Three-dimensional wafer-scale copper chemical–mechanical planarization model

Dipto G. Thakurta, Donald W. Schwendeman, Ronald J. Gutmann, Sadasivan Shankar, Lei Jiang, William N. Gill

Department of Chemical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
Department of Integrated Electronics and Electronic Manufacturing, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
Integrated Process Application Group, Technology CAD Department, Intel Corporation, Hillsboro, OR 97124, USA

Received 15 March 2002; received in revised form 10 April 2002; accepted 10 April 2002

Abstract

Models based on slurry hydrodynamics, mass transport, and reaction kinetics are developed in order to predict the removal rate of copper on a wafer-scale during chemical–mechanical planarization (CMP). The steps in the copper removal model include: mass transport of the oxidizer to the wafer surface; reaction of oxidizer with copper to form a reacted layer; and subsequent removal of the reacted layer by mechanical abrasion. The rates of the chemical reaction and mechanical abrasion steps are described by separate kinetic parameters. For low oxidizer concentrations the chemical step controls the process, while for high concentrations the mechanical step controls. The model shows that mass transport of the oxidizer to the surface controls the removal process at higher removal rates and can cause wafer-scale non-uniformity. Copper CMP experiments with potassium dichromate based slurry are compatible with the proposed surface kinetics steps and showed that the surface kinetics controlled the removal process for this slurry.

Keywords: Chemical–mechanical planarization; Copper; Wafer-scale; Lubrication theory; Surface chemistry

1. Introduction

Chemical mechanical planarization (CMP) has the capability to achieve adequate local and global planarization necessitated by increasingly stringent on-chip interconnects requirements in integrated circuit (IC) device manufacturing [1–3]. CMP has been widely used for planarization of interlevel dielectric (ILD), shallow trench isolation (STI), and the Damascene metallization process [1,2,4–6]. In spite of extensive use of the CMP process, many aspects of CMP are not well understood. Hence, physics-based modeling is useful in providing fundamental insight, optimization of the process, and performing sensitivity analyses of operating parameters [7]. Several CMP models have been published in the literature, with Nanz and Camilletti [8] and Subramanian et al. [9] presenting comprehensive summaries. In this paper, we develop a copper CMP model based on mass-transport theory and wafer surface kinetic steps, which predicts the removal rate on a wafer-scale. The concentration of the chemical reactant in the slurry is computed using a convective diffusive mass transport equation. The surface kinetic steps model the chemical reaction and mechanical abrasion processes at the wafer surface.

A conventional CMP process is depicted in Fig. 1. The slurry containing chemical reagents and abrasive particles is drawn beneath the wafer by the rotating pad. The slurry forms a lubricating film between the wafer and pad [10]. The film thickness, \( h \), ranges between 20 and 60 \( \mu \)m depending upon operating conditions [11]. Fig. 1b shows a side profile of the pad surface which has asperities of randomly varying heights [12]. The
surface roughness of the pad, defined as the standard deviation of the height of the asperities, ranges from 10 to 30 μm depending on the pad properties [13,14]. When \( h \) is of the same order as the mean height of the asperities, a significant fraction of the asperities are in contact with the wafer surface, while a thicker slurry film gives lesser contact. The applied load on the wafer is carried by the hydrodynamic pressure developed in the slurry film and by the asperities in contact.

For purposes of slurry flow modeling, the slurry is considered to flow between the finger-like protrusions (asperities) on the pad surface. The resistance to the slurry flow from these asperities has been modeled by Jiang and Shankar [15]. They obtain the slurry pressure and velocity distributions for flow with pad-asperity/wafer contact by using a statistical-averaging approach analogous to Patir and Cheng's [16] model for rough bearing lubrication. Here, we will use the 3-D wafer-scale slurry flow model developed by Thakurta et al. [17], which treats the pad as a smooth surface. In our approach, the effect of contact of the asperities with the wafer is brought into our mass-transport model through a boundary condition describing the surface kinetics at the wafer surface. The flux of the reactant at the wafer surface is computed from its concentration distribution, which is related to the local removal rate through a modified Langmuir–Hinshelwood model developed for copper CMP. The model includes a chemical reaction at the wafer surface with formation of a reacted layer, followed by its removal by mechanical abrasion. The chemical reaction and the mechanical abrasion are modeled by separate kinetic parameters. Copper CMP experiments are performed to evaluate the model.

Sundararajan et al. [18] analyzed the mass transport in copper CMP using an alkaline slurry. A 2-D lubrication model was used to analyze the mass-transport in the slurry film and was coupled to a chemical reaction at the wafer surface. The model agreed reasonably well with experiments predicting the effects of slurry chemistry, relative speed between the wafer and pad, and applied pressure on the wafer. The primary improvement of our present work is that the mass transport and flow models are extended to 3-D, which enables one to predict the wafer-scale local material removal rate. In addition, a more complete surface reaction kinetics scheme, including both chemical and mechanical effects, is presented for an acidic slurry, which is a more commonly used slurry for copper.

Subramanian et al. [9] have investigated transport phenomena issues in CMP from a somewhat different viewpoint. The pores of the pad are modeled as rectangular cells, which carry the slurry underneath the wafer. A mass-transport equation is used to describe the reactant concentration distribution in the cell, and the flux at the wafer surface is used to obtain the material removal rate. According to their model, the removal rate increases with pad speed, saturating at larger speeds, which agrees with trends observed experimentally. The transport model presented in our work differs from that in Subramanian et al. [9] in that we consider the pad fibers as protrusions from the pad surface and hence, do not restrict the slurry flow within a cell. Also, we include an additional kinetic rate parameter, which accounts for mechanical abrasion at the wafer surface.

In the work presented here we begin by briefly summarizing the flow model of Thakurta et al. [17] in Section 2. The flow results are needed as input to the mass-transport model developed in Sections 3.1 and 3.2. These two sections describe the model for the mass-transport in the gap between the wafer and the pad and the boundary condition given by our surface kinetics model, respectively. The model, as discussed in Section 3.3, determines the flux of copper at the wafer surface and this flux may be integrated to obtain the removal rate of copper. An effectiveness factor is defined in Section 3.3, which measures the mass-transport resistance to the overall removal process. The kinetic rate parameters used in the model are discussed in Section 3.4, followed by a discussion of the numerical solution procedure in Section 3.5. Results of the model for a representative flow configuration is presented in Section 4. The experiments used to test the model and determine the kinetic parameters are described in Section 5 and conclusions relating the model and experiments are given in Section 6.

