Theoretical and Experimental Investigation of ILD Removal Rates, Coefficient of Friction, and Pad Flattening Ratio

L. Borucki
Intelligent Planar, Mesa, AZ USA
H. Lee, Y. Zhuang and A. Philipossian
The University of Arizona, Tucson, AZ USA

Introduction

It is well established that pad conditioning facilitates polish rate stability in chemical-mechanical polishing (CMP) of interlayer dielectic (ILD) silicon dioxide films and that in the absence of conditioning, polish rates decay over time [Stein *et al*., Lawing]. ILD rate decay is also associated with changes in pad surface morphology [Lawing, Kojima *et al.*]. In [Lawing], it was found that the pad surface height probability density function (PDF), as measured with optical interferometry, develops a secondary peak that grows in magnitude and approaches or merges with the primary surface height peak with increasing polishing time in the absence of conditioning. In [Kojima *et al*.], polish rate decay was shown to be associated with an increase in the pad flattening ratio (PFR), a digitally extracted measure of the proportion of incident light from a source that is directly reflected back rather than being scattered by surface roughness. Both optical topographical measures may be interpreted as indicating the flattening of pad asperities that are tall enough to contact the wafer. Flattening may be due abrasive wear or irreversible plastic deformation. In [Lawing], it was found that the secondary peak grows more quickly if irregular fumed slurry particles are used rather than spherical colloidal particles under otherwise identical loading conditions. This suggests that pad abrasive wear by slurry particles may dominate the change in the surface. In [Borucki, Borucki *et al*.], models of surface abrasive wear, polish rate decay and conditioning have been formulated that quantitatively agree with the abrasive wear interpretation of rate decay in ILD polishing.

We study here ILD polish rate stability and rate decay in experiments that involve *ex-situ* conditioning with a diamond disc, pad cleaning without conditioning using a high-pressure microjet (HPMJ) and KOH, and in the absence of pad treatment. We measure during the experiments the PFR, the coefficient of friction (COF) between the polished wafers and the pad, and the material removal rate (RR). Using insights from the experiments, we extend a previously developed Langmuir-Hinshelwood model [Thakurta et *al*., Sorooshian *et al.*] for silicon dioxide removal and combine it with the abrasive wear and conditioning model in [Borucki *et al.*]. The integrated model is then shown to be in quantitative agreement with measurements under all three conditions.

Experimental Conditions and Results

All polishing experiments were performed on Rohm and Haas IC-1000^{TM} perforated polishing pads with no sub-pad using Fujimi PL-4217 slurry (12.5% solids by weight) delivered at 80 cc/min at the center of a scaled 100 mm polisher mounted on a sliding table [Philipossian *et al*.]. Pad and wafer were rotated in the same direction at the same rate at a relative sliding speed of 0.62 m/sec, the pressure applied to the wafer was 20.7 kPa (3 PSI), and polishing times were 2 min. For the diamond conditioning series of wafers, dressing was performed *ex-situ* for 30 sec with a TBW 5.08 cm (2”) diameter 100 grit tool using a disc pressure of 3.5 kPa

(0.5 PSI), a rotation rate of 30 RPM and a full sweep frequency of 20/sec over a 7.62 cm (3") center-to-center constant speed track. This produced a pad cut rate of 5.4 µm/min. For the microjet series, pad cleaning with HPMJ was performed *ex-situ* for 10 sec at a jet pressure of 10 MPa using a fan angle of 40 degrees, an actuator angle of 90 degrees, a 5 mm nozzle to pad distance, and a flow rate of 0.77 L/min. The HPMJ nozzle is shown in Fig. 1. KOH was added to the ultra pure water (UPW) feed to the HPMJ to match the pH of PL-4217 slurry. HPMJ produces no measurable cut rate under these conditions. Prior to all series (diamond, HPMJ or no conditioning), the pad was broken in by with 30 minutes of diamond conditioning in UPW as described above followed by a 2 minute polish of a silicon wafer. Removal rate, PFR and COF were then measured during or after polishing of each of 50 subsequent ILD wafers. COF was measured at 10,000 Hz using a load cell attached to the polisher table; standard deviations were typically less than 10% of measured values. PFR was measured with an optical tool whose principle of operation is indicated in Fig. 2. Flat areas of the pad surface return bright reflections due to low dispersion of the reflected image. The pad flatness ratio is then the ratio of the bright area to the total area in the field of view. Details of the operation of the PFR tool can be found in [Kojima *et al*.].

