Do You Know What Your Retaining Ring and Wafer are Doing?

Len Borucki, Yasa Sampurno, and Ara Philipossian

Araca, Inc., Tucson AZ 85718 USA; email: lborucki@aracainc.com

INTRODUCTION

Commercial CMP polishers have for decades used wafer retaining rings; however, the separate effects of the ring and wafer on a process and their interaction can still be opaque. For example, without special instrumentation, it may be impossible to isolate the shear force acting on the wafer from that on the ring, thereby complicating the study and interpretation of important removal rate and thermal phenomena.

We describe here a method for estimating individual ring and wafer normal and shear forces and coefficients of friction (COFs) on a 200 mm R&D polisher with a retaining ring, the Araca RDP-500 [1]. The tool measures and reports the net forces, the platen and carrier rotation rates and the pad temperature at a trailing edge spot. The method requires only an experiment already typical of the polisher's use. We then show an example in which separated wafer shear force data are used to model the removal rates of a very non-Prestonian slurry.

AN EXAMPLE EXPERIMENT

We explain the data separation method using an experiment run at 3 wafer pressures (1,2,3 PSI) and 3 platen speeds (60,90,120 RPM or 0.8,1.2 and 1.6 m/s) using a constant 6 PSI retaining ring pressure. Fig. 1 shows 0.5 s running averages of the 1000 Hz normal and shear force data and the platen rotation rate.

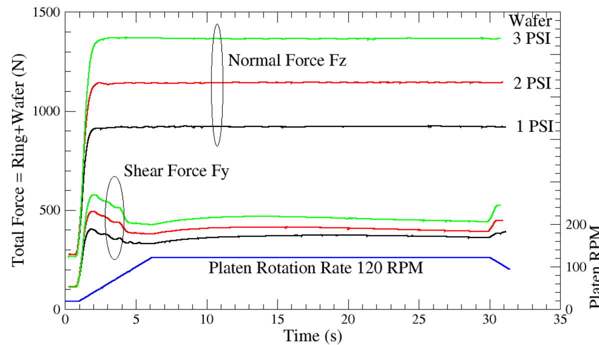

Fig. 1: Normal and shear forces at 120 RPM with a 6 PSI ring and the wafer at 1, 2 or 3 PSI using blanket W wafers, a W buff slurry and an IC-1000 pad.

The experiment begins with the platen at 20 RPM, the wafer at 0.5 PSI and the ring at 1.5 PSI. At 1 s, the platen begins ramp-up to 120 RPM, reaches polishing speed at 6 s, then maintains it until 30 s, when there is a brief ramp-down. The total normal force Fz also ramps up at 1 s but reaches steady state at 3 s and remains constant until the end of the process.

The shear force Fy is more interesting. Its evolution parallels the normal force until 2 s, when it begins to decrease. Since the down force is constant for t > 3, the decline in shear force is related to the continuing platen ramp-up and indeed stops when the platen speed stabilizes. Conversely, when the platen ramps down, the shear force increases at constant down force. This ramp behavior is likely due to lubrication of the pad/wafer interface at summit contacts since a lubrication layer inevitably forms there and speed is necessary to maintain it.

MODELING THE WAFER AND RING NORMAL FORCES

The polisher controls the ring and wafer down forces but does not report them. However, they are easy to separate. Suppose that the wafer normal force is ramped from F_w^0 to F_w^1 and the ring from F_r^0 to F_r^1, where these set points are known. The total normal force can be decomposed into the sum of the wafer and ring contributions, where the two have similar forms:

$$F_z(t) = F_w(t) + F_r(t) \quad (1)$$

$$F_w(t) = F_w^1 - (F_w^1 - F_w^0) \cdot f(t) \quad (2a)$$

$$F_r(t) = F_r^1 - (F_r^1 - F_r^0) \cdot f(t) \quad (2b)$$

Here, $f(t)$ is a transition function that is 1 initially and should go to 0 after the ramp-up. Since we know the set points, we have from (1) and (2) that $F_z(t) = A - B \cdot f(t)$ where A and B are known. Since $F_z(t)$ is measured, we can then solve for $f(t)$. The transition functions for Fig.1 are shown in Fig. 2. They are translated slightly upward because the actual total force is ~3.3% below the set points. The results of the normal force analysis are in Fig. 3.