2. Flow model

The purpose of the flow model is to compute the slurry film thickness and the velocity distribution of the slurry in the gap between the wafer and pad. These computed quantities are required by the mass-transport model to be discussed in the next section. As was pointed out earlier, the mass-transport model can be used with any available 3-D wafer-scale slurry flow model. For the purposes of this paper, we will use the 3-D wafer-scale slurry flow model developed by Thakurta et al. [17], which is briefly summarized here.
Fig. 2 shows the geometry of the wafer-pad system. The wafer radius is \( R_1 \), and the distance between the centers of the wafer and pad is \( R_2 \). The wafer and pad rotate about their respective centers with angular speeds \( \omega_1 \) and \( \omega_2 \), respectively. The co-ordinate system is fixed in space with the origin at the pad surface directly below the center of the wafer. The wafer height is given by \( z = h(x, y) \) measured relative to the pad surface (the \( x-y \) plane) and includes a global convex curvature and angles of attack. The curvature is represented in terms of the protrusion at the center of the wafer, also known as the wafer dome height \([10]\). The wafer dome height is denoted by \( \delta_0 \) and is of the order of 10 \( \mu m \) for a 200-mm diameter wafer. (The curvature shown in Fig. 2b is exaggerated for the purpose of illustration.) The wafer surface is taken to be

\[
h(x, y) = h_0 + S_x \left( \frac{x}{R_1} \right) + S_y \left( \frac{y}{R_1} \right) + \delta_0 \left( \frac{x^2 + y^2}{R_1^2} \right)
\]

where \( h_0 \) is the wafer height at the origin, and \( S_x \) and \( S_y \) are slopes associated with the angles of attack in the \( x- \) and \( y- \) directions, respectively. The parameters \( h_0, S_x \), and \( S_y \) in the expression for \( h \) are adjusted to balance the applied pressure and to satisfy zero moments on the wafer about the center \([10, 17]\).

The flow model in \([17]\) considers the lubrication approximation of the Navier–Stokes equations, which is solved numerically to obtain the slurry pressure in the gap and the adjustable parameters in the expression for \( h \). Once the pressure is found, the velocity distribution can be computed. Under typical CMP conditions, the computed slurry film thickness ranges from 20 to 60 \( \mu m \), and the slurry flow speeds range from 0.3 to 1.4 m s\(^{-1}\). Fig. 3 shows velocity-arrow plots of the computed slurry velocity field between the wafer and pad at three different normalized heights \( \zeta = 0, 0.75 \) and 0.5, where \( \zeta = z/h \), for a sample case with \( P_{app} = 25 \) kPa and \( \omega_1 = \omega_2 = 60 \) r.p.m. The length of the arrow represents the magnitude of slurry velocity. For reference, the arrows at the wafer edge in Fig. 3a denote a velocity of magnitude 0.63 m s\(^{-1}\). Mass-transport calculations for the same sample case will be presented in Section 4.1.

For the work presented here, pad bending and porosity effects are considered to be secondary and are neglected. Also, no contact of wafer and pad is assumed while solving for the slurry flow field. It is noted that contact and flow models may be coupled as demonstrated in Refs. \([15, 19]\). Here, the effect of pad contact is included only through kinetic rate parameters, which describe the rate of mechanical abrasion.

3. Copper removal model

A copper CMP slurry primarily consists of an oxidizer and abrasive particles. Several oxidizers, such as hydrogen peroxide \([20]\), ferric nitrate \([9]\), potassium ferricyanide, nitric acid \([6]\), and potassium dichromate \([21]\), have been used. Typically, sub-micrometer alumina particles are used as abrasives. A copper removal model is presented which has three steps: (1) mass transport of oxidizer to the wafer surface; (2) chemical reaction of the oxidizer with copper to form a reacted layer at the wafer surface; and (3) removal of both (a) the reacted layer and (b) copper by mechanical abrasion. These steps are depicted in Fig. 4 and are discussed in Sections 3.1 and 3.2 below.

3.1. Mass transport of oxidizer to the wafer surface

The thin slurry-film region between the wafer and pad may be visualized as a chemical reactor in which the reaction is occurring at the wafer surface. The fresh slurry consisting of an oxidizer, \( R \), enters this region as shown in Fig. 1a. The copper at the wafer surface, Cu, reacts with \( n \) moles of \( R \) forming a reacted layer, \( L \), which then is removed from the surface by mechanical abrasion, thus effecting material removal. The overall reaction of copper with the oxidizer may be written as...
follows:

$$\text{Cu} + nR \rightarrow L$$ \hspace{1cm} (2)

where the symbols with under-bars denote species present on the wafer surface. The slurry leaves the region depleted in $R$ and enriched in $L$. The rate of material removal depends upon the availability of $R$ in the fluid phase at the wafer surface. Both convection and diffusion transport $R$ to the wafer surface. Mass transport of $R$ to the surface is step 1 in the overall material removal process (see Fig. 4). A mass-transport model is used to solve for the concentration distribution of $R$ under the wafer region, which is used to compute the rate of consumption of $R$ at the wafer surface and subsequently, the local copper removal rate.

Assuming a dilute system with constant liquid density and diffusivity, we can write the following governing equation, from Bird et al. [22], for the concentration, $C(r, \theta, z)$, of oxidizer $R$:

$$\frac{\partial C}{\partial r} + v_r \frac{\partial C}{\partial r} + v_\theta \frac{\partial C}{\partial \theta} + v_z \frac{\partial C}{\partial z} = D \frac{\partial^2 C}{\partial z^2}$$ \hspace{1cm} (3)

where $D$ is the diffusivity of $R$ in the slurry and $v_r, v_\theta, v_z$ are functions of $(r, \theta, z)$ given by the slurry flow model. In Eq. (3), the contribution to diffusion in $r$- and $\theta$-directions is neglected because $h \ll R$.

