Journal of The Electrochemical Society, **151** (9) G559-G563 (2004)
0013-4651/2004/151(9)/G559/5/$7.00 © The Electrochemical Society, Inc.

Experimental and Theoretical Investigation of Heating and Convection in Copper Polishing

L. Borucki,[a,*,z] Z. Li,[b] and A. Philipossian[b]

[a]*Intelligent Planar, Mesa, Arizona 85205, USA*
[b]*Department of Chemical and Environmental Engineering, University of Arizona, Tucson, Arizona 85721, USA*

Pad thermal data taken with an infrared video camera during chemical mechanical polishing of copper at different slurry flow rates is analyzed using a model developed to explain pad heating during oxide polishing. After minimal recalibration to match thermal data as a function of time and pad radius at one flow rate, the model is able to correctly predict the variation in temperature as the flow rate is changed. Periodic thermal oscillations in the data that are not present in oxide polish data are also modeled and are attributable to *in situ* conditioning. Theoretical estimates of the wafer temperature suggest that the temperature rise of the wafer above ambient may be significantly higher than what is measured on the pad near the leading edge of the wafer.
 [DOI: 10.1149/1.1774489]

Manuscript submitted November 10, 2003; revised manuscript received January 31, 2004. Available electronically August 5, 2004.

The rate of copper removal in CMP depends on both physical and chemical processes. It may consequently be either mechanically limited or reaction-rate limited.[1,2,4] Because the reaction for copper oxide formation has an activation energy of only 0.5–0.6 eV,[3,4] a temperature increase of 10 K during a 1 min process is sufficient to double the removal rate in a reaction-rate-limited process, and fluctuations of only 1 K at 0.5 eV can change the polish rate of a reaction-rate-limited, room-temperature process by nearly 7%. Thus, it is important to understand factors that affect wafer temperature and wafer temperature uniformity during copper polishing.

Pad thermal measurements have been reported by many authors; see, for example, Ref. 4-8. However, the wafer temperature itself is much less accessible. Direct measurements of the temperature in the region between the pad and optically transparent glass wafers were reported in Ref. 9 and 10 using dual-emission laser-induced fluorescence (DELIF). In these studies, which involved convex or concave surfaces, the temperature rise in the pad-wafer region was found to be several degrees higher than in the bow wave. The pad-wafer temperature and the bow wave temperature were also found to be linearly correlated so that the wafer temperature can be empirically inferred from the bow wave temperature. Because DELIF cannot be applied to nontransparent substrates, there is a need for a method of estimating the wafer temperature. The need to know the wafer temperature rather than just the pad temperature in a reaction rate model is the main motivation for the current thermal model.

We have separately developed a thermal model for heat exchange between the wafer, pad, and slurry and have applied it successfully to explain pad thermal data taken during oxide polishing on Rodel, Inc., IC-1000 and JSR Corp. pads.[5] In this model, the majority of the frictional energy heats the wafer and the rest heats the pad. The slurry simultaneously acts as a coolant, taking up most of the thermal energy from the wafer and then redistributing it over and off the pad through radial convective transport. The model predicts that the temperature rise in the wafer is higher than can be observed on the pad. Here we apply the model to copper thermal data collected on an IC-1000 k-groove pad by making minimal changes in parameters to account for the differences in the slurry, wafer backing film, and the material being polished. We show that calibration at one flow rate is then sufficient to provide good predictions at other flow rates. The same calibration of the thermal model is applied in Ref. 6 in conjunction with a modified Langmuir-Hinshelwood reaction rate model to explain variations in copper removal rate behavior for the same slurry as a function of pressure, velocity, and flow rate.

