z direction, corresponding to the conventional “wafer face up” mathematical representation. Here, w(x,y) is used to describe the z-coordinate of the normal separation point between the bulk and the asperities of the pad, and the wafer surface. The distance between the wafer surface and the nominal bulk pad position is w(x,y) − z(x,y).

Modeling of bulk.— The bulk is elastic and can be modeled using a contact wear model.5,7 The relationship between the bulk surface displacement w(x,y) and pressure P(x,y) satisfies the following convolution

\[ w(x,y) - w_0 = F(x,y) \otimes P(x,y) \]  \[ \text{[1]} \]

where \( F(x,y) \) is the deformation response to a point pressure. This deformation is proportional to \( 1/E \), where \( E \) is the effective modulus of the bulk pad. Here, \( w_0 \) is the relative reference plane of the bulk surface when there is no pressure applied. For mathematical convenience, \( w_0 \) is set to zero. The boundary condition applied to Eq. 1 is

\[ \frac{1}{2} \int_S P(x,y) dxdy = P_0 \]  \[ \text{[2]} \]

where \( P_0 \) is the reference pressure applied to the bulk and \( S \) is the contact area. Equation 2 indicates that the bulk response force is equal to the force applied from the outside.

Modeling of asperities.—The asperities are assumed to have a negligible width and an exponential height distribution.8 Equation 3 defines the distribution with a characteristic asperity height \( \lambda \)

\[ l(\xi) = \frac{1}{\lambda} e^{-\xi/\lambda} \]  \[ \text{[3]} \]

where \( \xi \) is the asperity height. At location \( (x,y) \), the distance between the wafer surface \( z(x,y) \) and the nominal plane \( w(x,y) \) is
as shown in Fig. 2, Eq. 4 implies that the up-area pressure is given in Eq. 3, known, and it is a strictly decreasing function and approaches zero to the compressed amount. The expected value of the exerting force is proportional to the compressed amount. The expected value of $P(x,y)$ can be estimated by averaging across all the asperities as follows

$$P(x,y) = \int_{w(x,y)=z(x,y)}^{w(x,y)} k(\xi - [w(x,y) - z(x,y)])h(\xi)d\xi$$

$$= k\Psi[w(x,y) - z(x,y)]$$

where $k$ is a spring constant and $\Psi(z)$ is a derived asperity height distribution function, defined as $\Psi(z) = \int_{w(x,y)=z(x,y)}^{w(x,y)} h(\xi)d\xi$. $\Psi(z)$ can be calculated once the probability distribution of asperity height is known, and it is a strictly decreasing function and approaches zero at infinity. Because we assume the asperity height distribution as given in Eq. 3, $\Psi(z)$ is given by

$$\Psi(z) = k e^{-z/\lambda}$$

When a feature of step height $h(x,y)$ is pressed against the pad, as shown in Fig. 2, Eq. 4 implies that the up-area pressure is

$$P_u(x,y) = k\Psi[w(x,y) - z_u(x,y)]$$

and the down-area pressure is

$$P_d(x,y) = k\Psi[w(x,y) - z_d(x,y) + h(x,y)]$$

The total local pressure is the sum of the two pressures weighted by pattern density $\rho(x,y)$, which is the area fraction of the up area. So, we have

$$P(x,y) = \rho(x,y)P_u(x,y) + [1 - \rho(x,y)]P_d(x,y)$$

$$= \rho(x,y)k\Psi[w(x,y) - z_u(x,y)] + [1 - \rho(x,y)]k\Psi[w(x,y)$$

$$- z_d(x,y) + h(x,y)]$$

Combining Eq. 5-8, we relate the pressures to the step height and characteristic asperity height as follows

