Journal of ELECTRONIC MATERIALS, Vol. 42, No. 1, 2013
DOI: 10.1007/s11664-012-2250-z
© 2012 TMS

Impact of Pad Conditioning on Thickness Profile Control in Chemical Mechanical Planarization

S. KINCAL[1] and G.B. BASIM[2,3,4]

1.—Department of Chemical Engineering, Middle East Technical University, Ankara, Turkey. 2.—Department of Mechanical Engineering, Ozyegin University, Istanbul 34662, Turkey. 3.—Texas Instruments Incorporated, Dallas, TX 75265, USA. 4.—e-mail: Bahar.Basim@ozyegin.edu.tr

Chemical mechanical planarization (CMP) has been proven to be the best method to achieve within-wafer and within-die uniformity for multilevel metallization. Decreasing device dimensions and increasing wafer sizes continuously demand better planarization, which necessitates better understanding of all the variables of the CMP process. A recently highlighted critical factor, pad conditioning, affects the pad surface profile and consequently the wafer profile; in addition, it reduces defects by refreshing the pad surface during polishing. This work demonstrates the changes in the postpolish wafer profile as a function of pad wear. It also introduces a wafer material removal rate profile model based on the locally relevant Preston equation by estimating the pad thickness profile as a function of polishing time. The result is a dynamic predictor of how the wafer removal rate profile shifts as the pad ages. The model helps fine-tune the pad conditioner operating characteristics without the requirement for costly and lengthy experiments. The accuracy of the model is demonstrated by experiments as well as data from a real production line. Both experimental data and simulations indicate that the smaller conditioning disk size and extended conditioning sweep range help improve the post-CMP wafer planarization. However, the defectivity tends to increase when the conditioning disk sweeps out of the pad radius; hence, the pad conditioning needs to be designed by considering the specific requirements of the CMP process conducted. The presented model predicts the process outcomes without requiring detailed experimentation.

Key words: Chemical mechanical planarization (CMP), conditioning, pad profile modeling, defectivity

INTRODUCTION

Chemical mechanical planarization (CMP) has been widely adopted as the method of choice in semiconductor manufacturing to provide a planar wafer surface for subsequent photolithography steps. Decreasing critical dimensions (CDs) of microelectronic devices with increasing wafer sizes continuously require better control of the CMP process variability to achieve a desired metal or dielectric film thickness and to enable both within-wafer (WIW) and within-die (WID) thickness consistencies of the planarized layers.[1] Irregularities in film thickness are driven by local and global density differences on the wafers, resulting in local CD variations that reduce the specified CD budget as the dimensions become smaller.[2] These variations make the inline process control difficult and may be responsible for end-of-line failures that result in defective devices and loss of revenue.

In addition to the incoming pattern layout and the deposited film thickness profile variations, the operating conditions during the CMP process also affect the postpolish thickness profiles on the

(Received April 17, 2012; accepted August 28, 2012; published online October 2, 2012)

wafers. Since the incoming factors are generally better controlled than the CMP-related process variability, there are standard methods implemented in process flows, such as adding dummy structure fields to even out the pattern density distributions.[3] The main difficulty in mitigating the CMP process variations is that they are often consumable-age dependent. A number of wafer-to-wafer[4–7] and run-to-run[5,6] controllers have been outlined in the literature, but these schemes address wafer- or lot-level average thickness control challenges and do not account for the shifts in removal rate (RR) profiles due to the aging of consumables. To account for the shifts in RR profiles, one needs more complex polishing equipment (i.e., tools with multizone polishing heads) that can employ multivariable control strategies to adjust the pressure distribution within multiple zones to achieve an optimum thickness range.[8,9] However, such methods also have limitations, particularly near the very edge of the wafer and in regard to the complexity of both the implementation and the equipment maintenance.

This work focuses on the optimization of the conditioner sweep configuration with the objective of optimizing the pad thickness profile to obtain a desired RR profile. Pad conditioning serves three main functions in CMP: (i) prevent glazing of the pad surface by inducing surface roughness through abrading the polymeric pad material with diamonds embedded in the conditioner surface, (ii) reduce defectivity by cleaning the pad fibers from the polishing residue, and (iii) maintain a stable material removal rate by distributing the polishing slurry onto the pad surface and throughout the pad radius.[10,11] There have been studies focusing on modeling and improving the pad conditioning in an effort to understand how different pad materials respond to conditioning[12] and how the conditioning changes the pad surface microstructure.[13] Furthermore, there have been modeling studies both aimed toward understanding how the pad profile changes during pad conditioning via characterization of the changes in pad surface morphology[13] and using the perspective of a single diamond abrasive trajectory on the polishing disk.[14] In this investigation, pad wear due to conditioning is studied experimentally by using three different conditioning sweeps and evaluating the impact on postpolish wafer profiles for shallow trench isolation (STI) CMP. Additionally, the defect performance evaluations are presented. Moreover, a physical model that can predict the postpolish wafer profile through pad wear characterization is introduced. The model is also evaluated against the experimental results; accordingly, an optimized pad conditioner parameter is presented. The final section of the paper focuses on implementation of the developed model in a high-volume manufacturing (HVM) environment along with a demonstration of its benefits.

MATERIALS AND METHODS

Experimental Procedures

CMP experiments were carried out on an AMAT Mirra 3400 model 200-mm platform tool using a silica-based slurry (SS25, supplied by Cabot Microelectronics Corporation) with a high-selectivity additive on blanket tetraethylorthosilicate (TEOS) wafers and plasma-enhanced chemical vapor deposition (PECVD) oxide-deposited patterned (STI level, 130-nm node) wafers. The slurry flow was maintained at 140 mL/min on the platen. The wafer polishing was conducted on a single platen to observe the change in wafer profiles independent of platen variations. IC1000/Suba IV stacked polishing pads were used combined with TBW Industries Grid-Abrade perforated 4-inch diamond pad conditioner disks. A 2-inch TBW disk was also used to demonstrate the effect of disk diameter on the variations of the pad profile. Figure 1 shows a schematic of the polishing setup.