* Electrochemical Society Active Member.
[z] E-mail: lenborucki@intelligentplanar.com

Experimental

Experiments were performed on a scaled version of a Speedfam-IPEC 472 polisher using Rodel's IC-1000 k-groove pads and Fujimi's PL-7102 slurry containing hydrogen peroxide as the oxidant. The experimental apparatus is shown in Fig. 1a and is described in detail elsewhere.[11] Copper disks, having a nominal diameter of 100 mm and a purity of 99.99%, were used for substrates. Conditioning was performed *in situ* using a diamond conditioner at an applied pressure of 0.5 psi, a rotational velocity of 30 rpm, and a sweep frequency of 20/min. Pad conditioning was followed by a 5 min pad break-in while polishing a disk. Because this is a flow experiment, the wafer pressure was held at 2.5 psi and the corotation rate of the pad and wafer was held at 140 rpm (1.09 m/s at the wafer center). The flow rates considered were 60, 80, and 140 cm^3/min. An infrared (IR) video camera with 0.1°C resolution was used to record pad temperatures at 0.2 s intervals at ten fixed locations near the pad leading edge (shown in Fig. 1b). Pad sample locations 6–10 are outside of the bow wave and therefore provide a more accurate estimate of the pad temperature because the fluid is thinner and less affected by ripples.[7] Polar coordinates of these points relative to an axis from the pad center to the wafer center were determined by geometric analysis of IR photographs. Real-time coefficient of friction (COF) data were taken simultaneously with the infrared thermal measurements using load cells as described in Ref. 11. Removal rates were calculated by determining the weight of the copper disk prior to and after polishing using a precise balance with a readability of 0.01 mg.

Figures 2, 3, and 4 show the temperature history measured at sample locations 6 (nearest the pad center) and 10 (nearest the perimeter) at 80, 140, and 60 cm^3/min, respectively. Temperatures at locations 7-9 fall between these curves and are not shown for clarity. It can be seen that after an initial transient increase requiring ~20 s, temperatures continue to rise nearly linearly. The duration of the initial transient depends on the pad type and thickness and the thermal masses of objects in contact with the pad; it is well within the range of times observed elsewhere.[4,5,7,8] There is also a 0.33 Hz oscillation in the temperature that has the maximum amplitude at sample location 6. The frequency of the oscillation is the same as the half-sweep frequency of the conditioner, suggesting that this is a thermal signature related to the conditioning process. Also shown in each figure is the measured COF. Because the sampling frequency of the COF measurement is high, each point plotted is the mean of 1000 measurements. The COF for the 80 cm^3/min run has a low-frequency oscillation and a general upward trend that is consistent with the gradual increase in pad temperature. While the upward trend is visibly correlated with the thermal data, the low-frequency oscillation is not and we do not find it to be present in any other experimental runs. In this and other Cu COF data sets, only a piece-

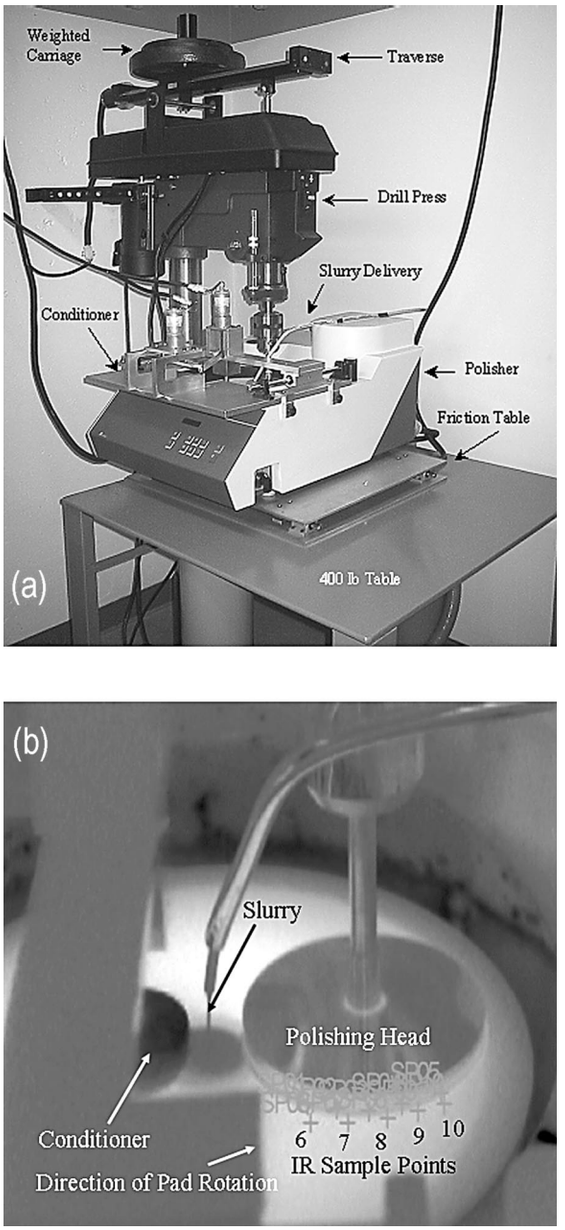

Figure 1. (a) Image of scaled polisher showing the diamond disk conditioner, drill press, and friction table. (b) Locations of leading edge sample points at which thermal data were measured.

wise linear regression fit to the COF data was consequently used for further analysis.