$$P(x,y) = k[\rho(x,y)\rho(x,y)P_u(x,y) + [1 - \rho(x,y)]P_d(x,y)$$

$$= k[\rho(x,y)e^{-h(x,y)/k}\lambda e^{-w(x,y)}/H20849]$$

$$P_d(x,y) = \frac{1}{1 + \rho(x,y)[e^{-h(x,y)/k} - 1]}P(x,y)$$

$$P_d(x,y) = \frac{e^{-h(x,y)/k}}{1 + \rho(x,y)[e^{-h(x,y)/k} - 1]}P(x,y)$$

**Modeling of CMP process.** — The physical CMP model can be obtained by combining the effects of the two parts above together: the elastic pad bulk, which is described by Eq. 1, and the asperities with exponential height distribution, which is described by Eq. 9. The pressure and deflection interactions between wafer surface topography and CMP pad are therefore described by

$$P(x,y) = k[\rho(x,y)\rho(x,y)P_u(x,y) + [1 - \rho(x,y)]P_d(x,y)$$

$$w(x,y) = [F(x,y) + P(x,y)] + w_0$$

To run the model, pattern density $\rho(x,y)$ needs to be extracted from the die layout. With initial values of step height and up-area coordinate, the two unknowns $P(x,y)$ and $w(x,y)$ can be calculated by solving Eq. 10 iteratively. Once $P(x,y)$ is solved, $P_u(x,y)$ and $P_d(x,y)$ can be obtained by Eq. 9. Then, the model employs Preston's equation with local pressure $P(x,y)$ to calculate the instantaneous material removal rates of the up and down areas as follows

$$\frac{dz(x,y)}{dt} = -K_pP(x,y)u$$

where $K_p$ is the Preston coefficient and $u$ is the linear velocity of the pad relative to the wafer. So, $z_u(x,y)$ and $h(x,y)$ are dynamically updated during the simulation of CMP.

**Model parameters.** — There are three key parameters in the physical die-level CMP model: blanket removal rate $K_{BP}$, effective modulus $E$, and characteristic asperity height $\lambda$. The blanket removal rate is affected by many CMP tools, consumable, and process parameters, such as the CMP system reference pressure. The effective modulus is related to the properties of the pad bulk and is hypothesized to most strongly impact the feature scale step height reduction. The pad property related parameters, $E$ and $\lambda$, are independent of each other; in the model we assume the CMP pad can be divided into two parts, pad bulk and asperities. $E$ is only affected by the pad bulk stiffness, while $\lambda$ is only affected by the pad surface asperity height distribution. Thus, extracting these two independent parameters from experimental data enables separate investigations of the pad bulk and surface properties.
Experimental

To evaluate the effect of pad stiffness, three water soluble particle (WSP) pads\textsuperscript{10-12} (produced by JSR Corporation) were engineered to have different stiffnesses (low, standard, and high) and were used in patterned wafer polishing. In the conditioning disk effect study, three conditioning disks with different diamond shapes (sharp, standard, and blocky) were applied on JSR standard pads. These diamond shapes were expected to “cut” into the pad surface differently, generating different pad asperity surface structures.

SKW7-2 oxide wafers patterned with a Massachusetts Institute of Technology CMP test layout\textsuperscript{13} were polished using JSR WSP pads. We varied the pad stiffness and the conditioning disk diamond shape separately, while the CMP machine recipe was kept all the same. Each test wafer was polished under a reference pressure of 5 psi for different times accumulative to more than 2 min: 0, 20, 40, 70, 90, 110, and 130 s.

Optical measurements of film thickness and step height were performed after polishing. On each wafer, we selected a center die, a middle die, and an edge die to test, as shown in Fig. 3. Figure 4 shows the pattern type and local pattern density within each die of the SKW7-2 wafer. For each testing die, the oxide thickness and step height were measured on five points in the step blocks with local pattern densities of 30, 70, and 50% (Fig. 5). Oxide thickness data were used to extract the necessary model parameters and simulate polishing performance of each pad using the physical die-level CMP model. We extracted model parameters by fitting to experimental data and by minimizing the fitting error from time split experiment data. The hypothesized pad property effects are seen clearly in the pad parameter and model simulation results, as discussed below.