Three different types of conditioner sweeps were evaluated in the experiments: (i) a commonly used sinusoidal sweep (SS), (ii) custom sweep 1 (CS1) in which the sweep is adjusted so that the conditioner disk spends a fixed time in each zone, and (iii) custom sweep 2 (CS2) in which the conditioner disk spends an equal amount of time per unit pad area; i.e., the residence time per zone is proportional to the area of the zone, so that the disk spends more time at the outer radius of the pad, where the total area is larger, and less time close to the center, where the total area is smaller. Figure 2 illustrates the time trends of the conditioner displacements for the three types of conditioner sweeps. The frequency of the conditioner sweeps was 10 times per minute for 10 zones. The conditioner was confined to the surface area of the pad to prevent diamond loss

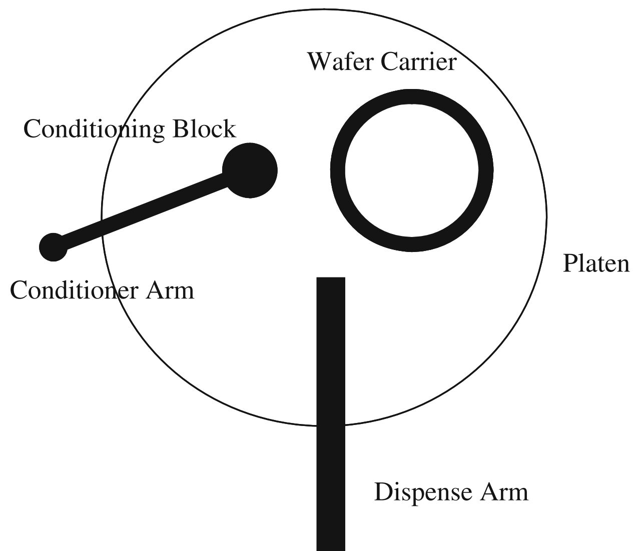

Fig. 1. Schematic representation of a CMP process showing the relative locations of the wafer, pad, conditioner, and slurry feeder.

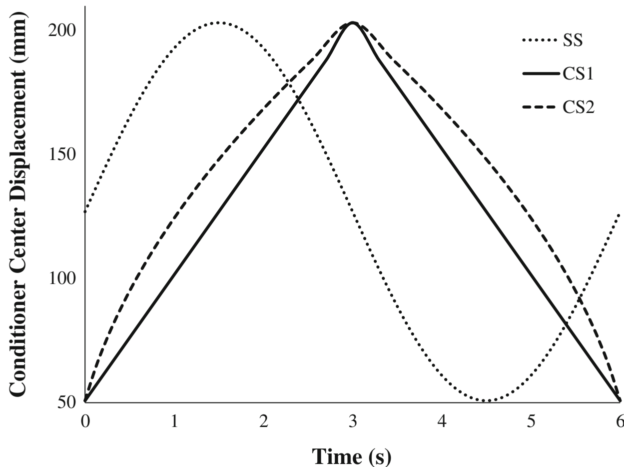

Fig. 2. Conditioner sweep profiles for the sinusoidal sweep (SS), custom sweep 1 (CS1), and custom sweep 2 (CS2).

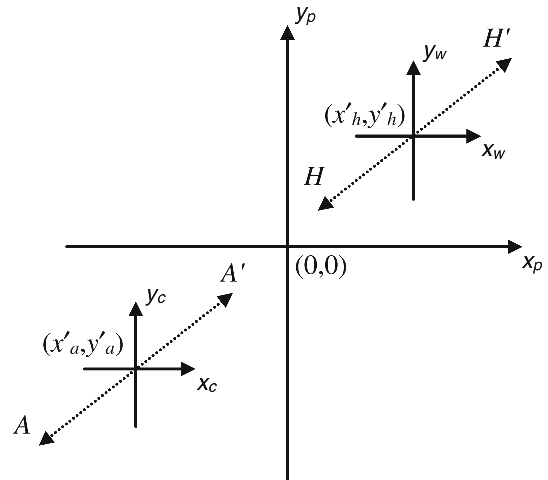

Fig. 3. The coordinate systems used to develop the pad wear model.

and edge pad wear, which may result in more defects. The conditioner downforce was set to 5 psi, and the polish head downforce was set to 6 psi. The polish head was oscillated with a sinusoidal profile at 10 sweeps per minute. The platen rotation was set to 85 rpm, while the head rotation was maintained at 100 rpm.

A set of three unpatterned TEOS wafers with a blanket oxide thickness of 18,000 Å and patterned wafers with a PECVD oxide thickness of 6200 Å were polished on new and aged pads after 1 h of conditioning with the conditioner sweep profiles mentioned above. All wafers were characterized for thickness, pre- and postpolish, to quantify the removal rate profile and the defectivity performance. The defectivity analyses were conducted with a KLA TENCOR SFS SP1 tool for the blanket oxide wafers and with a KLA 213X tool for the patterned wafers. The thickness measurements were collected using a KLA TENCOR ASET F5. The polishing pads were removed and cut in half, and then the Suba IV subpad was removed. The pads were analyzed using Vernier caliper measurements through the radius of the pads to establish postpolish profiles. The experimental results were simulated based on the model developed, and the simulation results were compared with experimental observations.

Model

Pad conditioning is a critical component of the CMP process. Its purpose is to clear out the polishing byproducts from the holes and grooves of the pad while continuously regenerating the surface of the pad to maintain a uniform polishing rate. In doing so, some pad material is removed during each pass of the conditioner. The absolute amount of pad material lost depends on various factors, including the diamond grain shape and protrusion, the rotation rate of the conditioning disk, and the pressure applied to the back of the disk. These factors are determined based on the specific requirements of the film being polished. However, the radial dependence of the wear rate depends on the sweep profile and the geometry of the conditioner. The model that is introduced in this study relates the wafer-level parameters to the aforementioned conditioner characteristics.

The model is based on the locally relevant Preston equation that relates the removal rate (RR) to the pressure (P) and the velocity (V) through the Preston constant, K[12], all of which can be expressed as a function of position in the most general form of Eq. 1.

$$\mathrm{RR}(r) = \frac{\mathrm{d}h(r)}{\mathrm{d}t} = K(r)P(r)V(r). \tag{1}$$

Earlier modeling studies outline the application of this expression to the wafer-polishing pad subsystem[10,11] and the conditioner pad subsystem to optimize the conditioner operation.[15,16] The model in this study combines the pad wear rate and the resulting radial pad profile into the Preston equation for the wafer–pad system using a modified Preston constant.[17] The constant takes into account the pad thickness and what the wafer is exposed to at any point at a given time.

