Chemical mechanical polishing (CMP) is currently being used in the fabrication of integrated circuits and has been identified as an enabling technology for the semiconductor industry in its drive toward gigabit chips and subquarter micron feature sizes in the near future. At the present time, it appears that the global planarization necessary for establishing reliable multilevel interconnects can only be achieved by using CMP. As with any process with such potential commercial impact, the technology has moved forward into production even though a complete model based on first principles is nonexistent. In the long run, the availability of such models will help optimize the operation of CMP and permit the user to define the best operating conditions for each specific application. Our objective in this article is to present a model which deals with some crucial transport phenomena issues in CMP.

In a typical CMP tool, the wafer is mounted on a suitable device and is held above a polishing pad made of a porous polymeric material. In the case of smaller wafers, it is common to mount more than one wafer on a holder called a polishing pad. The pad and the wafer holder can be rotated at specified rates while pressing the wafer against the pad surface by applying a load which can be varied. A liquid is distributed over the pad and occupies the space between the wafer and the pad when the equipment is operated. The liquid contains a colloidal suspension of abrasive particles such as alumina or silica as well as specific chemicals chosen for polishing. In another type of tool, the pad moves linearly on a belt while the wafer is pressed against it, but the idea is the same, namely, that relative motion occurs between the wafer and the pad in the presence of a liquid containing abrasive particles.

The state of the literature on the modeling of CMP in the semiconductor field is reviewed by Nanz and Camilletti. A review of CMP in glass polishing is provided by Cook, who discusses the mechanism by which material is removed. Cook suggests that polishing occurs due to softening of the glass surface by the chemicals with subsequent removal of this softened layer by the particles through abrasion. The abrasive process is described using a model of a sphere imbedded to a certain depth into a surface by the application of a load, with the particle gouging out material from the surface in proportion to this depth. One key aspect of this picture is that the rate of removal is proportional to the product of the relative velocity between the two surfaces and the applied pressure, leading to an equation which is traced to Preston and which is currently used in the industry for describing the rate of removal of silicon dioxide. Also, the amount removed is independent of the size of the abrasive particles in this picture. The Preston equation is used to describe overall removal rates, but cannot account for higher removal rates from elevated surfaces, which is the basis of planarization. Also, it is seen from the present work that when transport phenomena coupled with chemical reactions are accommodated, the linear dependence on the relative velocity implied by the Preston equation is not realized. That linear dependence arises from models of abrasive removal and does not directly pertain to transport processes which are the focus of the present article.

There are very few articles devoted to the fundamental aspects of CMP in the semiconductor wafer polishing process. The first reported modeling effort was made by Runnels and Eyman, who constructed a fluid mechanical model for the flow of the liquid film between the wafer and the pad. In this model, both the wafer and the pad are treated as planar while permitting a slight tilt of the wafer with respect to the surface of the pad. As is well known from the work of Reynolds, who formulated the principles of lubrication theory, such a tilt with convergence of the two surfaces in the direction of motion produces a positive pressure in the region between the two surfaces supporting the applied load. Runnels and Eyman numerically solved the steady Navier-Stokes equation applicable to the liquid film along with no-slip boundary conditions at the two surfaces, assuming a stress-free boundary where the liquid between the wafer and the pad encounters liquid exposed to the atmosphere on the pad surface. The results obtained by these authors were in the form of liquid film thickness and tilt angle for a given load and rotation speed. Predicted thicknesses of the liquid film ranged from approximately 45 to 95 μm. Runnels explored flow over a two-dimensional step of submicrometer height and length of the order of tens of micrometers on the wafer surface using a numerical solution of the steady Navier-Stokes equation. He used an empirical model for relating the calculated stress distribution on the surface of the step to the removal rate and made predictions regarding the time-dependent evolution of the shape of the step. The liquid film was assumed to be of thickness in the range 10-50 μm. Bramono and Racz presented a two-dimensional model of flow in the liquid layer on a pad and in the air above, assuming the pad to be smooth and planar. The governing equations were solved using the commercial package Fluent assuming a liquid film thickness of 40 μm. Results are presented in the form of sample velocity vector plots and the pressure distribution underneath the wafer in one case. Very recently, a two-dimensional lubrication liquid film model was used by Sundararajan et al. to analyze the mass transport in the liquid film coupled with a chemical reaction at the wafer surface. These authors also accommodated abrasive removal in their model through an empirical approximation. They assume that the relative enhancement in removal rate due to abrasion is linearly proportional to the product of the concentration of abrasive particles and the shear stress at the wafer surface.

In an effort to determine the thickness of the liquid film, Levert et al. used an apparatus specifically designed to measure the displacement of the wafer in the vertical direction using the capacitance of a suitable air gap for this purpose. Wafer displacements were measured relative to that at no rotation. For making these measurements, the wafer was held fixed while the pad was rotated. When using very low normal loads between 2.7 and 5.8 kPa, Levert et al. indeed found upward displacements of the wafer up to about 50 μm. The thickness of the liquid film increased with increasing relative speed between the wafer and the pad and with increasing viscosity of the liquid in a manner consistent with predictions from lubrication theory. However, under these conditions the authors also found that
material removal rates from a silicon dioxide film were negligible. It would be difficult to imagine a mechanism for abrasion of the wafer surface in such a thick liquid film since the abrasive particles are typically 0.3 μm in size in aggregated form. Significant rates of material removal require that the particles be caught between the wafer and pad mechanically. Interestingly, when Levert et al. used a perforated conditioned pad and normal loads between 34.5 and 62.1 kPa which are more typical of CMP conditions, they found that the wafer moved downward, with the displacement increasing in magnitude with relative speed. The authors also used a model wafer fitted with pressure taps and measured pressure gradients of the order of 1.4 kPa/mm beneath the wafer which exhibited complex spatial dependence. Pressures measured in the liquid film were less than atmospheric pressure with the difference increasing with relative speed, consistent with the displacement measurements. Interestingly, when a glazed pad whose pores were filled with debris was used at a normal load of 48.1 kPa, the wafer displacement was slightly positive, ranging from approximately 1 to 6 μm. Very recently, Tichy et al. have explained the subatmospheric pressures measured by Levert et al. using a lubrication liquid-film model in which the thickness of the film is determined by the deformation of the asperities and the pad under load. Their model, based on contact mechanics theory, predicts that the liquid film will typically be thin near the edges of the region of contact between the wafer and the pad, and thicker in the central portion. In a natural manner, this leads to pressure distributions which are of the correct form observed by Levert et al.