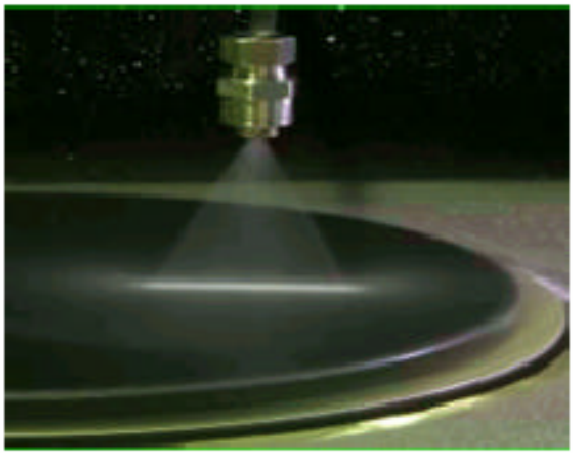


Figure 1: High pressure microjet nozzle.

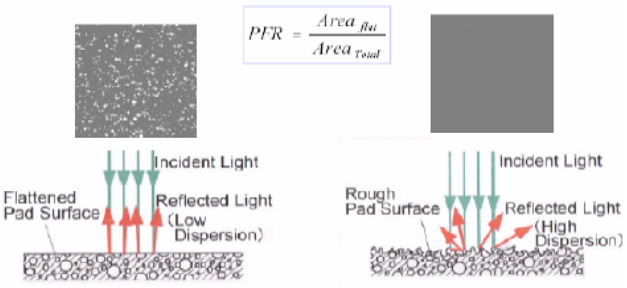
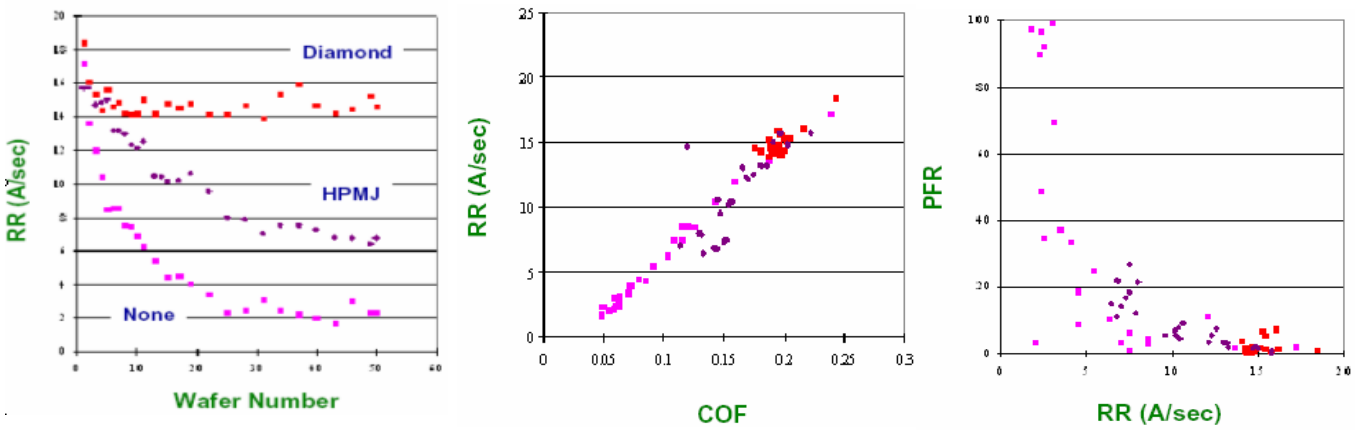


Figure 2: Images from the PFR tool for a pad with some surface flattening (left) and none (right). PFR is the ratio of the bright area to the total area.