Thermal Model and Modification for Copper

Thermal model summary.—Thermal analysis of our experimental data was performed with a model described in Ref. 5, in which the emphasis was on heating of the pad during oxide polishing. We provide a synopsis of the model and its assumptions here and then focus on the features that are important for copper thermal modeling.

The CMP tool thermal model is just the heat equation, which in fixed spatial coordinates is

$$\rho C_p\left(\frac{\partial T}{\partial t} + \mathbf{V}\cdot\nabla T\right) = \nabla\cdot(\kappa\nabla T) + Q \qquad [1]$$

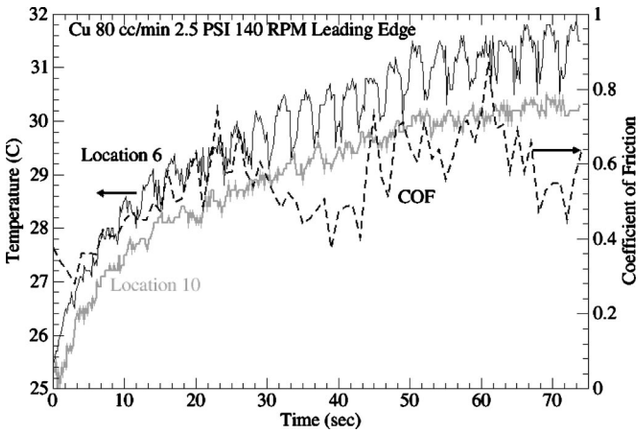

Figure 2. Pad temperatures at sample points 6 and 10 and the measured COF at 2.5 psi/140 rpm, 80 cm^3/min.

In Eq. 1, ρ, C_p, and κ are the local material density, heat capacity, and thermal conductivity, which are different for each component of the system. The term Q on the right side of Eq. 1 represents a source of frictionally and chemically generated heat that is nonzero only in a thin layer between the pad and wafer. As discussed in Ref. 7, the exothermic chemical contribution to Q is small relative to the frictional contribution in copper polishing under typical commercial conditions. The vector $\mathbf{V}$ in Eq. 1 is the local velocity of the rotating pad, the rotating wafer and carrier, or the slurry. At a fixed point in space, the $\mathbf{V}\cdot\nabla T$ term on the left side accounts for advective (horizontal) heat-transfer due purely to rotation or flow.

Equation 1 presents two significant computational difficulties: it must be solved in three dimensions (3D), and the advective term $\mathbf{V}\cdot\nabla T$ must be discretized carefully so that any artificial diffusion introduced by the numerical method itself does not dominate the physical thermal diffusion. This is particularly critical for the pad, which has both the highest velocity and the lowest thermal diffusivity of the tool components. As described in more detail in Eq. 1, we use two main ideas to simplify the thermal problem. These also provide more insight than a purely numerical solution. First, we use a tribological approximation,[12] described in further detail later, to partition the heat Q between the wafer and pad, with a fraction γ_1 going to the pad. This allows the wafer, carrier film, and carrier assembly to be treated as a separate but smaller 3D unit using Eq. 1. In this unit, the thermal diffusivities of the wafer and carrier are large enough that the advective problem (Eq. 1) can be solved directly with accuracy. The second main idea is that for the pad, we switch from spatial to material coordinates and calculate the thermal histories at points that are fixed to the rotating pad surface. A radial rotating array of fixed point on the pad is chosen for this, as shown in Fig. 5. The switch to the material view automatically captures rotational advection without introducing numerical diffusion. For each pad point in the array, we then solve only a one-dimensional (1D) heat equation in the vertical dimension z