Another technique for measuring liquid film thicknesses was recently described by Coppetta et al. In this method, the authors introduced two dyes, each fluorescing at a different wavelength, in order to obtain information on the thickness of the liquid film between a transparent wafer and the pad and information on the mixing of the liquid. The authors found that the thickness of the liquid on a conditioned pad was significantly larger than that on an unconditioned pad. No actual thicknesses are reported and the authors indicate that calibration work was in progress.

The observations of Levert et al. regarding the thickness of the liquid film suggest that in the presence of an open pore structure in the pad, the liquid film between the top of the pad and the wafer can indeed be very thin with the thickness limited only by the size of the abrasive particles caught between the top of the pad and the wafer. This is consistent with results reported by Bhushan et al. who performed experiments on the polishing of silicon oxide surfaces. Noting that in a typical application, the wafer holder presents a cylindrical surface to the pad and that the wafer projects out of this surface, the authors varied the distance by which the wafer projects by using inserts. They determined that when the wafer is slightly recessed within the holder surface, the polish rate is virtually zero. This implies that a lubrication liquid film of thickness much larger than that of the colloidal particle dimensions was present and that there was virtually no removal of material. When the authors used a wafer of nominal thickness 279 μm in a pocket of nominal depth 229 μm which permitted the wafer to project approximately 50 μm out of the surface, significant rates of removal of material were observed. The rate seemed insensitive to the relative velocity and the applied load within some range in each case. When the wafer was recessed by which the wafer projected was increased further to approximately 100-175 μm, the removal rates were larger but also showed substantial variation with the load and relative velocity. This is the configuration in which industrial CMP is normally performed. The experimental studies of Levert et al. and Bhushan et al. are crucial in formulating a good physical model since they clearly identify the need to accommodate the prospect of physical contact between the wafer surface and the abrasive particles under a load. In the following section we use these ideas in formulating a model that can be used to analyze transport phenomena aspects in CMP.

**Physical Description of the Model**

*Pad surfaces and the liquid film thickness.—*Scanning electron micrographs of certain types of soft pads show a collection of fibers with pores which are of the same length scale as the thickness of the fibers. The fibers are of different lengths so that the tops of the fibers present an uneven surface morphology. With this type of pad, when a load is applied to the wafer-pad interface, the fibers can bend and the porous surface presented to the wafer will be complex in geometry. Some harder pads appear under the microscope as a nearly planar surface with pores. The pores are of a variety of dimensions and the spacing between the pores also varies from place to place. Slurry-free pads such as those manufactured by the 3-M Company show a regular geometric structure with pyramidal mounds projecting out of a surface. Abrasive particles are imbedded in the pad. Finally, we note that pads are sometimes grooved or perforated with the objective of supplying slurry to the region where it is needed. In light of these observations of the surface regions in different types of pads, we proceed to provide a physical picture of the liquid film between a wafer and a pad which forms the basis of our modeling effort.

First consider an idealized soft pad which has a planar surface everywhere except where fibers project upward. A liquid film of thickness comparable to the height of a fiber is present between the wafer holder and the pad surface wherever a fiber is not present. However, in the region where the wafer is pressed into the pad by a load, we suggest that the wafer surface pushes down on the fibers of the pad, perhaps bending them if they are sufficiently flexible. The pad underneath the wafer then appears as a porous medium with a complex collection of pockets containing liquid. Abrasive particles are caught between the wafer surface and the top of the surface formed by the bent-over fibers and constitute the principal reason why the wafer does not physically contact the tops of the fibers everywhere. They are likely imbedded deeply into the fibers in the pad where they are held between the wafer and the top surface of the bent-over fibers under load. The liquid trapped in the pores of the pad provides a supply of chemicals which is not significantly replenished from the instant when a pore enters under the wafer until the pore leaves from underneath the wafer and encounters a supply of fresh liquid. Liquid held in the extremely narrow region between the wafer and the top surface of the fibers, of thickness of the order of the abrasive particle size, provides a mechanism for some exchange of chemicals from one pore to the next. However, the contribution from this is very small because the pores are typically 50 to 100 times deeper than the thickness of this liquid film. This picture would need to be modified appropriately when mechanisms exist to provide fresh slurry to the pores while the pores are underneath the wafer. The abrasive particles held between the wafer and the tops of the fibers mechanically remove material from the wafer surface. Chemical action also can remove material from the wafer surface. In addition, chemicals in the liquid can contribute to other processes, such as the formation of passivating layers, and modify the surface removal rates caused by particle abrasion. In the region where there is liquid held in a pore between the wafer surface and the pad, there is no removal of material by abrasive action. As long as the concentration of particles is small, chemical action occurs over the entire wafer surface, since screening of the wafer surface by the particles is minimal.

In the case of hard pads containing pores, the same basic scenario can be used with the idea that liquid is contained in the pores, and the abrasive particles are caught between the pad and the wafer in the regions on the surface of the pad where there are no pores. A liquid film whose thickness is consistent with this picture is present between the wafer and the pad where there are no pores. With the slurry-free pads described in Ref. 13, the abrasive is imbedded in the pad and the portion of it which projects from the tops of the mounds would be in contact with the wafer surface. Liquid would be present in the valleys surrounding the mounds.

