Investigating Slurry Transport beneath a Wafer during Chemical Mechanical Polishing Processes

Coppeta, J., Rogers, C., Racz, L., Philipossian, A., Kaufman, F. B.

1 Abstract

In order to better understand the role of the fluid behavior in CMP, we measured slurry transport beneath a glass wafer. We quantified the slurry transport using Residence Time Distribution techniques and two measures of slurry transport efficiency defined as the percentage of new slurry beneath a wafer. Slurry transport efficiency depended on platen speed, flow rate, and the conditioning method. We found that the average fluid residence times under the wafer decreased roughly linearly with platen speed. A threefold increase in platen speed decreased slurry mean residence times by three and a half times. Changing the flow rate from 20 cc/min to 50 cc/min decreased the slurry mean residences by 80%. In situ conditioning generally increased the slurry mean residence times and the amount of slurry mixing. In situ conditioning also decreased gradients in the slurry composition across the wafer. Pad topography had a large effect on the slurry gradients that developed across the wafer. Finally, we found that the slurry mixing history can be accurately modeled using a simple continuous function.

2 Introduction

Since its introduction into integrated circuit (IC) manufacturing by IBM corporation in the mid 1980s, chemical mechanical planarization (CMP) has become a key enabling technology to the semiconductor industry. All of the major IC manufacturers including Intel, Motorola, and IBM now incorporate CMP in the
production of its chips. In addition to IC production, CMP applications have spread into other manufacturing processes including DRAM chips, hard drives, and modem chips.

Given the magnitude of capital invested in this technology, there is a large impetus to develop a fundamental understanding of the process. Research with this goal in mind is being performed, however, an overall understanding of the process remains elusive because of the multi-disciplinary nature of CMP. Researchers have focused on individual aspects of the process such as slurry chemistry [1, 2, 3, 4, 5], wafer-pad dynamics [6, 7, 8], wear mechanisms [9, 10, 11, 12, 13, 14], and numerical simulations of the slurry fluid mechanics [15, 16, 17]. There has been, however, little experimental research regarding slurry fluid mechanics.

Several researchers have commented on the importance of slurry flow and slurry distribution beneath the wafer although to date no one has shown it experimentally. Stavreva et al. [18] discussed how the pad’s ability to transport slurry could affect the polishing rate and uniformity during copper CMP. Parikh found that slurry flow rate had a large effect on the polishing uniformity on an orbital polisher [19]. Ali et al [20] stated that the slurry composition, flow rate and direction of slurry impingement onto the polishing pad all play important roles in interdielectric removal rates. Singer [21] reported that the manner in which slurry is transported from the outside of the wafer to its center is critically important. Sugimoto et al. [22] showed that slurry transport in grooved pads is important in reducing thermal gradients across a wafer. Since polishing rates are temperature dependent, a reduction in thermal gradients across the wafer is believed to reduce the within-wafer non-uniformity. Ali et al. [23] postulated that the degradation in the removal rates of pads without conditioning is due to the decrease in the pad’s slurry holding capacity. Liang et al. [24] postulated that Cabot’s new open cell pads do not need macroscopic surface topography because of the pad’s efficiency in channeling the slurry. Despite the fact that slurry flow generally is considered to be an important factor in the CMP process, there has not been an experimental study of the slurry flow or a numerical simulation sophisticated enough to examine the slurry behavior under realistic conditions.

Slurry transport and mixing could influence the polishing performance two ways; 1.) transport of polished material and 2.) non-uniform slurry transport. The first mechanism was postulated by Cook in his research on glass polishing [4]. Cook suggested polishing removal rate is influenced by the transport of polished
material away from the wafer’s surface and substrate material is removed only if it is transported away from the wafer or chemically bound in some manner. Therefore, slower slurry entrainment rates should yield slower polishing rates. The same physical argument applies to slurry mixing. A reduction in slurry mixing should increase the polishing rate (provided that the slurry is efficiently transported to and from the wafer surface) because polished material is not mixing with the new slurry, resulting in more new slurry delivery to the wafer surface.