The equation for $C$ is solved in the region between the wafer and pad given by $r \in [0, R_1], \theta \in [0, 2\pi], \text{ and } z \in [0, h]$, with appropriate boundary conditions at the inlet, pad and wafer surfaces. The inlet is defined as the regions at $r = R_1$ where the slurry, assumed ‘fresh’ in $R$, enters the space between the wafer and pad. Thus, the inlet boundary condition for $C$ is given by:

$$C(R_1, \theta, z) = C_0 \text{ when } v_r(R_1, \theta, z) < 0$$ \hspace{1cm} (4)

where $C_0$ is the molar concentration of $R$ in fresh or as prepared slurry (assumed to be known), and the velocity criterion in Eq. (4) identifies the inlet regions. The pad is considered to be inert and impermeable, so the diffusive flux is set to zero at the pad surface.

The boundary condition at the wafer surface $z = h(r, \theta)$ is based on the subsequent surface kinetic steps as described in the next section.

3.2. Surface kinetic steps

The mass transport of $R$ (step 1) is followed by steps 2 and 3 both occurring at the wafer surface. In step 2, the copper film, $\text{Cu}$, reacts with $n$ moles of $R$ (in the fluid phase) to form a reacted layer, $L$, on the wafer surface:

$$\text{Cu} + nR \rightarrow L$$ \hspace{1cm} (6)

assuming that the adsorption of $R$ on the copper surface is sufficiently fast and that the backward reaction rate of Eq. (6) is negligible. The reacted surface layer, $L$, is an oxidized copper compound depending upon $R$. According to our model, part of the wafer surface is

Fig. 3. Computed slurry velocity vectors at different $\zeta = z/h$ surfaces between the wafer (at $\zeta = 1$) and pad (at $\zeta = 0$). Flow simulation parameters: $P_{\text{app}} = 25 \text{ kPa}, \omega_1 = \omega_2 = 60 \text{ r.p.m.}, R_1 = 100 \text{ mm}, R_2 = 178 \text{ mm}, \delta_0 = 10 \mu \text{m} \text{ and } \mu = 0.001 \text{ Pa s}$. 

Fig. 4. Steps in the removal model for copper.
covered by the reacted layer, \( L \), while the rest is unreacted copper, \( \text{Cu} \). In Step 3, topographically higher regions of the reacted surface layer are mechanically removed by a combined action of the pad and abrasive particles, and the abraded material, \( L \), goes into the slurry (Step 3a). Direct removal of the copper surface by mechanical abrasion occurs simultaneously (Step 3b). It is assumed here that the abraded material does not re-deposit onto the wafer surface. The mechanical abrasion steps are given as follows:

\[
L \rightarrow L 
\text{Step 3a} 
\]

\[
k_3 \rightarrow \text{Cu} 
\text{Step 3b} 
\]

The chemical reaction in Step 2 and the mechanical abrasive actions in Steps 3a and 3b are described by the rate parameters \( k_1, k_2 \) and \( k_3 \), respectively. The rate parameter \( k_1 \) is determined by the chemistry of reactant \( R \) with copper and is assumed to be constant independent of the CMP operating conditions. The rate parameters \( k_2 \) and \( k_3 \), on the other hand, are dependent on a number of CMP parameters, in particular the pad and wafer speed and applied pressure. The rate parameters are discussed further in Section 3.4.

A first order chemical reaction with respect to \( R \) with a kinetic rate parameter \( k_1 \) is assumed in Eq. (6). The local rate of formation of \( L \) per unit surface area by Step 2, \( r_1 \), is obtained by using the principles of kinetics of fluid–solid catalytic reactions [23]. The rate is proportional to the product of the concentration of \( R \) in the fluid phase at the wafer surface and the probability that the surface site is \( \text{Cu} \)

\[
r_1 = k_1 \theta_{\text{Cu}} C(r, \theta, h) 
\tag{9}
\]

where \( \theta_{\text{Cu}} \) is the probability that the surface site is \( \text{Cu} \) and \( C(r, \theta, h) \) is the local concentration of \( R \) in the slurry at the wafer surface. Similarly, the rates (per unit surface area) of \( L \) and \( \text{Cu} \) going into the slurry (by Steps 3a and 3b, respectively) are given by

\[
r_2 = k_2 \theta_L 
\tag{10}
\]

and

\[
r_3 = k_3 \theta_{\text{Cu}} 
\tag{11}
\]

respectively, where \( \theta_L = 1 - \theta_{\text{Cu}} \) is the probability that the surface site is \( L \).

We wish to eliminate \( \theta_{\text{Cu}} \) (and \( \theta_L \)) from the rate expressions in favor of \( C \) so that the wafer surface boundary condition can be expressed in terms of \( C \). Steps 2 and 3a are in series. To deduce the fractions of the surface covered by copper and \( L \), it is assumed that the rate of formation of \( L \) is equal to the rate of depletion which gives

\[
r_1 = r_2 
\tag{12}
\]

Substituting Eqs. (9) and (10) into Eq. (12) and solving for \( \theta_{\text{Cu}} \) gives

\[
\theta_{\text{Cu}} = \frac{1}{1 + \frac{k_1}{k_2} C(r, \theta, h)} 
\tag{13}
\]

The above expression for \( \theta_{\text{Cu}} \) may be substituted back into Eq. (9) to obtain \( r_1 \) (and \( r_2 \)) in terms of the kinetic parameters and the wafer surface concentration of \( R \):

\[
r_1 = r_2 = \frac{k_1 C(r, \theta, h)}{1 + \frac{k_1}{k_2} C(r, \theta, h)} 
\tag{14}
\]

Eq. (14) refers to the rates at which the reacted layer is created and removed. From stoichiometry of the reaction in Eq. (6), the flux of \( R \) normal to the wafer surface is given by

\[
-D \frac{\partial C}{\partial z}(r, \theta, h) = nr_1 = \frac{nk_1 C(r, \theta, h)}{1 + \frac{k_1}{k_2} C(r, \theta, h)} 
\tag{15}
\]

(On the wafer-scale, the curvature of the wafer surface is small; hence, the \( z \)-axis is approximately normal to the surface.) Eq. (15) indicates how the flux of \( R \) and the rate of formation of \( L \) are related to one another. Eq. (15) forms the boundary condition for Eq. (3) at the wafer surface. Once the concentration distribution of \( R \) is computed, the removal rate of copper may be computed as described in Section 3.3.