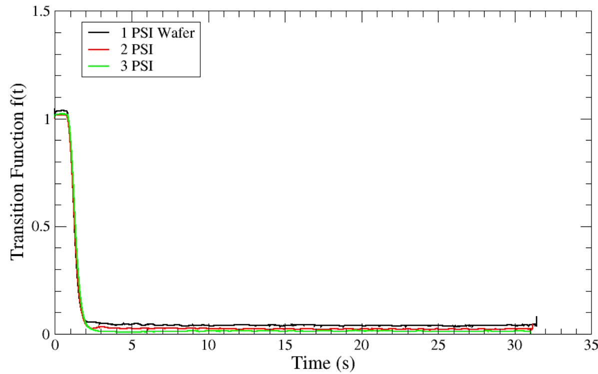

Fig. 2: Transition functions for the data in Fig. 1

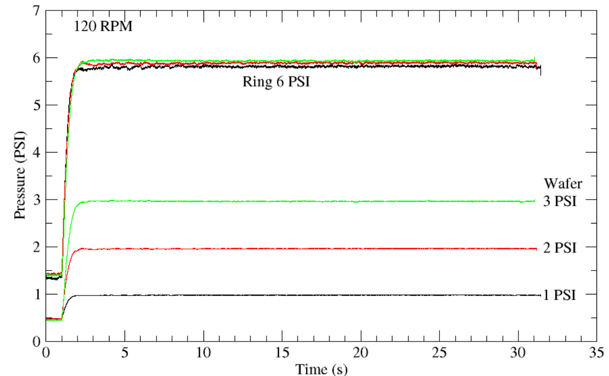

Fig. 3: Normal force pressures for Fig. 1

WAFER AND RING SHEAR FORCE SEPARATION

Unlike the normal force, the shear force is not controlled, so it requires a different analysis. We use two tools for this: extrapolation and differencing.

Fig. 4 illustrates the extrapolation step, whose purpose is to estimate the shear force on an isolated ring. This cannot be measured directly because the tool is designed to halt if there is no wafer in the carrier. In the figure, at each time, the total shear force at the three wafer pressures decreases approximately linearly with pressure. The blue curve in the figure is then the least squares linear extrapolation of the shear force to 0 PSI. Since the shear force measurements include not only solid-solid contact forces but also hydrodynamic forces, the extrapolation should zero out all of these for the wafer. Therefore, we refer to the result as an approximation of the shear force with *no* wafer rather than the shear force with a wafer in the pocket at 0 PSI. It is perhaps interesting that the increase in total shear force seen during ramp-down is not present for the ring, suggesting that it is associated with the wafer.

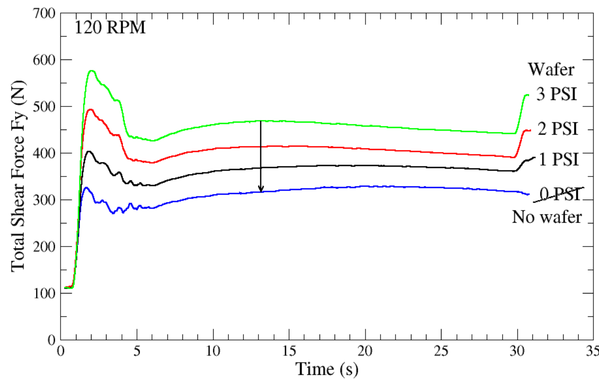

Fig. 4: Extrapolate to estimate the ring shear force

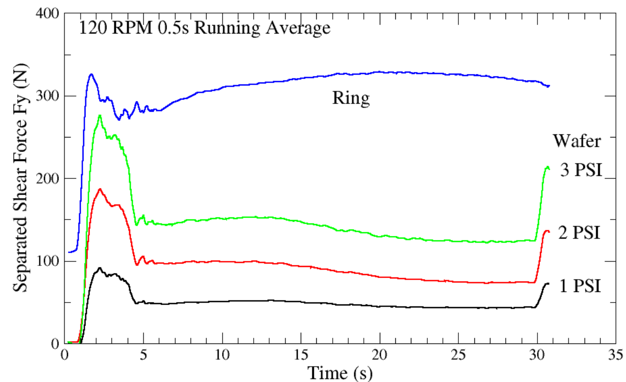

Fig.5: Separated ring and wafer shear forces.

Given the ring shear force, the wafer shear force at each pressure can be estimated by differencing; *i.e.,* by subtracting the ring shear force from the total shear force. The result for Fig. 1 is shown in Fig. 5.