Using the physical picture described above, we now proceed to develop a two-dimensional model of transport phenomena in CMP. In doing so, we assume the volume fraction occupied by the particles to be relatively small, as is common in CMP operations. When concentrated dispersions are used, consideration must be given to restricted access by the chemicals to the wafer surface as well as to slurry rheology. Also, we confine the model to the polishing of blan ket unpatterned wafers.
The repeating cell model.—Even though the industrial CMP process provides the motivation for this research, one must begin with a simpler system for the purpose of developing an understanding from first principles. A schematic of a CMP process is shown in Fig. 1. The key element of a polishing pad is the surface layer that is in contact with the wafer and the liquid. In a first two-dimensional description, this can be approximated as a series of pockets (pores) separated by planar regions modeled as shown in the exploded view of a small region underneath the wafer in the figure. This pattern can be regarded as a series of “cells.” We have assumed at this first level of modeling that the essential transport aspects can be captured by this repeating cell model depicted in the figure. By performing a detailed analysis of transport in a single cell, and by integrating over all the cells underneath the wafer surface, it is possible to make macroscopic predictions of the removal rate. Figure 2 is used to illustrate this concept. It shows a plan view of the region underneath the wafer at a given instant. The area enclosed by the circle consists of many unit cells, one of which is displayed in the inset. In some CMP tools, the pad is on a rotating table and revolves at a rate in the approximate interval 10-100 rpm. The wafer is on a holder which is offset from the axis of rotation of the pad and can rotate about the axis of the holder at a specified rate. It can be shown that when both the wafer and the pad are rotated at the same angular velocity, the relative velocity between any point on the wafer surface and a corresponding point below it on the pad surface is the same regardless of the location of the point in question. In Fig. 2, the $x^*$ direction has been chosen to point in the direction of this constant relative velocity $V = \frac{\Delta V}{\Delta t}$, where $i$ is a unit vector in the $x^*$ direction. In the present article we do not consider orbital polishing tools in which the relative velocity between the wafer and the pad is different at each point on the wafer surface.

The steady removal rate from a blanket wafer can be calculated by integrating the removal rate in a strip of width $\Delta z^*$, shown in Fig. 2, over the permissible range of values of $z^*$. This strip itself consists of many unit cells, each of which has been exposed for a different amount of time under the wafer. Therefore, we can approximate the summation over a discrete number of unit cells in the strip to an integration over exposure time. When a blanket wafer is polished under constant load, using a steady supply of chemicals and a polishing pad whose surface characteristics are assumed to remain unchanged, from the perspective of the wafer, the removal rate is steady. We found it convenient to work from the perspective of a unit cell in the pad over which the wafer surface moves at a constant velocity $V$. From the point of view of the unit cell, the transport problem is unsteady. A typical unit cell enters under the wafer at point $Q$ and spends a certain amount of exposure time underneath the wafer, exiting at the other end of the strip marked in Fig. 2. The unit cell shown in the sketch occupies the location identified $Q'$ in the figure and has been exposed to the wafer for a period of time $t = Q'O' N$, where $QQ'$ is the distance between $Q$ and $Q'$. The concentration of chemical species is presumed to be uniform within the unit cell as it enters under the wafer at $Q$. As the chemical species react at the surface, and possibly in the bulk liquid, and diffuse, concentration gradients evolve with exposure time, $t$. Therefore, if one follows a given unit cell, the concentrations of various species in the unit cell change with time spent underneath the wafer. The concentration fields can be modeled using an appropriate version of the unsteady conservation equation for each species, along with suitable boundary conditions. For this purpose, it is necessary to have a detailed knowledge of the velocity field within the unit cell. This velocity field also changes with exposure time underneath the wafer. A simple order-of-magnitude estimate reveals that the velocity field within the unit cell becomes independent of exposure time by the time the cell has moved a distance which is a small number times its own length. This number is a fraction of the Reynolds number which defined in the next section. The Reynolds number can be expected to lie in the range 0.1-50 in a typical CMP operation. The fraction multiplying the Reynolds number in this time estimate depends upon the geometrical aspect ratios used in defining the unit cell and is typically of the order $1/2$. Considering the strip of width $\Delta z^*$ shown in Fig. 2, this means that the velocity field changes from one unit cell to the next in perhaps as many as the first 25 unit cells starting from point $Q$, after which it is identical in the remaining unit cells in the strip. The number of unit cells in the strip ranges from a few hundreds to perhaps as many as a thousand, with the exception of relatively small regions close to the edge of the wafer in the $z^*$ direction, where the strips are short. Therefore, it is reasonable to neglect the error caused by assuming that the velocity field is the same in all the unit cells beginning from point $Q$. In a reference frame attached to the unit cell, this is the steady velocity field which prevails after the wafer surface has been moving over it for a sufficient amount of time.

Using this picture of the liquid film in the region between the pad and the wafer, we now proceed to pose and solve the relevant transport problems for predicting the steady removal rate of material from a blanket wafer due to chemical action.

**Theoretical Development**

The important features of a unit cell, displayed in Fig. 3, are the presence of liquid in a cavity of length $L_1$ and height $H$, and in a slit of length $L_2$ and thickness $d$. The cavity in the unit cell is placed at the center with half a slit on either side of it. The magnitudes of $L_1$ and $L_2$ would depend on the pad structure. Here they are assumed to be comparable to each other and of the order of the thickness of a single fiber in a soft pad or the average width of a pore on a hard pad. This typically lies in the range 10-50 μm, even though other values are possible. The value of $d$ can be expected to be very small compared to $H$ and of the order of the size of the abrasive particles. Typically, this is of the order 0.5 μm or less, even though aggregation of particles can
make it somewhat larger in some cases. We define one aspect ratio for the cavity as \( \alpha = L_1/H \) and two for the slit as \( \lambda = L_2/H \) and \( D = d/H \). It is in the slit that the abrasive particles are pressed against the wafer surface by the pad under load. Abrasion should be of negligible consequence in the cavity portion of the cell. Chemical reactions can occur over the entire wafer surface both in the cavity and in the slit. Also, as mentioned earlier, it is possible for chemical reactions to take place in the bulk of the liquid in both regions of the unit cell. Note that we assume the same values of the width and depth of the cavity and the slit to hold in every cell. Also, we assume the depth \( d \) of the slit to be uniform within a unit cell. In practice, depending upon the manner in which abrasive particles are caught between the wafer and the pad and the way in which the pad is deformed, these parameters vary from place to place. The model should be regarded as presenting an idealized simple picture in this regard.

While the pores in the pad are interconnected, the main effect of such connection is to support a very slow flow of fluid from one pore to another due to rotation of the pad, and the influence of this flow can be regarded as being negligible at this level of modeling. The principal mechanism that drives flow of liquid in the cavity and the slit is the relative movement of the wafer surface with respect to the pad. It is possible to account for the role of global pressure gradients that develop in the liquid film as experimentally demonstrated by Levert \textit{et al.}, but we neglect the effects of such gradients in this first-level model. A larger natural pressure gradient arises within each cell due to the flow in this geometry, which we illustrate using sample pressure profiles in the next section.

