

Modeling and calculation of slurry-chemistry effects on chemical–mechanical planarization with a grooved pad

Yao-Chen Wang · Tian-Shiang Yang

Received: 7 September 2007 / Accepted: 30 July 2008 / Published online: 15 August 2008
© Springer Science+Business Media B.V. 2008

Abstract During chemical–mechanical planarization (CMP), a rotating wafer is pressed against a rotating pad, while a slurry is dragged into the pad–wafer interface. Here, taking into account the dependence of local material removal rate (MRR) on the slurry’s chemical activity, the effects of pad groove geometry and various other process parameters on the spatial average and non-uniformity of MRR are examined. Technically, the slurry flow is calculated by following an existing approach that integrates two-dimensional fluid-film lubrication theory and contact-mechanics models. A slurry impurity transport equation is then used to calculate the impurity concentration that determines the slurry’s chemical activity and hence the local MRR. The numerical results obtained here indicate that the presence of pad grooves generally decreases the average slurry impurity concentration, and increases the average contact stress on the pad–wafer interface. However, as a grooved pad has less contact area for effective interaction with the wafer surface, the average MRR may or may not be increased, depending upon the specific setting of process parameters. Meanwhile, it appears that the retaining ring generally used to keep the wafer in place also plays an important part in reducing the MRR non-uniformity.

Keywords Chemical–mechanical planarization · Contact mechanics · Grooved pads · Lubrication theory · Material removal rate · Slurry impurity

1 Introduction

Chemical–mechanical planarization (CMP) is the enabling technology for eliminating topographic variations and achieving near-perfect planarity of interconnection and metal layers in ultralarge-scale integrated (ULSI) semiconductor devices [1]. During the CMP process, a rotating wafer is pressed facedown against a rotating pad, while a slurry containing chemicals and abrasive particles is dragged into the pad–wafer interface. Polishing is then accomplished by the interaction of the pad and slurry with the wafer surface or by direct contact between the wafer and the pad [2]. In order to ensure stable and high performance of CMP, especially for wafers of larger diameters (say, 12 inches), it is important to optimize the slurry, pad, and other consumables [3,4]. This has thus necessitated the use of grooved pads that help discharge debris and prevent subsequent “particle loading effects” [5,6].

Y.-C. Wang · T.-S. Yang (✉)
Department of Mechanical Engineering, National Cheng Kung University, 1 University Road, Tainan 70101, Taiwan, ROC
e-mail: tsyang@mail.ncku.edu.tw

With or without grooves on the pad, CMP is a complicated process where contact mechanics, fluid dynamics, and slurry chemistry interact synergistically. Optimizing the use of CMP consumables therefore requires more than just trial and error, and would benefit greatly from theoretical modeling and numerical simulations, to which we wish to make an additional contribution in this paper. In particular, here we shall take into account the dependence of local material removal rate (MRR) on the chemical activity of the slurry, and examine how the spatial average and non-uniformity of MRR are affected by the pad groove geometry (*viz.* the groove width, depth, and spacing), the filtration capability of the pad, and the sensitivity of the slurry's chemical activity to its impurity concentration.

To put the present work in perspective, let us first briefly review some of the previous studies that are related to the theoretical modeling and numerical calculation of slurry flow in CMP with a grooved pad. It is worth noting that Muldowney and Tselepidakis [7] used the commercial software FLUENT 6.1 to deal with the complicated three-dimensional surface topography of a grooved pad. In their calculations, the grooves were modeled as fluid-filled cells and the asperity layer on the land area of the pad as porous cells, whose equivalent porosity and characteristic length were determined experimentally [8]. The approach of Muldowney and coworkers [7,8] revealed many detailed features of the slurry flow, but the necessity of resolving the slurry flow in each of the numerous grooves on the entire pad generally would require a very large amount of computational resources. To reduce the numerical computation costs, simpler yet reasonably accurate simulation tools for CMP slurry flow with a grooved pad therefore still are highly desirable.

Attempting to devise one such tool, Eaton et al. [9] used fluid-film lubrication theory to calculate the slurry flow. To account for the finite Reynolds number and "direction effects" on the slurry flow, a number of correction factors, which were deduced from numerical results obtained by solving the Navier–Stokes equations directly (using FLUENT), were introduced into their model. Since the fluid velocity and pressure variations across the slurry-film thickness were dealt with analytically using the approximation of lubrication theory, the numerical computations no longer needed to resolve the flow structure in the film thickness direction, and therefore were less expensive than solving the Navier–Stokes equations directly. Note, however, that since the interaction between the pad and wafer was not modeled, a constant slurry-film thickness had to be specified in the calculations of Eaton et al. [9].

A number of other theoretical models exist in the literature (e.g., [10–12]), each considering certain aspects of the CMP process. To our knowledge, one of the simplest models that calculates the contact stress on the pad–wafer interface, which then determines the slurry-film thickness, was proposed by Shan et al. [11] for CMP with a flat pad. Briefly, they considered a two-dimensional (2-D) model problem in which a rigid punch (modeling the wafer) slides over the surface of an elastic half-space (modeling the pad) at a constant relative velocity (see Fig. 1; but disregard the pad grooves and the retaining ring for now). Meanwhile, using a commercial bench-top polishing machine (Logitech Model PM4), and replacing the typically rotating wafer by a stationary pressure sampling fixture, they also measured the fluid pressure distribution on the pad–fixture interface. Despite being an oversimplification of the real CMP process, the model of Shan et al. [11] still produced results that compared favorably with their experimental data. One of their major findings was that the average fluid pressure on the pad–wafer interface typically was subambient; the contact stress on the pad–wafer interface therefore was increased by the presence of the fluid (slurry).

On the surface of the pressure sampling fixture used in the experiments of Shan et al. [11], there was a particular diameter that was perpendicular to the line connecting the centers of the rotating pad and the stationary fixture. The relative sliding speed between the pad and fixture surfaces was approximately constant along that diameter. As a constant sliding speed inevitably had to be assumed in the 2-D model of Shan et al. [11], it was the pressure measurements along that diameter that were compared with the model predictions. Note also that, in practice, the relative sliding speed can be, and usually is, rendered uniform across the entire wafer surface simply by rotating the wafer and pad at the same angular speed [2]. So, the predictions of the 2-D model hopefully still could provide certain insights into the slurry flow and other CMP characteristics for the real 3-D situation. We shall compare some numerical results of 2-D models with real experimental data to provide partial support for this argument (see Sect. 4).

In a previous work [13], aiming to incorporate the pad–wafer interaction into the calculation of CMP slurry flow with a grooved pad, we integrated the groove modeling of Eaton et al. [9] with the 2-D contact-stress modeling of

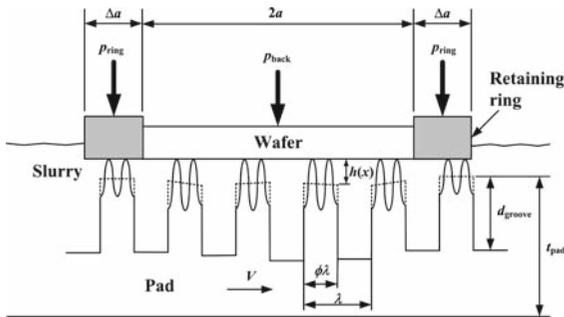


Fig. 1 Schematic of the model system (the wafer, retaining ring, grooves and compressed asperities on the pad are not shown to scale)

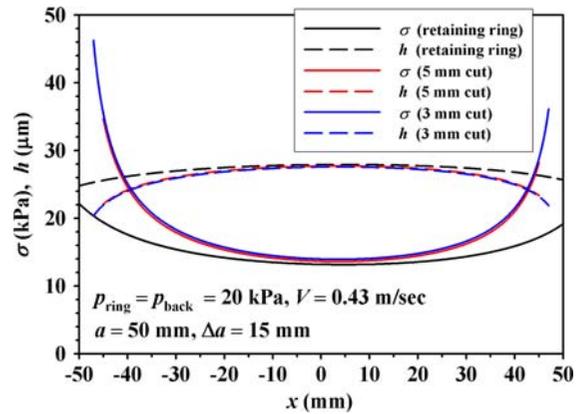


Fig. 2 Contact stress and slurry-film thickness distributions for a flat pad, with $p_{ring} = p_{back} = 20 \text{ kPa}$, $a = 50 \text{ mm}$ and $\Delta a = 15 \text{ mm}$ (other parameters are detailed in the text). Corresponding results of the “cut model” devised by Shan et al. [11] are also included for comparison

Shan et al. [11]. Note that, when grooves are placed periodically on the polishing pad, the geometry of the pad–wafer interface region would vary periodically with time as the wafer slides over the pad surface at a constant relative speed. So, from a technical viewpoint, calculating the CMP slurry flow on a grooved pad is more complicated than that on a flat pad. It was precisely for this reason that we chose to base our previous work [13] on the simple 2-D model of Shan et al. [11], expecting that, once the technical difficulties arising from the presence of pad grooves were resolved, similar approaches would also be applicable to other CMP models.