In the model, there are three separate coordinate systems, as shown in Fig. 3. The absolute origin (0, 0) is located at the center of the platen coordinate system, denoted by (x_p, y_p). The polish head that supports the wafer is denoted by the coordinate system (x_w, y_w) and rotates about the point $(x'_\mathrm{h}, y'_\mathrm{h})$, which is a function of time as defined by the carrier sweep. The arm that moves the conditioner block is similarly represented by the coordinate system (x_c, y_c), where the conditioner rotates about a point $(x'_\mathrm{a}, y'_\mathrm{a})$, which is a function of time as defined by the conditioning arm sweep. The system is a 200-mm Mirra platform on which the head generally sweeps back and forth along HH' and the conditioner arm sweeps along AA'. However, the equations

developed for this model do not impose any restriction on, but are generalized to fit, any profile that can be defined in time. They will be introduced in two parts as (i) the pad–conditioner system and (ii) the wafer–pad system.

(i) *Pad–Conditioner System*:

To characterize the pad wear rate via the conditioner movement as a function of pad radius, the Preston equation is used in the following form of Eq. 2 for the pad–conditioner system:

$$\mathrm{RR_p}(r_\mathrm{p}) = \frac{\mathrm{d}h_\mathrm{p}(r_\mathrm{p})}{\mathrm{d}t} = K_\mathrm{pc} P_\mathrm{c} V_\mathrm{pc}(t; r_\mathrm{p}), \tag{2}$$

where $\mathrm{RR_p}(r_\mathrm{p})$ is the removal rate of pad material by the conditioner, $h_\mathrm{p}(r_\mathrm{p})$ is the height of the remaining pad material, K_pc is the Preston constant between the pad and the conditioner, P_c is the pressure applied on the back of the conditioner, and $V_\mathrm{pc}(t; r_\mathrm{p})$ is the relative velocity of the conditioner with respect to the pad.

The major assumption in quantifying the pad wear is that the majority of the pad material is abraded by the conditioner, while the pad material lost due to the impact between the polishing head and the wafer is negligible. Therefore, the integration of Eq. 2 between two points in time with the appropriate initial conditions will determine the pad thickness at any given radius, as expressed in Eq. 3.

$$h_\mathrm{p}(r_\mathrm{p}) = K_\mathrm{pc} P_\mathrm{c} \int_{t_0}^{t_\mathrm{f}} V_\mathrm{pc}(t; r_\mathrm{p}) \mathrm{d}t. \tag{3}$$

Equation 3 can be evaluated at any number of desired radial points on the pad (more points result in higher resolution at the expense of computational power). When the constants are taken out of the integration operation, the remaining part relates the time-dependent relative velocity to the operational parameters of the polisher.

The relative velocity can be expressed as a combination of the x-y components of the velocity vector as shown in Eq. 4:

$$V_\mathrm{pc}(t; r_\mathrm{p}) = \sqrt{V_\mathrm{pc}^{x2}(t; r_\mathrm{p}) + V_\mathrm{pc}^{y2}(t; r_\mathrm{p})}. \tag{4}$$

$V_\mathrm{pc}^x(t; r_\mathrm{p})$ is the x-component of the relative velocity of the conditioner with respect to the pad, and $V_\mathrm{pc}^y(t; r_\mathrm{p})$ is the y-component of the relative velocity of the conditioner with respect to the pad.

The individual velocity terms can be further broken down to expressions (5) and (6):

$$V_\mathrm{pc}^x(t; r_\mathrm{p}) = \omega_\mathrm{p} y_\mathrm{p} - \omega_\mathrm{c} y_\mathrm{c}, \tag{5}$$

$$V_\mathrm{pc}^y(t; r_\mathrm{p}) = \omega_\mathrm{p} x_\mathrm{p} - \omega_\mathrm{c} x_\mathrm{c}. \tag{6}$$

ω_c is the rotational speed of the conditioner, ω_p is the rotational speed of the pad, x_p, y_p is the pad coordinate system, and x_c, y_c is the conditioner block coordinate system.

By using the following linear coordinate transformation,

$$y_\mathrm{c} = y_\mathrm{p} - y'_\mathrm{a}(t) \tag{7}$$

$$x_\mathrm{c} = x_\mathrm{p} - x'_\mathrm{a}(t) \tag{8}$$

$x'_\mathrm{a}(t), y'_\mathrm{a}(t)$ are the center coordinates of the conditioner block in the pad coordinate system the individual velocity terms appearing in Eqs. 5 and 6 are transposed to the pad coordinates, as shown in Eqs. 9 and 10.

$$V_\mathrm{pc}^x(t; r_\mathrm{p}) = (\omega_\mathrm{p} - \omega_\mathrm{c}) y_\mathrm{p} + \omega_\mathrm{c} y'_\mathrm{a}(t), \tag{9}$$

$$V_\mathrm{pc}^y(t; r_\mathrm{p}) = (\omega_\mathrm{p} - \omega_\mathrm{c}) x_\mathrm{p} + \omega_\mathrm{c} x'_\mathrm{a}(t). \tag{10}$$

Before inserting Eqs. 9 and 10 back into Eq. 3, the following coordinate system transformation is carried out:

$$x_\mathrm{p} = r_\mathrm{p} \cos(\theta_\mathrm{p}) = r_\mathrm{p} \cos(\omega_\mathrm{p} t), \tag{11}$$

$$y_\mathrm{p} = r_\mathrm{p} \sin(\theta_\mathrm{p}) = r_\mathrm{p} \sin(\omega_\mathrm{p} t), \tag{12}$$

where r_p, θ_p are the polar coordinates of the pad system.

The final form of Eq. 4 becomes

$$V_\mathrm{pc}(t; r_\mathrm{p}) = \sqrt{\left((\omega_\mathrm{p} - \omega_\mathrm{c}) r_\mathrm{p} \sin(\omega_\mathrm{p} t) + \omega_\mathrm{c} y'_\mathrm{a}(t)\right)^2 + \left((\omega_\mathrm{p} - \omega_\mathrm{c}) r_\mathrm{p} \cos(\omega_\mathrm{p} t) + \omega_\mathrm{c} x'_\mathrm{a}(t)\right)^2}. \tag{13}$$

Equation 13 can then be inserted into Eq. 3 and then integrated to determine the pad thickness profile.