The model is two-dimensional. We assume no variation in the \( z \) direction normal to the plane of the paper. The origin of coordinates is chosen at the left bottom corner of the cavity as shown in Fig. 3. In a reference frame attached to the pad, \textit{i.e.}, from the perspective of the cell, the flow in the cell is steady after a brief initial transient, and it is this steady flow that we determine and use in the mass-transport problem. The flow of the liquid is assumed incompressible and Newtonian. While a suspension of particles can behave as a non-Newtonian fluid, little has been published about the rheology of the slurries used commercially. Also, as noted earlier, the volume fraction occupied by particles in CMP slurries is typically not large. Therefore, we made the assumption of Newtonian flow here, with the understanding that in the long run, it would be useful to incorporate the correct slurry rheology into the model. Physical properties are assumed to be constant, which implies isothermal conditions. In practice, due to solid-solid fraction and viscous dissipation, the temperature underneath the wafer varies with position. In this first model, for simplicity we ignore the heat-transport aspects and assume constancy of temperature everywhere underneath the wafer.

The velocity field satisfies the continuity and Navier-Stokes equations\(^6\) which are written below in dimensionless form

\[
\frac{\partial V_x}{\partial X} + \frac{\partial V_y}{\partial Y} = 0 \tag{1}
\]

\[
Re\left(\frac{\partial V_x}{\partial X} + V_y \frac{\partial V_x}{\partial Y}\right) = -\frac{\partial P}{\partial X} + \frac{\partial^2 V_x}{\partial X^2} + \frac{\partial^2 V_x}{\partial Y^2} \tag{2}
\]

\[
Re\left(\frac{\partial V_x}{\partial X} + V_x \frac{\partial V_y}{\partial Y}\right) = -\frac{\partial P}{\partial Y} + \frac{\partial^2 V_y}{\partial X^2} + \frac{\partial^2 V_y}{\partial Y^2} \tag{3}
\]

The symbols \( X = x/H \) and \( Y = y/H \) represent scaled coordinates. Similarly, \( V_x = v_x/V \) and \( V_y = v_y/V \) represent scaled velocity components in the two coordinate directions. The scaled hydrodynamic pressure \( P = \rho H(\mu V) \) where \( \mu \) is the dynamic viscosity of the liquid. The Reynolds number \( Re \) is defined as

\[
Re = \frac{vH}{\nu} \tag{4}
\]

where \( \nu \) is the kinematic viscosity of the liquid. Since the relative velocity in CMP typically varies from 0.1 to 1 m/s, and the kinematic viscosity of the liquid varies from approximately 1-10 \( \times 10^{-6} \) m\(^2\)/s, this Reynolds number, based on the cavity height, varies roughly between 0.1 and 50. The boundary conditions are those of no slip at the top and bottom surfaces of the cavity and the slit in Fig. 3. Also, there is no slip at the sidewalls of the cavity. There is no normal flow at any of the rigid bounding surfaces. By using this boundary condition, we are neglecting the small velocity in the \( y \) direction at the wafer surface arising from mass transfer, which is a good assumption in dilute solutions. Finally, we assume equality of pressure and velocity at each value of \( y \) at the two ends of the cell, located at \( X = -\lambda/2 \) and at \( X = \alpha + \lambda/2 \). These are known as periodic conditions in the cell model. They are reasonable approximations when there are many cells, since there is no reason for one cell to be different from the next from a fluid mechanical perspective.

Equations 1 to 3, along with the associated boundary conditions, were solved by the method of finite differences using central difference approximations of the derivatives and a staggered grid for the pressure field.\(^17,18\) A variable grid was used so that flow in the narrow slit can be adequately resolved while accurately describing the flow in the cavity. The solution was tested for convergence with respect to grid size and validated against the limiting cases of flow in a lid-driven cavity and Couette flow in a slit. The computational grid is shown in Fig. 4. In the final calculations, we used 21 grids in the \( X \) direction from \( X = 0 \) to 1 and 21 grids from \( X = -\lambda/2 \) to 0, and similarly, 21 grids in the slit on the right. The region from \( Y = 0 \) to 1 in the cavity was divided into 21 grids. The region from \( Y = 1 \) to \((1 + D)\) in the cavity and in the slit was divided into 11 grids. As seen from the figure, the grids were not equally spaced but were designed to obtain the concentration distribution as accurately as possible.
possible. This means that where sharp concentration gradients were encountered, more mesh points were used.

The next aspect of the modeling effort is the handling of the rate of removal of material from the wafer surface. Removal occurs by two mechanisms. One is abrasion of substrate directly from the wafer surface. The state of modeling abrasive removal in CMP is unsatisfactory at the current time. A Hertzian penetration and microcutting picture of the removal of metals was presented by Brown et al. Brown and Cook suggested that this is a reasonable approximation for modeling glass polishing, even though there are substantial discrepancies between actual removal rates and those predicted by the model. Liu et al. obtained a result which accommodates different elastic moduli for the wafer surface and for the particles but which uses a Hertzian indentation picture. They suggested that this model can be used to describe abrasive removal in oxide polishing. The validity of using a Hertzian indentation, which is an elastic response in calculating the depth of a gouge made by an abrasive which leads to microcutting, is unclear. Also, it is quite possible that abrasive removal occurs by brittle fracture in some situations. We can accommodate any reasonable description of removal rates due to mechanical abrasion within the framework of the overall transport model. In particular, we can allow for coupling between chemical softening of the surface and abrasive removal if appropriate. However, in the absence of an accepted abrasion model, we restrict ourselves to a consideration of chemical removal of material in the present article.