Here we shall extend our previous work [13] in two aspects. First, the presence of a retaining ring that is generally used to keep the wafer in place will be modeled (see Sect. 2). The resulting slurry flow and contact-stress distribution then will be discussed in Sect. 3, and it will be seen that the retaining ring is crucial for producing uniform CMP results, as was also observed experimentally by Gitis et al. [14]. Second, we shall take into account the dependence of local MRR on the chemical activity of the slurry, and calculate the spatial average and non-uniformity of MRR—which are important performance indices of CMP. To that end, a simple model for the dependence of the local MRR on the slurry’s chemical activity is proposed here. Also, a slurry impurity-transport equation that accounts for the generation of impurity by local material removal, the convection of impurity by slurry flow, and the filtration of impurity by the porous pad is used to calculate the impurity concentration distribution (see Sect. 4). The slurry-impurity concentration then determines the chemical activity of the slurry, and hence the local MRR. In Sect. 4, we shall also calibrate some model parameters by fitting our numerical results with existing experimental data [15]. Then, in Sect. 5, we shall examine the effects of various pad-groove and process parameters on the slurry-impurity concentration and MRR. Finally, concluding remarks will be given in Sect. 6.

2 Slurry flow and contact-stress calculation

Figure 1 is the schematic of our 2-D model problem for CMP with a grooved pad. Here the 2-D rigid punch of width $2a$ models the wafer, and hence will be referred to as the “wafer” hereafter. The two rigid punches of width Δa on both sides of the wafer represent the retaining ring generally used to keep the wafer in place. The pressure p_{ring} applied on the retaining ring usually is different from the wafer back pressure p_{back} , and we shall discuss how the slurry flow and contact-stress distribution are affected by Δa and p_{ring} in Sect. 3.1. For now let us just assume that the retaining ring and wafer surfaces are coplanar, so that they effectively are a single rigid punch of width $2(a + \Delta a)$, below which the pad slides at a constant speed V with friction coefficient f .

Suppose also that the grooved pad has a fraction ϕ of land area in each repeating groove unit of length λ (see Fig. 1; hereafter λ is referred to as the “pitch” of the grooves, and ϕ the “contact ratio” of the pad). As in our previous work [13], here we shall use a “screening function” $\zeta(x)$ to mark the position of pad grooves with respect to the wafer at a certain reference instant $t = 0$: with $\zeta = 1$ if x corresponds to a point on the land area of the pad and $\zeta = 0$ otherwise; the origin of the coordinate system ($x = 0$) is located at the wafer center. The screening function $\zeta(x)$ therefore is periodic in x with period λ . Furthermore, since the pad is moving at a constant speed V with respect to the wafer, the groove position at any instant t is given by $\zeta(x - Vt)$, which for a fixed location is time-periodic with period $T = \lambda/V$.

The theoretical model and numerical procedures for calculating the contact-stress distribution and slurry flow are essentially the same as that devised in our previous work [13]. So, in what follows we shall only outline the major ingredients of the model. The reader is invited to consult our previous work for further details.

2.1 Contact-stress calculation

Results of finite-element analysis suggest that the contact stress over the pad grooves is negligibly small compared with that on the land area, where the pad is in direct contact with the wafer. Also, when extrapolated over the grooves, the contact stress on a grooved pad has approximately the same spatial distribution as that on a flat pad [16]. Assuming that this is the case, we use a classic result of 2-D contact mechanics [17, p. 41] to write the extrapolated contact stress distribution for $x \in [-a - \Delta a, a + \Delta a]$ as

$$\sigma_0(x, t) = \frac{\tilde{P}(t) \cos \pi \kappa}{\pi [(a + \Delta a)^2 - x^2]^{1/2}} \left(\frac{a + \Delta a + x}{a + \Delta a - x} \right)^\kappa, \quad (1)$$

where the parameter κ is calculated from

$$\cot \pi \kappa = -\frac{2(1 - \nu)}{f(1 - 2\nu)}, \quad (2)$$

ν being the Poisson ratio of the pad. Note that $\int_{-a-\Delta a}^{a+\Delta a} \sigma_0(x, t) dx = \tilde{P}(t)$, and hence the “strength” function $\tilde{P}(t)$ in (1) is the resultant compression force (per unit length) acting upon the wafer-retaining ring assembly by $\sigma_0(x, t)$. Moreover, multiplying $\sigma_0(x, t)$ by the screening function $\zeta(x - Vt)$ to render the contact stress zero over the grooves yields the contact-stress distribution for all $x \in [-a - \Delta a, a + \Delta a]$:

$$\sigma(x, t) = \frac{\tilde{P}(t) \cos \pi \kappa}{\pi [(a + \Delta a)^2 - x^2]^{1/2}} \left(\frac{a + \Delta a + x}{a + \Delta a - x} \right)^\kappa \zeta(x - Vt). \quad (3)$$

In the steady state, $\tilde{P}(t)$ is time-periodic, i.e., $\tilde{P}(t) = \tilde{P}(t + T)$.

In addition to the contact stress $\sigma(x, t)$, the wafer-retaining ring assembly is subjected to the back pressures p_{back} , p_{ring} , and fluid (slurry) pressure $p(x)$. As there typically are relatively wide grooves on the retaining-ring surface to facilitate slurry flow in and out of the pad–wafer interface, we shall neglect the flow resistance across the retaining ring, and take $p = 0$ for $a \leq |x| \leq a + \Delta a$. Then, invoking force balance on the wafer-retaining ring assembly, we may express the resultant compression force in terms of the back and fluid pressures:

$$P(t) = \int_{-a-\Delta a}^{a+\Delta a} \sigma(x, t) dx = 2a \cdot p_{\text{back}} + 2\Delta a \cdot p_{\text{ring}} - 2a \cdot p_{\text{AVG}}(t), \quad (4)$$

where

$$p_{\text{AVG}}(t) = \frac{1}{2a} \int_{-a}^a p(x, t) dx \quad (5)$$

is the average fluid pressure on the pad–wafer interface. Note that, when the retaining ring is not modeled (as was the case in our previous work [13]), the wafer half-width (or “radius”) a then replaces the total width $a + \Delta a$ of the wafer-retaining ring assembly in both (1) and (3). Also, the contribution of p_{ring} is absent in the force balance (4).

It was proposed in our previous work that

$$\tilde{P}(t) \approx P(t)/\phi. \quad (6)$$

The rationale behind such an approximation was quite simple: Since the land area of the extrapolated flat pad is approximately $1/\phi$ times that of the grooved pad, the corresponding resultant compression force $\tilde{P}(t)$ is expected to be magnified by approximately the same factor. We have checked that the errors incurred in such an approximation are negligibly small, and hence will also use (6) to calculate the strength function $\tilde{P}(t)$ appearing in both (1) and (3).

In practice, the wafer typically is pressed against the pad by a fixture with a ball joint, and therefore cannot sustain a moment (see [18] for example). So, strictly speaking, the wafer-surface tilting has to be taken into account in the model to ensure that the resultant moment with respect to the ball joint due to the contact stress, fluid pressure, and friction acting on the wafer surface is zero. However, since the tilting angle of the wafer surface usually is extremely small (on the order of a few micro radians in the calculations of Jin et al. [12]), for simplicity we will not explicitly consider the moment balance here.

In (3), and likewise (1) for the extrapolated flat pad, the contact stress tends to infinity at the outer edges of the retaining ring, $x = \pm(a + \Delta a)$, due to stress concentration. This does not cause difficulties in our subsequent calculation of the slurry flow in the pad-wafer interface region ($-a \leq x \leq a$), however, since the contact stress remains finite there. In contrast, as the retaining ring was not modeled in our previous work [13] and that of Shan et al. [11], a small length had to be excluded from both edges of the wafer in the numerical computations, in order to avoid infinite contact stress at the wafer edges. Accordingly, the zero-fluid-pressure condition ($p = 0$) was imposed at the boundaries of the truncated computational domain, and the average fluid pressure had to be evaluated for the truncated domain, rather than for the entire wafer surface—like (5) above. In Sect. 3.1, we shall compare some numerical results of the two models for a flat pad.

2.2 Slurry-film-thickness calculation

The equivalent slurry-film thickness is taken to be the distance between the wafer surface and the mean asperity plane of the pad (see Fig. 1). In particular, the equivalent slurry-film thickness $h_0(x, t)$ on the land area of the pad is simply that produced by the extrapolated contact-stress distribution $\sigma_0(x, t)$. Using the Greenwood–Williamson contact model [19] with an exponential asperity height distribution, we calculate that

$$h_0(x, t) = s \log \frac{\pi^{1/2} \eta E R^{1/2} s^{3/2}}{(1 - \nu^2) \sigma_0(x, t)}, \quad (7)$$

where E is the elastic modulus of the pad, η and R are the density and average radius of the asperities on the pad, and s the root-mean-squared average of the pad-surface roughness [11, 13].

To determine the total film thickness over the grooves, we view the land area in each groove unit in Fig. 1 as a column [17, pp. 104–106]. When subjected to the compressive contact stress σ_0 , the column height is shortened by an amount of σ_0/K , where the stiffness parameter K is related to the thickness t_{pad} and elastic modulus E of the pad. The compressed groove depth is therefore $d_{\text{groove}} - \sigma_0/K$, with d_{groove} being the uncompressed groove depth. Adding then the compressed groove depth to the extrapolated film thickness $h_0(x, t)$ for points over the grooves, one obtains the film-thickness distribution for the entire wafer:

$$h(x, t) = h_0(x, t) + \{1 - \zeta(x - Vt)\} \cdot \{d_{\text{groove}} - \sigma_0(x, t)/K\}. \quad (8)$$

Using the practical parameter values to be specified in Sect. 3, it can be estimated that the decrease in groove depth $\sigma_0/K \sim O(p_{\text{back}}/K)$ is on the order of a few μm 's, while the uncompressed groove depth d_{groove} typically is on the order of $100 \mu\text{m}$. So, σ_0/K usually is relatively small compared with d_{groove} , and one expects that the slurry-film-thickness distribution given above by (8) would be relatively insensitive to the specific way in which the pad substrate is modeled.