(ii) *Wafer–Pad System*:

Similar to the pad–conditioner system, the Preston equation is expressed as shown in Eq. 14 to model the wafer–pad system.

$$\mathrm{RR_w} = \frac{\mathrm{d}h_\mathrm{w}(r_\mathrm{w})}{\mathrm{d}t} = K_\mathrm{pw}(t; r_\mathrm{w}) P_\mathrm{w}(r_\mathrm{w}) V_\mathrm{pw}(t; r_\mathrm{w}), \tag{14}$$

where RR_w is the removal rate of material on the wafer, $h_w(r_w)$ is the remaining film thickness on the wafer, $K_{pw}(t;\ r_w)$ is the Preston constant between the pad and the wafer (note the time dependence due to pad thickness interactions), $P_{pw}(r_w)$ is the pressure applied on the back of the wafer (being a function of wafer radius), and $V_{pw}(t;\ r_w)$ is the relative velocity of the pad with respect to the wafer.

In the wafer–pad system, the breakdown of the constant term and the time-dependent terms is slightly different from that in the pad–conditioner system because the pad thickness changes continuously in the wafer–pad system as a result of conditioner movement, which affects its interaction with the wafer surface. Hence, the Preston constant needs to be broken down into a constant and a variable pad thickness as shown in Eq. 15.

$$\frac{dh_w(r_w)}{dt} = K'_{pw} h_p(t; r_p - r_w) P_w(r_w) V_{pw}(t; r_w). \quad (15)$$

K'_{pw} is the constant portion of the Preston constant between the pad and the wafer (which does not include pad thickness effects), $h_p(t;\ r_p - r_w)$ is the height of the remaining pad material (expressed in terms of the wafer coordinate system), and $P_w(r_w)$ is the pressure applied on the back of the wafer (being a function of wafer radius).

The amount of material removed from the wafer surface is determined in Eq. 16 through the integration of Eq. 15, where the integrand now includes the solution to Eq. 3, coupling the removal rate of the wafer with the thickness of the pad.

$$h_w(r_w) = K'_{pw} P_w(r_w) \int_{t_0}^{t_f} h_p(t; r_p - r_w) V_{pw}(t; r_w) dt. \quad (16)$$

The derivation of the relative velocity is identical to the previous section and starts with the breakdown into the orthogonal components of the velocity vector

$$V_{pw}(t; r_w) = \sqrt{V_{pw}^{x^2}(t; r_h) + V_{pw}^{y^2}(t; r_h)}. \quad (17)$$

$V_{pw}(t; r_w)$ is the relative velocity of the pad with respect to the wafer, $V^x_{pw}(t; r_w)$ is the x-component of the relative velocity of the pad with respect to the wafer, and $V^y_{pw}(t; r_w)$ is the y-component of the relative velocity of the pad with respect to the wafer.

The orthogonal components are expressed in Eqs. 18 and 19.

$$V^x_{pw}(t; r_w) = \omega_p y_p - \omega_w y_w, \quad (18)$$

$$V^y_{pw}(t; r_w) = \omega_p x_p - \omega_w x_w, \quad (19)$$

where ω_w is the rotational speed of the wafer.

Once again, the pad coordinate system is converted to the wafer coordinate system through Eqs. 20 and 21.

$$y_p = y_w + y'_h(t), \quad (20)$$

$$x_p = x_w + x'_h(t). \quad (21)$$

$x'_h(t), y'_h(t)$ are the center coordinates of the polishing head that holds the wafer in the pad coordinate system.

From these, the following equations result:

$$V^x_{pw}(t; r_w) = (\omega_p - \omega_w) y_w + \omega_w y'_h(t), \quad (22)$$

$$V^y_{pw}(t; r_w) = (\omega_p - \omega_w) x_w + \omega_w x'_h(t). \quad (23)$$

$V^x_{pw}(t; r_w)$ is the x-component of the relative velocity of the pad with respect to the wafer, and $V^y_{pw}(t; r_w)$ is the y-component of the relative velocity of the pad with respect to the wafer.

The equations finally yield the explicit form of Eq. 17 as expressed in Eq. 24

$$V_{pw}(t; r_w) = \sqrt{\left((\omega_p - \omega_w) r_w \sin(\omega_w t) + \omega_w y'_h(t)\right)^2 + \left((\omega_p - \omega_w) r_w \cos(\omega_w t) + \omega_w x'_h(t)\right)^2}, \quad (24)$$

through the transformation into the polar coordinate system by Eqs. 25 and 26.

$$x_w = r_w \cos(\theta_w) = r_w \cos(\omega_w t), \quad (25)$$

$$y_w = r_w \sin(\theta_w) = r_w \sin(\omega_w t). \quad (26)$$

x_w, y_w is the wafer coordinate system, and r_w, θ_w are the polar coordinates of the wafer system.

Simulation

The main simulation and visualization program was coded in MATLAB 7.1. There were supporting routines written in Visual Basic 6.0 to define the sweep profiles using a similar interface to the AMAT Mirra platform. The approach was expression of the three subsystems of the polisher as