Results and Discussion

We applied two different sets of experimental conditions to emphasize either the pad bulk dependence or the pad surface and asperity structure dependence. The reference pressure $P_0$ is 5 psi for both sets. In the first set of experiments, we polished wafers using pads with different stiffnesses (low, standard, and high), while we used the same conditioning disk. In the second set of experiments, we polished wafers using a standard pad, while we applied three different conditioning disks (sharp, standard, and blocky). The first set investigates pad stiffness effect, and the second set verifies pad conditioning effects.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig3}
\caption{(Color online) Optical measured die positions on a SKW7-2 dielectric wafer.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig4}
\caption{(Color online) Pattern type and pattern density within a die of SKW7-2 wafer: (a) Layout of a die on SKW7-2 wafer. A “F” preceding a number indicates a pitch structure with 50% density, with the number following in micrometers. All other numbers are localized densities, with the number indicating the density. Density structures have a fixed 100 µm pitch. (b) Topography of the 70% STEP array in a die.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig5}
\caption{(Color online) Optical measurement positions in the STEP blocks in a die of SKW7-2 dielectric wafer. Dots indicate both thickness and step height measurements. Crosses indicate only step height measurement.}
\end{figure}
Within-die nonuniformity results from faster removal rates in low pattern density regions on the die compared to high pattern density areas: The local pad force is applied to a smaller number of up features in the low pattern density region, generating much higher local pressure and resulting in a large "recessed" or "eroded" region on the die. Across lateral distances of several millimeters, this within-die nonuniformity can be substantial, with oxide thickness differences of several hundred nanometers. To combat this effect, pads with a higher stiffness are of interest, as they are subject to less long-range bending. Across the oxide step structures, hard pads deflect less into recessed low pattern density areas, reducing the regional polish rates compared to high pattern density regions, resulting in a smaller final nominal range, as shown in Fig. 7. In this figure, the creation of within-die nonuniformity in the first 50–70 s due to pattern density removal rate dependence is seen, with subsequent reduction as polishing progresses. At all times, the harder pads create a less nominal range between 90 and 10% region thicknesses.

We can also understand the pad stiffness effect vs asperity height effect by comparing step height evolution in the polishing process. Figure 8a shows the data fitting results of step height vs time on the center point of a middle die in the 50% pattern density array. We see that the high stiffness pad maintains a linear step height reduction region to a comparable remaining step height at which point an exponential decay in time occurs across all three pads, corresponding to our expectation that bulk pad stiffness is not the primary factor in individual feature step height removal or down-area polish. Comparing the step height evolution at the center (Fig. 8a), left edge (Fig. 8b), and right edge (Fig. 8c) points of the 50% pattern density array on the middle die, however, we see the pattern density dependence of polishing as expected. As indicated by the straight dashed lines, step height reaches to 200 nm with the remaining height moving slowly (at about 95 s) at the left edge point (next to 70% pattern density array) and moving quickly (at about 66 s) at the right edge point (next to 10% pattern density array). This is the result of the long-range pad bending, which results in different “effective” pattern densities at different points within the same local pattern density region, due to spatial “averaging” of the applied pad pressure across the die.

**CMP pad conditioning effect**—JSR standard pads with different conditioning disk diamond types (blocky, standard, and sharp) were tested. Because we used the same type of pad, the extracted effective modulus is similar for each conditioning disk (Fig. 9a), as expected. Especially for the middle die, the modulus differences are very small (less than 2.5%). But the characteristic asperity height is consistently different (Fig. 9b); the characteristic asperity height varies according to the diamond shape. The blanket removal rate does not change substantially (Fig. 9c).

In the conditioning test, the simulated nominal range (or within-die oxide thickness uniformity) based on the extracted model parameters is not significantly impacted by the characteristic asperity height, and thus, the range evolutions are similar for different conditioning disks (Fig. 10). However, there is a difference in step

---

**Figure 6.** (Color online) Extracted model parameters for different pad stiffnesses: (a) Effective modulus, (b) characteristic asperity height, and (c) blanket removal rate.