### 3.3. Copper removal rate and effectiveness factor

Our model assumes that copper is removed simultaneously by the following two processes: (i) conversion into \( L \) followed by abrasion of \( L \) at the rate \( r_2 \) and (ii) direct abrasion at the rate \( r_3 \). Since copper is removed by two mechanisms, the instantaneous rate of copper removal at any point \( (r, \theta) \) on the wafer surface is given by

\[
RR(r, \theta) = \frac{M_w}{\rho} (r_1 + r_3) = \frac{M_w}{\rho} \left( \frac{k_3 [k_1 C(r, \theta, h)]}{1 + \frac{k_1}{k_2} C(r, \theta, h)} \right) 
\tag{16}
\]

where \( M_w \) and \( \rho \) denote the molecular weight and density of copper, respectively. Since the wafer rotates about its center and the duration of time over which the wafer is polished is typically much greater than the period of revolution, the average rate of copper removal for points on the wafer at a distance \( r \) from its center is given by

\[
RR(r) = \frac{1}{2\pi} \int_0^{2\pi} \frac{M_w}{\rho} \left( \frac{k_3 + k_1 C(r, \theta, h)}{1 + \frac{k_1}{k_2} C(r, \theta, h)} \right) d\theta 
\tag{17}
\]
We will refer to $RR(r)$ as the local removal rate. It is also useful to define

$$RR_{avg} = \frac{2}{R_1} \int_0^{R_1} R_r(r) r dr$$

(18)

which is the average removal rate for the entire wafer surface.

An effectiveness factor, $\eta$, is defined by

$$\eta = \frac{RR_{avg}}{RR_{avg,0}}$$

(19)

where $RR_{avg,0}$ is given by Eqs. (17) and (18) with $C(r, \theta, h)$ replaced by $C_0$ and with $k_2$ and $k_3$ replaced by the reference values $k_{20}$ and $k_{30}$, respectively, so that

$$RR_{avg,0} = \frac{M_w}{\rho} \left( \frac{k_{30} + k_1 C_0}{1 + \frac{k_1}{k_{20}} C_0} \right)$$

(20)

The average removal rate would take the value of $RR_{avg,0}$ if the entire wafer surface were exposed to the inlet concentration. This would be the case if mass transport occurred very fast relative to the time scale for the surface kinetics so that the surface concentration would be $C_0$ to a good approximation. The effectiveness factor, $\eta$, is thus a measure of the mass-transport resistance to the overall removal process. That is, $\eta$ is very close to unity if the mass transfer rate of $R$ to the surface is fast enough to feed the chemical reaction and maintain the concentration at the slurry–surface interface near the inlet concentration, $C_0$.

3.4. Kinetic rate parameters

In Section 3.2, the removal mechanism of copper was modeled using three surface kinetic parameters $k_1$, $k_2$ and $k_3$. The parameter $k_1$, which describes the rate of the chemical reaction at the wafer surface [Eq. (6)], is primarily a function of the oxidizer $R$ present in the slurry. The local temperature, pressure and concentration of $R$ are also known to affect the rate of chemical reaction. In CMP, the abrasive particles also can enhance the rate of reaction by weakening the surface bonds by microcutting or brittle fracture. Here, we assume that $k_1$ depends only on the oxidizer used and it is independent of position on the wafer surface.

The rate parameters $k_2$ and $k_3$ describe the rates of mechanical abrasion of $L$ and $Cu$, respectively. The contact pressure, pad, abrasive size and concentration have strong effect on $k_2$ and $k_3$. The value of $k_2$ also depends strongly on the type of the oxidized layer, $L$, formed by the oxidizer. In the case of oxidizers such as potassium dichromate ($K_2Cr_2O_7$), the reacted layer passivates the surface [21] and prevents further corrosion of the underlying copper. Other oxidizers, such as nitric acid ($HNO_3$) [6] and ferric nitrate [Fe(NO$_3$)$_3$] [24], do not form an effective passivating film. The properties of the reacted layer may depend critically upon the concentration of the oxidizer in the slurry as in the case of hydrogen peroxide ($H_2O_2$) [24]. Lower concentrations of $H_2O_2$ give a porous reacted layer while higher concentrations produce a passivating layer. A reacted layer, which is passivating, is more difficult to abrade than the case when it is not passivating or porous in nature. It is noted that $k_3$ is the rate at which copper is abraded and it is not a function of the slurry oxidizer.

The values of the abrasion rate parameters $k_2$ and $k_3$ vary locally as a function of the local contact pressure at the wafer surface. To illustrate the effect of variation of the contact pressure on the local removal rate, we assume that $k_2$ and $k_3$ are directly proportional to the local contact pressure between the wafer and pad in some of our simulation runs. The mechanical rate parameters $k_2$ and $k_3$ are assumed to be proportional to a shape function $\phi(r)$, which describes the shape of the contact pressure distribution,

$$k_2(r) = k_{20} \phi(r), \quad k_3(r) = k_{30} \phi(r)$$

(21)

The shape function captures the positional dependency of $k_2$ and $k_3$ across the wafer, while the proportionality parameters $k_{20}$ and $k_{30}$ determine the scale, and they are functions of applied pressure, wafer/pad speed, abrasives, back pressure, wafer-backing film and the type of pad.

Several articles have been published which present models to compute the contact pressure distribution [19,25–29]. Baker’s 2-D contact model [26], for example, treats the pad as a beam supported by an elastic foundation which leads to an analytical expression for the shape, $\phi(r)$, of the contact pressure distribution which takes the form

$$\phi(r) = 1 + \{1 + 4\cos(-\beta(R_1-r))/\exp(-\beta(R_1-r))\}$$

(22)

where $\beta$ is a parameter which depends on the elastic properties of the pad. The shape function, $\phi(r)$, near the wafer edge is plotted in Fig. 5 for a typical value of $\beta$ given by Baker [26]. The contact pressure is constant in the central region of the wafer and there is a six-fold increase at the edge. The parameter $\beta$ primarily determines the distance from the edge at which the contact pressure starts increasing. For the choice of $\beta$ used here, the contact pressure starts increasing from approximately 4 mm away the edge.