The shear force separation procedure is somewhat more general than the example suggests. While it does require the measurement of the total shear force, preferably with a force transducer, it can be done with just two wafer pressures, with one wafer pressure and several ring pressures or by pulsing the ring pressure.

FRICTION COEFFICIENTS

The friction coefficient for the ring or wafer is just the ratio of the shear force in Fig. 5 to the normal force underlying Fig. 3. The result for the example is shown in Fig. 6. The COF of the ring dominates that of the wafer. The wafer COF is similar to a classical COF in that it is almost independent of the applied pressure, as indicated by the black, red and green curves. The naïve COF for the ring and wafer together, Fy/Fz, is also shown in light grey. It is quite different than the wafer COF and might suggest misleading conclusions; *e.g.* that the wafer COF decreases with increasing pressure.

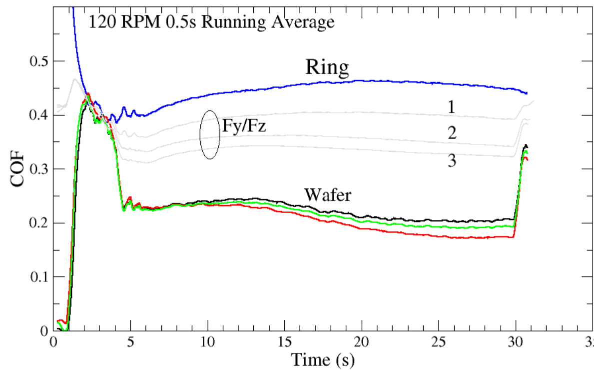

Fig. 6: Friction coefficients for the data in Fig. 1.

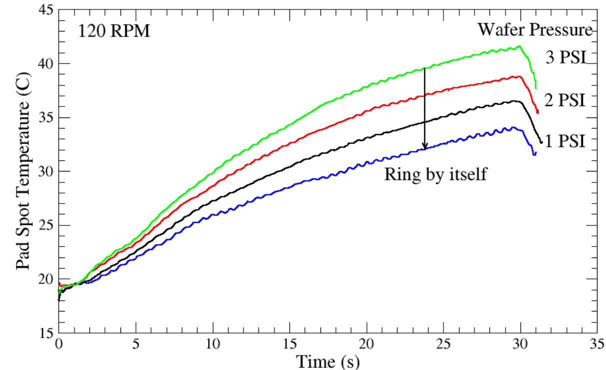

Fig. 7: Estimate of the spot temperature for the ring alone at 120 RPM using extrapolation.

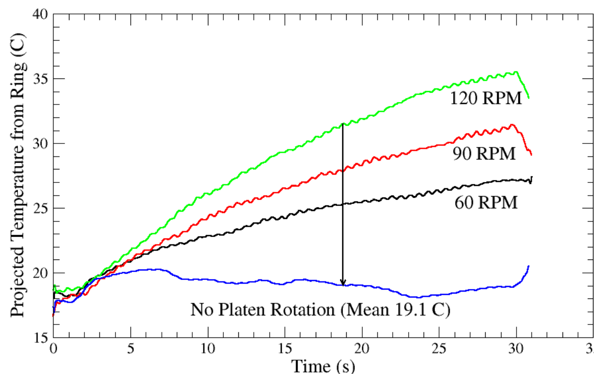

Fig. 8: The lab temperature by extrapolation.

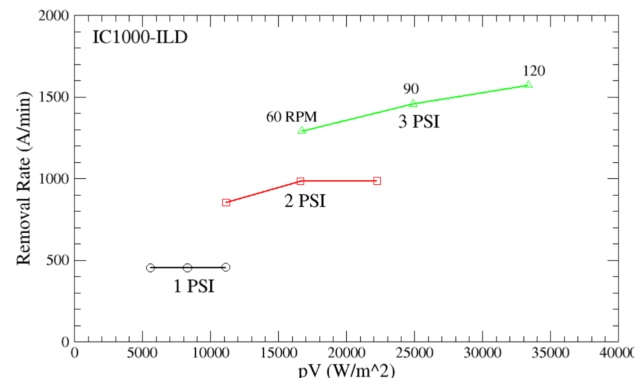

Fig. 9: Preston plot of ILD removal rates from a non-Prestonian slurry using the same experiment as in Fig. 1.