$$\rho C_p\frac{\partial T_p}{\partial t} = \kappa\frac{\partial^2 T_p}{\partial z^2} \qquad [2]$$

where in Eq. 2, T_p refers to the pad temperature and the material parameters are for the pad. This decoupling of nearby material points on the pad is a good approximation if the radial spacing of the points is larger than the thermal diffusion length, $\sqrt{Dt_p}$, where t_p is the total process time and $D = \kappa/(\rho C_p)$ is the thermal diffusivity. The thermal diffusion length at 60 s is ~2 mm for the IC-1000 pad. In Eq. 2, the temperature at the platen interface is held at the platen temperature, while on the slurry side of the pad the 1D heat equations for the radial array of material points are coupled by a simpli-

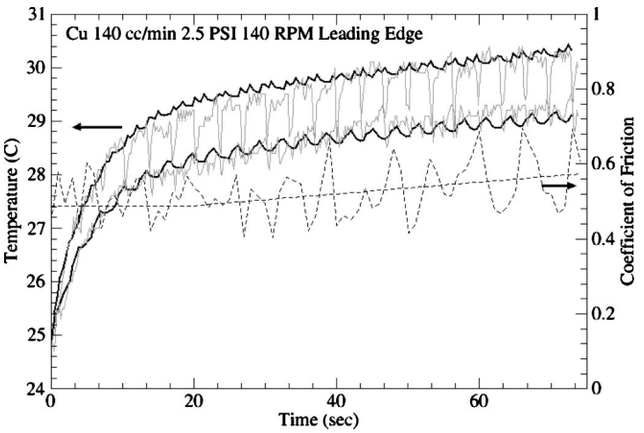

Figure 3. Measured (light solid curves) and simulated (dark solid curves) pad temperatures at pad locations 6 and 10 at 140 cm^3/min flow rate.

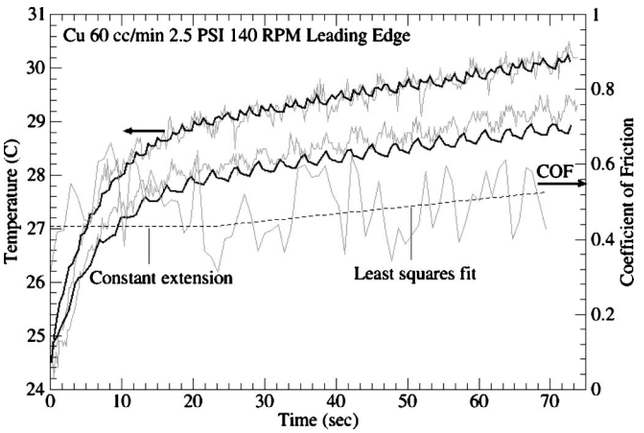

Figure 4. Measured (light solid curves) and simulated (dark solid curves) pad temperatures at pad locations 6 and 10 at 60 cm^3/min flow rate.

fied model for slurry flow and heat convection. The boundary condition at the slurry interface is discussed later. The slurry thermal model, which is obtained by expressing Eq. 1 in cylindrical coordinates and averaging over the thickness of the fluid, is

$$\rho C_p\left(\frac{\partial T_s}{\partial t} + (1 - \chi(r, t))\frac{V_{slurry}(r)}{h_s}\frac{\partial (h_s T_s)}{\partial r}\right) = \frac{\kappa}{r h_s}\frac{\partial}{\partial r}\left(r\frac{\partial (h_s T_s)}{\partial r}\right) + \frac{Q_s}{h_s} \quad [3]$$

where the material parameters are now for the slurry and $T_s = T_s(r, t)$ is the local slurry temperature. The function h_s is the local fluid thickness, taken to be a constant h_s^0 in the high or "land" areas of the pad and to be the groove depth otherwise for concentrically grooved pads. In Eq. 3, V_{slurry} is the local radial flow speed, which is grossly approximated using only mass conservation

$$V_{slurry}(r) = \frac{f_s}{2\pi h_s^0 r} \quad [4]$$

where f_s is the volumetric slurry flow rate. For a material point at radius r, the function $\chi(r, t)$ in Eq. 3 is one when the point is under the wafer and zero otherwise; the factor $1 - \chi(r, t)$ thus turns off convection, entraining the fluid at radius r to the pad rotation when the fluid is under the wafer. The convective term on the left side of Eq. 3 is necessary to describe the variation of the pad surface temperature with radius, as shown in Ref. 5. Finally, the source/sink function Q_s uses laminar flow correlations[5,13] and expressions similar to the second term on the right side of Eq. 5 to model the exchange of heat between the slurry and the pad, wafer, and surrounding air.