3.5. Solution procedure

A numerical method of solution is needed to obtain the concentration distribution of the oxidizer $C$ in the region between the wafer and the pad. We can compute the copper removal rate using Eq. (16) once $C$ is computed. Eq. (3) is solved numerically for $C(r, \theta, z)$.
in the region between the wafer and pad defined by \( r \in [0, R], \ \theta \in [0, 2\pi], \) and \( z \in [0, h], \) with boundary conditions at inlet, pad and wafer surfaces given by Eqs. (4), (5) and (15), respectively.

Since the wafer surface is curved, transformation of the \( z \)-coordinate is convenient for the purpose of discretization. The following co-ordinate transformation is applied

\[
\zeta = z/h(r, \theta) \tag{23}
\]

which changes the solution domain from \( z \in [0, h] \) to \( \zeta \in [0, 1], \) Eq. (3) and its boundary conditions are changed to the \( (r, \theta, \zeta) \) co-ordinate system and then discretized using a finite volume approach [30]. A method of iteration (Newton’s method) is used to solve the system of discrete equations for the oxidizer concentration. The accuracy of the numerical results was checked by using standard grid refinement techniques. In the results reported, the number of grid points in the \( r-, \theta- \) and \( \zeta- \)directions are 75, 50 and 50, respectively, for a 200 mm diameter wafer.

4. Model results

4.1. Sample results

As described in the previous section, a numerical approximation to the concentration distribution of the oxidizer, \( C(r, \theta, z), \) and the local removal rate of copper, \( RR(r), \) may be computed once the flow and mass-transport problems are solved. Three sample run conditions and key results are presented in Table 1, with more results presented elsewhere [30]. A typical mass diffusivity of a species in liquid media of \( 10^{-9} \text{ m}^2 \text{ s}^{-1} \) is used along with a stoichiometric coefficient, \( n, \) of unity for these runs. The direct abrasion rate coefficient, \( k_{\text{sb}}, \) is set to zero for simplicity and the slurry flow shown in Fig. 3 is used as input to the mass-transport equations. The results listed in Table 1 show that the average removal rate decreases with decreasing \( k_1, \) a measure of the rate of chemical reaction at the wafer surface. The behavior of \( \eta \) and the average concentrations of the oxidizer at the wafer and pad, denoted by \( C_s \) and \( C_b, \) respectively, as functions of \( k_1 \) are also indicated in Table 1. The values of \( k_1 \) have been varied by a factor of more than 20 to show how the chemical reaction affects the results. Higher chemical rate constants cause more rapid consumption of oxidizer at the wafer surface so that mass transfer to the wafer surface is more rapid than at the edge caused, in part, by \( R(r) \) being larger at the edge, and also by mass transfer being rapid at the inlet. The small dip at point (B) corresponds to the dip in the value of \( k_2 \) at \( r=96 \text{ mm} \) as seen in Fig. 5. (The change in slope of \( RR(r) \)

![Fig. 5. Shape function for the contact pressure distribution (plotted near the wafer edge) from Baker’s model [26].](image)

Table 1. Summary of results of the three different cases

<table>
<thead>
<tr>
<th>Case</th>
<th>( k_1 \times 10^3 ) (m/s)</th>
<th>( RR_{\text{avg}} ) (nm/min)</th>
<th>( \eta )</th>
<th>( C_s/C_0 )</th>
<th>( C_b/C_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>1256</td>
<td>0.5</td>
<td>0.3</td>
<td>0.85</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>677</td>
<td>0.7</td>
<td>0.65</td>
<td>0.93</td>
</tr>
<tr>
<td>3</td>
<td>0.6</td>
<td>215</td>
<td>0.9</td>
<td>0.9</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Simulation parameters: \( C_0=0.1 \text{ mol/l}, \) \( D=10^{-9} \text{ m}^2 \text{ s}^{-1}, \) \( k_{\text{sb}}=1 \times 10^{-8} \text{ mol m}^{-2} \text{ s}^{-1}, \) \( k_0=0, \) \( \beta=655 \text{ m}^{-1} \) and \( n=1. \) Flow simulation parameters are the same as in Fig. 3.
Fig. 6. Case 1: (a)–(c) contours of normalized concentration distribution \((C/C_0)\) of reactant, \(R\), at different \(\zeta = \) constant surfaces between the wafer (at \(\zeta = 1\)) and pad (at \(\zeta = 0\)). Wafer edge is shown by the dashed line. (d) \(C/C_0\) contours of \(R\) at the \(y-z\) co-ordinate plane. Wafer surface is shown by the thick line. (c) Copper removal rate as a function of radial position. \(k_1 = 15 \times 10^{-3}\) m/s. Other simulation parameters listed in Table 1.

near (B) is very rapid and is not a slope discontinuity as it appears in the plot. The high removal rate consumes more oxidizer near the edge, contributing to the decrease in the removal rate towards the left of (7B).

The effect of Baker’s contact model on the radial variation in the removal rate is shown in Fig. 7. The solid curve in the plots is Case 1, which uses \(\phi(r)\) from Eq. (22) (Baker model), while the dashed curve is obtained using no contact model, i.e. \(\phi(r) = 1\). The removal rate variation in the central region is due purely to effects of mass transport and hydrodynamics with Baker’s model only affecting the results near the edge (enlarged in Fig. 7(b)).

Cases 2 and 3 have lower values of \(k_1\) as compared to Case 1 (see Table 1). As \(k_1\) decreases the rate of chemical reaction decreases at the wafer surface, with a corresponding decrease in removal rates as listed in Table 1. Also, a higher value of \(k_1\) results in a lower average surface concentration of the oxidizer, \(C_s\). A low value of the oxidizer concentration near the wafer surface indicates (i) a fast wafer surface reaction rate and (ii) that mass transport of oxidizer to the wafer surface plays a role in the removal process of copper.