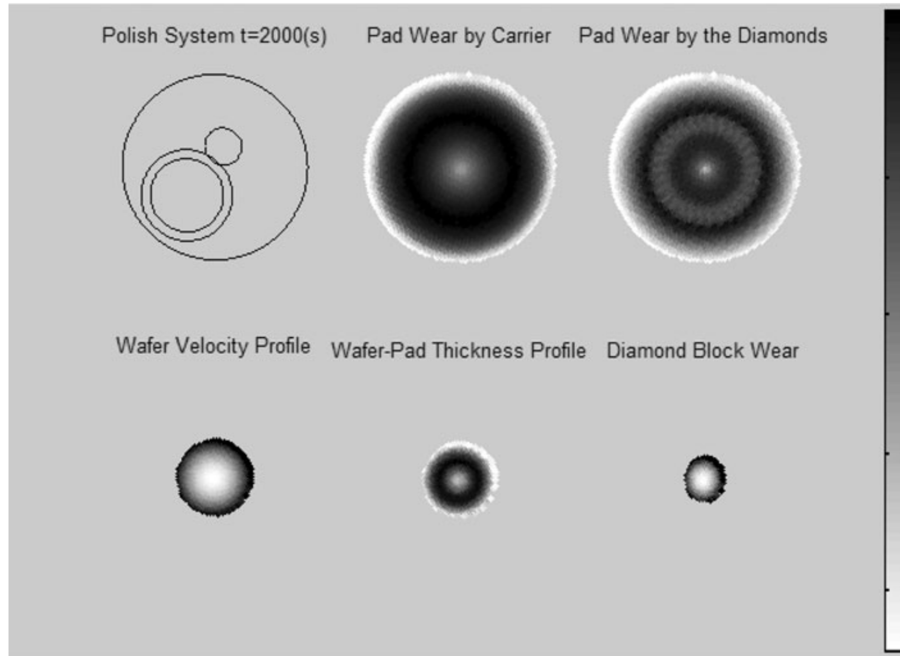

Fig. 4. A snapshot of the visualization routine output that enables monitoring the real-time progress of the polishing system and all relevant variables. Light colors represent areas exposed to least amount of conditioning, and darker colors represent a higher degree of conditioning.

meshes of the appropriate resolution to ensure the stability and convergence of the simulations.[20] The integrals in Eqs. 3 and 16 were then numerically evaluated for each point on the mesh, and the parameters of interest were stored in memory for later manipulation.

There was also a comprehensive visualization routine used to demonstrate the real-time progress of the variables of interest, shown in Fig. 4. Based on the simulation, lighter colors represent less-conditioned regions, and darker colors represent a higher degree of conditioning. The information was conveyed by averaging the two-dimensional (2-D) mesh results into a one-dimensional (1-D) plot as a function of radius via numerical integration with respect to angle (as all variables of interest in this work are symmetric with respect to radial position). Various nonuniformity metrics were also calculated and reported. The most general form of the algorithm is numerically inefficient due to the lack of restrictions in the simulation parameters. However, the code is modular; the execution can be optimized with analytical expressions for sweep profiles, resulting in closed-form analytical solutions where possible (i.e., a sinusoidal sweep for the conditioner and/or the polishing head). Simulations for this study were carried out on a computer driven by a 2.41-MHz Intel Pentium processor.

RESULTS AND DISCUSSION

As outlined in the “Introduction,” pad conditioning in CMP processes prevents pad glazing and defectivity while enabling continuous material removal by opening the pad fibers to the slurry. In the process of conditioning, the pad surface is abraded by the conditioner disk, which results in profile changes on the pad as well as on the wafer. In this section, the impact of the pad conditioning sweep changes on the pad and wafer profiles will be discussed. This is followed by an analysis of the impact of pad conditioning on the defectivity and a description of how to optimize the pad conditioning to improve postpolish profiles of patterned wafers in an HVM setting.

Pad and Wafer Profile Analyses

Initial studies on the pad profile change due to conditioning were conducted on the blanket oxide wafers. Figure 5a illustrates the pre and postpolish wafer thickness profiles from diameter–thickness scans of the wafers with an 18,000-Å-thick deposited blanket oxide. This set of experiments was run after the polishing pads were broken in using the designated conditioner sweep profiles for 10 min. Two blanket oxide wafers were run on the pad to stabilize the pad’s surface before the three wafers used for testing were polished. The wafers were polished for a total of 90 s. The change in the conditioning sweep profile did not affect the postpolish wafer profile in the beginning of the pad life. All of the wafers had a faster polish rate at the wafer center with more material removal and a slower polish rate at the wafer edges with less material removal (and hence a thicker edge profile). The material removal rate was higher for the sinusoidal sweep at the wafer center (~1200 Å/min) compared with CS1 and CS2 (~650 Å/min). However, the postpolish wafer profiles were comparable in shape.

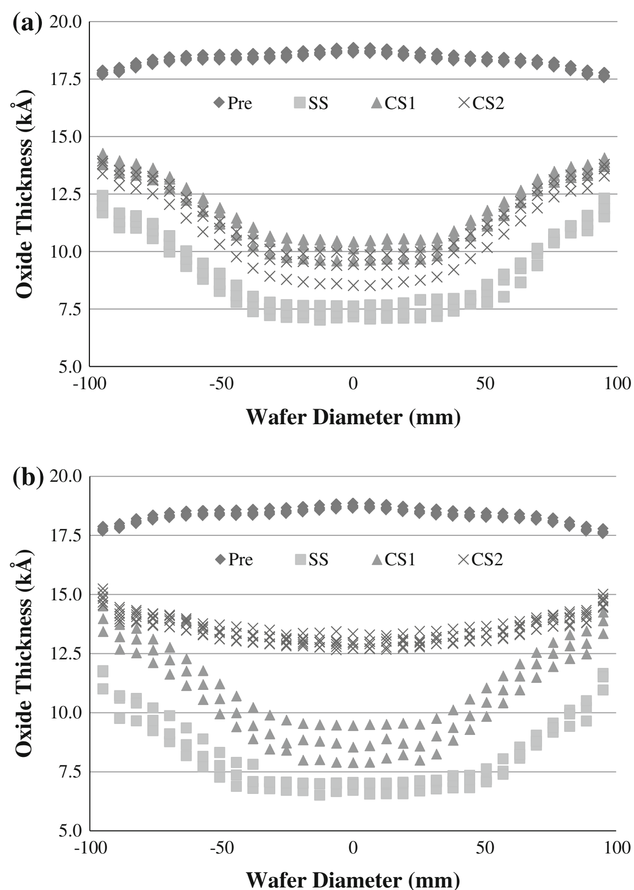

Fig. 5. (a) Diameter profiles of the pre-CMP and post-CMP blanket TEOS wafers with the SS, CS1, and CS2 sweep profiles at the beginning of the pad life; the profiles obtained with all three sweep profiles are comparable. (b) Diameter profile of the pre-CMP and post-CMP blanket TEOS wafers with the SS, CS1, and CS2 sweep profiles at the end of the pad life, after 1 h of continuous conditioning with the designated conditioning sweeps; the profiles obtained demonstrate differences.