Model features important for copper.—In the model in Ref. 5, pad material points receive a fraction of the frictional heat generated at the pad/wafer interface and exchange heat with the slurry through the boundary condition

$$\kappa\frac{\partial T_p}{\partial z} = \chi(r, t)\gamma_1\mu_k(t)pV_s - h_{ps}(T_p - T_s) \quad [5]$$

where κ is the pad thermal conductivity, μ_k is the regression fit to the COF data, p is the applied pressure, V_s is the pad-wafer relative sliding speed, γ_1 the fraction of the total frictional energy that is partitioned to the pad, and $\chi(r, t)$ switches the heating term on and off at the material point at radius r as described previously. The coefficient h_{ps} is a convective (vertical) heat-transfer coefficient for exchange between the pad and slurry.

An important factor that distinguishes oxide from copper is γ_1. The heat partition fraction depends on the thermal conductivities, shape, and relative sliding velocity of the contacting surfaces. For a planar surface sliding under a stationary disk, the fraction of heat to the planar surface is approximately[12]

$$\gamma_1 = 1 - \frac{\kappa_w}{\kappa_w + 0.627\kappa\sqrt{\frac{V_s r_w}{D}}} \quad [6]$$

In Eq. 6, r_w is the wafer radius, κ is the pad thermal conductivity, and D is the pad thermal diffusivity. In Ref. 5, κ_w is taken to be the thermal conductivity of silicon, 124 W/m K. We then find that 25% of the frictional energy enters the pad, in agreement with an optimization of γ_1 that is independent of Eq. 6. Because the polishing surface is now different, we therefore again need to optimize this parameter, and the result is equivalent to a choice of k_w.

The second important factor that appears to distinguish oxide from copper is the pad-slurry heat exchange coefficient h_{ps} in the second term on the right side of Eq. 5. This coefficient is modeled with an empirical correlation[13] of the form

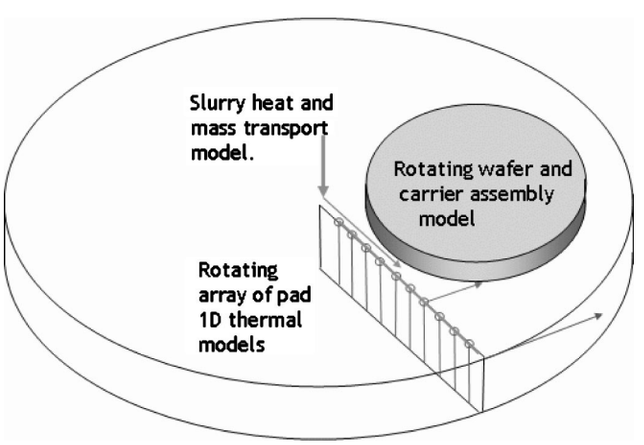

Figure 5. Schematic of the numerical approach in Ref. 5 for simulating heat-transfer in a rotary CMP tool.

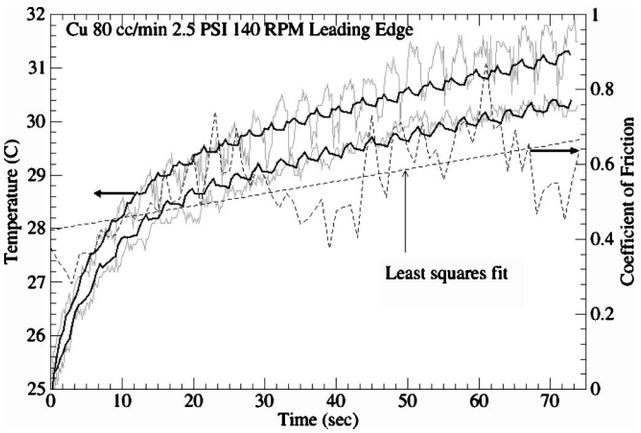

Figure 6. Measured (light solid curves) and simulated (dark solid curves) pad temperatures at pad locations 6 and 10 at 80 cm^3/min flow rate. Also shown is the measured COF and the linear regression to the COF that was used in the simulations.