We assume the presence of a chemical substance in the liquid that reacts with the wafer surface and forms a product species. The reactant species is labeled A. The slurry flow over the pad provides the pockets in the pad with a fresh supply of the reactant A so that a unit cell begins its journey under the wafer with an initial concentration, $C_0$, of A. As the cell spends time under the wafer, the species A is depleted by reaction at the wafer surface, and product species B is formed. Thus, the mass-transport problems for the reactant and for the product B are unsteady from the perspective of the unit cell. Assuming a dilute system in both the reactant and the product and constant liquid density and diffusivities, we can write the following dimensionless governing equations, from Ref. 16, for the scaled concentrations of A and B in the unit cell

$$\frac{\partial C_A}{\partial t} + \frac{\partial}{\partial x} \left( V_x \frac{\partial C_A}{\partial x} + V_Y \frac{\partial C_A}{\partial Y} \right) = \frac{\partial^2 C_A}{\partial x^2} + \frac{\partial^2 C_A}{\partial Y^2} \quad [5]$$

$$\frac{\partial C_B}{\partial t} + \frac{\partial}{\partial x} \left( V_x \frac{\partial C_B}{\partial x} + V_Y \frac{\partial C_B}{\partial Y} \right) = \Omega \left( \frac{\partial^2 C_B}{\partial x^2} + \frac{\partial^2 C_B}{\partial Y^2} \right) \quad [6]$$

In these equations, $Pe = VH/\delta A$ is a Péclet number. The symbol $\delta A$ stands for the diffusivity of the reactant species in the liquid, and the scaled concentrations of A and B are defined as $C_A = c_A/C_0$ and $C_B = c_B/C_0$, where $C_0$ is the initial molar concentration of A. The scaled time $T = \delta A \eta H^2$ where $\eta$ is exposure time. The dimensionless parameter $\Omega = \delta A^2 \Omega A$.

The bottom boundaries in the cavity and slit, including the sidewalls, are impermeable to the reactant or product. Therefore, the diffusive flux at these boundaries can be set equal to zero. This is equivalent to setting the component of the concentration gradient normal to each surface equal to zero. In addition, we must write conditions on the concentration fields at the surfaces in the fluid which correspond to the two ends of the slit and at the top surface over the cavity and the slit. For the former, we continue to use periodic boundary conditions even though there should be a slight change in concentration between the inlet to the cell and the exit to the cell. Given the length scale of the cell and the large number of cells underneath the wafer surface, this is a reasonable first approximation which needs to be refined in an improved future model. At the top surface corresponding to the surface of the wafer, we assume a chemical reaction consumes A and produces B. The reaction rate is assumed to depend upon the concentration of A. In the present model, for illustration purposes we assume irreversible first-order kinetics at the surface of the wafer in the concentration of A, but other kinetic forms and more complex reactions involving passivation can be accommodated by the model. The boundary conditions at the entire wafer surface then become

$$-\frac{\partial C_A}{\partial Y}(T, X, 1 + D) = KC_A(T, X, 1 + D) \quad [7]$$

$$\Omega \frac{\partial C_B}{\partial Y}(T, X, 1 + D) = KC_A(T, X, 1 + D) \quad [8]$$

Here, the dimensionless reaction rate parameter $K = kH/\delta A$. It represents the relative importance of the surface reaction rate to the diffusive transport rate in the cavity and is called the Damkohler number.

Equations 5 and 6, along with the associated boundary conditions, were solved by the alternating direction implicit method. The same grid was employed as that used in solving the Navier-Stokes equation so that the velocity field calculated previously could be used. The time step had to be chosen with care so as to produce accurate results using a reasonable amount of computing time. The calculations were made with a scaled time step $\Delta T = 10^{-5}$. A given run was terminated when the cell was completely depleted of chemical A. Alternatively, when the maximum time a cell could spend underneath the wafer for a given set of conditions was reached, the calculations were stopped. The resulting solution was used to calculate the average flux of species A to the wafer surface in the cell by integrating from $X = -\lambda/2$ to $(\alpha + \lambda/2)$. This is the main output from the mass transport/reaction model. Also, detailed information on the concentration distributions of A and B in the cavity and the slit were obtained as needed.

The final step is the integration of the results from the cell model to obtain the steady removal rate of material from the entire wafer surface, which is an experimentally measurable quantity. We assume the wafer to be a circular disk of radius $R$. The reader will recall that the mass-transport problem is unsteady from the perspective of any given unit cell. The path for obtaining the steady removal rate from the wafer was discussed earlier in the Theoretical development section, by making reference to Fig. 2, which shows the location of an arbitrary unit cell underneath the wafer surface. This cell entered underneath the wafer at point $Q$ at a certain instant which corresponds to zero exposure time and is located at $Q^*$ after the passage of an exposure time $t$. Since unit cells with a fresh supply of the chemical continue to arrive at the point $Q$, the concentrations of A and B in the liquid within any unit cell that arrives at the point $Q^*$ underneath the wafer are steady. They correspond in the unsteady model of transport in the unit cell to the concentrations calculated at exposure time $t = QO^*/V$. Therefore, the steady concentration fields within the various unit cells positioned along the line $QO^*$ correspond to concentration fields calculated in the unsteady model for various exposure times lying in the interval 0-$t$. Since the model is two-dimensional, it assumes uniform conditions to prevail across any given unit cell in the direction designated $z^*$ in Fig. 2. A first integration is carried out with respect to $x^*$ from the leading edge to the trailing edge of the wafer using a strip of width $\Delta z^*$. This result is then integrated with respect to $z^*$ across the wafer to obtain the following final expression for the overall rate, $F_A$, at which species A is transported to the wafer surface.

$$F_A = -\frac{2VHc_0}{\alpha + \lambda} \int_0^R \int_0^{QO^*/V} \int_{\alpha + \lambda}^{\alpha + 2\lambda} \frac{\partial C_A}{\partial Y}(T, X, 1 + D) dX dY dT d\alpha^*$$

The removal rate of material from the wafer surface can be inferred from this rate using the stoichiometry of the chemical reaction. For presentation purposes, we have converted it to the usual form of nm/min when displaying results in the next section.
In the limiting situation when the reaction rate parameter $K$ is small, the process is kinetically limited by the reaction rate at the surface. One can then approximate the unit cell as a stirred tank reactor, since composition gradients are then negligibly small. In this case, an unsteady mass balance can be written for the contents of the cell and integrated immediately to yield an exponential dependence of the concentration of $A$ in the cell upon exposure time to the wafer. Since the concentration of $A$ is uniform within the unit cell, the rate of disappearance of $A$ at the wafer surface in the cell is the product of the reaction-rate constant, surface area, and the instantaneous concentration of $A$ in the cell. When this rate is integrated in the same fashion as outlined above, we obtain the following final result for $F_A$.

$$F_A = 2KHVC \int_0^R \frac{2D_A}{VH} \left( R^2 - z^2 \right) \exp \left( -\beta KT \right) dt dz$$

$$= \frac{2C_0 HV}{\beta} \left( R - \int_0^R \exp \left( -\frac{2\beta k}{VH} \left( R^2 - z^2 \right) \right) dt dz \right) \quad [10]$$

where

$$\beta = \frac{\alpha + \lambda}{\alpha + D(\alpha + \lambda)} \quad [11]$$

Since some diffusional resistance is always present, the result in Eq. 10 serves as a theoretical maximum removal rate by chemical reaction at the surface.