Non-uniform slurry transport may also influence the polishing uniformity. Uneven or asymmetrical mixing across the wafer’s surface may cause some areas of the wafer to be constantly exposed to a different slurry environment than other areas of the wafer. We have noted that our polishing rates correlate with the non-linear slurry transport across the wafer’s surface. That is, the center of the wafer entrains new slurry more slowly than the edges of the wafer and the polishing rates are faster on the edge of the glass compared with the center of the glass. Although our results do not prove that the edge fast polishing is caused by the uneven slurry transport, further investigation is warranted. This paper focuses on factors that influence slurry transport beneath a wafer while polishing.

3 Experimental Apparatus

We have created a functional small scale version of an industrial silicon dioxide (oxide) chemical mechanical polishing platform in order to develop a fundamental understanding of the CMP process. Although there are a variety of other polishing configurations including linear and orbital polishers, the rotary polisher is the least complicated and therefore was adopted as a starting point. Likewise, our setup was built to mimic oxide substrates because it the most common type of polishing and the simplest in terms of slurry chemistry and consumable sets. Figure 1 shows an example of typical polishing rates obtained on the tabletop polisher. The polishing substrate was BK7 optical glass polished on an IC1000 pad with Cab-O-Sperse SC1 slurry. Note that the polishing rates are approximately 2000 Å/min. Our system consistently has higher removal rates on the edges of the wafer compared to the center of the wafer (edge-fast polishing). Experiments

3
showed that the polishing rates are not affected by the presence of fluorescent dyes.

In order to create a 1:2 scale model of an industrial polisher, there were a number of parameters that had to be scaled. These parameters include platen diameter, wafer diameter, slurry flow rate, platen speed, head speed and wafer load. Since the IPEC 472 industrial oxide polisher is commonly found in IC circuit manufacturing centers, this platform was chosen as our industrial model. Scaling factors for each of these parameters are chosen so that the polishing fluid mechanics mimic that of an IPEC 472 polisher. Table 1 shows the appropriate scaling factor for each parameter as well as a numerical comparison of our parameter values versus typical IPEC 472’s parameter values. The ratio of platen to wafer diameters and the slurry flow rate normalized by the platen area were set equal to values found on an IPEC 472. The platen and head speeds are scaled by keeping the range of slurry Reynolds numbers equivalent between our model and the IPEC 472. An equivalent Reynolds number can be achieved by making the relative velocity between pad and wafer on our model equal to the relative velocity between the pad and wafer on the IPEC 472. This assumes that the kinematic viscosity of the slurry and the pad-wafer gap on the two polishers is equivalent.

Figure 2 shows our modified CMP setup used to study the slurry flow beneath a wafer. A tabletop Struers RotoPol-31 polisher is used to rotate a 30.5 cm (12”) polishing pad. The standard RotoPol head has been replaced with a 20” industrial rated drill press that both rotates and applies down force to a 76 mm (3”) wafer. A weighted carriage mounted on a traverse provides a variable (7-70 kPa +/- 1 kPa or 1-10 psi +/- 0.2 psi) down force to the wafer. The traverse is attached to the drill press at two locations; the traverse’s pivot point is located above the drill press supporting column while the other end of the traverse is prevented from pivoting by the drill press shaft. This arrangement allows the down force to be transmitted to the wafer without creating a moment about the drill press itself. The wafer is gimbaled so that it remains roughly parallel to the pad surface.
Since we are measuring fluid parameters using an optical technique known as dual emission laser-induced fluorescence or DELIF (described in Section 4) [25], the wafer must be transparent and a pure silicon wafer cannot be used. Instead, a glass wafer that is transparent to visible light is used. Two high-resolution spatially aligned 12 bit digital cameras are used to measure the fluorescence beneath the wafer. In situ or ex situ pad conditioning can be performed using the device shown in Figure 3. A 76 mm (3”) diamond grit wafer co-rotating with the platen conditions the pad by periodically sweeping across the pad radius. Many of the polishing parameters including platen speed, down force, slurry delivery, and conditioning speeds are computer controlled and monitored. In addition, the computer synchronizes the camera to the polishing process so that we can interrogate the wafer at any point in the polishing process.