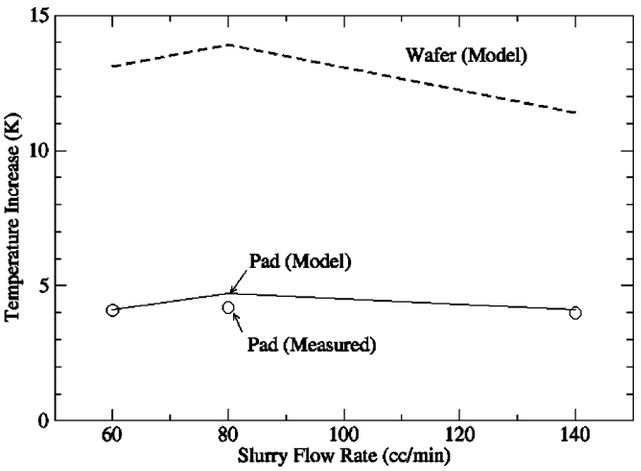

Figure 8. Comparison of overall mean (——) calculated and (○) measured pad temperature increase at sample points 6-10 *vs.* flow rate. (— — —) The calculated mean wafer temperature increase.

$$h_{ps} = h_{ps}^{1}\sqrt{V_{slurry}} \quad [7]$$

where h_{ps}^{1} is the exchange coefficient at unit flow velocity. The parameter h_{ps}^{1} may be affected by the slurry chemistry and by the size, solid fraction, and selection of material used for the slurry abrasive particles. Because an identical term with the opposite sign appears in Q_s in Eq. 3, h_{ps} influences the radial variation of the pad temperature distribution. In order to match the observed radial distribution in the baseline case at 80 cm^3/min during copper polishing, we find it necessary to recalibrate h_{ps}^{1}. At any other flow rate f_s, we then simply combine Eq. 7 and 4 to deduce that h_{ps} rescales as the ratio of the square root of the flow rates, $\sqrt{f_s/80}$.

The final recalibration concerns the wafer backing film. The backing film is composed of a soft open-cell foam bonded to a solid neoprene backing with an adhesive layer that is used to hold the film on the carrier. During use, the open cell foam absorbs slurry. Thus, its thermal properties are not straightforward to estimate from individual material properties. For copper, we simply adjust the thermal conductivity of the film until the calculated temperature on the top surface of the carrier matches estimates from IR images.

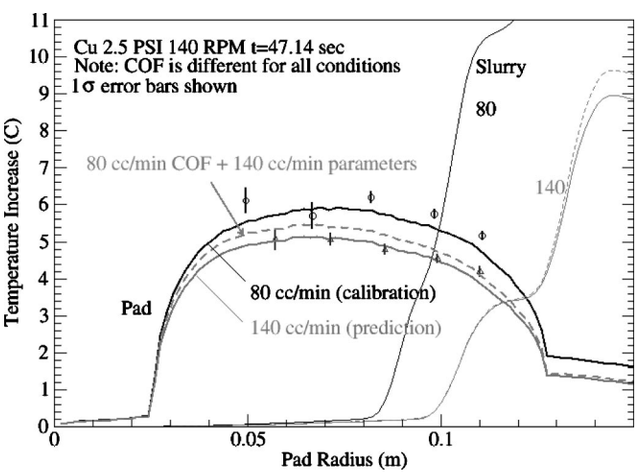

Figure 7. Simulated (solid curves) and measured (points with error bars) pad and slurry temperatures at 80 and 140 cm^3/min slurry flow rates along a radial line running approximately through sample points 6-10. Also shown are the slurry temperature distributions and a curve that suggests the relative influences of the COF and the pad-slurry heat-transfer coefficient.