After 1 h of continuous conditioning of the pads with the designated conditioning sweeps, the polishing tests were repeated on three blanket oxide wafers. In this case, the postpolish profiles showed significant differences as a function of the applied conditioner sweep profile, as shown in Fig. 5b. The postpolish pad surface profile obtained with CS2, where the conditioner spends an equal amount of time per unit area (i.e., more conditioning time is spent at the wafer edges and less time at the center), was flatter than the other two profiles. The postpolish profiles obtained using SS and CS1 remained similar to the profiles obtained at the beginning of the pad life. To understand the changes in the postpolish profile as a function of pad conditioning, the radial profile of the pads (because the polishing head only sweeps within the radius of the pad) used for the experiments were measured and compared with the pre and post profiles of the blanket TEOS wafers, as shown in Fig. 6. It is clear that CS2 created a concave-down abrasion on the pad radius which is similar to the prepolish profile of the TEOS wafers as they face towards the pad surface. This combination resulted in the most uniform postpolish oxide thickness profile on the blanket wafer. SS and CS1 polished the protruded wafer center faster than they polished the edges, which resulted in a more planar abrasion on the pad surface. It is believed that the initial silicon profile was matched on the oxide surface by polishing with the CS1 abraded pad surface, achieving a more uniform oxide thickness profile.

Figure 7 shows the same comparison with the STI-patterned wafers at a 110-nm node with the PECVD oxide. It can be observed that the prepolish deposition profile had a hump at the wafer center, which is typical of a shallow trench oxide deposition profile. In this case, SS provided the best match between the pad surface profile and the prepolish wafer profile. SS and CS1 created a planar surface finish, with a dip towards the edges of the wafers (which is consistent with the similar pad profiles they generated); CS2, which created the most planar profile on the blanket oxide wafers, unexpectedly created the least planar profile for the patterned wafers, with extreme overpolishing at the edges and underpolishing at the center. Because CS2 resulted in a dipped profile at the mid-radius of the pad surface, there was a reduced removal rate at the wafer center. The enhanced overpolishing at the wafer edges was caused by the elevated pad profile at the wafer edges, exposing the edges of the wafer to a higher pressure, which increased the edge removal rate. Additionally, the changes in the pattern density at the wafer edges (incomplete patterns and edge exclusion) added to the polish rate differences between the edge of the patterned and unpatterned wafers during CMP. It should also be noted that the change in the material removal rates at the oxide/nitride interface after the bulk oxide was removed and the sacrificial nitride layer was exposed also could have affected the material removal rates. The response of the patterned wafers to the pad profile changes indicates that the overall pattern layout, product design, and process integration need to be evaluated before the pad profile is changed using a customized sweep for CMP during high-volume manufacturing (HVM). However, as an example, the bulk of the material can be planarized on a pad installed on the first platen of the polisher that is shaped with CS2 before the interface is uncovered at a subsequent platen using SS. Overall, it is critical to be able to predict the pad profile as a function of pad life for better control of the postpolish wafer profile.

Figure 8 shows the measured pad profiles compared with the profiles predicted using the pad wear rate model. There is good agreement between the experimental and predicted results, and it is clear that the difference in the pad thickness profile that is exposed to the wafer is translated to the postpolish wafer profile. To optimize the pad wear