Fig. 7. (a) Normalized local removal rate variation with and without Baker’s model [26]. (b) Enlarged view near the edge showing the difference.
The impact of mass transport on the removal rate is quantified by the efficiency factor, $\eta$, as defined in Eq. (19) which varies from 0 to 1. The efficiency factor for Case 1 is lowest indicating that mass-transport resistance of the oxidizer in reaching the wafer surface plays an important role in the removal process. In Case 3, a high value of $\eta = 0.9$ indicates that mass transport plays a less significant role as compared to Cases 1 and 2. Table 1 also lists the average concentration of the oxidizer at the pad surface $C_s$, which does not vary as much as $C_b$ because it is further away from the reacting surface and in effect is the source from which diffusion occurs to feed the reaction.

4.2. Effect of the rate parameters and inlet concentration on removal rate

Fig. 8 shows the effect of varying $k_1$, $k_{20}$, $k_{30}$ and $C_0$ on the average removal rate. Values for the other input parameters are fixed and are taken to be $D = 10^{-9}$ m$^2$s$^{-1}$, $\beta = 655$ m$^{-1}$ and $n = 1/3$, and the remaining flow simulation parameters are the same as those used previously. Fig. 8a indicates that the average removal rate increases steadily with $C_0$ at small values of $C_0$ and then plateaus to a maximum at high values of $C_0$. At small values of $C_0$, the removal rate is limited by the rate of chemical reaction [Eq. (6)], whereas at larger values of $C_0$ it is limited by the rate of mechanical abrasion [Eq. (8)]. The chemical rate parameter, $k_1$ determines the slope of the curve of $RR_{avg}$ vs. $C_0$ at small values of $C_0$, while the mechanical rate parameter, $k_{20}$ determines the height of the plateau at larger values. This behavior is made clear by Eq. (20) taking the limits of small and large values of $C_0$.

In Fig. 8a, the three curves shown have different initial slopes which are directly proportional to their respective values of $k_1$. All three curves approach a maximum value of $RR_{avg}$ at large values of $C_0$ (not shown in the figure) which is approximately the same value as $k_{20}$ and is the same for all the curves. In Fig. 8b, the removal rate curves have the same initial slope and reach different plateau heights which are directly proportional to their respective $k_{20}$ values. Fig. 8c shows the direct abrasion of copper (i.e. parameter $k_{30}$) causes an approximately uniform increment in removal rate at all oxidizer concentrations. Direct abrasion of copper is the only mechanism of removal in our model when the slurry has zero oxidizer concentration ($C_0 = 0$), so that $k_{30}$ can he determined using a slurry with $C_0 = 0$.

4.3. Dimensionless removal rate and effectiveness factor

The removal rate and the effectiveness factor [Eq. (19)] depend on the parameters that govern the slurry flow, e.g. $P_{app}$, $\omega_1$, $\omega_2$, $R_1$ and $R_2$, and those that control the surface kinetics, e.g. $k_1$, $k_2$ and $k_3$. For the case
when the resistance of mass transport of oxidizer to the wafer surface, either by convection or diffusion, is small, the rate determining step is either chemical (Step 2 as given by Eq. (6)) or mechanical abrasion (Step 3a as given by Eq. (8)) at the wafer surface. Our model may be used to determine the behavior of the removal rate and the effectiveness factor as functions of \( k_1 \) (for Step 2) and \( k_{20} \) (for Step 3a) while holding the remaining input parameters fixed as indicated in Fig. 9.

In order to motivate the choice of the dimensionless quantities used to plot Fig. 9, we must define some new quantities that are related to the magnitudes of the various parameters in the process. We begin with a length scale, \( z_0 \), which is proportional to the distance between the pad and the wafer. This quantity is related to known quantities by balancing scales for the horizontal pressure gradient and the vertical component of the shear stress in the horizontal momentum equation, i.e. Eq. (9) of Thakurta et al. [17]. Appropriate scales for these quantities are \( P_{\text{app}}/R_1 \) and \( \omega_z R_2/z_0 \), respectively. Equating these scales leads to

\[
  z_0 = \left[ \frac{\mu \omega_z R_2 R_1}{P_{\text{app}}} \right]^{1/2}
\]

which gives \( z_0 \) in terms of known quantities.

We now consider the left-hand-side of the flux boundary condition in Eq. (15), and approximate it in terms of the difference in concentration at the wafer and at the pad as follows:

\[
  - D \frac{\partial C(r, \theta, h)}{\partial z} \approx - D \left[ \frac{C_s - C_b}{z_0} \right] \frac{C_b - C_s}{z_0}
\]

When the concentration at the wafer surface is small, \( C_s = 0 \), the process is mass transfer limited because the rate of consumption of the active ingredient by the surface reaction is rapid enough to consume essentially all of it that reaches the surface by mass transfer. In this case, we may define a dimensionless average removal rate by

\[
  RR_{\text{avg}}^* = \frac{RR_{\text{avg}}}{\frac{1}{n \rho} \frac{DC_0}{z_0}}
\]

where the concentration at the pad, \( C_p \), is taken as the inlet concentration, \( C_0 \). When \( RR_{\text{avg}}^* \) is small, substantially less than one, the surface reaction is relatively slow and the effectiveness factor is close to unity. Under these conditions the concentration of active ingredient remains close to \( C_0 \) throughout the slurry.