Conditioning model.—In order to confirm whether the conditioner is responsible for the observed 0.33 Hz oscillations, an *in situ* thermal model was created for the conditioner. By analogy with the wafer and carrier assembly in Ref. 5, the conditioner is treated as a rotating circular steel disk. The nominal contact pressure at the disk surface is adjusted to reflect the actual contact area. Frictional energy is again partitioned between the conditioner and pad, and heat is exchanged with the slurry model each time that the latter passes under the conditioner. The conditioner nearly touches the pad center and perimeter at the extremes of the sweep. The sweep velocity is sinusoidal because of the mechanics of the sweep mechanism and has a half-period of 3.2 s.

Comparison of Theory and Experiments

Simultaneous recalibration of the two slurry model parameters γ_1 and h_{ps}^{1} described was performed with a downhill simplex algorithm[14] using the 80 cm^3/min flow rate data from sample point locations 6-10 as the calibration case. Except for the wafer backing film thermal conductivity, all other parameters in the thermal model reported in Ref. 5 were held constant. The optimized γ_1 was 0.18, indicating that less energy is entering the pad and more is partitioned to the wafer than in oxide polish. The effective wafer thermal conductivity in Eq. 6 is then κ_w = 217 W/m K, which lies between silicon (124 W/m K) and pure copper (385 W/m K). The intermediate value of κ_w may be conjectured to be related to the formation of a low-thermal-conductivity oxide on the copper surface during polishing. The heat-transfer coefficient was simultaneously reduced by the optimization algorithm from h_{ps} = 19.8 W/m^2 °C, as used in Ref. 5 for oxide polishing, to 11.7 W/m^2 °C. This reduction is related to the fact that the spread of the thermal profiles in the previous oxide experiment is tighter than the spread of the copper thermal profiles. More fundamental physical reasons for the reduction of h_{ps} were not investigated. The thermal conductivity of the wafer backing film was separately adjusted prior to the simplex optimization to bring the predicted temperature of the top surface of the carrier into agreement with temperatures estimated from thermal images; it was changed from 0.048 W/m K for the oxide carrier to 0.035 W/m K. Because of the high thermal resistance of the backing film, heat loss through the film is minor. Subsequent optimization of γ_1 and h_{ps}^{1}, therefore, had little effect on the calculated carrier surface temperature.

Figure 6 compares simulated with measured pad temperatures at

sample point locations 6 and 10 in the calibration case. The figure also shows the measured COF and the linear regression that was used in the simulation. The general rise in temperature and the temperature spread between the innermost and outermost sample points are accurately captured by the model. For oxide polish, pad temperatures had been observed to reach equilibrium within 15 to 20 s.[5] By contrast, pad temperatures do not reach equilibrium within this time in the copper polishing data. The continued increase in temperature after 20 s appears to be correlated with a general upward trend in COF, as indicated by the regression line. We also see from Fig. 6 that the conditioning model produces oscillations in the pad temperature that agree approximately in frequency and magnitude with the measurements at the outer point location 10. At inner point location 6, however, the simulated amplitude is much smaller that the measured amplitude. It is likely the measurements at this point are influenced by ripples in the slurry[7] and by direct slurry ''dragging'' as the conditioner passes in and out of the flow at the center of the pad.

Following calibration, the model was run with 140 and 60 cm^3/min flow rates without any further changes to test the predictive capability of the model. The results are shown in Fig. 3 and 4. We see that the model correctly estimates the temperature rise as a function of time and predicts the radial spread, but again, does not capture the magnitude of the 0.33 Hz excursions that occur at location 6. The rapidity and magnitude of the downward temperature shifts again suggest that direct slurry transfer occurs when the conditioner interacts with the fluid near the slurry stream at the center of the pad.

Figure 7 shows simulated slurry temperature profiles at 80 and 140 cm^3/min along a radial line running approximately through sample point locations 6-10 after nearly 50 s of polishing. Because the trend in the data is nearly linear between 30 and 70 s, a least-squares regression line was fit as a function of time at each sample location and used to calculate temperature estimates and 1σ error bars. As seen in Fig. 7, the radial variation in temperature across sample locations is adequately described at both flow rates. Also shown in this figure is an analysis using the model of the relative influence of COF and flow rate. In the analysis, the 140 cm^3/min simulation was rerun using the measured 80 cm^3/min COF instead of the 140 cm^3/min COF; all other parameters were unchanged. We see that the radial temperature distribution shifts about halfway between the two flow rate simulations. This suggests that the COF accounts for about half of the change and that convective heat-transfer by the slurry accounts for the rest in this data set.