[Figure 3 about here.]

4 DELIF Measurement Technique

The DELIF technique uses the fluorescence from two different dyes each fluorescing at different wavelengths to measure mixing, fluid depth, or temperature and is described in detail in Coppeta and Rogers [25]. The fluorescence from one dye contains the information about the parameter of interest as well as the laser distribution and film thickness information. If mixing is the desired parameter, fluorescence from the second dye will contain the laser distribution and film thickness information. Normalizing the fluorescence of the first dye by the fluorescence from the second dye causes the resulting ratio to be a function of mixing only. Our previous work (Coppeta et al. [26] - [27]) has shown that by injecting tagged commercially available slurry with two fluorescent dyes, slurry mixing, entrainment beneath the wafer, residence time on the pad, and film thickness can be measured. For the purposes of this paper, we are concerned with the slurry mixing only.

Repeated calibration curves have shown that the uncertainty in the DELIF measurements of mixing are less than 2% of the volumetric mixing. Mixing measurements beneath the wafer are an exponentially weighted average of the slurry mixing values at each fluid depth between the pad and wafer with the slurry
closest to the wafer having the greatest weighting factor. For the system we are investigating, fluid at depths
greater than 400 μm below the slurry surface does not contribute to the average measurement (although
this number varies depending upon the dye concentrations and the excitation and emission wavelengths).
Therefore, we are unable to measure the slurry mixing in pad grooves greater than 400 μm (15 mils).

5 Methodology

The two-camera system acquires a series of images whenever a pulse of "new" slurry is injected onto the
pad where "new" slurry refers to the fluorescent dye used to tag the slurry. Unless otherwise noted, slurry
tagged with Dihydroxyphthalonitrile (green fluorescence) is considered "new" slurry and slurry tagged with
Coumarin (blue fluorescence) is considered "old" slurry. New slurry was pulsed onto the pad for a 60 second
duration. Images were typically acquired at 0.5 Hz for up to two minutes in order to capture the entire
mixing history. The exposure time for each image was 700 ms.

Figure 4 shows an example of the mixing history beneath the wafer; note the two runs are very repeatable.
Time zero is defined by the moment when the new slurry reaches the pad. Since mixing measurements were
made under the wafer, one can see the time delay between when the new slurry is first introduced onto
the pad (time zero) and when the new slurry is entrained beneath the wafer as evident in the rise of the
mixing curve (Figure 4). Through image processing, we aligned the two images to within a fraction of one
pixel in the vertical and horizontal directions. Since each pixel maps onto 50 μm in the image plane, our
spatial resolution is less than 100 μm in each direction. Image processing was also used to select a subsection
from each image for analysis. The image subsections were then used to calculate statistics such as the mean
concentration of new slurry, standard deviation, and mixing gradients. The image subsection was selected by
the user for each case, and all images within a case were analyzed using the same image subsection. Because
the camera may have been moved over a period of days, the image subsection had to be re-selected between
cases so that the same part of the wafer was analyzed. The subsection was centered about the center of
the wafer and was approximately 2.25 by 3.8 cm, as shown in Figure 5. This image area is focused onto
approximately 260,000 pixels of the CCD. Therefore, each point in the mixing history represents the average of 260,000 data points.

[Figure 4 about here.]

[Figure 5 about here.]

The mixing curves were analyzed using Residence Time Distribution techniques [28]. The mean residence time was used to measure the rate of slurry transport. The mean residence time (MRT) is the average time the slurry spends beneath the wafer and is

\[
MRT = \frac{\int_0^\infty t \frac{dF}{dt} dt}{\int_0^\infty \frac{dF}{dt} dt},
\]

where \( F \) is the mixing curve obtained for a step input (60 second pulse here) and \( t \) is time measured in seconds. The variance term is the amount of slurry mixing beneath the wafer and is

\[
\sigma^2 = \frac{\int_0^\infty (t - MRT)^2 \frac{dF}{dt} dt}{\int_0^\infty \frac{dF}{dt} dt}.
\]

Therefore, the variance is a measure of the amount of slurry mixing and the mean residence time is a measure of the rate of slurry transport beneath the wafer.