The dimensionless average removal rate, \( RR_{\text{avg}}^* \), is plotted in Fig. 9 as a function of dimensionless rates \( k_1^* \) and \( k_{20}^* \). These quantities are defined below by a further consideration of Eq. (15). If we divide Eq. (15) (with \( \theta(r) \) taken to be 1) by the mass transfer limited flux, \( DC_0/z_0 \), we obtain

\[
  - \frac{\partial C^*(r, \theta, h)}{\partial z^*} = \frac{k_1^* C^*(r, \theta, h)}{1 + \frac{k_1^*}{k_{20}^*} C^*(r, \theta, h)}
\]

where \( C^* = C/C_0 \) and \( z^* = z/z_0 \), and the dimensionless rates \( k_1^* \) and \( k_{20}^* \) are given by

\[
  k_1^* = \frac{n k_1 z_0}{D} \quad \text{and} \quad k_{20}^* = \frac{n k_{20} z_0}{DC_0}
\]

The surface rate limited flux occurs when \( C(r, \theta, h) = C_0 \) in which case

\[
  - \frac{\partial C^*(r, \theta, h)}{\partial z^*} = \frac{k_1^*}{1 + \frac{k_1^*}{k_{20}^*}}
\]

The right-hand-side of Eq. (29) gives the ratio of the surface rate limited flux to the mass transfer rate limited flux, the ratio of the two limiting cases of importance. It also suggests the definitions of the dimensionless rate parameters \( k_1^* \) and \( k_{20}^* \). The quantity, \( k_1^* \), is the ratio of the rate of chemical reaction to the mass transfer rate and \( k_{20}^* \) is the ratio of the rate of the mechanical abrasion to mass transfer.

Similarly, if we divide Eq. (16) by the mass transfer limited rate, we obtain

\[
  RR'(r, \theta) = \frac{k_{30}^* + k_1^* C^*(r, \theta, h)}{1 + \frac{k_1^*}{k_{20}^*} C^*(r, \theta, h)}
\]

where

\[
  RR'(r, \theta) = \frac{RR(r, \theta)}{\frac{1}{n \rho} \frac{DC_0}{z_0}} \quad \text{and} \quad k_{30}^* = \frac{n k_{30} z_0}{DC_0}
\]

Performing the integrations indicated in Eqs. (17) and
(18) gives us the complete expression for $RR_{avg}^*$, and if $\eta = 1$ this reduces to

$$RR^* = \frac{k_{30}^* + k_1^*}{1 + \frac{k_1^*}{k_{20}^*}} \quad (32)$$

In Fig. 9a we show the behavior of a dimensionless average removal rate, $RR_{avg}^*$. Contours of the dimensionless average removal rate are plotted as a function of the dimensionless rate parameters $k_1^*$ and $k_{20}^*$. Also, contours of the effectiveness factor, $\eta$, are plotted in Fig. 9b as a function of $k_1^*$ and $k_{20}^*$. For these plots, $k_{30}^*$ is taken to zero (as in Table 1), because the contribution of direct abrasion of copper (Step 3b) to the removal rate is not significant unless the inlet concentration of oxidizer is small.

Plots such as those in Fig. 9 provide a means of interpreting experimental data. If the removal rate has been measured, one can compute $RR_{avg}^*$ from Eq. (26). For example, suppose $RR_{avg}^* = 0.2$, then the contour for this removal rate can be located on Fig. 9a and the $k_1^*$ and $k_{20}^*$ values corresponding to that contour on Fig. 9b indicate that $\eta$ is between 0.96 and 0.98. For this case, one would conclude that surface kinetics limit the rate of removal, as one would want in CMP because then the concentration is close to $C_0$ everywhere between the pad and the wafer and the removal rate is more uniform. Furthermore, if $\eta = 1$, Eq. (20) determines the removal rate to a very good approximation. However, if the range of $\eta$ includes values that are not close to unity then the full slurry flow/mass transfer model must be solved to determine the removal rate. In addition, if $\eta$ is not close to unity, then local variations in removal rate may be significant and can be determined from Eq. (16) by using the mass transfer model given here.

5. Copper CMP experiments with potassium dichromate based slurry

Copper CMP experiments are performed with a $K_2Cr_2O_7$ slurry, with the applied pressure, wafer/speed and oxidizer concentration varied. The objectives are to (i) evaluate the model assumptions, (ii) determine values of the kinetic parameters $k_1$, $k_{20}$ and $k_{30}$ by fitting the model predicted average removal rate to the experimentally measured value, and (iii) determine the behavior of $k_{20}$ and $k_{30}$ as functions of applied pressure and speed.

For our CMP experiments, blanket copper films are sputtered on 125 mm diameter unpatterned wafers. The slurry consisted of an oxidizer ($K_2Cr_2O_7$), alumina particles (50 nm nominal diameter) and de-ionized (DI) water, with the abrasive concentration fixed at 3 wt.%. An IPEC 372M polishing tool is used with a Pan-W pad manufactured by Freudenberg. The diameter of the polishing table is 57 cm and the eccentric distance between the polishing head and the center of the polishing table, $R_e$, is 173 mm. Both the platen and carrier are rotated at equal speed in all experiments, with the slurry delivery rate fixed at 150 ml/min. The polish time varied from 1 to 2 min. Sheet resistance is measured using a four-point probe before and after the polishing experiment at four locations in the central region of the wafer. The sheet resistance measurements are used to obtain the local removal rates, which are then averaged to compute an average removal rate, $RR_{avg}$.

The $K_2Cr_2O_7$ slurry is ideal to evaluate our model because the formation of a reacted layer on the surface with this slurry has been confirmed by potentiodynamic and surface Auger electron spectroscopy (AES) measurements [21]. The overall oxidation reaction at the wafer surface can be written as follows

$$3Cu + Cr_2O_7^{2—} + 14H^+ \rightarrow 3Cu^{2+} + 2Cr^{3+} + 7H_2O \quad (33)$$

Three moles of copper react per mole of the oxidizer, giving a stoichiometric coefficient $n = 1/3$. (For comparison with Eq. (2), $R$ is $Cr_2O_7^{2—}$ and $L$ is $Cu^{2+}$.) The applied pressure, wafer/speed and $K_2Cr_2O_7$ concentration in the slurry are varied to study their effects on the average removal rate. The following pressure (kPa)/ speed (r.p.m.) combinations are used: 14/45; 28/45; 42/45; 28/30; and 28/60. $K_2Cr_2O_7$ concentration in the slurry, $C_0$, is varied between 0 and 0.35 mol/l for each pressure/speed combination. The experimental data plotted in Fig. 10 show that the average removal rate increases with an increase in either pressure or speed. A measurable removal rate exists at zero $K_2Cr_2O_7$ concentration, which is also a function of pressure and speed. For any given pressure/speed combination, $RR_{avg}$ initially increases with $C_0$ and then plateaus with further increase in $C_0$, in agreement with our model.