VALIDATION

The polisher cannot be run without a wafer, so indirect validation methods have to be used. The pad temperature increase should be linear in the shear force since shear force friction causes the heating. By analogy with the shear force analysis, we should then be able to use temperature data to estimate the spot temperature using the ring alone. We can't measure the latter, but we can further use the ring estimates at the three available platen speeds to project the spot temperature when there is *no platen rotation*. We *know* that the latter should be the lab temperature, so this provides a weak check on the procedure.

Fig. 7 shows the projection of the spot temperature from the ring alone for the 120 RPM data in Fig. 1. Fig. 8 shows similar projections for the runs at 60 and 90 RPM along with an extrapolation to 0 RPM. The mean of the latter is 19.1 +/-1 C, close to the measured lab temperature and validating the procedure in an unusual way.

APPLICATION TO REMOVAL RATE MODELING

We next discuss a significant application that shows why we need to know the wafer shear force. Fig. 9 shows a Preston plot of removal rate vs pV for an experiment using the same kinematics and slurry as in Fig. 1 but on blanket ILD wafers. Since Preston's Law [2] asserts that the removal rate is proportional to pV, $RR = c_p \cdot pV$, where c_p is a constant, the slurry in the figure is non-Prestonian for several reasons: 1) The data do not lie on a straight line passing through the origin, nor do they project to 0 removal at 0 speed for any pressure; and 2) There are values of pV at which there are *two* removal rates! How should we understand such a slurry? Part of the problem is that we are viewing the data in the wrong way.

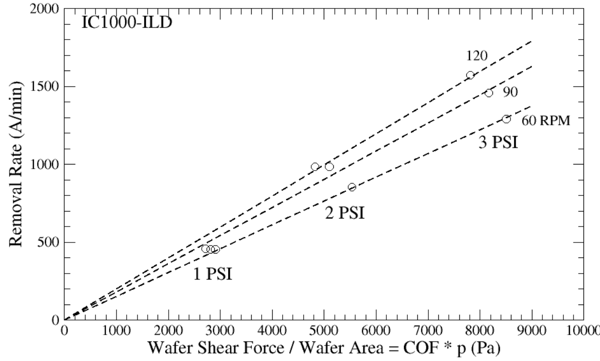

Fig. 10: The data in Fig. 9 plotted as a function of the mean wafer shear force between 5 and 30 s divided by the wafer area.

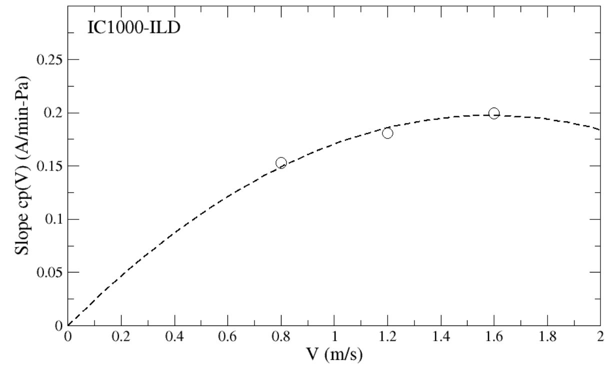

Fig. 11: Slopes of the lines in Fig. 10. The slope is also necessarily 0 when the platen is not rotating

No CMP tool removes material without applying shear force to the wafer, so it must be fundamental. Fig. 10 shows the data from Fig. 9 plotted against the mean wafer shear force normalized by the wafer area (which is the same as the mean wafer COF times p). Means for each condition are taken between 5 and 30 s, when the process is stable. At each speed, we see that the removal rate data fall on or near a line that passes through the origin, so that the removal rate is zero when the shear force is zero. The slopes $c_p(V)$ of the lines are analogous to Preston constants. They are plotted in Fig. 11.