2.3 Fluid-pressure calculation

Slurry flow at the pad–wafer interface typically has a small Reynolds number, and the fluid (slurry) pressure $p(x, t)$ can be calculated from the Reynolds equation of fluid-film lubrication theory [20, Chap. 7]:

$$\frac{\partial h}{\partial t} = -\frac{\partial q}{\partial x} = -\frac{\partial}{\partial x} \left\{ \frac{1}{2} Vh - \frac{1}{12\mu} \psi(h) h^3 \frac{\partial p}{\partial x} \right\} \quad (-a \leq x \leq a), \quad (9)$$

where μ is the viscosity of the fluid and

$$q(x, t) = \frac{1}{2} Vh - \frac{1}{12\mu} \psi(h) h^3 \frac{\partial p}{\partial x} \quad (10)$$

is the volumetric slurry flowrate per unit wafer breadth. The first term on the right-hand side of (10), $Vh/2$, is contributed by the relative sliding motion between the wafer and pad, and the second term by the pressure gradient whose strength is reduced by a “flow factor” $\psi(h) = 1.0 - 0.9 \exp(-0.56 h/s)$ that was determined in a numerical study by Patir and Cheng [21] to account for the additional flow resistance caused by isotropic surface roughness.

The numerical procedures for calculating the fluid pressure (and other quantities of interest at the same time) can be found in our previous work [13], and the reader is referred to that work for a detailed discussion of the technicalities. Note, however, that for flat pads having contact ratio $\phi = 1$ and screening function $\zeta(x, t) = 1$ identically, the slurry-film thickness would not depend on time, so $\partial h/\partial t = 0$, and (9) reduces to the steady Reynolds equation used by Shan et al. [11]. The geometry of the computational domain then is invariant with time, thus greatly simplifying the numerical procedures for calculating the fluid pressure. To illustrate how the various process variables are interrelated, we give a brief description of the solution procedures for a flat pad here. Specifically, one could start with a properly guessed average fluid pressure, p_{AVG} , and then calculate the corresponding resultant compression force P and contact-stress distribution on the pad from (4) and (1) in turn. [For flat pads, $\tilde{P} = P$ exactly, $\sigma_0(x, t) = \sigma(x)$, $h_0(x, t) = h(x)$, and all other process variables also are independent of time in the steady state.] Next, the slurry-film-thickness distribution is calculated from (7), and the result is then used for integrating the steady version of (9)—discretized using central differences and subjected to the boundary conditions $p = 0$ at $x = \pm a$ —to obtain the fluid-pressure distribution $p(x)$. Once $p(x)$ is calculated, the average fluid pressure p_{AVG} can be re-evaluated using (5) and corrected by Newton’s method. The above procedures are then repeated until the value of p_{AVG} converges.

3 Features of the slurry flow and contact-stress distribution

For low slurry-impurity concentrations, it would be appropriate to assume that the slurry flow simply transports the impurity originating from material removal, and is not affected by the presence of the impurity. The slurry flow therefore can be calculated without considering the impurity transport. So, before we proceed to explain the modeling and calculation of slurry-impurity transport and MRR, let us first discuss the basic features of the slurry flow and contact-stress distribution.

In all computations, the pad properties (listed in Table 1) are taken to be the same as those used in the work of Shan et al. [11], and the viscosity of water at room temperature $\mu = 0.001$ Pa s is used for the slurry. The half-width of the wafer is taken to be $a = 50$ mm, the wafer back pressure $p_{\text{back}} = 20$ kPa, and the relative sliding speed $V = 0.43$ m/s, as in the work of Shan et al. [11]. Furthermore, following Tichy et al. [10], the stiffness parameter for calculating the compressed groove depth is taken to be $K = 2.5$ MPa/mm (corresponding to a pad thickness on the order of a few millimeters). The friction coefficient between the pad and wafer (and retaining ring) surfaces is $f = 0.8$.

Here we divide the computational domain into 20,000 equal-sized panels (the grid size therefore is $\Delta x = 2a/20,000 = 5$ μm) so that even the narrowest grooves encountered in the computations can be adequately resolved. Moreover, for grooved pads, the discrete time step is chosen to be $\Delta t = \Delta x/V = 11.63$ μs , so that in each time step the pad moves a distance of Δx exactly. It has also been checked that reducing the grid size (and the corresponding time step) does not alter the numerical results significantly.

Table 1 Pad properties (corresponding to Rodel IC 1000 pad; adapted from [11])

Parameters	Values
Elastic modulus	$E = 12.01 \text{ MPa}$
Surface roughness	$Ra = 5 \text{ }\mu\text{m}$, rms $s = 6 \text{ }\mu\text{m}$
Poisson's ratio	$\nu = 0.3$
Asperity density	$\eta = 400 \text{ mm}^{-2}$
Average radius of asperities	$R = 0.1 \text{ mm}$

3.1 Flat pads

First we shall discuss the results for a flat pad. As a particular example, suppose that the retaining ring width $\Delta a = 15 \text{ mm}$ and back pressure $p_{\text{ring}} = 20 \text{ kPa}$. It is then calculated that the average fluid pressure $p_{\text{AVG}} = -1.0526 \text{ kPa}$, producing a resultant compression force of $P = 2.7053 \text{ kN/m}$ on the pad, as calculated from (4) using $p_{\text{ring}} = p_{\text{back}} = 20 \text{ kPa}$. The negative average fluid pressure (i.e., suction) is consistent with the findings of Shan et al. [11]. The corresponding contact stress $\sigma(x)$ on the pad and slurry-film thickness $h(x)$ are plotted in Fig. 2. It is seen that, due to friction, the contact-stress distribution is asymmetric with respect to the center of the wafer. Also, the contact stress near the wafer edges still is higher than that near the wafer center. The fluid-pressure distribution $p(x)$ is plotted in Fig. 3, and the volumetric slurry flowrate (per unit wafer breadth) is calculated to be $0.058 \text{ cm}^2/\text{s}$. Note that the fluid pressure is negative near the left wafer edge, and positive near the right wafer edge, thus helping drive the slurry in and out of the pad–wafer interface.

It is also of interest to compare the results of the present model with that of the previous “cut model,” so in Figs. 2 and 3 the corresponding results of the cut model are presented as well. (Recall that the cut model for a flat pad was first proposed by Shan et al. [11], and then extended in our previous work [13] to include the presence of pad grooves.) In particular, results obtained with a 5 mm edge cut and that with a 3 mm edge cut are plotted. It is seen in Fig. 2 that the contact-stress distributions for the two different edge cuts do not differ much in the central region. However, when the boundary of the truncated computational domain is moved closer to the wafer edge (where the contact-stress singularity exists), a larger contact stress at the truncated domain boundary is obtained. This then produces a smaller slurry-film thickness there, and the correspondingly higher flow resistance results in more profound fluid-pressure variations, as can be clearly seen in Fig. 3. The average fluid pressure is calculated to be $p_{\text{AVG}} = -2.0506 \text{ kPa}$ with a 3 mm edge cut, and $p_{\text{AVG}} = -1.5483 \text{ kPa}$ with a 5 mm edge cut (while $p_{\text{AVG}} = -1.0526 \text{ kPa}$ for the present model with a retaining ring). Compared with the much higher wafer back pressure $p_{\text{back}} = 20 \text{ kPa}$, the variation of p_{AVG} with the edge cut size is rather insignificant, and this explains why the contact-stress and slurry-film-thickness distributions plotted in Fig. 2 for the two cut sizes are almost identical except near the wafer edges. However, the fluid pressure plotted in Fig. 3 does exhibit much more sensitive dependence upon the edge cut size and the presence of a retaining ring. How this dependence would affect the material removal and slurry impurity transport will be discussed in Sect. 5.

The results discussed above point to the fact that the presence of a retaining ring reduces the contact stress at the wafer edges, thereby giving a more uniform contact-stress distribution on the wafer surface (see Fig. 2). Among other factors, the local MRR of the wafer is determined by the local contact stress (to be explained in Sect. 4), so a more uniform contact-stress distribution implies a more uniformly distributed local MRR, which is as important as having a large overall (i.e., spatially averaged) MRR. Of course, the reduction of contact stress at the wafer edges results from the assumption that the retaining ring and wafer surfaces are properly aligned so that they are coplanar. We now proceed to examine how the retaining ring width and back pressure affect the contact-stress uniformity, in order to see if it is always possible to align the wafer and retaining ring surfaces.