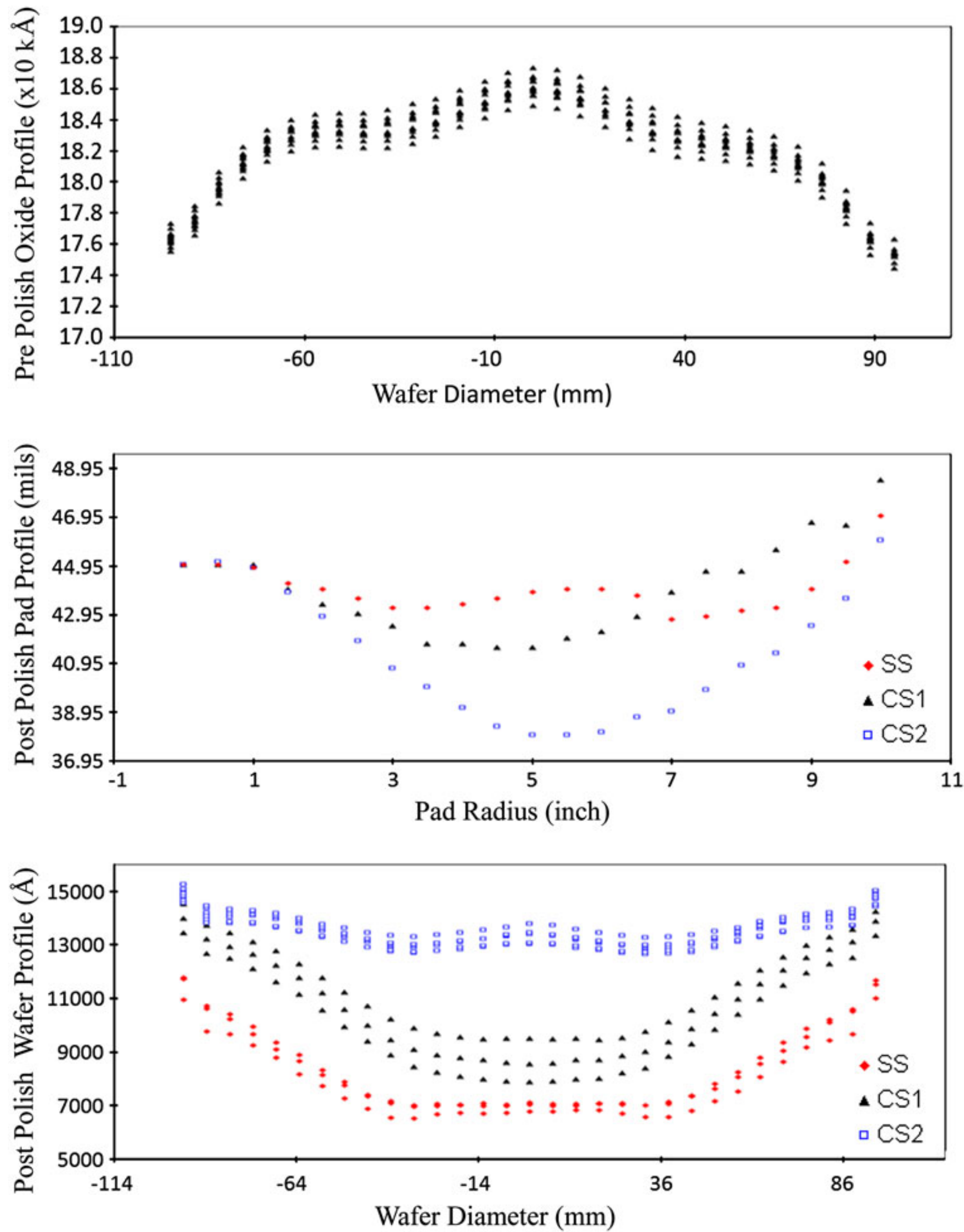

Fig. 6. Post-CMP pad surface radial profiles obtained with the three different conditioning sweeps compared with the pre-CMP and post-CMP wafer diameter profiles on blanket TEOS wafers at the end of the pad life. CS2 provided the most complementary pad surface shaping, resulting in the most planarized postpolish wafer profile.

profile to obtain the most planar post-CMP wafer profiles, the modeling approach introduced previously was utilized to observe the impact of conditioner radius and sweep range variations on postpolish profiles.

Conditioner Radius Optimization

The conditioner radius has been reported to be a significant modulator of the pad wear profile in literature.[18–20] The simulations that resulted from the model developed in this work also confirmed the impact of the disk radius on the pad wear profile for the three sweep profiles studied. Figure 9 illustrates the simulated profiles for the SS, CS1, and CS2 pad wear profiles as the disk diameter is changed from 4 inches to 3 inches to 2 inches to 1 inch. Among the three profiles, CS1 is the most sensitive to the conditioner radius. Both SS and CS1 converge to a constant wear rate in the center zone of the pad (excluding approximately 25 mm from either edge) when the radius reaches 1 inch. The CS2 exhibits the same trend for all conditioner diameters considered, although the peak is less pronounced and closer to the edge as the conditioner disk gets smaller. As expected, more uniform pad

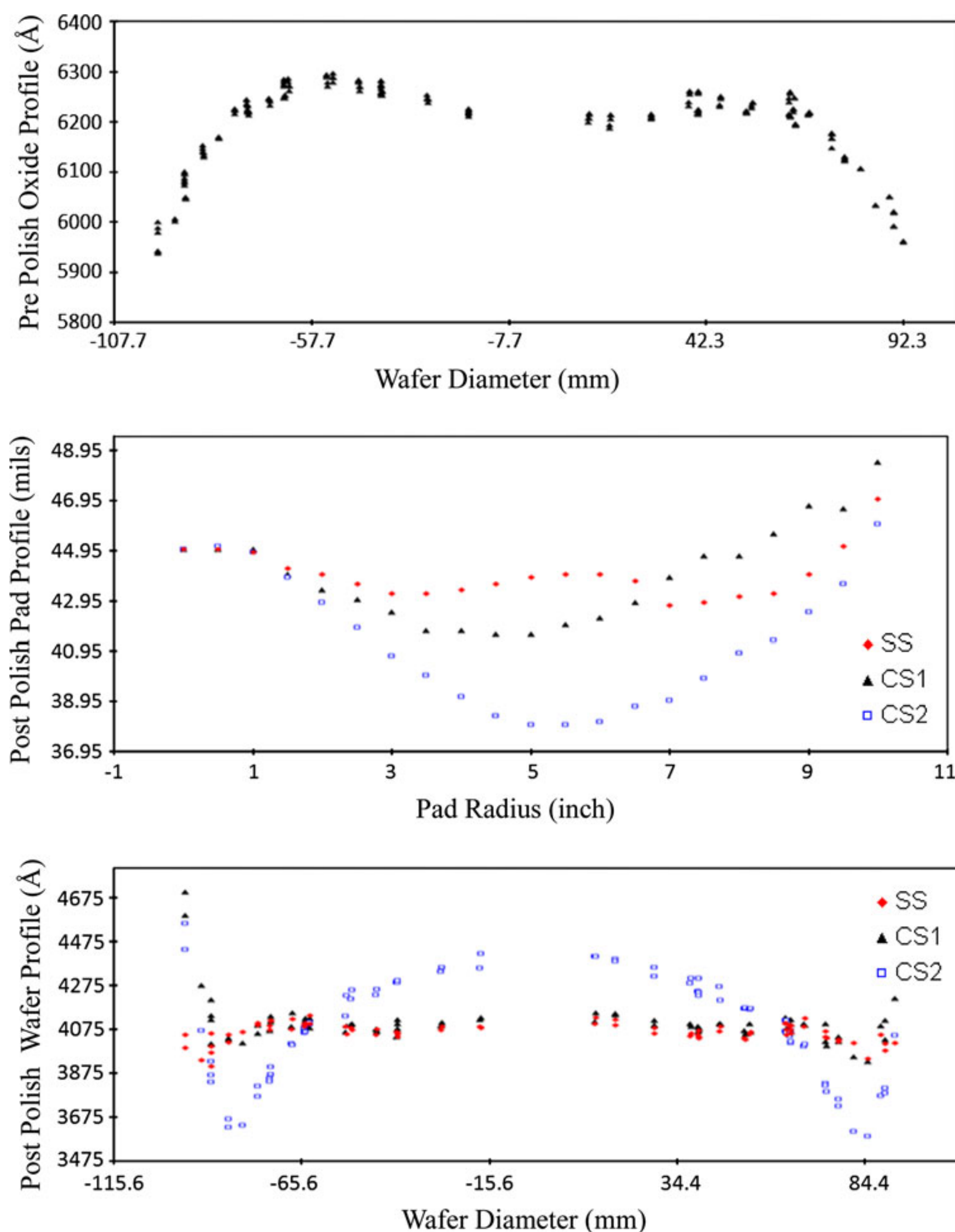

Fig. 7. Post-CMP pad surface radial profiles obtained with the three different conditioning sweeps compared with the pre-CMP and post-CMP wafer diameter profiles on patterned STI-level 130-nm-node wafers at the end of the pad life.