Measured material removal rates under all three conditions are shown in Fig. 3(a). Under *ex-situ* diamond conditioning (red symbols), the rate initially drops slightly and then stabilizes after approximately 10 wafers. When no conditioning is applied (magenta), the removal rate declines continuously to a very low level in the classical manner of polish rate decay. When HPMJ is used to clean the pad between wafers (blue), there is also a decline in rate but to a value intermediate between diamond conditioning and no conditioning. Figure 3(b) plots the measured removal rate against COF using the same color coding as in Fig. 3(a). We see that regardless of conditioning procedure, the removal rate is linearly correlated with the COF on a universal curve. Figure 3(c) shows the PFR plotted against removal rate using the same color coding. For diamond conditioning, the PFR is very low and the removal rate is high, HPMJ has a PFR that is on the whole larger and also a lower removal rate, and finally with no conditioning, PFR is highest and rate is lowest. There is approximately an inverse relationship between RR and PFR.



(a) (b) (c)

Figure 3: (a) Measured oxide removal rates, (b) RR vs. COF, (c) PFR vs. removal rate.

Theory

The theory used in the analysis of the above experiment will be summarized here. The overall model consists of three elements: (1) a material removal rate model, (2) a rough surface contact model, and (3) a model for conditioning and abrasive wear. Details of the derivations of the conditioning and abrasive wear models may be found in [Borucki] and [Borucki *et al.*].

Removal rates are described with a Langmuir-Hinshelwood model supplemented with a simple flash heating model. In the Langmuir-Hinshelwood model, a softened layer is formed on the surface of the ILD film by a chemical process at rate k_1 and the softened layer is then removed by a mechanical process at rate k_2. The abraded material is carried away by the slurry and is not re-deposited. The local removal rate in this mechanism is

$$RR = \frac{M_w}{\rho}\frac{k_1 C}{1+\frac{k_1 C}{k_2}} \qquad (1)$$

where M_w is the molecular weight of silicon dioxide, ρ is the density, and C is the local molar concentration of reactant. It is assumed that there is little reactant depletion so that C remains constant. This allows C to be absorbed into k_1 and effectively set to 1. In (1), the chemical reaction rate is taken to have an Arrhenius form, $k_1 = A\exp(-E/kT)$, and the mechanical removal rate taken to be proportional to the shear stress, $k_2 = c_p \mu_k pV$, where c_p is an assumed proportionality constant, p is the applied pressure, V is the sliding speed (constant across the wafer in these experiments) and μ_k is the COF. In the mechanically-limited extreme, the polish rate is $RR = (\mu_k c_p M_w/\rho)pV$ so that the rate is proportional to the COF as in Fig. 3(b). The reaction temperature T in the chemical model is taken to have the form

$$T = T_a + (\beta/V^a)\mu_k pV \quad . \qquad (2)$$

In this formula, T_a is the ambient temperature and β/V^a is a factor that approximates hydrodynamic effects that are generally present on plain and perforated pads. The reaction temperature increase over ambient is seen to be proportional to the frictional power density. As explained in [Sorooshian *et al.*], this model may be justified theoretically as a pad asperity tip flash heating model. For our purposes, the reaction rate model (1) and flash heating model (2) constitute an empirical model with five fitting parameters (A,E,c_p,β,a), the other parameters being known. Values for all parameters except c_p are taken from [Sorooshian *et al.*]; due to the difference in slurry used here, c_p must be adjusted to give the correct initial removal rate.

We interpret the PFR here as being proportional to the actual contact area fraction A_f (actual contact area divided by nominal area) between the pad and wafer. Figs. 3(a) and 3(b) then imply that the COF is inversely related to A_f. To be specific, we assume that

$$\mu_k = \frac{\mu_k^1}{A_f} \tag{3}$$

where μ_k^1 is the COF at unit contact area fraction. This assumption says that the sliding COF of the wafer against the pad decreases as the pad surface becomes smoother. While intuitively appealing, there is no theoretical basis for (3); we take it as being mandated by observations. This assumption introduces no new additional fitting parameters into (!) and (2) since μ_k^1 may be combined with c_p and β, but it does introduce a new function $A_f(t)$.