First, for $p_{\text{ring}} = p_{\text{back}} = 20 \text{ kPa}$, we calculate the slurry flow and contact-stress distribution for a number of retaining ring widths. The calculated average fluid pressure, p_{AVG} , and the corresponding ratio between the maximum and minimum contact stresses, $\sigma_{\text{max}}/\sigma_{\text{min}}$, are plotted in Fig. 4 against the retaining width Δa . Note that the $\sigma_{\text{max}}/\sigma_{\text{min}}$ ratio is a measure of the contact-stress uniformity. We see that, as Δa increases (so that the wafer

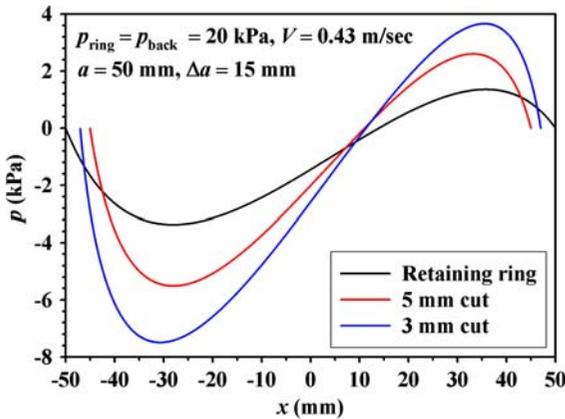


Fig. 3 Fluid-pressure distribution for a flat pad, with $p_{\text{ring}} = p_{\text{back}} = 20 \text{ kPa}$, $a = 50 \text{ mm}$ and $\Delta a = 15 \text{ mm}$. Corresponding results of the “cut model” devised by Shan et al. [11] are also included for comparison

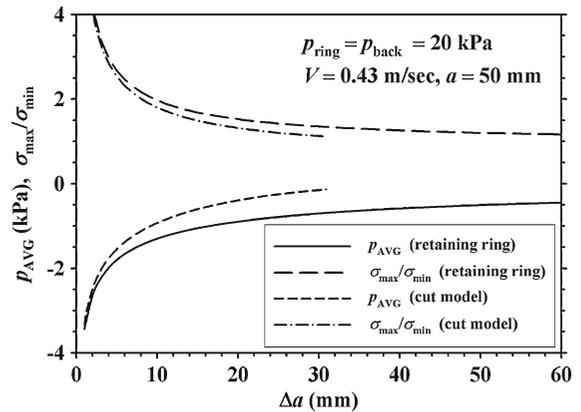


Fig. 4 Dependence of the average fluid pressure, p_{AVG} , and the corresponding ratio between the maximum and minimum contact stresses, $\sigma_{\text{max}}/\sigma_{\text{min}}$, on the retaining ring width Δa (with $p_{\text{ring}} = 20 \text{ kPa}$) for a flat pad. Corresponding results of the “cut model” devised by Shan et al. [11] are also included for comparison; for such results the abscissa Δa represents the edge cut size

edges become farther away from the edges of the wafer-retaining ring assembly where the contact-stress singularity exists), the $\sigma_{\text{max}}/\sigma_{\text{min}}$ ratio decreases towards unity. Hence the contact stress and MRR uniformities are expected to be improved by using a wider retaining ring. Meanwhile, the reduced contact stress at the wafer edges produces a greater slurry-film thickness and hence lower flow resistance. Accordingly, the magnitude of the (negative) average fluid pressure is decreased, as seen in Fig. 4. For the present parameter settings, it appears that using a retaining ring width of $\Delta a = 15 \text{ mm}$ reduces the $\sigma_{\text{max}}/\sigma_{\text{min}}$ ratio to a sufficiently small level, so we shall use such a width in all subsequent computations.

In Fig. 4, the average fluid pressure p_{AVG} and contact-stress ratio $\sigma_{\text{max}}/\sigma_{\text{min}}$ calculated using the cut model are also plotted against the edge cut size. The important observation there is that both $\sigma_{\text{max}}/\sigma_{\text{min}}$ and $|p_{\text{AVG}}|$ decrease with increasing cut size. Note, however, that it is indeed an artificial way to remove the contact-stress singularity by truncating the computational domain; the edge cut size therefore has to be kept small compared with the wafer size. Also, the present model clearly is superior to the cut model, because it includes the physical presence of a retaining ring and removes the contact-stress singularities at the wafer edges simultaneously. So, for grooved pads, only the results of the present model will be discussed (in Sect. 3.2).

We have also calculated the slurry flow and contact-stress distribution for various values of the retaining ring back pressure p_{ring} , fixing the retaining ring width at $\Delta a = 15 \text{ mm}$. The calculated results (not shown here to save some space) of the average fluid pressure p_{AVG} and contact-stress ratio $\sigma_{\text{max}}/\sigma_{\text{min}}$ turn out to be relatively insensitive to p_{ring} . Specifically, increasing p_{ring} from 0 to 100 kPa causes p_{AVG} to increase by about 1 kPa only, and the variation of $\sigma_{\text{max}}/\sigma_{\text{min}}$ ratio is even more insignificant. This seems to suggest the possibility of fine-tuning the retaining ring back pressure (over a rather wide range) to align the retaining ring and wafer surfaces, without affecting the slurry flow significantly. In all our subsequent computations, it is therefore assumed that the retaining ring back pressure can always be fine-tuned to produce coplanar retaining ring and wafer surfaces, and the value of the retaining ring back pressure is simply set to be $p_{\text{ring}} = p_{\text{back}} = 20 \text{ kPa}$, as the numerical results are not sensitive to its exact value.

3.2 Grooved pads

As a specific example, consider a grooved pad having contact ratio $\phi = 0.9$, uncompressed groove depth $d_{\text{groove}} = 200 \mu\text{m}$, and 40 grooves under the wafer ($N = 40$). The groove pitch then is $\lambda = 2a/N = 2.5 \text{ mm}$, and the period of slurry flow $T = \lambda/V = 5.81 \text{ ms}$. Despite that a retaining ring has been added here, the calculated results turn out

Fig. 5 Variation of the spatial average of fluid pressure with time, for a grooved pad with $N = 40$, $\phi = 0.9$, and $d_{\text{groove}} = 200 \mu\text{m}$

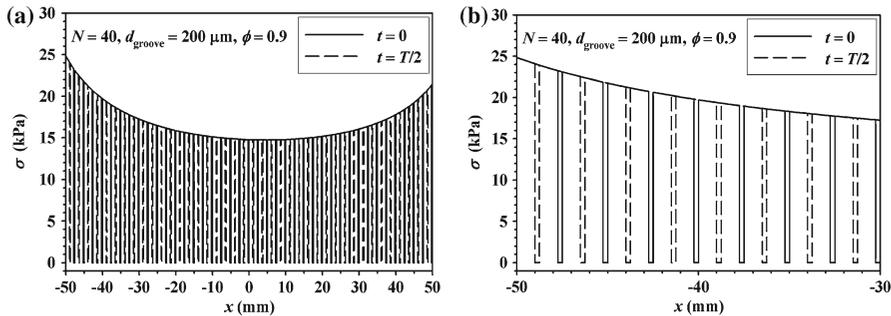
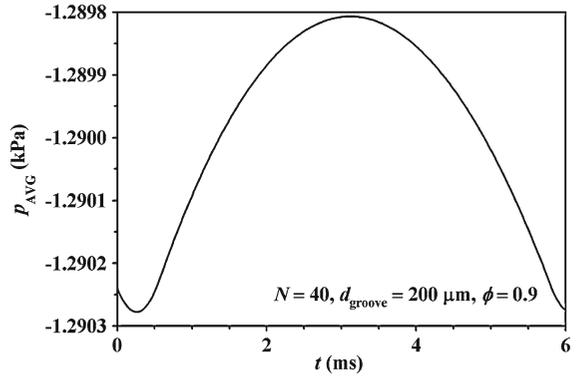


Fig. 6 Contact-stress distributions at $t = 0$ and $T/2$, for $N = 40$, $\phi = 0.9$, and $d_{\text{groove}} = 200 \mu\text{m}$. (a) Wafer-wide distribution. (b) Close-up of the first few grooves

to be qualitatively similar to those discussed in our previous work [13]. So, to avoid repetition, in what follows we shall only briefly summarize the major slurry-flow features that are essential for understanding the slurry impurity transport and material removal characteristics to be discussed in subsequent sections. The interested reader is invited to consult our previous work for a more detailed discussion.

Now, using the parameter values specified above, the temporal variation of the spatially averaged fluid pressure p_{AVG} is calculated, and the result is shown in Fig. 5. The temporal mean of $p_{\text{AVG}}(t)$ is calculated to be $\langle p_{\text{AVG}} \rangle = \int_0^T p_{\text{AVG}}(t) dt / T = -1.2900 \text{ kPa}$, higher (in magnitude) than the time-invariant $p_{\text{AVG}} = -1.0526 \text{ kPa}$ for the flat pad discussed in Sect. 3.1. Note that the fluid pressure plotted in Fig. 5 is much less than the back pressure $p_{\text{ring}} = p_{\text{back}} = 20 \text{ kPa}$, indicating that the back pressure is balanced mainly by interfacial contact stress.

The contact-stress distributions at $t = 0$ and $T/2$ are shown in Fig. 6. The “dips” having zero contact stress there correspond to the groove locations. Note also that the reference instant $t = 0$ does not have an absolute meaning; it is assigned here to the instant when a groove is about to go under the wafer at $x = -a$ (see Fig. 6). Comparing Figs. 2 and 6, it is seen that the presence of pad grooves increases the contact stress on the land area by a factor of about $1/\phi \approx 1.1$. The increase in contact stress therefore is caused mainly by the fact that a grooved pad has less land area in contact with the wafer-retaining ring assembly.

We have also calculated the slurry-film thickness, pressure, and flowrate distributions for the grooved pad specified above. Briefly, due to the increased contact stress on the land area, the slurry-film thickness there is somewhat smaller than that on the flat pad. The slurry-film thickness over the grooves, however, is substantially greater. Note that a smaller slurry-film thickness implies higher local flow resistance. The local pressure gradient $|\partial p / \partial x|$ in the land area on the grooved pad therefore is higher than that on the flat pad. The increased local pressure gradient then accumulatively develops a greater pressure variation across the wafer surface, consistent with the aforementioned observation that the average fluid pressure for the grooved pad, $\langle p_{\text{AVG}} \rangle = -1.2900 \text{ kPa}$, is higher than that for the flat pad ($p_{\text{AVG}} = -1.0526 \text{ kPa}$). Note also that the presence of pad grooves increases the average slurry-film

thickness, and hence decreases the average flow resistance. Consequently, the calculated temporal mean of slurry flowrate for the grooved pad, $\langle q \rangle = \int_0^T q(x, t) dt / T = 0.140 \text{ cm}^2/\text{s}$, is much higher than the time-invariant slurry flowrate $q = 0.058 \text{ cm}^2/\text{s}$ for the flat pad [as can be readily deduced from (9), $\langle q \rangle$ does not depend on x].