**Figure 7.** (Color online) Simulated middle die nominal range evolution for different pad stiffnesses.
height evolution. Figure 11 shows the data fitting result of step height vs time of the middle die 50% pattern density array. The pattern density dependence of the polish is also verified by comparing the step height evolution at the center (Fig. 11a), left edge (Fig. 11b), and right edge (Fig. 11c) points of the 50% pattern density array on the middle die. The sharp diamond disk data go to a non-linear region slightly earlier (at a larger remaining step height) than the other disks. This transition in the linear step height reduction in time occurs when asperities start to polish both the up and down feature areas rather than just the up areas; we thus expect this transition to depend on the pad asperity height. A good way to understand the linear to nonlinear transition in step height is shown in Fig.
In the beginning of the process, the down-area removal amount is zero. When the step height reduction goes to nonlinear, down area removal amount starts to increase above zero. In Fig. 12, the sharp disk curve goes to above zero at the largest remaining step height among the three disks, while the blocky disk curve goes to above zero at the lowest remaining step height. The characteristic asperity heights vary about 15% for the three conditioning disks, so we see a clear relationship verifying the physical model sensitivity to pad properties.

Simulation of CMP die-level uniformity and end point.— CMP is used to planarize the wafer surface to remove topography resulting from previous patterning and deposition processes. In practice, perfect planarization is difficult to achieve due to the pattern density differences on the die, where planarization of low density areas is faster than that of high density areas. In general, the CMP process must be stopped once a certain end-point strategy is fulfilled. In one approach, the polishing should stop when the original step height has been reduced to some value that is “small enough” to meet product requirements, which is called the step height target strategy. Another commonly used approach is thickness target strategy, in which the process should stop when the up-area thickness has been reduced to some value as product requirements.

To understand the die-level uniformity and impact on the process end point, a full chip simulation (middle die of Fig. 3) was run for pad conditioning effects using the JSR standard pad. The end point was defined as occurring when a step height target of 100 nm at the 50% pattern density area is reached, i.e., the polishing stops when the remaining step height is 100 nm at the center point of the 50% pattern density area. Table I lists the simulation results for the different diamond conditions. We can see the differences in final within-die and step height uniformity with these numbers. While the 50% pattern density area reaches the end point, the 10% pattern density area is overpolished and the 90% pattern density area still has a remaining step height substantially larger than the end-point target. From the end-point time, we see that the standard pad with a blocky conditioning disk planarizes faster, which corresponds to Fig. 11a. Table I also tells us that within-die oxide thickness uniformity is not substantially impacted by conditioning disks because both the nominal range (difference between the up-area oxide thickness of the 90% pattern-density area and that of the 10% pattern-density area) and the full chip range (difference between the maximum up-area thickness and the minimum up-area thickness across the entire chip) are similar for the three different disks. The full chip simulation of the middle die surface at the end point for the JSR standard pad with a standard conditioning disk is shown in Fig. 13. We see that the uniformity is affected significantly by pattern density. Although the STEP-90 array has a 90% pattern density, its final thickness and step height are not high. The reason is that its neighboring arrays have lower pattern densities (10 and 30%). The end point of thickness target strategy was defined as occurring when an up-area thickness of 1000 nm at the 50% pattern density area is reached; i.e., the process stops, and then the remaining up-area
thickness is 1000 nm at the center point of the 50% pattern density area. Table II lists the simulation results of the thickness strategy for the different diamond conditioning disks. It is very easy to see the pattern-density-dependent within-die nonuniformity. The planarization efficiency of the blocky disk is higher than the other two because at the end point the remaining step height of the blocky disk is less than that of the others.