**Results and Discussion**

In this section, we present representative results from the model and discuss their implications. First, considering the fluid mechanical model, typical flow patterns in the cavity and the slit are illustrated in Fig. 5 for an aspect ratio $\alpha = 1$. The vectors represent scaled velocities at various points. The magnitude maximum occurs at the wafer surface and this can be used for comparison purposes in determining the strength of the flow at various locations. The flow field behaves as expected from intuition, with recirculating flow in the cavity and nearly Couette flow, modified by the presence of a pressure gradient, in the slit. In the insets in Fig. 5, we have magnified the region from $Y = 1$ to $(1 + D)$ so that the flow can be conveniently visualized. One inset shows the slit on the left side of the cavity from $X = -\lambda/2$ to 0 and the other shows the top of the cavity from $X = 0$ to 1. Note that the velocity gradient at the wafer surface is relatively large in the slit portion and small over the cavity. The two smaller insets at the top left and right show the entrainment flow near the leading edge of the cavity and the corresponding re-entry flow into the cavity at the trailing edge. We have investigated the flow patterns for an aspect ratio $\alpha = 2$ and also for other values of the Reynolds number up to 30 and observed the same qualitative behavior. It is possible to get secondary recirculating eddies in a sufficiently deep cavity, that is, when $\alpha < 1$. Shankar reports that in Stokes flow ($Re = 0$), a second recirculating eddy appears at the bottom of the cavity when $\alpha$ changes from 0.625 to 0.588, and additional eddies are formed as the aspect ratio is decreased further. The significance of this observation is as follows. The reactant in such secondary eddies is not used efficiently, because it first has to move into the primary recirculating eddy through diffusion. Of course, the estimates for the transitions are only approximate since the Reynolds number is not always negligible in CMP situations. However, we can use them to suggest that an optimal pad design, from the perspective of efficient use of the reactant, would tend to keep the aspect ratio $\alpha$ at a value above the branch where the transition to a two-eddy structure occurs.

The pressure would be uniform in flow between long parallel rectangular surfaces when one surface is moved. Here, as displayed in Fig. 6, the pressure distribution at the wafer surface $Y = 1 + D$ is not uniform. This gradient in pressure arises from the variation that must occur over the cavity portion of the unit cell. We have made calculations for a simple lid-driven cavity to confirm this to be the case. The flow in the slit portion of the unit cell provides a means of adjusting the pressure distribution so that the pressure variation can be periodic from one cell to the next. In physical units, the variations noted here range from 0.025 to 75 kPa depending upon the parameters. However, they occur over length scales of the order of 10-100 $\mu$m in the direction of relative motion. The resulting pressure gradient is of the order of 0.25-7500 kPa/mm. Values in the lower range are comparable to the global gradients across the wafer surface, which were measured by Lever et al. over lengths of the order of millimeters. However, the range here spans much larger values. The origin of the global variation reported in Ref. 9 can be different from the source of the pressure variation over the length scale of the cavity calculated here, as suggested by the analysis in Tichy et al. Since the present model may not apply to the physical situation in which the measurements reported in Ref. 9 were made; we refrain from speculation regarding this issue.

In Fig. 7, we present typical contours of constant concentration of $A$ obtained from the mass-transport model assuming $\Omega = 1$. For this purpose, we have used ferric nitrate as a model chemical which reacts with a copper surface. The overall oxidation reaction can be written as follows.

$$Cu + 2Fe^{3+} \rightarrow Cu^{2+} + 2Fe^{2+} \quad [12]$$

![Figure 5](image-url)

*Figure 5.* Sample velocity vector field plots for $Re = 1$, $\alpha = 1$, $\lambda = 1$, and $D = 1/50$. Shown in the inset bottom left is a magnified view of the flow in the slit on the left side of cavity in Fig. 3. The inset above the main figure shows the flow in the region $Y = 1$ to $(1 + D)$ above the cavity portion. Also included are two smaller insets which show the details of the flow near the leading and trailing edges of the cavity. The velocity vectors are magnified by a factor of ten in these two insets for clarity.
Figure 7. (a-e) Normalized concentration contour plots for the chemical removal of copper by first-order surface reaction using 0.1 M ferric nitrate. $Re = 1$, $\alpha = 1$, $\lambda = 1$, and $D = 1/50$. Shown in the inset is a magnified view of the contours in the region $Y = 1$ to $(1 + D)$. (a-e) Sequence of contours as scaled exposure time $T$ is increased from $10^{-4}$ to 2.
The dissolution rates of a copper coupon in a solution of ferric nitrate in a stirred reactor were measured at various temperatures by Ramarajan and Babu. These authors measured the rate as a function of the concentration of ferric nitrate and found that their results at each temperature could be fitted to a first-order process in the molar concentration of Fe$^{3+}$. The removal rate at 45°C was found to be approximately 1.8 times its value at 25°C. Therefore, the kinetic constant at the higher temperature would be larger by the same factor. This gives a rough estimate of the effect of temperature variation on chemical removal rates in this system. Here we consider Fe$^{3+}$ as a species A and use the kinetic constant measured at room temperature, which is $k = 3.2054 \times 10^{-5}$ m/s. Using this value of $k$, $H = 30 \mu$m, and a diffusivity $\bar{D}_A = \bar{D}_B = 1.812 \times 10^{-9}$ m$^2$/s, the dimensionless parameter $K = 0.5306$. This is the case to which the contours in Fig. 7 apply. In making the calculations, we have used the same values of the Reynolds number and the parameters $\alpha$, $\lambda$, and $D$ as in Fig. 5. Since the Schmidt number corresponding to a kinematic viscosity $v = 10^{-6}$ m$^2$/s is $Sc = 700$, the corresponding Péclet number for the computations also is 700. We have made similar calculations for $Re$ up to a value of 30, corresponding to $Pe$ up to 21,000. The computations at large values of the Péclet number are more difficult and subject to more numerical error. The qualitative features of the contours in this range of parameters are adequately represented by the sample case shown. We note that even in the case shown, the sharp slope discontinuities observed in several contours are due to numerical resolution limitations and are not physical. Examination of the evolution of the concentration contours shows that the entrainment flow at the leading edge of the cavity (on the left) brings out fresh reactant and the re-entry of fluid into the cavity near the trailing edge reduces concentrations in the right side of the cavity with the passage of time. Initially, the concentration in the cell is uniform and gradients gradually evolve with exposure time to the wafer surface as chemical reaction consumes the reactant A. It can be seen from Fig. 7a at $T = 10^{-4}$ s that convection has not yet had time to distort the concentration profile resulting from diffusive transport to the surface. The corresponding physical exposure time of the unit cell is 49.7 $\mu$s. As time progresses, we see that the concentration contours in the slit portion become virtually normal to the surface of the wafer, indicating nearly uniform conditions in the $Y$ direction within the slit. Since the average velocity is one-half the velocity at the wafer surface for Couette flow, one can estimate the ratio of the travel time of the fluid to the diffusion time across the slit as approximately $2\alpha X/(\bar{D}_A^2 Pe)$. In the example shown, this ratio is 7.1, which is sufficiently large compared to unity, thus explaining the observation. Significant gradients of concentration evolve within the cavity in approximately 25 ms of exposure to the wafer as noted in Fig. 7c. After this period, the concentration decreases with exposure time but the qualitative structure of the contours does not change with the passage of time. Ultimately, the concentration distribution approaches nearly uniform conditions as the reactant is depleted in the cell. In the example shown, the diffusion time scale over the height of the cavity $H^2/2\bar{D}_A$ is 0.497 s. After about 2 s of exposure, the reactant has been virtually depleted from the unit cell. Concentration contours for the product species B exhibit similar behavior, with the difference that concentrations increase with time. Therefore, they are not shown.