The effects of varying the contact ratio ϕ , number of grooves under the wafer $N = 2a/\lambda$, and uncompressed groove depth d_{groove} on the slurry-flow characteristics have also been investigated. Basically, decreasing ϕ and increasing N (with all other parameters fixed) produce the same qualitative effect of reducing the land area on a grooved pad. (Note that a flat pad may be thought of as having $\phi = 1$ or $N = 0$.) So, for the same reason given above, it is observed that the (negative) average fluid pressure $\langle p_{\text{AVG}} \rangle$ increases in magnitude with decreasing ϕ and, equivalently, with increasing N . Moreover, the mean slurry flowrate $\langle q \rangle$ increases with decreasing ϕ and increasing N . In other words, increasing the number and/or width of the grooves on the pad would increase the slurry flowrate. As a larger flowrate would facilitate debris discharge, one therefore concludes that placing grooves on the pad is beneficial for the CMP process. Meanwhile, as the relatively small groove size does not affect the fluid pressure significantly, the value of $\langle p_{\text{AVG}} \rangle$ is rather insensitive to the uncompressed groove depth d_{groove} . The mean slurry flowrate $\langle q \rangle$, however, increases dramatically with d_{groove} . This makes sense, because (10) indicates that the volumetric slurry flowrate $q(x, t)$ has a cubic dependence on $h(x, t)$ and hence is much more sensitive to the groove depth variation (which dominates the slurry-film thickness variation) than the fluid pressure is.

4 Slurry-impurity transport and MRR calculations

To examine the effects of slurry chemistry on MRR, here we calculate the slurry-impurity concentration using a slurry-impurity transport equation—which accounts for the generation of impurity by local material removal, the convection of impurity by slurry flow, and the impurity filtration by the porous pad. Note that, since the slurry-impurity concentration generally is relatively low, the effects of diffusion on impurity transport is neglected here. In what follows, we shall first derive the slurry-impurity transport equation, and then use the experimental data of Thagella et al. [15] to calibrate the values of the parameters appearing in that equation.

4.1 Slurry-impurity transport equation

Let us denote the local MRR distribution by $\dot{R}(x, t)$; the temporal mean of spatially averaged MRR, $\langle \dot{R}_{\text{AVG}} \rangle$, then is a measure of the overall MRR. A simple model frequently used to estimate the overall MRR in CMP is the Preston equation [22]. It relates the overall MRR to the product of the wafer back pressure and the relative sliding speed between the wafer and pad in a linear fashion:

$$\langle \dot{R}_{\text{AVG}} \rangle = C_P \cdot p_{\text{back}} \cdot V, \quad (11)$$

where C_P is a constant that implicitly absorbs the dependence of $\langle \dot{R}_{\text{AVG}} \rangle$ upon the properties of consumables and various other process parameters. Although the Preston equation (11) has been used with some success, there also are situations in which it does not well correlate the experimental data, thus motivating several experimental and modeling works aiming to modify (11) (see [23–28] for example).

In particular, Runnels and Eyman [23, 24] postulated that the local MRR is a function of the local normal and tangential stresses acting on the wafer surface by the flowing slurry. As a universal functional form for the local MRR was (and still is) unavailable, Runnels [24] suggested some simple functional forms based on heuristic arguments. Using such forms, the erosion profiles of the wafer surface were calculated, and the results agreed reasonably well with experimental data. Meanwhile, combining Runnels' postulate [24] with some heuristic arguments regarding the material removal process, Tseng and Wang [25] derived an expression relating the overall MRR to $p_{\text{back}}^{5/6} V^{1/2}$, instead of the $p_{\text{back}} V$ dependence of the Preston equation (11). It is also worth mentioning that, to account for the deterioration in the shearing efficiency of abrasive particles, Tseng et al. [26] replaced the constant C_P in (11) by an exponentially decaying function of the sliding velocity V .

It appears that all the attempts to modify the Preston equation discussed above more or less had some empirical flavor, and none of them resulted in universal mathematical models for calculating MRR, reflecting the fact that the mechanism of material removal in CMP still is not very well understood. Also, as long as having a number of tuning parameters, all models seem to serve well the purpose of fitting experimental data obtained under particular process conditions. Since our purpose here is not to derive new models for MRR calculation, we shall simply use the Preston equation (11), and calibrate the relevant parameter values by comparing the numerical results with experimental data. Note however that (11) only gives the overall MRR, and to calculate the local MRR we shall replace the wafer back pressure on the right-hand side of (11) by the local contact stress, yielding

$$\dot{R}(x, t) = C \cdot \sigma(x, t) \cdot V. \quad (12)$$

Now, to incorporate slurry-chemistry effects into our model, Eq. (12) still needs further modification. To that end, let us consider the scenario that material removal is limited by chemical reaction of the wafer surface with the slurry; the mechanical interaction between the pad and wafer is much more efficient, so that chemically reacted layer on the wafer surface is immediately polished off. The removed material then increases the impurity concentration in the slurry, and reduces its chemical activity. In particular, it will be assumed that the slurry's chemical activity decays exponentially with its impurity concentration. The constant C in the local Preston equation (12) is then replaced by $C_0 \exp(-k\gamma)$, where the parameter k measures the sensitivity of the slurry's chemical activity to its impurity concentration $\gamma(x, t)$ —defined as the local volumetric fraction of the impurity present in the slurry. This results in a modified expression for calculating the local MRR:

$$\dot{R}(x, t) = C_0 e^{-k\gamma(x, t)} \cdot \sigma(x, t) \cdot V. \quad (13)$$

Of course, as the sensitivity parameter $k \rightarrow 0$, the local Preston equation (12) is recovered. Recall also that the constant C in (12) accounts for mechanical, chemical, and all other effects that are not explicitly modeled. With all other factors invariant, a different value of C may need to be used when the spatial distribution of slurry impurity concentration $\gamma(x, t)$ is substantially altered, in order for the model predictions to correlate well with experimental observations. By specifically using $C_0 \exp(-k\gamma)$ in place of C , we hope to capture some chemical effects on material removal, so as to render the new constant C_0 less dependent on local details of $\gamma(x, t)$. However, to provide sufficient justification of the above argument, it remains to be done in future work to collect more experimental data, and systematically compare them with the predictions of various models.

To proceed, we shall also assume that the porous polishing pad is capable of filtering the impurity out of the slurry. In particular, the filtration rate is assumed to be linearly proportional to the local impurity concentration. Putting things together, we may then write the slurry-impurity transport equation as follows:

$$\frac{\partial}{\partial t}(h\gamma) + \frac{\partial}{\partial x}(q\gamma) = C_0 e^{-k\gamma(x, t)} \cdot \sigma(x, t) \cdot V - C_1\gamma \quad (-a \leq x \leq a), \quad (14)$$

where the constant C_1 is a measure of the pad's filtration capability. The first and second terms on the left-hand side of (14) are the local and convective rates of change of the impurity, respectively.

Since $h(x, t)$, $q(x, t)$, and $\sigma(x, t)$ can be calculated independently, the only unknown in the transport equation (14) is the impurity concentration $\gamma(x, t)$. To calculate $\gamma(x, t)$, (14) is discretized using standard finite-difference formulae. It is also assumed that the slurry entering the pad-wafer interface is free of impurity, so that the boundary condition

$$\gamma(0, t) = 0 \quad (t \geq 0)$$

is imposed. In addition, the initial condition

$$\gamma(x, 0) = 0 \quad (-a \leq x \leq a)$$

is used to start the calculations, although it does not really affect the steady periodic impurity concentration and local MRR distributions.