In the present theoretical model, it is necessary to know how A_f changes with time under conditioning and abrasive wear. First we need to relate $A_f(t)$ to the surface height PDF $\phi(z,t)$ of the pad and the applied load. In the Greenwood and Williamson contact model [Greenwood *et al.*], A_f is related to $\phi(z,t)$ by an integral transform,

$$A_f(t) = \Lambda \int_{d(t)}^{\infty} K_A(z,d)\phi(z,t)dz\,, \tag{4}$$

where $\Lambda = \pi\eta_s/\kappa_s$, η_s is the asperity area density, κ_s is the mean asperity tip curvature, the kernel K_A is $K_A(z,d) = z-d$ and $d(t)$ is the location (height) of the wafer surface at time t. The height of the wafer $d(t)$ in (4) is determined by load balance,

$$p = G\int_{d(t)}^{\infty} K_p(z,d)\phi(z,t)dz\,. \tag{5}$$

In Greenwood and Williamson theory, $G = 4E\eta_s/((1-\nu^2)\kappa_s^{1/2})$, E is the bulk Young's modulus, ν is the Poisson ratio and the kernel is $K_p(z,d) = (z-d)^{3/2}$. In a strict interpretation of (4) and (5), the PDF of the pad surface summits should be used instead of the surface height PDF. The theory also requires the mean asperity tip curvature κ_s, which is notoriously difficult to characterize and increases as the asperities wear. An alternative is to use a "spring" model, which treats the surface as being composed of independent linear springs. In this model, Λ=1,

K_A=1, G is a compressive modulus, and $K_p(z,d) = (z-d)/(z-(\mu - h_a/2))$ is the strain that occurs when a point on the surface at height $z>d$ is compressed to height d. The denominator in K_p is the unstrained height of the surface measured relative to an imaginary reference at height μ-h_a/2, where μ is the surface mean height and h_a is the thickness of the rough surface. The spring model has the advantages that all of the parameters can be characterized with simple surface loading and staining experiments and that it uses the visible surface height PDF. We use it in this study and identify the bearing area (4) with the PFR. We ignore for simplicity the fact that the nominal contact pressure p in (5) is non-uniform on the wafer surface due to friction-induced tilting of the wafer carrier. Instead, we will take p in (5) as the pressure at the wafer center and $d(t)$ as the corresponding height.

The third element of the overall model is the evolution model for the surface height PDF $\phi(z,t)$. Instead of being stated directly in terms of $\phi(z,t)$, this portion of the model uses the smoother complementary cumulative density function (CCDF), defined as

$$q_f(z,t) = \int_z^\infty \phi(\zeta,t)d\zeta \tag{6}$$

and related to the PDF by $\phi(z,t) = -\partial q_f(z,t)/\partial z$. The subscript f stands for "foamed" and refers to the fact that the IC-1000^{TM} is a closed cell polyurethane foam. At any height z, $q_f(z,t)$ is the probability of finding an optically visible point on the foamed surface (*i.e.*, not hidden within a void) at or above z at time t. A basic fact about the CCDF $q_f(z,t)$, explained in detail in [Borucki *et al.*], is that it can be related to the CCDF $q(z,t)$ of an *identically conditioned solid pad* by an integral transform,

$$q_f(z,t) = \int_z^0 q(\zeta,t)\Phi(z-\zeta)d\zeta\,. \tag{7}$$

The phrase "identically conditioned" means that the solid pad is treated with exactly the same sweep schedule as the foamed pad but that the load is adjusted to achieve the same cut rate. The kernel $\Phi(z)$ of the transform (7) is called the *intrinsic* PDF of the foamed pad. It is the optically visible surface height distribution that one would measure after slicing the foam with a perfectly planar cut. For IC-1000^{TM}, experimentally measured distributions suggest that $\Phi(z)$ is composed of an exponential function that describes the portion of the material that is visible though open voids and a delta function that describes the proportion of exposed area lying on the cut (Fig. 4). It is convenient to take the height origin $z=0$ to be at the cut plane. With this background, the CCDF of the solid pad is governed by the partial differential equation

