Task ID: 425.032
Task Title: Fundamentals of Advanced Planarization: Pad Micro-Texture, Pad Conditioning, Slurry Flow, and Retaining Ring Geometry

Deliverable: Report on the extended die-level model incorporating pad-micro-structure and slurry dependencies in chip-scale prediction of dishing and erosion across each die.

I. Summary/Abstract

We are investigating the effect of CMP pad properties on the resulting chip-scale planarity and uniformity. In our previous work, a physical die-level CMP model was developed based on contact mechanics to understand and optimize the planarization process. To understand the interactions between pad asperities and the wafer, a physically-based particle-level model is introduced to predict the contact area. Then an extended die-level model is proposed that integrates these two models together. The extended die-level model includes four main parameters we are interested in: pad bulk modulus, asperity modulus, asperity size and characteristic asperity height.

II. Technical Results and Data

Physical Die-Level CMP Model

Our physical die-level CMP model describes the dependence of pressures on 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 Figure 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 pad bulk bending.

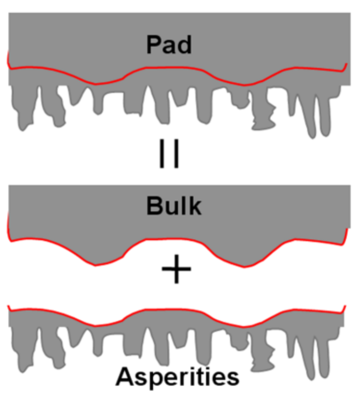


Figure 1: Pad structure assumption in physical die-level CMP model. The whole pad is comprised of bulk and asperities.

Figure 2 illustrates the model framework. The wafer is assumed to sit face down, and the wafer surface is pressed down onto the polishing pad. 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 nominal separation point between the bulk and asperities of the pad, and $z(x, y)$ is used to describe the wafer surface. The distance between the wafer surface and nominal bulk pad position is $w(x,y)-z(x,y)$.

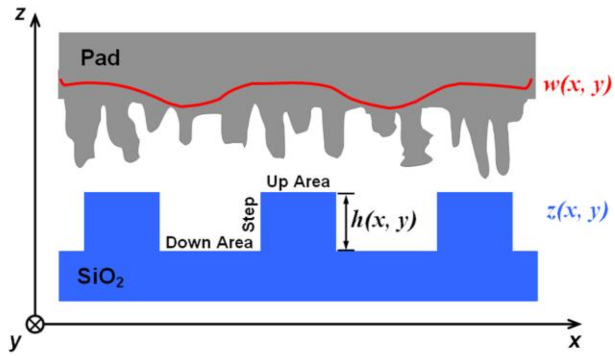


Figure 2: Framework of physical die-level CMP model.

Modeling of Pad Bulk

The bulk is elastic and can be modeled using a contact wear model [1, 2]. The relationship between bulk surface displacement $w(x,y)$ and pressure $P(x,y)$ satisfies the following convolution:

$$w(x,y)-w_0=F(x,y)\otimes P(x,y) \tag{1}$$

$F(x,y)=\frac{1}{\pi E_b}\int d\xi \int d\eta \frac{1}{\sqrt{(x-\xi)^2+(y-\eta)^2}}$ is the deformation response to a point pressure, where E_b is the effective modulus of the pad bulk. 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 Equation 1 is

$$\frac{1}{S}\int_S P(x,y)\cdot dx\cdot dy=P_0 \tag{2}$$

where P_0 is the reference pressure applied to the bulk and S is the chip area. Equation 2 indicates that the bulk response force equals the force applied from outside.

Modeling of Asperities

The asperities are assumed to have a fixed width and an exponential height distribution [3]. Equation 3 defines the distribution with characteristic asperity height λ,

$$l(\xi)=\frac{1}{\lambda}e^{-\frac{\xi}{\lambda}} \tag{3}$$

where ξ 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 $w(x,y)-z(x,y)$, so asperities of height ξ larger than $w(x,y)-z(x,y)$ will contact the wafer surface and the amount of compression of these asperities is $\xi-[w(x,y)-z(x,y)]$. All the asperities are assumed to obey Hooke's law, i.e., the exerting force is proportional to the compressed amount. The expected value of $P(x,y)$ can be estimated by averaging across all of the asperities as follows:

$$
\begin{aligned}
P(x,y) &= \int_{w(x,y)-z(x,y)}^{\infty} k \cdot \{\xi - [w(x,y) - z(x,y)]\} \cdot l(\xi) d\xi \\
&= k \cdot \Psi(w(x,y) - z(x,y))
\end{aligned} \tag{4}
$$