Finally, since a polishing pad is an open system, it is useful to characterize slurry transport in terms of an entrainment efficiency. This efficiency can be defined as the volume of new slurry entrained beneath the wafer compared to the total amount of slurry entrained beneath the wafer for a given pulse length of new slurry. That is, the amount of new slurry that displaces the old slurry beneath the wafer within a certain process window where the process window is defined as a change in the process that lasts 60 seconds. This
efficiency can refer to just the wafer or the system as a whole. The first, a wafer efficiency, treats the wafer as a separate system independent of the platen and can be written as

\[ \eta_{\text{wafer}} = \frac{V_{\text{New Entrained}}}{V_{\text{Total Entrained}}} \times 100, \]  

(3)

where \( \eta_{\text{wafer}} \) is the transport efficiency, \( V_{\text{New Entrained}} \) is the volume of new slurry that becomes entrained beneath the wafer, and \( V_{\text{Total Entrained}} \) is the total amount of slurry entrained beneath the wafer in a given period of time. For \( \eta_{\text{wafer}} \) the amount of new fluid is compared to the total fluid beneath the wafer when the new slurry first begins to be entrained. The second efficiency that can be defined includes the slurry transport to the wafer and is written as

\[ \eta_{\text{system}} = \frac{V_{\text{New Entrained}}}{V_{\text{Total Entrained}}} \times 100, \]  

(4)

where \( \eta_{\text{system}} \) is the slurry transport across the pad and beneath the wafer. The two efficiencies \( \eta_{\text{wafer}} \) and \( \eta_{\text{system}} \) differ by the slurry fluid mechanics that transport the slurry to the wafer and are shown schematically in Figure 10. Note that \( \eta_{\text{system}} \) will be less than \( \eta_{\text{wafer}} \) because of the additional slurry entrained beneath the wafer before the new slurry is transported to the wafer.

[Figure 6 about here.]

Both formulations of efficiency are time dependent since a longer pulse of new slurry will have a higher efficiency; a longer pulse length will cause the transient mixing to become less significant. We feel the time dependence in the efficiencies is important because wafers are typically polished for a relatively short period of time. In cases where the transients in the mixing curves can be on the same order of magnitude as the polishing time (minutes), the transient behavior is important.
Experimental Results

This section shows how slurry transport is influenced by platen speed, flow rate, conditioning, and pad grooving. Plots with the prefix "cond" in the legend refer to data taken with in situ conditioning; those without refer to runs with ex situ conditioning. Time zero in the plot is when the new slurry first hits the center of the pad.

A single example of the wafer mixing history (Figure 7) illustrates a number of important concepts. First, the 60 second mixing curves are generally shaped like an asymmetrical sigmoid. The rate of new slurry entrainment is very fast initially but then slows down as the concentration of new slurry beneath the wafer approaches 100 percent. This is due to the fact that initially all new slurry is displacing pure old slurry both on the pad and beneath the wafer, causing a fast increase in the concentration of new slurry. As time progresses, the new fluid displaces a mixture of old and new slurry. As the slurry beneath the wafer and on the pad approaches 100 percent new slurry, the new slurry introduced onto the pad is displacing slurry that is very nearly the same concentration. After the 60 second period, no more "new" or tagged slurry is introduced onto the pad and therefore the mixing plot drops back to zero.