Figure 8 shows the calculated and measured mean pad temperature increase above ambient at all three flow rates at locations 6-10 along with the calculated mean wafer temperature increase. We see from this figure that the calculated and measured mean pad temperatures agree well. The mean wafer temperature is also predicted to be higher than the mean pad temperature, in qualitative agreement with Ref. 9 and 10. The mean pad and wafer temperature changes from the model may also be seen to be linearly correlated, also in agreement with Ref. 9 and 10. More extensive data demonstrating this correlation at other pressures and velocities can be found in Ref. 6.

Conclusions

Through minimal recalibration and addition of a thermal model for *in situ* conditioning, we have applied a model developed for simulation of pad temperatures during oxide polishing to copper polishing. Given the COF, the model correctly provides the pad temperature as a function of time and position, including a portion of the thermal oscillation signal caused by the conditioner. The model also gives an estimate of the wafer temperature, suggesting that the rise in wafer temperature may be higher than the observable rise in pad temperature. The wafer and pad mean temperatures are also found to be linearly correlated, in qualitative agreement with direct measurements on a similar system.

Acknowledgments

The authors express their gratitude to Rodel, Incorporated, and JSR Corporation for pad donation and to Professor Duane Boning of MIT for allowing the authors to borrow MIT's infrared camera for thermal analysis. Special thanks are also due to Nikko Materials for the generous donation of copper disks used in the experiments. This work was financially supported by the NSF/SRC Engineering Research Center for Environmentally Benign Semiconductor Manufacturing.

The University of Arizona assisted in meeting the publication costs of this article.

References

1. P. Renteln, S. Srivatsan, and K. Y. Ramanujam, in *Proceedings of the 2003 CMP-MIC*, IMIC, pp. 344-348 (2003).
2. D. G. Thakurta, D. W. Schwendeman, R. J. Gutmann, S. Shankar, L. Jiang, and W. N. Gill, *Thin Solid Films*, **414**, 78 (2002).
3. J. Sorooshian, D. DeNardis, L. Charns, Z. Li, D. Boning, F. Shadman, and A. Philipossian, in *Proceedings of the 2003 CMP-MIC*, IMIC, pp. 43-50 (2003).
4. P. Renteln and T. Ninh, *Mater. Res. Soc. Symp. Proc.*, **566**, 155 (1999).
5. L. Borucki, L. Charns, and A. Philipossian, Abstract 918, The Electrochemical Society Meeting Abstracts, Vol. 2003-2, Orlando, FL, Oct 12-16, 2003.
6. Z. Li, L. Borucki, and A. Philipossian, Abstract 909, The Electrochemical Society Meeting Abstracts, Vol. 2003-2, Orlando, FL, Oct 12-16, 2003.
7. D. White, J. Melvin, and D. Boning, *J. Electrochem. Soc.*, **150**, G271 (2003).
8. H.-J. Kim, D.-H. Kwon, H.-D. Jeong, E.-S. Lee, and Y.-J. Shin, in *Proceedings of the 2002 CMP-MIC*, IMIC, pp. 201-208 (2002).
9. J. Cornely, C. Rogers, V. Manno, and A. Philipossian, in *Chemical Mechanical Planarization*, S. Seal, R. L. Opila, C. Reidsema-Simpson, K. Sundaram, H. Huff, and I. I. Suni, Editors, PV 2002-1, pp. 185-192, The Electrochemical Society Proceedings Series, Pennington, NJ (2002).
10. J. Cornely, C. Rogers, V. Manno, and A. Philipossian, *Mater. Res. Soc. Symp. Proc.*, **767**, 33 (2003).
11. A. Philipossian and S. Olsen, *Jpn. J. Appl. Phys., Part 1*, **42**, 6371 (2003).
12. R. Cowan and W. Winer, in *ASM Handbook*, Vol. 18, P. J. Blau, Editor, pp. 39-44, ASM International, Materials Park, OH (1992).
13. Y. A. Cengel, *Introduction to Thermodynamics and Heat Transfer*, McGraw-Hill, New York (1996).
14. W. H. Press, S. A. Teulkolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed., p. 994, Cambridge University Press, New York (1992).