where k is a spring constant and $\Psi(z)$ is a derived asperity height distribution function, defined as $\Psi(z) = \int_z^{\infty} (\xi - z) \cdot l(\xi) d\xi$. $\Psi(z)$ can be calculated once the probability distribution of asperity height is known, and it is a strictly decreasing function and approaches zero at infinity. Since we assume the asperity height distribution as given in Equation 3, $\Psi(z)$ is given by

$$
\Psi(z) = \lambda \cdot e^{-\frac{z}{\lambda}} \tag{5}
$$

When a feature of step height $h_s(x,y)$ is pressed against the pad, as shown in Figure 2, Equation 4 implies that the up area pressure is

$$
P_u(x,y) = k \cdot \Psi(w(x,y) - z_u(x,y)) \tag{6}
$$

and the down area pressure is

$$
P_d(x,y) = k \cdot \Psi(w(x,y) - z_u(x,y) + h_s(x,y)) \tag{7}
$$

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

$$
P(x,y) = \rho(x,y) \cdot P_u(x,y) + (1 - \rho(x,y)) \cdot P_d(x,y) \tag{8}
$$

Combining Equations 5, 6, 7 and 8, we relate pressures to step height and characteristic asperity height as follows:

$$
\begin{cases}
P(x,y) = k\left[\rho(x,y) + (1-\rho(x,y)) e^{-\frac{h_s(x,y)}{\lambda}}\right] \cdot \lambda e^{-\frac{w(x,y)-z_u(x,y)}{\lambda}} \\[2ex]
P_d(x,y) = \dfrac{1}{1 + \rho(x,y) \cdot \left(e^{\frac{h_s(x,y)}{\lambda}} - 1\right)} P(x,y) \\[2ex]
P_u(x,y) = \dfrac{e^{\frac{h(x,y)}{\lambda}}}{1 + \rho(x,y) \cdot \left(e^{\frac{h_s(x,y)}{\lambda}} - 1\right)} P(x,y)
\end{cases} \tag{9}
$$

Modeling of CMP Process

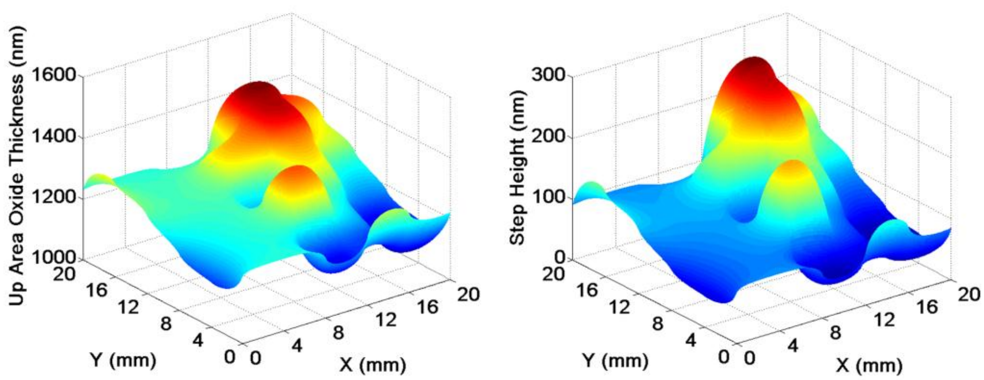
The physical CMP model can be obtained by combining the effects of the two parts above together: the elastic pad bulk, which is described by Equation 1, and the asperities with exponential height distribution, described by Equation 9. The pressure and deflection interactions between wafer surface topography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

$$\begin{cases} P(x,y) = k\left[\rho(x,y) + (1-\rho(x,y))e^{-\frac{h_s(x,y)}{\lambda}}\right]\cdot \lambda e^{-\frac{w(x,y)-z_u(x,y)}{\lambda}} \\ w(x,y) = F(x,y)\otimes P(x,y) + w_0 \end{cases} \tag{10}$$

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 coordinates, the two unknowns $P(x,y)$ and $w(x,y)$ can be calculated by solving Equation 10 iteratively. Once $P(x,y)$ is solved, $P_u(x,y)$ and $P_d(x,y)$ can be obtained by Equation 9; this pressure then translates into local material removal rate depending on the pressure-rate relationship used. In the example below, we employ the classic Preston equation [4] with local pressure $P(x,y)$ to calculate the instantaneous material removal rates of up area and down area as follows:

$$\frac{dz(x,y)}{dt} = -K_p \cdot P(x,y)\cdot v \tag{11}$$

where K_p is the Preston coefficient and v is the linear velocity of the pad relative to the wafer. In other applications (e.g., STI or copper CMP), non-Prestonian relationships that may include slurry chemistry effects can be used. Given the dynamic relationships above, $z_u(x,y)$ and $h(x,y)$ are updated in a time-stepping fashion during the simulation of CMP. Figure 3 shows a full chip simulation for the MIT/Sematech (SKW7-2) CMP test pattern [5].