If the static etch rate is close to zero, then we also *know* that when the platen is not turning, the removal rate is 0 regardless of the pressure. The function $c_p(V)$ must therefore always pass through the origin. The simplest quadratic function that meets the requirements is $c_p(V) = (\alpha - \beta \cdot V) \cdot V$. For the example, the best fit in Fig. 11 is

$$c_p(V) = (0.249 - 0.079 \cdot V) \cdot V \quad (3)$$

THE SHEAR FORCE LAW

We find then that the removal rate has the form

$$\begin{aligned} RR &= c_p(V) \cdot WaferShearForce \,/\, WaferArea \\ &= c_p(V) \cdot COF \cdot p \end{aligned} \quad (4)$$

where V is a factor of $c_p(V)$. We call this the *Shear Force Law*. It is a *generalization* of Preston's Law as can be seen by moving the factor V in $c_p(V)$ next to p. It reduces to Preston's Law when $\beta = 0$ if the COF is constant. In general, we find that the COF may also have some velocity and pressure dependence; however, this is irrelevant to the model's applicability. Fig. 12 compares the model with the data. It reproduces all of the non-Prestonian features of the data with an RMS error of 46 A/min (< 10%).

With our formulation of $c_p(V)$, the model also makes predictions that go beyond Preston's Law. The first is that the removal rate is 0 at *two* speeds: $V = 0$ and $V = V_h = \frac{\alpha}{\beta}$. The latter can be interpreted as the

hydroplaning velocity. For the example, it is 3.2 m/s. At half of the hydroplaning velocity, the model also predicts that *the removal rate has a maximum that is linear in the pressure*. For the example, the maximum should occur at 1.6 m/s. There are suggestions in Fig. 12 that this may indeed be the case.

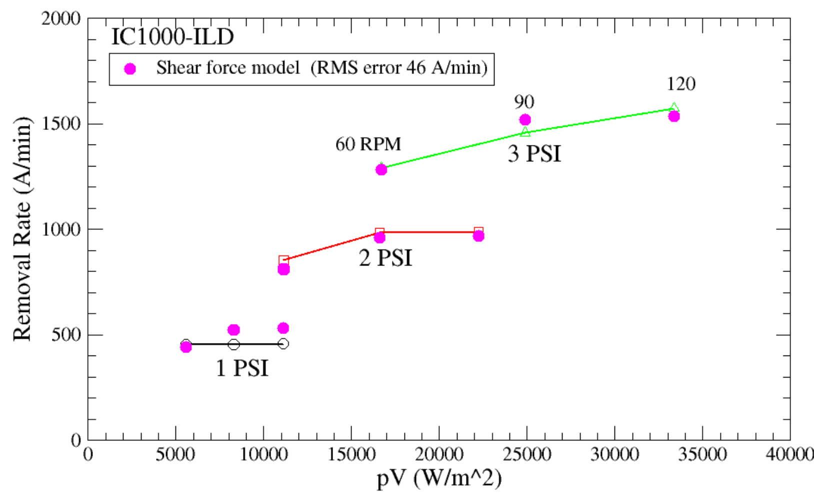

Fig.12: Comparison of the shear force model with removal rate data.

The shear force law, which we have verified for more than 16 examples, has potentially useful implications for polisher design and operation. When the removal rate is proportional to the shear force, the former can be directly controlled by monitoring and manipulating the latter. For example, instead of controlling the pressure on the wafer and letting the shear force drift with time as in Fig. 5, the wafer pressure could be varied so that the shear force (removal rate) is constant. In a process that would ordinarily exhibit polish rate decay, like *ex situ* oxide polishing, this would automatically counter the decay and shorten the polishing time. Shear force control might also be a more explicit way to reduce scratch defects. These and other applications are the subject of [3].

SUMMARY

We have developed methods for teasing apart the wafer and retaining ring shear and normal forces on polishers that can measure the total force. We find that the wafer shear force, which is always present in CMP, is particularly helpful for understanding removal rates. In particular, the shear force law empirically seems to be what underlies the behavior of both Prestonian and non-Prestonian slurries. It further predicts phenomena like hydroplaning and might be a viable basis for an alternative method of running polishing machines.

REFERENCES

1. RDP-500 polisher details can be found at http://aracainc.com/equipment/r-and-d-polishers.
2. Preston, F.W. (1927) The Theory and Design of Plate Glass Polishing Machines. Journal of the Society of Glass Technology, 11, 214-256. Curiously, this important paper contains no supporting data for Preston's Law but says only that *"… there is good experimental evidence to believe that the amount of polishing done in time t is proportional to* pvt*."*
3. 2022. CHEMICAL MECHANICAL PLANARIZATION (CMP) APPARATUS AND METHOD. US 63397932 filed 08/15/2022. Provisional Patent.

Corresponding Author:
Len Borucki
Tel: +1 480-748-5105
E-mail: lborucki@aracainc.com
2550 E. River Rd., Tucson, AZ 85718 USA