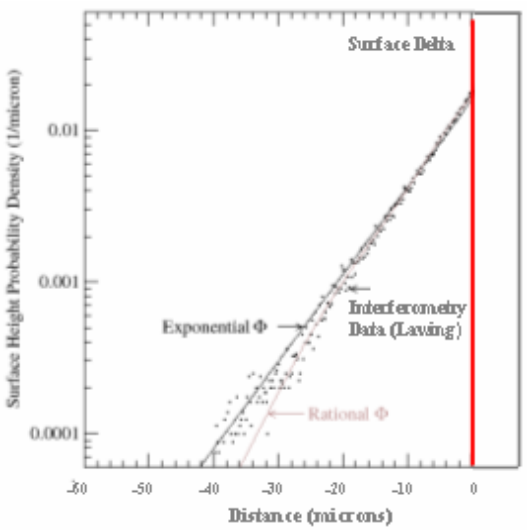


Figure 4: Intrinsic PDF of an IC-1000^{TM} pad.

$$\frac{\partial q}{\partial t} = W\frac{\partial q}{\partial z} - Cq\,. \tag{8}$$

On the right side of (8), the first term describes abrasive wear and the second models conditioning. The factor *W* is the *wear function* and is defined as

$$W(r,t) = \eta(r)C_a V\Psi K_W(z,d)\,. \tag{9}$$

In Greenwood and Williamson theory, $\Psi = 4E\kappa_s^{1/2}/(3\pi(1-\nu^2))$, $\eta(r)$ is the fraction of each rotational period that an asperity on the pad at radius *r* spends under the wafer and is therefore subject to abrasion, and $K_W = (z-d)^{1/2}$. In the spring model, Ψ=*G* and K_w=K_p. We evaluate $\eta(r)$ here at the radius of the wafer center. The factor C_a in (9) is an empirical wear coefficient. Since wear occurs only when there is wafer contact, the wear function is nonzero only when *z>d(t)*; otherwise it is taken to be zero in (8). Note that (9) introduces only one new parameter, C_a.

The coefficient of the second term on the right side of (8) is the *conditioning function C*, defined as

$$C(t) = \frac{\Omega}{2\pi}\frac{w(z-h(t))}{\ell}\,. \tag{10}$$

In (10), Ω is the pad rotation rate in radians/sec and *h(t)*, for diamond conditioning, is the location of the bottom of the mean furrow at time *t*. The mean furrow, described by *w(z)*, will be defined presently. We take *h(t)* to be linear in the measured pad cut rate *c, h(t)=-ct*. To understand ℓ and *w(z)*, imagine that the diamonds on the conditioning disc are removed from the disc and remounted with even spacing on a straight bar whose length is the endpoint center-to-center distance of the conditioning sweep. Then ℓ is the spacing between diamonds in this arrangement. If the linear arrangement of diamonds is dragged over the surface of the solid pad at sufficient force so that they collectively cut a single deep furrow, then *w(z)* is the width of the furrow at height *z* after vertical translation of the furrow shape so that the bottom is at *z=0*. We model the mean furrow shape as having a Gaussian-like width $w(z) = A_0\exp(-((z-A_1)/A_2)^2)$ for *z* < A_1 and constant width A_0 for z > A_1. The mean furrow shape can be extracted from the surface PDF of a solid pad. It is possible to obtain a good description of the measured PDF of a freshly conditioned IC-1000™ pad using (8) and (10) (Fig. 5). Note that the conditioning function *w(z-h(t))* is defined only for $z \geq h(t)$; otherwise, we take it to be zero. For HPMJ, evidence suggests that no actual material