(a) (b)

Figure 3: Full-chip simulation for the SKW7-2 test pattern at the time point when a 100 nm step height endpoint in the 50% pattern density region is reached. Simulation is for a JSR standard pad with standard conditioning disk: (a) Up area oxide thickness. (b) Step height.

Model for Asperity-Wafer Contact

To model the mechanical response of the pad asperities to the wafer, the pad surface can be considered as a nominally flat surface covered with asperities of various shapes and different heights. It is typical to assume some asperity shape and use the height of the asperity as a characteristic parameter. The problem can then be broken down into two steps.

First, we assume a certain asperity shape and solve the elastic deformation problem of a single asperity with height h when pressed upon the wafer surface. If the asperity deformation is δ, we can express the following terms as functions of δ, as illustrated in Figure 4: the contact area $a(\delta)$, the single asperity load as $L(\delta)$, and the pressure distribution in the contact area $P(x, y; \delta)$.

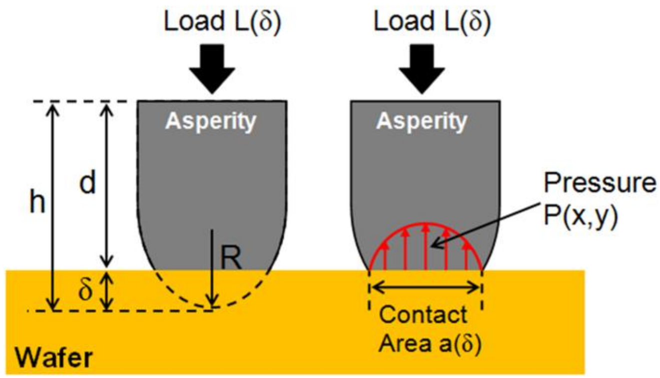


Figure 4: Diagram of a single asperity being compressed.

Second, we assume an asperity height distribution or probability density function *ξ(h)*, i.e., the number of asperities per unit area with height between h and $h + dh$ is $\xi(h)dh$. If the distance between the wafer and the nominal surface of the pad is d, the asperities with height greater than d will be in contact with the wafer surface. The number of asperities in contact is

$$n = N\int_d^\infty \xi(h)\,dh \tag{12}$$

where N is the total number of asperities.

For the asperity with height $h > d$, the deformation is $\delta = h-d$. The total contact area is

$$A = N\int_d^\infty a(h-d)\xi(h)\,dh \tag{13}$$

For an applied force of F_0, the distance d can be obtained as

$$F_0 = N\int_d^\infty L(h-d)\xi(h)\,dh \tag{14}$$

Greenwood [6] assumes that the asperities have spherical surfaces, all with the same radius R, and the contact is Hertzian. Based on the same assumptions as Greenwood and using the Hertzian results, we consider the pressure distribution and find:

$$\begin{cases} a(\delta) = \pi R\delta \\ L(\delta) = \dfrac{4}{3}E_a R^{\frac{1}{2}}\delta^{\frac{3}{2}} \\ P(x,y;\delta) = P_c\left(1-\dfrac{\pi\left(x^2+y^2\right)}{a(\delta)}\right), \quad \pi\left(x^2+y^2\right) \le a(\delta) \end{cases} \tag{15}$$

where E_a is the reduced modulus of the asperity and $P_c = \dfrac{3L(\delta)}{2a(\delta)}$ is the pressure at the center of the contact area. Here it is assumed that the wafer material is much more rigid than the pad asperity.

Measurements have shown that the asperity height distribution approximately follows an exponential decay for large asperity heights [3],

$$\xi(h) = \frac{1}{\lambda} e^{-\frac{h}{\lambda}} \tag{16}$$

where λ is the characteristic asperity height. Then the number of asperities in contact n, total contact area A, and the applied force F_0 can be determined as

$$\begin{cases} n = N e^{-\frac{d}{\lambda}} \\ A = N\pi R \lambda e^{-\frac{d}{\lambda}} \\ F_0 = E_a N \sqrt{\pi R \lambda^3} e^{-\frac{d}{\lambda}} \end{cases} \tag{17}$$

The total area occupied by N asperities is $A_0 = 4NR^2$ if we assume that all asperities stand close to each other without separation. We then get a reference pressure

$$P_{a0} = \frac{F_0}{A_0} = \frac{E_a}{4} \sqrt{\frac{\pi \lambda^3}{R^3}} e^{-\frac{d}{\lambda}} \tag{18}$$

Since the polishing only happens in the contact area [**Error! Reference source not found.**], an important pad surface property is the contact area percentage $f(P_{a0}) = \frac{A}{A_0}$ under the applied reference pressure P_{a0}. Using Eq. 17 and 18, we have

$$f(P_{a0}) = \frac{P_{a0}}{E_a} \sqrt{\frac{\pi R}{\lambda}} \tag{19}$$

Thus, we can now relate the contact area percentage to key pad surface geometry and mechanical properties. If we find the characteristic asperity height, asperity radius and asperity modulus, we can predict the pad-wafer contact area percentage. Figure 5 shows that the contact area percentage decreases when the characteristic asperity height increases [7]. This is because a smaller number of high asperities contact the wafer, and bear the load with aggregate smaller contact area.