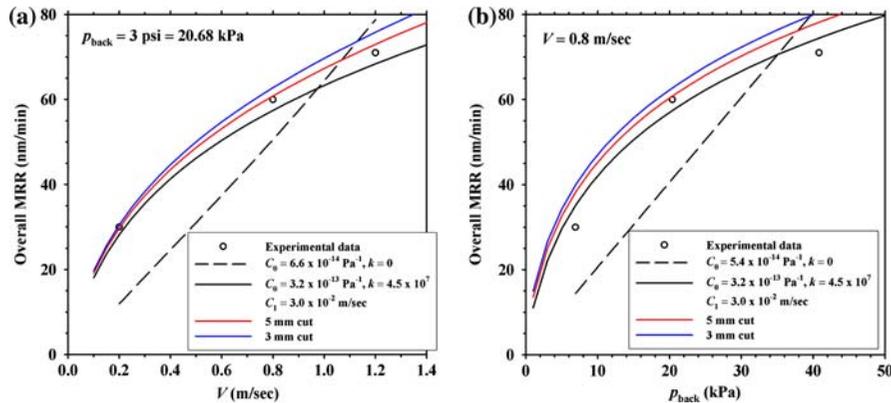


Fig. 7 Comparison of the calculated overall MRR for a flat pad, \dot{R}_{AVG} , with the experimental data of Thagella et al. [15]. Corresponding results for the slurry flow produced by the cut model are also included for comparison

4.2 Calibration of parameter values

Let us now compare the experimental data of Thagella et al. [15] (for silicon oxide polished on a flat IC1000-B4/SubaIV pad) with the numerical results of our model. The slurry used in their experiments was Klebesol 1501 (with $\text{pH} \approx 10.5$), and the volumetric flowrate was 75 ml/min. Also, the wafer-back pressure p_{back} was varied between 1 and 6 psi (or 6.9 and 41 kPa), and the relative sliding speed between the wafer and pad surfaces V was varied in the range of 0.2–1.2 m/s [15]. As shown in Fig. 7, the experimental data can be fitted with excellent accuracy by $C_0 = 3.2 \times 10^{-13} \text{ Pa}^{-1}$, $k = 4.5 \times 10^7$, and $C_1 = 3.0 \times 10^{-2} \text{ m/sec}$. In fact, since there are three parameters in our model, it is possible to pass the fitting curves in Figs. 7a and b through all of the three data points therein. The resulting parameter values, however, will be slightly different for Fig. 7a and b; so we fine-tune the parameters a little to render the fitting in both figures nearly equally good. The best fitted results (having minimized root-mean-squared errors) with $k = 0$ are also shown in Figs. 7a and b. (The MRR then does not depend on the impurity concentration, so that the filtration parameter C_1 becomes irrelevant.) It is clearly seen that incorporating slurry chemistry effects into our model allows for more free parameters, and hence greatly improves the overall fitting of the experimental data. As the back pressure and sliding speed ranges in Figs. 7a and b cover the particular combination of $p_{\text{back}} = 20 \text{ kPa}$ and $V = 0.43 \text{ m/s}$ that has been used in our calculations, the parameter values calibrated above will be used in all our subsequent calculations of slurry-impurity transport and MRR for grooved pads.

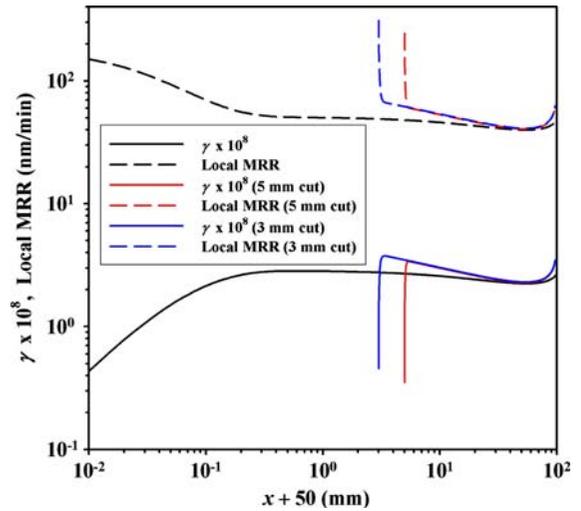
For the slurry flow produced by the cut model discussed in Sect. 3.1, we have also computed the corresponding overall MRR, using the same values of C_0 , k , and C_1 determined above; the results are included in Figs. 7a and b for comparison. It is seen that the overall MRR for the cut model generally is higher than that for the retaining-ring model. This observation will be interpreted in the following section, where we shall discuss the spatial distributions of the slurry-impurity concentration and local MRR. The quantitative difference between the two models, however, is not that large to warrant further fine-tuning of the parameter values, especially when our purpose here is to compare the numerical results of the two models.

5 Results of slurry-impurity concentration and MRR

5.1 Typical results for flat and grooved pads

First let us discuss the calculated slurry-impurity concentration and local MRR distributions, for both the flat and the grooved pads discussed in Sect. 3. The time-independent results for the flat pad are shown in Fig. 8. It is observed that the fresh slurry (with $\gamma = 0$) entering the pad-wafer interface is contaminated immediately by the material removed from the wafer surface, resulting in a sharp increase in the impurity concentration. (To resolve this sharp

Fig. 8 Spatial variations of the slurry-impurity concentration $\gamma(x)$ and local MRR $\dot{R}(x)$ for a flat pad, with $p_{\text{back}} = 20$ kPa and $V = 0.43$ m/s. Corresponding results for the slurry flow produced by the cut model are also included for comparison



increase, the spatial axis of Fig. 8 is shown on a logarithmic scale. Otherwise the curve for $\gamma(x)$ would display a sharp rise near the leading edge of the wafer, $x = -50$ mm, and look rather like the corresponding curves for the cut model included in Fig. 8.) At about 0.2 mm from the leading edge of the wafer, the impurity concentration rises to a level of about 2.6×10^{-8} , and varies only slightly after that. As a result, the local MRR also drops sharply from about 165 nm/min at the wafer’s leading edge to about 56 nm/min within a distance of 0.2 mm, and stays about the same after that. However, since the contact stress is larger near the exit of the interface region ($x = 50$ mm), the local MRR and slurry-impurity concentration there rise slightly. The average impurity concentration and MRR are calculated to be $\gamma_{\text{AVG}} = 2.38 \times 10^{-8}$ and $\dot{R}_{\text{AVG}} = 42.5$ nm/min, which will be compared with the corresponding values for the grooved pad below. One can also evaluate the spatial non-uniformity of MRR by comparing the standard deviation of the local MRR distribution with the overall MRR:

$$\text{NU} = \left(\frac{1}{2a} \int_{-a}^a \{ \dot{R}(x) - \dot{R}_{\text{AVG}} \}^2 dx \right)^{1/2} / \dot{R}_{\text{AVG}}. \tag{15}$$

For the present case, we calculate that $\text{NU} = 8.0\%$.

The slurry-impurity concentration and local MRR distributions for the slurry flow produced by the cut model are also plotted in Fig. 8 for comparison. Basically, the curves are qualitatively similar to those for the present model. However, as the contact stress near the wafer edges in the cut model is higher than that in the present model, the cut model produces a higher local MRR and hence a higher rising rate of the slurry-impurity concentration there. We have also calculated that, with a 5 mm edge cut, $\gamma_{\text{AVG}} = 2.50 \times 10^{-8}$, $\dot{R}_{\text{AVG}} = 44.8$ nm/min, and $\text{NU} = 12.6\%$. Meanwhile, with a 3 mm edge cut, $\gamma_{\text{AVG}} = 2.57 \times 10^{-8}$, $\dot{R}_{\text{AVG}} = 46.0$ nm/min, and $\text{NU} = 15.0\%$. It then transpires that the cut model generally produces a higher slurry-impurity concentration, overall MRR, and MRR non-uniformity. Also, the increases in such quantities become more significant as the boundary of the truncated computational domain is moved closer to the wafer edge. It is clear that such increases all result from the higher local contact stress near the wafer edges in the cut model. We thus see once again that the presence of a retaining ring improves the MRR uniformity, and that the present model is an improvement over the cut model. So, hereafter, only the results of the present retaining-ring model will be discussed.

Let us now examine the slurry-impurity concentration and local MRR distributions for the grooved pad (with $N = 40$, $\phi = 0.9$, and $d_{\text{groove}} = 200 \mu\text{m}$). In Figs. 9a and b, respectively, we plot the distributions of the slurry-impurity concentration, $\gamma(x, t)$, and local MRR, $\dot{R}(x, t)$, at $t = 0$ and $T/2$. Note that we have plotted the left and right halves of the figures using different scales for x , to show both the wafer-wide and groove-scale variations of $\gamma(x, t)$ and $\dot{R}(x, t)$ simultaneously. Like the results for the flat pad discussed above, here the slurry-impurity concentration increases—and the local MRR drops—sharply at the wafer’s leading edge. Looking more closely at the first few cycles in Figs. 9a and b, one can also see that, since the contact stress at points over a groove is zero, the local

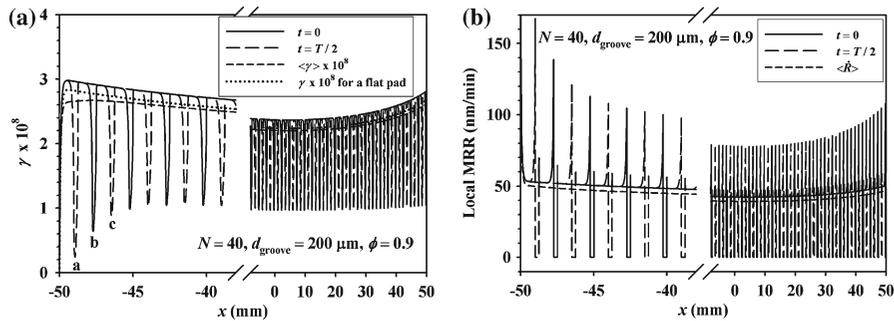


Fig. 9 Distributions of (a) slurry-impurity concentration, $\gamma(x, t)$, and (b) local MRR, $\dot{R}(x, t)$, at $t = 0$ and $T/2$. Here $N = 40$, $\phi = 0.9$, and $d_{\text{groove}} = 200 \mu\text{m}$. The temporal means of $\gamma(x, t)$ and $\dot{R}(x, t)$, i.e., $\langle \gamma \rangle$ and $\langle \dot{R} \rangle$, and the time-independent slurry-impurity concentration for a flat pad are also shown for comparison

MRR there is zero accordingly, and hence the slurry-impurity concentration over a groove is substantially lower than that on the land area. Meanwhile, from a Lagrangian point of view, as a particular parcel of the relatively purer slurry in a groove (e.g., at the groove location marked by “a” in Fig. 9a) is convected downstream, it passes over the wafer surface and picks the removed material, so that its impurity concentration increases. So, we see in Fig. 9a that the impurity concentration at the groove location “b” is higher than that at “a”, and that at the groove location “c” is even higher. Moreover, although the local MRR is zero over a groove, the substantially lower slurry-impurity concentration there results in much higher local MRR at the edges of the land area than that further inland, so that the local MRR distributions shown in Fig. 9b exhibit sharp increases at the edges of the land area before dropping to zero over a groove.