In Fig. 8 we show a typical plot of the distribution of the scaled concentration of A at the wafer surface over the cell at an intermediate value of the scaled time. Since the reaction rate is proportional to this surface concentration in the model, the figure also can be interpreted as displaying the flux of A to the surface of the wafer. The most important feature in the figure is the sharp increase in concentration that occurs right after the leading edge of the cavity ($X = 0$). The reason is the entrainment of a fresh supply of reactant from the cavity which also was noted earlier in Fig. 7. In the slit portion, the reactant is depleted as the fluid moves down the slit. In practice, there is a slight difference in concentration between the locations $X = -\lambda/2$, which is the leading edge of the unit cell, and $X = \alpha + \lambda/2$, which is the trailing edge. This difference is ignored in the present model since accommodating it would involve an unwarranted increase in complexity.

Now we illustrate the ability of the model to predict removal rates from wafers in Fig. 9. We consider the removal of copper from a blanket wafer of diameter 5 in. A 0.1 M ferric nitrate solution is assumed to be fed in the fresh slurry for the purposes of this illustration. Since the reaction is of first order, the results can be immediately used for any other concentration by using a proportional scale factor. Results for the removal rate plotted as a function of the relative velocity $V$ are displayed in Fig. 9. We have kept the parameter $\lambda$ at a value of unity in all the calculations. Illustrative results are provided for a set of values of the physical depth of the cavity $H$. In each case, the aspect ratio $\alpha$ is varied to show its influence. We recognize that the predicted removal rates are smaller than actually observed rates in a CMP tool at this concentration of ferric nitrate. One reason is the fact that the model predictions only apply to chemical removal and do not include the contribution to material removal made by abrasive particles. Also, we have used a kinetic constant measured at room temperature, and it is known that the fluid film and the surfaces with which it is in contact are warmer, primarily due to energy dissipated by friction, with some additional contribution from heat released due to chemical reaction. In fact, the rate of removal of copper per unit area from a coupon exposed to ferric nitrate solution in a stirred reactor has been found by Ramarajan and Babu to be smaller than the corresponding removal rate in a CMP tool.
tool, even in the absence of abrasive particles. This is likely due to the effect of the elevated temperatures and possibly due to some abrasive removal of softened material by the pad itself. As one would expect, the results for removal rates predicted by the model are indeed consistent with the results from the stirred reactor experiment from which the kinetic constant was obtained. In spite of the differences from the rates observed in a CMP tool, the present model predictions prove worthwhile in identifying several key features regarding the dependence of the removal rate on the various parameters in chemical removal of material from blanket wafers.

First, for a given set of parameters, the removal rate increases rapidly with velocity for small values of the velocity and levels off at a sufficiently high velocity. The slope is not constant but changes with velocity. Therefore, for purely chemical removal the Preston equation, which was developed for abrasive removal should not be used. In fact, when abrasive and chemical actions are combined, it is unlikely that a linear dependence of the removal rate on the relative velocity can be expected under all conditions. The trend observed here for chemical removal can be explained by noting that at small velocities, the chemical A initially present in the unit cell entering under the wafer is depleted by the time the cells move partway. In principle, since this is a first-order reaction, such complete depletion occurs only at infinite time. In practice, as can be seen from Fig. 7, the chemical is mostly used up by the diffusion time scale over the height of the cavity. During the rest of its passage under the wafer, the unit cell makes no contribution to removal of copper. The situation improves with increasing relative velocity until at sufficiently large velocity, the chemical concentration remains fairly high in the unit cell throughout its passage under the wafer, providing nearly maximum rates of reaction at the wafer surface. The theoretical limiting value is 688 nm/min, which corresponds to the maximum concentration of ferric nitrate being maintained at the entire wafer surface. It can be seen from Fig. 9 that this value is being approached asymptotically at high relative velocities.

The second feature displayed in Fig. 9 is the effect of the aspect ratio $\alpha$. For a given cavity depth, at very small velocities there is a small difference, and as the velocity is increased to intermediate values, we see that decreasing the aspect ratio results in substantially reduced removal rates. Finally, as the velocity is increased further, the aspect ratio has only a slight effect. These observations are consistent with the concept that reactant in the secondary eddy, which is present when $\alpha = 0.5$, has to be transferred to the primary eddy mainly through diffusion before it is able to participate in the reaction at the wafer surface. At large velocities, this is unimportant because most of the chemical is virtually intact in the cavity in the unit cell when it completes its journey under the wafer. In contrast, at intermediate velocities, the availability of the reactant to the primary eddy is important. At very low velocity, the time spent by the unit cell under the wafer is large and diffusive exchange is not a limiting factor. Rather, depletion of the entire reactant supply tends to control the removal rate.