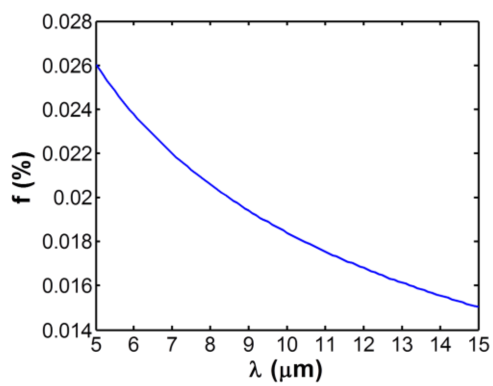


Figure 5: Contact area percentage vs. characteristic asperity height at reference pressure 4 psi.

Extended Die-Level CMP Model

With the particle-level model for asperity-wafer contact, we can now extend the die-level model to include asperity properties which are assumed to be different from pad bulk properties. Here Equation 18 can be utilized to calculate the local pressures needed in the die-level model. The up area pressure is thus

$$P_u = \frac{E_a}{4}\sqrt{\frac{\pi\lambda^3}{R^3}}e^{-\frac{w(x,y)-z_u(x,y)}{\lambda}} \tag{20}$$

and the down area pressure is

$$P_d = \frac{E_a}{4}\sqrt{\frac{\pi\lambda^3}{R^3}}e^{-\frac{w(x,y)-z_u(x,y)+h_s(x,y)}{\lambda}} \tag{21}$$

Then Equation 20 and 21 can replace Equation 6 and 7 in the die-level model derivation, giving

$$\begin{cases} P(x,y) = \rho(x,y)\cdot P_u(x,y) + (1-\rho(x,y))\cdot P_d(x,y) \\ w(x,y) = F(x,y)\otimes P(x,y) + w_0 \end{cases} \tag{22}$$

This die-level model is thus extended to include asperity size R and asperity modulus E_a. The two unknowns $P(x,y)$ and $w(x,y)$ can be calculated by solving Equation 22 iteratively. Once up area pressure and down area pressure are solved, Equation 11 is utilized to calculate the instantaneous material removal rates of up area and down area and update the chip surface evolution during CMP process.

Future Work

The extended die-level CMP model includes additional pad properties than our previous model, including parameters that can be directly measured. Using information about the pad surface structure, the new model enables us to estimate the bulk surface displacement $w(x,y)$ in a more refined way than in the previous physical model. However, a perhaps surprising result is that the pressure distribution across the chip does not change when the asperity structure is taken into account as above. There are two approaches to improve the model to include the effect of asperity properties on pressure distribution. The first is introducing an asperity size distribution and taking into account layout pattern size. The down area pressure may only be affected by a certain asperity size range; very wide asperities cannot touch some down areas. A second model modification under consideration is to relate the distance between the wafer and the bulk surface to slurry flow rates, which might affect blanket removal rates.

Adding some system effects can also improve the die-level model. In existing wafer-level models, both pressure and relative velocity are non-uniform spatially across the pad and wafer. Slurry flow rate and abrasive particle average resident time are strongly affected by pad groove design and conditioning. A future opportunity is to couple and transfer information about wafer-level non-uniformity into the die-level model to estimate both time-averaged effects on any given die of interest, and to calculate nonuniformity in planarization for different die across the wafer.

References

[1] O. G. Chekina and L. M. Keer, J. Electrochem. Soc., 145, 2100 (1998).

[2] T. Yoshida, Electrochem. Soc. Proc. 3rd Internl. Symp. CMP in IC Device Manufacturing, PV 99-37, 593, Honolulu, HI (1999).
[3] L. Borucki, R. Zhuang, T. Sun, Y. Zhuang, A. Philipossian, and D. Slutz, "Mechanical and Optical Analysis of Pad Surface Micro-texture Differences Caused by Conditioning," International Conference on Planarization Technology, Foster City, CA (2006).
[4] F. Preston, J. Soc. Glass Tech., 11, 214 (1927).
[5] W. Fan, D. Boning, L. Charns, H. Miyauchi, H. Tano and S. Tsuji, J. Electrochem. Soc., 157(5), H526 (2010).
[6] L. Charns, M. Sugiyama and A. Philipossian, Thin Solid Films, 485, 188 (2005).
[7] D. Boning and W. Fan, Mater. Res. Soc. Symp. Proc., 1249, E05-04 (2010).