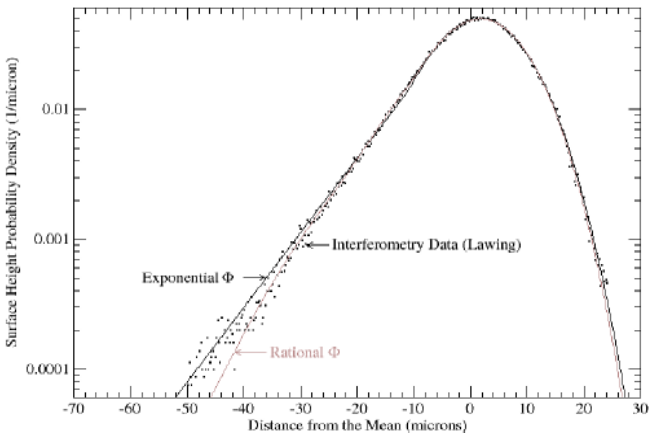
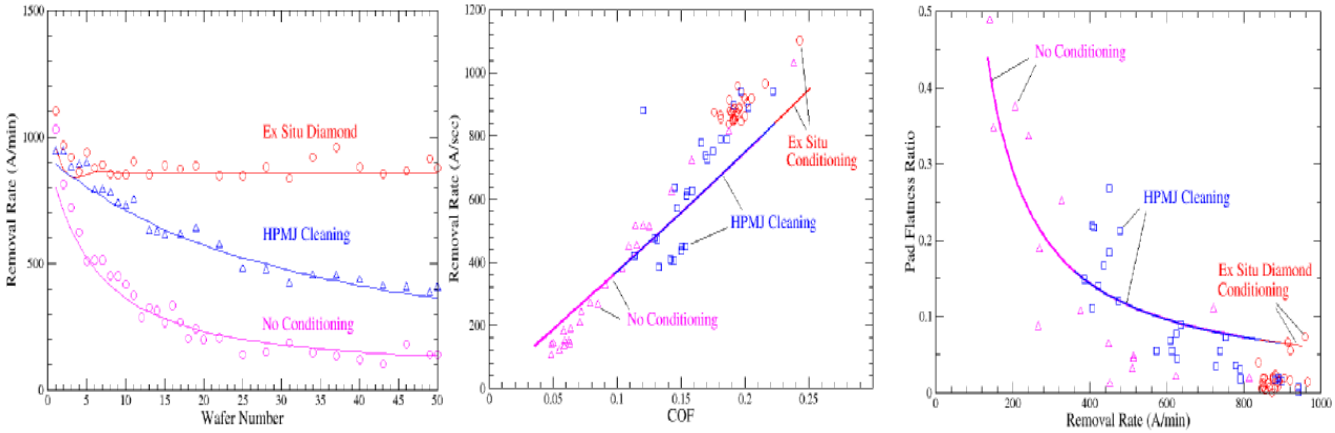


Figure 5: Fit of the conditioning model to a measured foamed pad PDF.

removal occurs at the jet pressure used but that the jet does remove abrasive particles. We therefore take the conditioning function to be zero for HPMJ and appropriately reduce the abrasive wear coefficient C_a.

Data Analysis

The above model was applied in the following way. First, a conditioning simulation was run for the 30 min pad break-in time at the measured cut rate. The initial PDF obtained was then used as the starting point for all subsequent runs. For polishing with no conditioning, the asperity abrasive wear parameter C_a was adjusted to provide agreement between the measured and modeled removal rates. By applying the conditioning model between wafers, a prediction was then obtained for the steady removal rate during *ex-situ* conditioning. Finally, for HPMJ, C_a was reduced from the no conditioning value to obtain agreement with measured rate data. The best match was found at 5X reduction. Figure 6 shows the result of this procedure. We see that the model is in good quantitative agreement with the data; in particular, the model correctly reproduces the time (wafer number) dependence in Fig. 6(a) and the variation of rate with PFR in 6(c). The linear variation of rate with COF (Fig. 6(b)) was built into the model.



(a) (b) (c)

Figure 6: (a) Model removal rates (solid), (b) removal rate vs. COF, (c) PFR (A_f) vs. rate.

Figure 7 shows the simulated PDFs. Starting with the intrinsic PDF, characteristic of a fresh, unused pad, the surface PDF evolves over the break-in period into a shape similar to that in Fig. 5. During polishing with *ex-situ* diamond conditioning, the PDF retains the same basic shape but develops a barely discernable swelling in the right hand tail due to abrasive wear. This is associated with an increase in contact area and the slight drop in rate seen in the data. At steady state, abrasive wear and pad