**Conclusion**

A physical die-level CMP model has been applied to the pad property and polish process analysis. The experimental results confirm the decomposition of the model between the pad bulk and pad asperity components. We are able to relate the pad stiffness and the conditioning disk effects to the model parameters and data fitting results. Quantitative relationships between pad properties and model parameters were established, which can be used to evaluate pad performance and to assist in pad design and optimization. Model prediction is the same as our expectation: A higher pad stiffness gives better within-die uniformity, and a conditioning disk with blocky diamonds keeps the linear polishing region longer. This physical CMP model correctly simulates pattern density dependence and within-die uniformity.

**Table I. Simulation results of step height target strategy for pad conditioning effects on JSR standard pad.**

<table>
<thead>
<tr>
<th>Conditioning disk diamond shape</th>
<th>Sharp</th>
<th>Standard</th>
<th>Blocky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial step height (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial up-area thickness (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End-point time (s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remaining step height (nm)</td>
<td>Center point of 10% pattern density area</td>
<td>114</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Center point of 50% pattern density area</td>
<td>36.5</td>
<td>34.5</td>
</tr>
<tr>
<td></td>
<td>Center point of 90% pattern density area</td>
<td>186.7</td>
<td>192.6</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>259.5</td>
<td>267.9</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>29.5</td>
<td>27.4</td>
</tr>
<tr>
<td>Nominal range (nm)</td>
<td>272.6</td>
<td>275.9</td>
<td>285.0</td>
</tr>
<tr>
<td>Full chip range (nm)</td>
<td>385.8</td>
<td>389.5</td>
<td>398.7</td>
</tr>
<tr>
<td>Up-area thickness (nm)</td>
<td>Center point of 10% pattern density area</td>
<td>1089.5</td>
<td>1097.6</td>
</tr>
<tr>
<td></td>
<td>Center point of 50% pattern density area</td>
<td>1237.2</td>
<td>1246.2</td>
</tr>
<tr>
<td></td>
<td>Center point of 90% pattern density area</td>
<td>1362.1</td>
<td>1373.4</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1447.3</td>
<td>1458.8</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>1061.4</td>
<td>1069.4</td>
</tr>
</tbody>
</table>

**Figure 12.** (Color online) Middle die 50% pattern density array center point step height vs down area removal amount for different conditioning disk diamond shapes. Take down area removal amount at 1 nm as the linear to nonlinear transition point. The remaining step heights of the blocky, standard, and sharp disks at the transition point are 436.0, 497.7, and 533.5 nm, respectively.

**Figure 13.** (Color online) Full chip simulation for the SKW7-2 test pattern at the time point when a 100 nm step height end point in the 50% pattern density region is reached. Simulation is for a JSR standard pad with a standard conditioning disk: (a) Up-area oxide thickness and (b) step height. The x-y coordinate corresponds to the layout shown in Fig. 4a.
Acknowledgments

The authors thank the GRC/SEMATECH Joint Engineering Research Center for Environmentally Benign Semiconductor Manufacturing for providing support for this work.

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

References


Table II. Simulation results of up-area thickness target strategy for pad conditioning effects on JSR standard pad.

<table>
<thead>
<tr>
<th>Conditioning disk diamond shape</th>
<th>Sharp</th>
<th>Standard</th>
<th>Blocky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial step height (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial up-area thickness (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End-point time (s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remaining step height (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal range (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full chip range (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up-area thickness (nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conditioning disk diamond shape</th>
<th>Center point of 10% pattern density area</th>
<th>Center point of 50% pattern density area</th>
<th>Center point of 90% pattern density area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>65.3</td>
<td>59.7</td>
<td>23.2</td>
</tr>
<tr>
<td>Minimum</td>
<td>5.0</td>
<td>3.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conditioning disk diamond shape</th>
<th>Center point of 10% pattern density area</th>
<th>Center point of 50% pattern density area</th>
<th>Center point of 90% pattern density area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>1079.6</td>
<td>1077.2</td>
<td>1078.7</td>
</tr>
<tr>
<td>Minimum</td>
<td>875.7</td>
<td>877.8</td>
<td>882.7</td>
</tr>
</tbody>
</table>