Figure 7 also shows that in situ conditioning reduces the rate of new slurry entrainment beneath the wafer for grooved pads. This is reflected in the slurry MRT which is 60% higher with in situ conditioning (22 seconds) compared with ex situ conditioning (13.5 seconds). Flow visualization shows that the conditioner pulls fluid out of the grooves to the top of the pad. This action would reduce the rate of new slurry entrainment by impeding the slurry flow in the grooves and mixing the new slurry with old slurry on the pad and in the pocket that holds the conditioning wafer. The conditioner’s motion significantly enhances slurry mixing as measured by the variance in the slurry mixing. With in situ conditioning the variance is 0.74 and compared to 0.41 with ex situ conditioning. In addition, the oscillations in the conditioning mixing curves match the oscillation period of the conditioner sweeping across the pad. As the conditioner moves away from the pad center more slurry is drawn onto the pad surface where it accumulates in a standing wave at the leading edge of the wafer (bow wave). Since the slurry deposited on the pad surface by the conditioner is a mixture and is older than the new slurry introduced onto the pad, the rate of new slurry...
decreases as the conditioner moves away from the pad center. The wafer's slurry transport efficiency is 13% higher with ex situ conditioning (84%) compared with in situ conditioning (71%) because the conditioner is diluting the new slurry with old slurry. The system’s transport efficiencies are 79% and 67% for ex situ and in situ respectively. The system efficiencies are characteristically lower than the wafer’s efficiency since pure old slurry is entrained beneath the wafer until the new slurry is transported to the wafer. The difference between \( \eta_{wafer} \) and \( \eta_{system} \) are similar between the in situ and ex situ conditioning cases indicating that the conditioner does not play a large role in bulk slurry transport to the wafer for grooved pads.

While in situ conditioning reduces new slurry entrainment rates for grooved pads it has the opposite effect for flat pads as shown in Figure 8 (Rodel IC1000 pad, 60 rpm, 35 cc/min, 4 psi, no grooving). The slurry MRT for in situ conditioning is 30% lower than ex situ conditioning (13.8 versus 19.5 seconds respectively). Flow visualization shows that in the absence of in situ conditioning, the slurry forms a pool in the center of the pad and the wafer precludes the slurry from efficiently spreading out across the entire pad. With in situ conditioning, the conditioner breaks up the pool of new slurry in the center of the pad and distributes it across the entire pad thus increasing the slurry entrainment efficiency. This motion causes a 40% increase in the slurry mixing compared to the ex situ conditioning. Note that the conditioner still causes oscillations in the mixing curves. The first two oscillations cause a large increase in the mixing rate as each time the conditioner moves to the center of the pad it distributes the pool of new slurry beneath the wafer. The wafer’s slurry transport efficiency is approximately the same for both conditioning methods. The wafer’s transport efficiencies were 77% and 78% for ex situ and in situ conditioning respectively. A significant difference is evident in the \( \eta_{system} \) which is 70% and 78% for ex situ and in situ conditioning respectively. This difference indicates that unlike grooved pads, in situ conditioning plays a role in bulk slurry transport for flat pads. The higher \( \eta_{system} \) with in situ conditioning indicates that the conditioner sweeps slurry from the center of the pad to the wafer faster than the slurry would normally be transported without the conditioner.
In situ conditioning can reduce the rate of mixing for flat pads when the pad is flooded with slurry. This occurs when the platen is rotating slowly and the flow rates are moderate to high. Figure 9 shows an example of this situation (Cabot pad, 30 rpm, 35 cc/min, 4 psi, no grooving). A flooded pad stores a lot of old slurry because the slurry does not leave the pad very quickly. In situ conditioning forces the new slurry to mix with all of the old slurry on the pad thereby reducing the entrainment rate. With ex situ conditioning new slurry can displace the old slurry without having to completely mix with it, although recirculation patterns on the pad do cause some mixing to occur. In situ conditioning causes the slurry MRT to double compared to ex situ conditioning. These recirculation patterns are responsible for the oscillations evident in the ex situ conditioning mixing curves. The transport efficiencies were $\eta_{\text{wafer}}=62\%$, $\eta_{\text{system}}=57\%$ and $\eta_{\text{wafer}}=45\%$, $\eta_{\text{system}}=45\%$ for ex situ and in situ conditioning respectively. These values of $\eta_{\text{wafer}}$ and $\eta_{\text{system}}$ are the lowest of all of the conditions investigated and indicate that low platen speeds entrain new fluid inefficiently. Again in situ conditioning plays a large role in the bulk slurry transport to the wafer and dilutes the new slurry on the pad.