The third feature shown in Fig. 9 that merits discussion is the role of the depth of the cavity. For this consider the case of a fixed aspect ratio, as for example $\alpha = 1$. We see that for a cavity of small depth, not only are the removal rates low but it takes a large relative velocity to achieve the asymptotic situation wherein the removal rate is not as sensitive to changes in the relative velocity. As the depth is increased (and the width along with it), we observe larger removal rates. Also, the removal rate levels off at a smaller velocity. Clearly, a cavity of small depth is depleted of its chemical content more readily than one which is deeper. The relatively lower concentrations that prevail in this situation explain the smaller removal rates. In order to avoid depleting the chemical substantially, the cavity must pass rapidly under the wafer. This is the explanation for the observed behavior.

The results presented in Fig. 9 were calculated with a fixed value of the scaled depth of the slit, $D = 1/50$. In Fig. 10 we show illustrative results for one case in which all the parameters are held fixed while varying the parameter $D$ by a factor of 2. It is evident that the effects are not dramatic, but there still is a significant improvement in removal rates at intermediate velocities when the relative depth of the slit is doubled. This is presumably due to the ability of the slit to hold more chemical as well as the more complex influence on the efflux of fluid containing chemical at a higher concentration from the leading edge of the cavity into the region near the wafer surface.

Finally, we compare the results from the detailed finite difference solution with those predicted by a simple stirred-tank model for a typical case in Fig. 11. The latter was calculated from Eq. 10 using Maple, which employs a numerical technique known as the Clenshaw-Curtis quadrature for performing the integration. It is evident from the figure that diffusional resistance is substantial at low velocities where depletion of the chemical in the cell is important. At higher velocities where the chemical concentration remains relatively high near the wafer surface and concentration gradients are weak, a stirred-tank model provides a reasonable first approximation to the removal rate. One must be cautious in extrapolating this observation to other situations. When the dimensionless reaction rate parameter is larger, diffusion plays a more important role. The stirred-tank model is recommended for use when $K$ is relatively small, and the relative velocity between the wafer and the pad is large.

It is possible to make calculations for other chemical reactions and the polishing of materials other than copper. However, our intent is not to embark on a parametric excursion but rather to present a fresh approach to dealing with the complex transport phenomena
issues that arise in CMP modeling. Therefore, we have refrained from providing a large amount of results for different situations, focusing instead on the physical explanation of the observed trends.

Conclusion

A new model accommodating transport phenomena coupled with chemical reactions in CMP is presented here, along with illustrative results for the chemical polishing of blanket copper films in a ferric nitrate slurry without any abrasive particles. The power of this model is that it is based on first principles, and by selecting certain key parameters based on observations of the pad architecture, predictions can be developed without making further assumptions. The model can be used to investigate competitive chemical processes, such as the formation of passivating films (Kaufman et al.,14), and phenomena such as dishing and erosion. Simple models of abrasion which are decoupled from the chemistry, as well as comprehensive models which couple abrasion with chemical softening of the surface of the film, also can be accommodated into the transport model as they become available. If we wish to include abrasive removal, one way is to assume, as suggested in Steigerwald et al.,27,28 that the abraded material, being in the form of small clusters of atoms, dissolves in the liquid adjacent to the surface by chemical reaction. In such a model, abrasive removal would appear as an additional known sink for the reactant at the wafer surface. In more complex models, abrasion would be coupled with the chemistry at the surface if removal follows softening of the surface.

Acknowledgments

This work was supported by a grant from the New York State Science and Technology Foundation to the Center for Advanced Materials Processing at Clarkson University and by NYSERDA through a subcontract from Ferro Corporation.

Clarkson University assisted in meeting the publication costs of this article.

List of Symbols

\[ C_0 \] molar concentration of chemical A in fresh slurry
\[ C_A \] scaled concentration of chemical A
\[ C_B \] scaled concentration of chemical B
\[ c_A \] molar concentration of chemical A
\[ c_B \] molar concentration of chemical B
\[ D \] ratio of the thickness of a slit to the height of a cavity
\[ \mathcal{D}_A \] diffusivity of chemical A
\[ \mathcal{D}_B \] diffusivity of chemical B
\[ d \] thickness of a slit
\[ F_A \] overall rate at which chemical A is consumed at the wafer surface
\[ H \] height of a cavity
\[ i \] unit vector in the \( x^* \) direction in Fig. 2
\[ K \] dimensionless chemical reaction rate parameter
\[ k \] chemical reaction rate constant for first-order surface reaction
\[ L_1 \] length of a cavity
\[ L_2 \] length of a slit
\[ \rho \] density of the liquid
\[ \rho_f \] density of the fluid
\[ \rho_s \] density of the solid
\[ \rho_v \] density of the vapor
\[ \Delta T \] scaled time step
\[ \tau \] time
\[ V \] relative velocity between the wafer and the pad
\[ V_x \] scaled velocity component in \( x \) direction
\[ V_y \] scaled velocity component in \( y \) direction
\[ v_x \] velocity component in the \( x \) direction
\[ v_y \] velocity component in the \( y \) direction
\[ X \] scaled rectangular coordinate
\[ x \] rectangular coordinate in the unit cell defined in Fig. 3
\[ x^* \] rectangular coordinate on the wafer scale defined in Fig. 2
\[ Y \] scaled rectangular coordinate
\[ y \] rectangular coordinate in the unit cell defined in Fig. 3
\[ y^* \] rectangular coordinate on the wafer scale defined in Fig. 2
\[ \Delta \] width of a strip shown in Fig. 2
\[ \alpha \] ratio of the length of a cavity to its height
\[ \beta \] constant defined in Eq. 11
\[ \lambda \] ratio of the length of a slit to the height of a cavity
\[ \Omega \] ratio of the diffusivity of chemical B to that of chemical A
\[ \mu \] dynamic viscosity of the liquid
\[ \nu \] kinematic viscosity of the liquid

References