The temporal means of the slurry-impurity concentration and local MRR, $\langle \gamma \rangle$ and $\langle \dot{R} \rangle$, are also shown in Figs. 9a and b as functions of x . Because each point on the wafer surface slides over both the groove and the land area in a groove pitch during one period T , the temporal averaging removes the groove-scale variations of $\gamma(x, t)$ and $\dot{R}(x, t)$. It is seen that the resulting wafer-wide variations qualitatively are the same as that for the flat pad (the slurry impurity concentration distribution for the flat pad is duplicated in Fig. 9a for comparison). Note that, since the grooved pad suffers a higher contact stress—and hence produces a higher local MRR—on its land area, the local slurry-impurity concentration on the land area is higher than that on the flat pad (see Fig. 9a). Due to the decreased impurity concentration in the grooves, however, the temporally averaged impurity concentration for the grooved pad still is lower than the time-independent impurity concentration for the flat pad. Meanwhile, the spatial averages of $\langle \gamma \rangle$ and $\langle \dot{R} \rangle$ are calculated to be $\langle \gamma_{\text{AVG}} \rangle = 2.33 \times 10^{-8}$ and $\langle \dot{R}_{\text{AVG}} \rangle = 41.7 \text{ nm/min}$. Comparison of such values with the corresponding values for the flat pad ($\gamma_{\text{AVG}} = 2.38 \times 10^{-8}$ and $\dot{R}_{\text{AVG}} = 42.5 \text{ nm/min}$) suggests that the presence of pad grooves slightly decreases the overall MRR, resulting in a slightly decreased average slurry-impurity concentration. A qualitative explanation for this will be offered in the next subsection.

From the temporally averaged local MRR distribution, $\langle \dot{R} \rangle$, we can also calculate the spatial non-uniformity of MRR, using the following expression that generalizes the definition (15) for flat pads:

$$\text{NU} = \left(\frac{1}{2a} \int_{-a}^a \{ \langle \dot{R} \rangle(x) - \langle \dot{R}_{\text{AVG}} \rangle \}^2 dx \right)^{1/2} / \langle \dot{R}_{\text{AVG}} \rangle. \quad (16)$$

We then calculate that $\text{NU} = 7.9\%$ for the grooved pad, which is slightly lower than the 8.0% NU value for the flat pad.

5.2 Effects of pad groove parameters on the average slurry-impurity concentration and MRR

It was observed above that the presence of pad grooves decreases the average slurry-impurity concentration, the overall MRR, and the MRR non-uniformity simultaneously. To see how the pad-groove parameters affect the aforementioned performance indices of CMP, the mean values of the spatially averaged slurry-impurity concentration and MRR (i.e., $\langle \gamma_{\text{AVG}} \rangle$ and $\langle \dot{R}_{\text{AVG}} \rangle$), and the MRR non-uniformity (NU), are calculated for the uncompressed groove

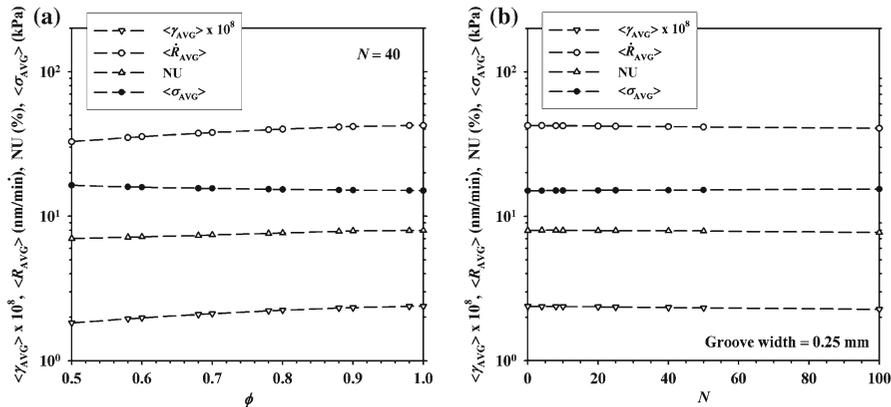


Fig. 10 Mean values of the spatially averaged contact stress, slurry-impurity concentration and MRR (i.e., $\langle \sigma_{AVG} \rangle$, $\langle \gamma_{AVG} \rangle$ and $\langle \dot{R}_{AVG} \rangle$), and the MRR non-uniformity (NU), for the uncompressed groove depth $d_{groove} = 200 \mu\text{m}$. (a) Dependence of the aforementioned quantities on the contact ratio ϕ for $N = 40$. (b) Dependence of such quantities on the groove number N for groove width $(1 - \phi)\lambda = 0.25 \text{ mm}$

depth $d_{groove} = 200 \mu\text{m}$. In Fig. 10a, the results are plotted against the contact ratio ϕ , with the groove number kept constant at $N = 40$. Meanwhile, in Fig. 10b the groove width $w = (1 - \phi)\lambda$ is fixed at 0.25 mm, and the groove number N is being varied. Recall that decreasing ϕ and increasing N produce the same qualitative effect of reducing the land area on a grooved pad.

It is seen in Fig. 10a that as ϕ decreases—and equivalently, as N increases in Fig. 10b, both $\langle \gamma_{AVG} \rangle$ and $\langle \dot{R}_{AVG} \rangle$ decrease slightly. Note, however, that when slurry chemistry is not considered (i.e., with $k = 0$), the presence of pad grooves would increase the average contact stress $\langle \sigma_{AVG} \rangle$ (see Figs. 10a and b), and hence the overall MRR. Here we attempt to offer a qualitative explanation for the opposite trends exhibited by the results with slurry chemistry modeled and that without: recall that, compared with a flat pad, a grooved pad has higher contact stress on its land area, but less contact area for effective polishing of the wafer surface. For $k = 0$, the increased contact stress has a slightly stronger effect on the overall MRR (which is proportional to $\langle \sigma_{AVG} \rangle$) than the reduced contact area does, so that the overall MRR is increased by the presence of pad grooves. Now, when the dependence of the slurry’s chemical activity on its impurity concentration is taken into consideration, since the impurity concentration on the land area of a grooved pad is higher than that on a flat pad (see Fig. 9a), the increased contact stress no longer increases the local MRR enough to compensate for the reduced contact area of the pad and wafer surfaces, and hence the overall MRR is reduced by the pad grooves’ presence.

Quantitatively, $\langle \gamma_{AVG} \rangle$ and $\langle \dot{R}_{AVG} \rangle$ are more sensitive to the contact ratio ϕ than to the groove number N , as can be seen by comparing Figs. 10a and b. Meanwhile, compared with $\langle \gamma_{AVG} \rangle$ and $\langle \dot{R}_{AVG} \rangle$, the MRR non-uniformity (NU) is less sensitive to the pad groove parameters (see Figs. 10a and b). This may be because the contact-stress non-uniformity—which makes a major contribution to the MRR non-uniformity—is not affected significantly by the pad grooves’ presence, although the average contact stress is. We have also calculated $\langle \gamma_{AVG} \rangle$, $\langle \dot{R}_{AVG} \rangle$, and NU for other uncompressed groove depths; the results turn out to be relatively insensitive to d_{groove} , and therefore are not presented here. Based upon the results discussed above, it may be concluded that the presence of pad grooves increases the slurry flowrate, thereby reducing the average slurry-impurity concentration. Depending upon the sensitivity of the slurry’s chemical activity to its impurity concentration (characterized by the parameter k), however, the average MRR may or may not be increased by the pad grooves’ presence.

5.3 Effects of parameters k , C_0 and C_1 on the average slurry impurity concentration and MRR

For $N = 40$, $\phi = 0.9$, and $d_{groove} = 200 \mu\text{m}$, we have also calculated the dependence of $\langle \gamma_{AVG} \rangle$, $\langle \dot{R}_{AVG} \rangle$, and NU on the parameters k , C_0 and C_1 . The results are not presented here to save some space, but a few observations are described qualitatively below.

First, it is found that both $\langle \gamma_{\text{AVG}} \rangle$ and $\langle \dot{R}_{\text{AVG}} \rangle$ decrease with increasing value of k , while NU does the same but to a much less extent. This trend is easy to understand indeed: Recall that with a larger value of k , the local MRR would decrease more rapidly with the slurry-impurity concentration ϕ , through the assumed exponential dependence of local MRR on the product $k\gamma$ in (13). As a result, the average MRR is reduced, thereby reducing the average slurry-impurity concentration as well. It is also intuitively clear that reducing the parameter C_0 would decrease the local MRR, resulting in reduced average slurry-impurity concentration and MRR. Meanwhile, as the value of parameter C_1 is a measure of the filtration (cleaning) capability of the porous pad, increasing the C_1 value clearly would reduce the slurry-impurity concentration. The reduced slurry-impurity concentration, in turn, increases the local MRR and results in increased average MRR. The MRR non-uniformity (NU) turns out to be less sensitive to the values of the parameters k , C_0 and C_1 than the average impurity concentration and MRR are. This is similar to the observation in Fig. 10 that NU is less sensitive to the pad groove parameters than $\langle \gamma_{\text{AVG}} \rangle$ and $\langle \dot{R}_{\text{AVG}} \rangle$ are, and the interpretation offered there should remain valid here as well.