[Figure 9 about here.]

In addition to conditioning, platen speed and flow rate have a large effect on the mixing curves. Increasing either flow rate or platen speed increases the entrainment rate, as seen in Figures 10 and 11 respectively. Figure 10 shows that tripling the platen speed (30 to 90 rpm) causes the slurry MRT to decrease by a factor of 3.5. The transport efficiencies are $\eta_{\text{wafer}}=88\%$, $\eta_{\text{system}}=84\%$ and $\eta_{\text{wafer}}=62\%$, $\eta_{\text{system}}=54\%$ for 90 rpm and 30 rpm respectively. Higher platen speeds decrease the slurry MRT and increase the efficiencies because the pad stores less old slurry on the pad and increases the slurry flow in the pad’s radial direction. As expected the platen speed has a large effect on the slurry transport from the center of the wafer to the wafer as seen in difference between the wafer and system efficiencies in the 30 rpm and 90 rpm cases.

Figure 11 shows that increasing the flow rate from 20 cc/min to 50 cc/min causes the slurry MRT to decrease by 80%. Increasing the slurry flow rate has a similar effect as platen speed in increasing the slurry flow in the radial direction. However, unlike platen speed, increasing the slurry flow rate can increase the storage of slurry on the pad that may dilute the new slurry. The transport efficiencies are $\eta_{\text{wafer}}=87\%$, $\eta_{\text{system}}=57\%$ for 50 cc/min and 20 cc/min respectively.
\(\eta_{\text{system}} = 81\%\) and \(\eta_{\text{wafer}} = 77\%\), \(\eta_{\text{system}} = 74\%\) for 30 cc/min and 20 cc/min respectively. It is evident that increasing the flow rate increases the slurry transport efficiency by increasing the rate of slurry flow in the radial direction.

[Figure 10 about here.]

[Figure 11 about here.]

In addition to the mean mixing profiles, one can examine gradients in the slurry composition across the wafer. Pad grooving and in situ conditioning have a large effect on the slurry gradients in the pad’s radial direction, that is, in the direction along a line drawn from the center of the pad to the edge of the pad passing through the center of the wafer. Figure 12 shows the average percentage of new slurry across the wafer shortly (20 seconds) after a pulse of new slurry. Each point in the graph is an average of the image column or an average of the percentage of new slurry in the direction perpendicular to the pad’s radial line. Note that the ex situ conditioning line has a steep gradient towards the edge of the wafer closest to the pad’s center. This gradient may be caused by the location of new slurry impinging on the wafer. The edge of the wafer closest to the center of the pad prevents the slurry pool from spreading out and is in direct contact with new slurry causing a gradient. The gradient may also be caused by the manner in which the pad conforms to the wafer. The wafer has a 10-15 micron bow across its diameter. In situ condition distributes and mixes the slurry, eliminating the gradient.

[Figure 12 about here.]

Grooved pads have a very different slurry gradient pattern for ex situ conditioning, as shown in Figure 13. The grooved pads support a linear gradient in the concentration of new slurry across the wafer. This is because the grooved pads channel the slurry in the radial direction and prevent slurry from mixing with its nearest neighbor. Like the flat pad, conditioning reduces slurry gradients by mixing and distributing the slurry across the pad.

[Figure 13 about here.]
7 Modeling

We performed 46 different mixing measurements in a statistically designed experiment to investigate the effects of the main CMP parameters including platen speed, flow rate, wafer down force, pad manufacturer, and pad topography. We found that all of the mixing curves could be described by the following simple function:

\[
F = \frac{kt^n}{1 + kt^n}
\]

where \(k\) and \(n\) are constants and \(t\) is time. This function has been proposed by Racz and Abedian (1998)[29] for polymer cure kinetics. Figures 14 and 15 shows that this function fits the mixing curves extremely well.

The correlation coefficients for the best fit lines are greater than 0.99 for the ex situ conditioning curves and greater than 0.97 for the in situ conditioning curves. The correlation coefficient is lower for the in situ conditioning curves because Equation 3 cannot fit the oscillatory behavior of the mixing curves. The high correlation coefficients imply a physical significance to Equation 3 explained below.