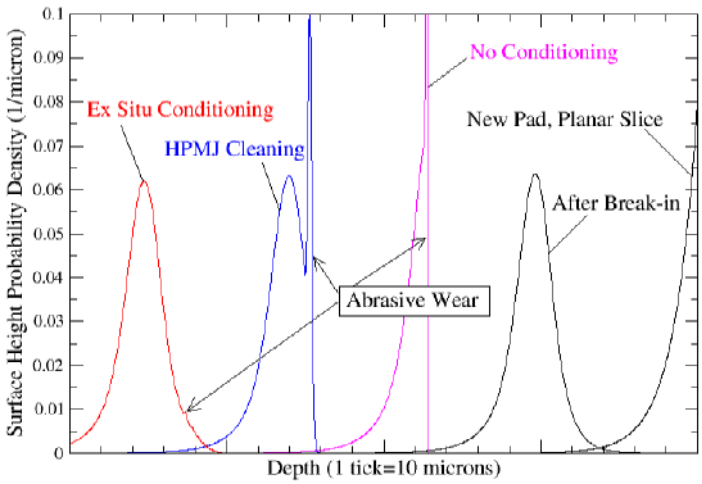


Figure 7: Comparison of simulated PDFs.

renewal are in equilibrium, so the PDF becomes stationary in shape and simply translates in depth at the cut rate. Without conditioning, there is no renewal and under abrasive wear the PDF evolves back to the intrinsic PDF. The contact area increases correspondingly and the removal rate drops. The intrinsic PDF is the stationary PDF in this case. With HPMJ cleaning, the rate of abrasive wear is much lower. Unlike the situation with no conditioning, the HPMJ PDF develops a secondary tail that remains distinct from the main peak. The contact area is therefore lower and the rate higher. We predict, however, that the HPMJ PDF will eventually evolve to the intrinsic PDF if there is no surface renewal.

Conclusions

The model for pad conditioning, asperity abrasive wear and silicon dioxide material removal described here is in good general agreement with measured rates for polishing with *ex-situ* conditioning, no conditioning and HPMJ cleaning. Two of the hypotheses that are critical in the theory are the proportionality of removal rate to COF and the inverse relation between COF and contact area fraction. The model agrees with the observed time dependence of rate variations and supports the hypothesis that the reduced rate under HPMJ cleaning is due to removal of slurry particles from the pad. As-conditioned PDFs and PDFs under mild and heavy abrasive wear may also be seen to be similar to those measured by [Lawing].

Acknowledgment

The authors gratefully thank H. Kojima for the kind loan of his PFR instrument.

References

Borucki, L. "Mathematical modeling of polish rate decay in chemical-mechanical polishing," *J. of Engineering Mathematics*, 43, pp. 105-114 (2002).

Borucki, L.J., Witelski, T. Please, C., Kramer, P.R. and Schwendeman, D. "A theory of pad conditioning for chemical-mechanical polishing," accepted by *J. of Engineering Mathematics*, (2004).

Greenwood, J.A. and Williamson, J.B.P, "Contact of nominally flat surfaces," *Proc. of the Royal Society of London A*, 295, pp. 300-319 (1966).

Kojima, H. and Nishiguchi, T., "A new method for evaluating CMP pad surface conditions using digital images," *Proc. Eighth International Chemical-Mechanical Planarization for ULSI Multilevel Interconnect Conference (CMP-MIC)*, Vol. 2003 IMIC-800P, pp. 37-42 (2003).

Lawing, A.S., "Polish rate, pad surface morphology and pad conditioning in chemical mechanical polishing," In: S. Seal, R. L. Opila, C. Reidseme Simpson, K. Sundaram, H. Huff and I.I. Suni (eds.), *Proc. of the Fifth International Symposium on Chemical Mechanical Polishing*. The Electrochemical Society Proc. Vol. PV2002-1, pp. 46-60 (2002).

Philipossian, A. and Mitchell, E., "Performing mean residence time analyses of CMP processes," *Micro Magazine*, 20(7) (2002).

Sorooshian, J., Borucki, L., Stein, D. Timon, R. Hetherington, D., and Philipossian, A., "Revisiting the removal rate model for oxide CMP," submitted to *ASME J. Tribology*, 2004.

Stein, D., Hetherington, D., Dugger, M. and Stout, T., "Optical Interferometry for surface measurement of CMP pads," *J. Electronic Materials* 25, pp. 1623-1627 (1996).

Thakurta, D., Schwendeman, D., Gutmann, R., Shankar, S., Jiang L., and Gill, W., "Three-dimensional wafer-scale copper chemical-mechanical planarization model", *Thin Solid Films*, 414, pp. 78-90 (2002).