Recall that here the baseline values of parameters k , C_0 and C_1 were determined by fitting the numerical results of our model with the experimental data of Thagella et al. [15] (see Fig. 7). The qualitative effects of k , C_0 and C_1 on the average slurry-impurity concentration and MRR outlined above were observed in the neighborhood of the baseline parameter setting in the (k, C_0, C_1) parameter space. In the absence of sufficient relevant experimental data, we could not validate the predictions of our model at this point. However, it would be interesting and extremely important to do so in a future work.

6 Concluding remarks

In this paper, we have extended our previous work [13] on the modeling and calculation of slurry flow in CMP with a grooved pad. The specific modifications here are the inclusion of a retaining ring in the geometrical setup and the consideration of slurry chemistry effects on MRR through a simple model. A lot more slurry-flow features and CMP performance indices—such as the spatial average and non-uniformity of MRR—therefore can be calculated and carefully examined. As demonstrated in Sect. 3.1 for flat pads, a retaining ring typically used to hold the wafer in place also reduces the stress concentration near the wafer edges, and hence helps improve the wafer-wide MRR uniformity.

We have also shown that the local MRR calculation model (13) proposed here enables us to correlate the experimental data better in the parameter space of interest (see Fig. 7). Despite the fact that some parameters of the present model still need to be tuned, it is expected that once the parameter values are calibrated by comparing the numerical predictions with relevant experimental data, our theoretical models and numerical methods can serve as an efficient tool for calculating the effects of various process parameters on the average MRR, MRR non-uniformity, and many other performance indices that are difficult or expensive to evaluate in practice. In this regard, the simulation methodology devised in this paper would be useful to the semiconductor manufacturing industry.

Using the simulation tool developed here, and varying the process parameters systematically, we have made a number of interesting observations. Specifically, it is found that the presence of pad grooves generally increases the contact stress on the land area of the pad, and increases the slurry flowrate at the pad–wafer interface. When slurry chemistry effects are not considered, the increased contact stress implies increased local MRR. Meanwhile, although a grooved pad also has less contact area for effective interaction with the pad, the overall MRR is increased as it turns out. The picture becomes different, however, when slurry-chemistry effects are taken into account. In fact, the numerical results suggest that the overall MRR then is decreased by the pad grooves' presence. This reflects the extreme sensitivity of CMP results to all of the process parameters, and reiterates the importance of having an efficient simulation tool to aid in selecting the optimal process parameters. Moreover, the numerical results suggest that the MRR uniformity is mostly determined by the contact-stress uniformity, which can be improved by use of a sufficiently large retaining ring.

In principle at least, one could calculate the slurry flow, contact mechanics, and chemical effects using relevant fundamental principles, thereby eliminating the need for parameter tuning completely. To really simulate the real

situation, a truly 3-D geometrical setup allowing for arbitrary groove pattern on the pad also will be needed. However, incorporating all such aspects in the numerical computations clearly will be extremely expensive, if it ever is a sensible thing to do. Our purpose in the previous work [13] and the present one therefore is to develop simpler models that still retain most of the relevant physics and chemistry and are capable of providing useful interpretations for results observed in real CMP processes. It is certainly true that the present model still is rather oversimplified (for one thing, the 2-D assumption excludes the possibility of considering more complicated groove patterns). As pointed out in the introduction, however, here we have started with the simplest possible models, and in future work we hope to incorporate more geometrical, physical, and chemical effects into the overall model—up to an extent that only modest computational resources are required. Note also that, despite the semi-empirical nature of our present model (and most other models in the literature), and the need to calibrate the parameter values, the results discussed above do demonstrate qualitative trends that appear to be intuitively correct, and therefore can be used for providing valuable insights into the optimization of CMP process parameters.

Acknowledgements The authors gratefully acknowledge generous support of this work by the Taiwan Semiconductor Manufacturing Company (TSMC) through grants NCKU-0403, NCKU-0603 and NCKU-0704. They would also like to thank Drs. S.-M. Cheng, T.-C. Tseng, Y.-W. Chou, and J.-C. Sheu of TSMC, and Prof. K.-S. Chen of NCKU, for a number of fruitful discussions on this work and other related topics. Three anonymous reviewers of this paper provided useful comments for improving the presentation of this work, which the authors greatly appreciate.

References

1. Singh RK, Bajaj R (2002) Advances in chemical-mechanical planarization. *MRS Bull* 27:743–747
2. Patrick WJ, Guthrie WL, Standley CL, Schiabe PM (1991) Application of chemical mechanical polishing to the fabrication of VLSI circuit interconnections. *J Electrochem Soc* 138:1778–1784
3. Singh RK, Lee S-M, Choi K-S, Basim GB, Choi W, Chen Z, Moudgil BM (2002) Fundamentals of slurry design for CMP of metal and dielectric materials. *MRS Bull* 27:752–760
4. Moinpour M, Tregub A, Oehler A, Cadien K (2002) Advances in characterization of CMP consumables. *MRS Bull* 27:766–771
5. Doy TK, Seshimo K, Suzuki K, Philipossian A, Kinoshita M (2004) Impact of novel pad groove designs on removal rate and uniformity of dielectric and copper CMP. *J Electrochem Soc* 151:G196–G199
6. Rosales-Yeomans D, Doi T, Kinoshita M, Suzuki T, Philipossian A (2005) Effect of pad groove designs on the frictional and removal rate characteristics of ILD CMP. *J Electrochem Soc* 152:G62–G67
7. Muldowney GP, Tselepidakis DP (2004) A computational study of slurry flow in grooved CMP polishing pads. In: Proceedings of the 2004 CMP-MIC Conference, Marina Del Ray, California, 24–26 February 2004, p 022504
8. Muldowney GP, James DB (2004) Characterization of CMP pad surface texture and pad–wafer contact. *Mat Res Soc Symp Proc* 816:147–158
9. Eaton JK, Elkins CJ, Burton TM, Miyaji A, Coon DP (2003) Simulation method for CMP slurry flow with a grooved polishing pad. In: Proceedings of the 2003 CMP-MIC Conference, Marina Del Ray, California, 19–21 February 2003, pp 143–150
10. Tichy J, Levert JA, Shan L, Danyluk S (1999) Contact mechanics and lubrication hydrodynamics of chemical mechanical polishing. *J Electrochem Soc* 146:1523–1528
11. Shan L, Levert J, Meade L, Tichy J, Danyluk S (2000) Interfacial fluid mechanics and pressure prediction in chemical mechanical planarization. *ASME J Tribol* 122:539–543
12. Jin X, Keer LM, Wang Q (2005) A 3D simulation of CMP. *J Electrochem Soc* 152:G7–G15
13. Wang Y-C, Yang T-S (2007) Effects of pad grooves on chemical mechanical planarization. *J Electrochem Soc* 154: H486–H494
14. Gitis NV, Xiao J, Kumar A, Sikder AK (2004) Advanced specification and tests of CMP retaining rings. In: Proceedings of the 2004 CMP-MIC Conference, Marina Del Ray, California, 24–26 February 2004, pp 252–255
15. Thagella S, Sikder AK, Kumar A (2004) Tribological issues and modeling of removal rate of low- k films in CMP. *J Electrochem Soc* 151:G205–G215
16. Yen J-L, Chen Y-C, Chen K-S, Yang T-S (2006) Contact stress uniformity analysis for chemical mechanical polishing application (in Chinese). In: Proceedings of the 23rd National Conference on mechanical engineering, Yung Kang, Taiwan, 24–25 November 2006, p C2-023
17. Johnson KL (1985) Contact mechanics. Cambridge University Press, Cambridge, UK
18. Borucki LJ, Ng SH, Danyluk S (2005) Fluid pressures and pad topography in chemical mechanical planarization. *J Electrochem Soc* 152:G391–G397
19. Greenwood JA, Williamson JBP (1966) Contact of nominally flat surfaces. *Proc R Soc London Ser A* 295:300–319
20. Hamrock BJ, Schmid SR, Jacobson BO (2004) Fundamentals of fluid film lubrication, 2nd edn. Marcel Dekker, Inc, New York

21. Patir N, Cheng HS (1978) An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication. *ASME J Lubr Technol* 100:12–17
22. Preston FW (1927) The theory and design of plate glass polishing machines. *J Soc Glass Technol* 11:214–256
23. Runnels SR, Eyman M (1994) Tribological analysis of chemical-mechanical polishing. *J Electrochem Soc* 141:1698–1701
24. Runnels SR (1994) Feature-scale fluid-based erosion modeling for chemical-mechanical polishing. *J Electrochem Soc* 141:1900–1904
25. Tseng W-T, Wang Y-L (1997) Re-examination of pressure and speed dependences of removal rate during chemical-mechanical polishing processes. *J Electrochem Soc* 144:L15–L17
26. Tseng W-T, Chin J-H, Kang L-C (1999) A comparative study on the roles of velocity in the material removal rate during chemical mechanical polishing. *J Electrochem Soc* 146:1952–1959
27. Homma Y, Fukushima K, Kondo S, Sakuma N (2003) Effects of mechanical parameters on CMP characteristics analyzed by two-dimensional frictional-force measurement. *J Electrochem Soc* 150:G751–G757
28. Homma Y (2006) Dynamical mechanism of chemical mechanical polishing analyzed to correct Preston's empirical model. *J Electrochem Soc* 153:G587–G590