For the ideal case of unrestricted mixing, the mixing rate will be

\[
\phi_i = kt,
\]

where \(k\) is a constant, \(t\) is time, and \(\phi_i\) is the ideal volumetric fraction containing fluid "A" (new fluid). The ideal volumetric fraction, \(\phi_i\), can be expressed as
where $V_a$ is the volume of fluid "A" and $V_t$ is the total volume. This ideal volumetric fraction $\phi_i$ will be a function of a number of parameters:

$$\phi_i = \frac{V_a}{V_t}.$$ (7)

Each of these independent variables may also be a function of time. We assume a simple model:

$$\phi_i = kt^n.$$ (9)

The amount of mixing beneath the wafer will be the amount of unrestricted mixing multiplied by some weighting factors. As more new slurry covers the pad and becomes entrained beneath a wafer, the mixing rate will decrease. This is because the new slurry (A) will have a higher probability than the old slurry (B) of replacing new slurry in a fluid element as new slurry becomes more abundant. Therefore the weighting factors should be a function of the amount of the amount of fluid "B" left on the pad and beneath the wafer. These weighing factors will have two contributing parts: (1) mixing with slurry in the bow wave and on the pad ($\frac{V_b}{V_t}$) and (2) mixing under the wafer itself (also $\frac{V_b}{V_t}$). The rate of restricted mixing, $d\phi$, can be written

$$d\phi = d\phi_i \frac{V_b}{V_t}.\frac{V_b}{V_t}.$$ (10)

where $V_b$ is the volume of fluid "B" and $d\phi$ is the change in the unrestricted mixing. Since $V_b$ can be written
\[ V_k = V_i - V_a, \quad (11) \]

Equation 10 can be written as:

\[ d\phi = d\phi_i (1 - \phi)(1 - \phi). \quad (12) \]

Integrating equation 12 and applying the boundary condition that the volume fraction of new slurry at time zero equals zero gives the final form:

\[ \phi = \frac{k t^n}{1 + k t^n}, \quad (13) \]

where \( \phi \) is the unrestricted mixing function that describes the response of a chemical reactor to a step input (F-curve). Equation 12 can be rewritten as a linear equation in order to perform a least squares fit to the data. The linear form of equation 13 is given by

\[ \ln \left( \frac{\phi}{1 - \phi} \right) = \ln k + n \ln t. \quad (14) \]

8 Discussion

CMP polishing rates have historically been described by the Preston equation, written as
\[
\frac{dh}{dt} = k \frac{ds N}{dt A}.
\]

where \( h \) is the thickness of the substrate being polished, \( s \) is the total sliding distance traveled by the substrate, \( N \) is the normally applied load on area \( A \), and \( t \) is time. Equation 15 shows that the removal rate increases linearly with the platen speed. This coincides with slurry transport rates beneath the wafer which also increases almost linearly with platen speed. Slurry transport would be expected to contribute to the increase in removal rates with platen speed if Cook’s model [4] of polishing is correct. Equation 15 also shows that the removal rate increases linearly with down force. Slurry transport was not significantly influenced by slurry transport, which implies that down force is a purely mechanical effect provided the wafer is operating within one lubrication regime. Finally, the Preston equation considers chemical and other effects as one lump constant \( k \). Slurry transport may influence the polishing chemistry and therefore the magnitude of this constant.