The values of $k_1$, $k_{20}$ and $k_{30}$ for this slurry are determined by fitting the average removal rate predicted by our model with that determined experimentally. We assume the following: (i) the value of $k_1$ is constant; and (ii) $k_{20}$ and $k_{30}$ are functions of pressure and speed. The shape function, $\phi$, is set equal to one as it only affects the removal rate in a small region at the wafer edge and does not significantly affect the average removal rate calculations.

Initial estimates for the $k$ values are obtained as follows. At a given pressure and speed, $k_{20}$ is estimated to fit the removal rate data in the plateau region, while $k_{30}$ is estimated to fit the data at zero oxidizer concentration. The value of $k_1(1-k_3/k_2)$ determines the rate of change of $RR_{avg}$ with $C_0$ for small values of $C$, is estimated by fitting to the initial slopes at all pressure/speed combinations of the experimental data shown in Fig. 10.
The efficiency factor, $\eta$, for all experimental conditions is very close to one implying that the mass transport of the oxidizer (Step 1) is fast enough to maintain the concentration near the wafer surface very close to $C_0$. Therefore, the removal process is controlled by the surface kinetic steps, and the average removal rate is well approximated by Eq. (20). The dots in Fig. 9b show the corresponding dimensionless rate parameters for these experiments.

The five sets of data plotted in Fig. 10 correspond to the five chosen pressure/speed combinations in the experiments. Given that $\eta$ is close to one, values for $k_1$, $k_{20}$ and $k_{30}$ may be determined from the data using a non-linear least squares fit to the curves given by Eq. (20) with the constraint that $k_1$ is held fixed for all data sets [30]. The fitted rate parameters are listed in Table 2.

In the experimental range of pressure and speed, both $k_{20}$ and $k_{30}$ are approximately proportional to the wafer-pad speed while $k_{20}$ has a stronger dependency on the applied pressure than that for $k_{30}$. Intuitively, the hardness of the copper layer relative to that of the reacted layer plays a significant role on the dependency of $k_{20}$ and $k_{30}$ on pressure, but not so much on speed. For the oxidizer used in the experiments, the dependency suggests that the copper layer is harder than the reacted layer. Also, at a given pressure and speed, the higher values of $k_{20}$ as compared to $k_{30}$ (both fitted from the experiments) again implies the reacted layer is easier to abrade than copper.

Removal by chemical reaction followed by the abraison of the reacted layer is preferred as a better surface finish is obtained. Wafers which were polished with $C_0 \geq 0.1 \text{ mol/l}$ had a shiny metallic finish, while polishing with $C_0 \leq 0.1 \text{ mol/l}$ resulted in a hazy finish with significant visible scratching.

### Table 2

<table>
<thead>
<tr>
<th>$P_{\text{app}}$ (KPa)</th>
<th>$\omega_1 = \omega_2$ (r.p.m.)</th>
<th>$k_1 \times 10^6$ (mol m$^{-2}$ s$^{-1}$)</th>
<th>$k_{20} \times 10^4$ (mol m$^{-2}$ s$^{-1}$)</th>
<th>$k_{30} \times 10^4$ (mol m$^{-2}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>45</td>
<td>2.82</td>
<td>1.83</td>
<td>0.68</td>
</tr>
<tr>
<td>28</td>
<td>45</td>
<td>2.82</td>
<td>4.88</td>
<td>1.21</td>
</tr>
<tr>
<td>42</td>
<td>45</td>
<td>2.82</td>
<td>8.61</td>
<td>1.40</td>
</tr>
<tr>
<td>28</td>
<td>30</td>
<td>2.82</td>
<td>2.96</td>
<td>0.75</td>
</tr>
<tr>
<td>28</td>
<td>60</td>
<td>2.82</td>
<td>6.95</td>
<td>1.64</td>
</tr>
</tbody>
</table>

### 6. Conclusions

A model for copper CMP was developed for an acidic slurry and evaluated with experiments. The model takes into account the effects of slurry hydrodynamics, slurry chemistry, chemical reaction and mechanical abrasion at the wafer surface. The steps in the copper removal process include: mass transport of the oxidizer to the wafer surface; reaction of oxidizer with copper to form a reacted layer with subsequent removal of the reacted layer by mechanical abrasion; in addition, copper may be removed by direct abrasion. The rates of the chemical reaction and mechanical abrasion steps are described by separate kinetic rate parameters. Variation in the removal rate results from mass transport effects and variation in the local contact pressure.

A good CMP process should be controlled by the surface processes, not mass transfer, and requires a balance between chemical and mechanical effects. For low oxidizer concentrations, the chemical reaction (Step 1) limits the removal rate; as a result, surface damage is possible due to direct mechanical abrasion (Step 2).

![Fig. 10. Behavior of the average removal rate as a function of K$_2$Cr$_2$O$_7$ concentration for (a) varying applied pressure and (b) varying speed. The marks represent experimental measurements. The error bars for the 45 r.p.m./14 kPa case indicate one standard deviation in the average removal rate measurements and are representative of the other cases. The solid curves show model predictions based on fitted values of the kinetic parameters listed in Table 2.](image-url)
and the removal rate is lower than can be achieved. At high oxidizer concentrations, mechanical abrasion controls the removal rate of the chemically-altered surface layer. Copper CMP experiments with K$_2$Cr$_2$O$_7$ oxidizer and alumina abrasive particles are compatible with the proposed surface kinetics model.

The importance of mass transport has been incorporated in our model and is measured by the efficiency factor, $\eta$. Mass transport effects are important when the removal rates are quite high; more specifically, mass transport of the oxidizer to the wafer surface not only limits the removal rate, but results in a wafer-scale radial non-uniformity. Unfortunately, this regime of operation could not be experimentally evaluated with our equipment and consumable set. In fact, any robust CMP manufacturing process requires a value of $\eta = 1$ to avoid wafer-scale non-uniformities in removal rate.

Acknowledgments

This work was supported by the Center of Interconnect Science and Technology (CAIST), Semiconductor Research Corporation (SRC) contract number 99-IC-448. The authors also thank Freudenberg Nonwovens for donating Pan-W polishing pads.

References