9 Conclusion

In this paper, we examined factors influencing slurry transport. Slurry transport beneath a wafer while polishing was quantified by four different measures: slurry mean residence time, slurry mixing, wafer slurry transport efficiency, \( \eta_{\text{wafer}} \), and system slurry transport efficiency, \( \eta_{\text{system}} \). The slurry transport efficiencies varied from approximately 45% to 90%. Both measures of slurry transport efficiency increased with increasing platen speed and flow rate. These efficiencies are useful for general physical insight into the slurry transport and may be replaced after further research findings. Mean slurry transport can be modeled with a simple continuous function for all of the conditions investigated. Significant differences were seen between slurry gradients that developed across a wafer surface on a flat pad as compared with a grooved pad, implying that pad topography plays a strong role in maintaining uniformity across the wafer. In situ conditioning was shown to have a large effect on the slurry transport. In general, the amount of slurry mixing increased and
the rate of slurry transport decreased with in situ conditioning. Changes in $\eta_{\text{system}}$ with in situ conditioning showed that in situ conditioning aided slurry transport to the wafer with flat polishing pads. Finally, in situ conditioning reduced gradients in the slurry age across the wafer. Future investigations will attempt to correlate the polishing performance to the slurry behavior observed. Also the results shown here will be reproduced on an IPEC 372 tool to ensure that all of the scaling arguments apply.

10 Acknowledgements

The authors would like to thank Intel and Cabot corporations for funding this research. We would also like to thank VEECO corporation for donation of a Dektak 200 Si diamond tip profilometer that enabled us to measure removal rates. We would also like to thank Freudenberg Nonwovens for donating FX-9 polishing pads.
References


List of Figures

1 Polishing results of BK-7 glass ........................................... 22
2 Experimental Apparatus .................................................. 23
3 In situ conditioner ......................................................... 24
4 Mixing history beneath wafer during polish .......................... 25
5 Mixing measurement interrogation region ............................... 26
6 Graphical representation of $\eta_{wafer}$ and $\eta_{system}$ ........... 27
7 Mixing history with in situ and ex situ conditioning ............... 28
8 Effects of in situ conditioning on flat pad mixing history .......... 29
9 In situ conditioning with a flooded flat pad ......................... 30
10 Effect of platen speed on mixing history ............................... 31
11 Effect of flow rate on mixing history .................................. 32
12 Effect of in situ conditioning on slurry gradients on a flat pad .... 33
13 Effect in situ conditioning on slurry gradients on a grooved pad .. 34
14 Model fit versus ex situ conditioning mixing curves .............. 35
15 Model fit versus in situ conditioning mixing curves .............. 36
Figure 1: Polishing results of BK-7 glass
Figure 2: Experimental Apparatus
Figure 3: In situ conditioner
Figure 4: Mixing history beneath wafer during polish
Figure 5: Mixing measurement interrogation region
Figure 6: Graphical representation of $\eta_{\text{wafer}}$ and $\eta_{\text{system}}$. 
Figure 7: Mixing history with in situ and ex situ conditioning
Figure 8: Effects of in situ conditioning on flat pad mixing history
Figure 9: In situ conditioning with a flooded flat pad
Figure 10: Effect of platen speed on mixing history
Figure 11: Effect of flow rate on mixing history
Figure 12: Effect of in situ conditioning on slurry gradients on a flat pad
Figure 13: Effect in situ conditioning on slurry gradients on a grooved pad.
Figure 14: Model fit versus ex situ conditioning mixing curves
Figure 15: Model fit versus in situ conditioning mixing curves
List of Tables

1  Comparison of Parameter Values ................................. 38
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Scaling Factor</th>
<th>IPEC472</th>
<th>Struers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down Force</td>
<td>1</td>
<td>4 psi (13.790 kPa)</td>
<td>4 psi (13.790 kPa)</td>
</tr>
<tr>
<td>Platen Speed</td>
<td>Reynolds Number</td>
<td>30 rpm</td>
<td>(linear velocity 0.5 m/s)</td>
</tr>
<tr>
<td>Platen Diameter</td>
<td>( \frac{D_{platen}}{D_{wafer}} )</td>
<td>22 inches (56 cm)</td>
<td>12 inches (30.5 cm)</td>
</tr>
<tr>
<td>Wafer Diameter</td>
<td>( \frac{D_{platen}}{D_{wafer}} )</td>
<td>6 inches (150 mm)</td>
<td>3 inches (76 mm)</td>
</tr>
<tr>
<td>Flow Rate</td>
<td>Platen Surface Area</td>
<td>100 ml/min</td>
<td>25 ml/min</td>
</tr>
</tbody>
</table>

Table 1: Comparison of Parameter Values