profiles result in more uniform wafer profiles. Furthermore, Fig. 10 demonstrates the postpolish profiles of the blanket TEOS wafers, which are polished with 2-inch and 4-inch conditioner disks using SS profiles after 1 h of conditioning. The postpolish wafer profile is more planar for the pad conditioned using the 2-inch-diameter disk, in agreement with the simulation from the modeling approach. Although the removal rates were lower with the smaller size disks, which may be attributed to the less effective slurry delivery at the wafer–pad interface (due to a reduced abrading disk size), the control of the postpolish profiles was significantly better, as demonstrated by the more uniform pad wear profiles in Fig. 9. Additionally, the larger conditioner blocks had a higher edge velocity. Thus, a 4-inch-diameter disk will be four times more effective than a 1-inch-diameter disk in wearing the pad material. Hence, when the material removal rate needs to be increased while keeping a planar post-CMP profile, more aggressive conditioning may be applied by increasing the conditioner sweep frequency, the block rotation speed, or the down force using smaller-diameter disks instead of increasing the disk size.

Conditioner Sweep Range Optimization

Another possible modulator to achieve material removal rate uniformity is optimization of the sweep

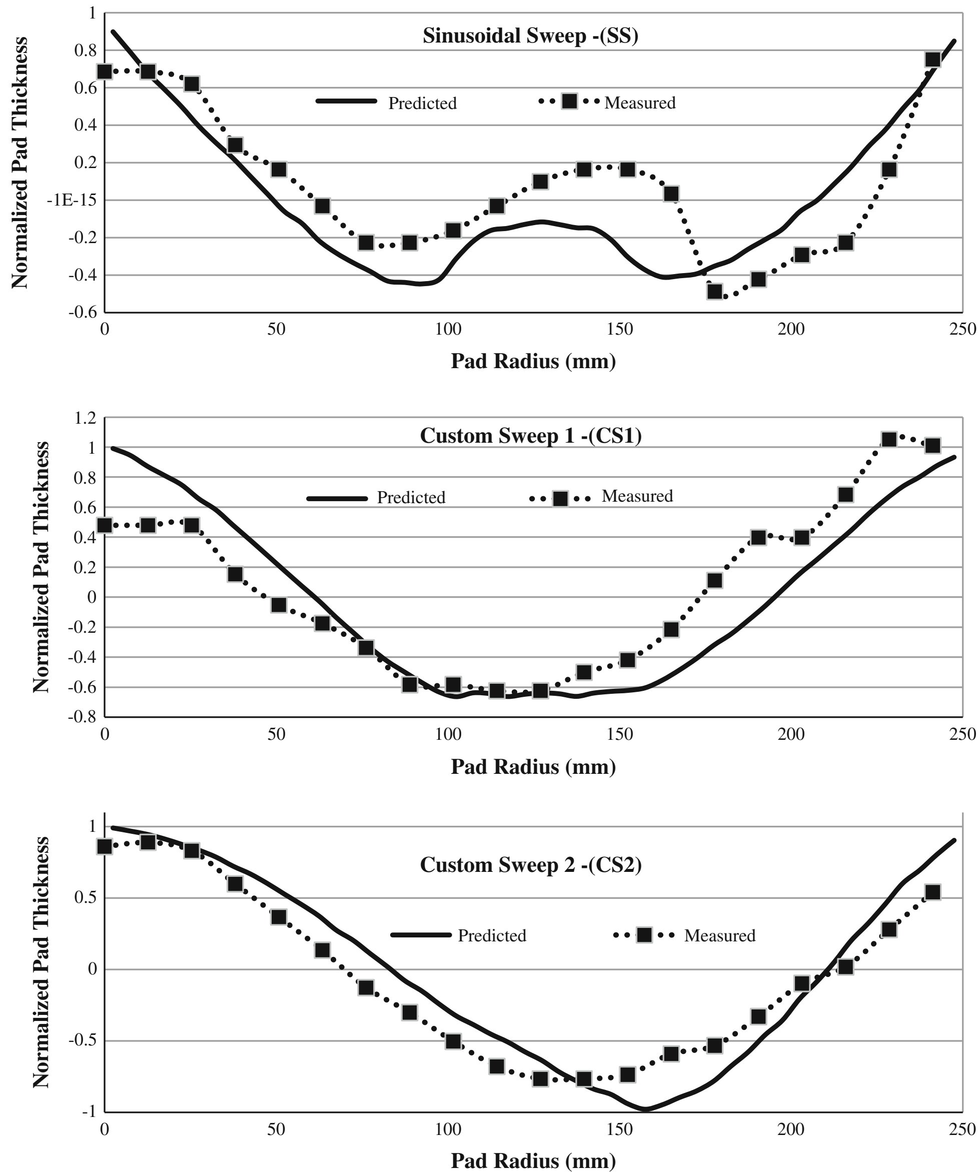

Fig. 8. Comparison of the simulated radial thickness profiles of the pads worn by 1 h of conditioning with SS, CS1, and CS2 versus the measured thickness profiles.

range. Smaller disks produce a more uniform profile because the sweeps in each zone are more independently controlled. This is especially pronounced at the edge of the pad. As an example, the edge of a 12-inch-size pad conditioned by a 4-inch-size disk is not effectively abraded when the disk sweep is extended to the pad edge, and the disk center reaches the edge of the pad, resulting in half of the disk rotating away from the pad surface. This effect can be mimicked by simulating the extension of the sinusoidal conditioner sweep at 0 inches, 1 inch, 2 inches, and 3 inches beyond the edge of the pad, as demonstrated in Fig. 11. The extension of the disk sweep beyond the pad radius helps to achieve a more uniform pad profile, which is also related to the aggressiveness of conditioning observed as a function of the disk size. Because the disk is 4 inches in diameter in the simulation, the advantage of a higher relative velocity at the edge of the disk is realized. A commonly observed downside of extending the conditioning sweep out of the pad radius is the probability of increased scratches on the wafer because the disk sweeping inside and outside of the edge of the pad may result in diamond loss as well as tearing apart of pad pieces, as discussed in the next section.

Defectivity Analyses

One of the main functions of pad conditioning is to reduce the CMP-related defectivity (typically microscratches) on the wafer surface by cleaning the

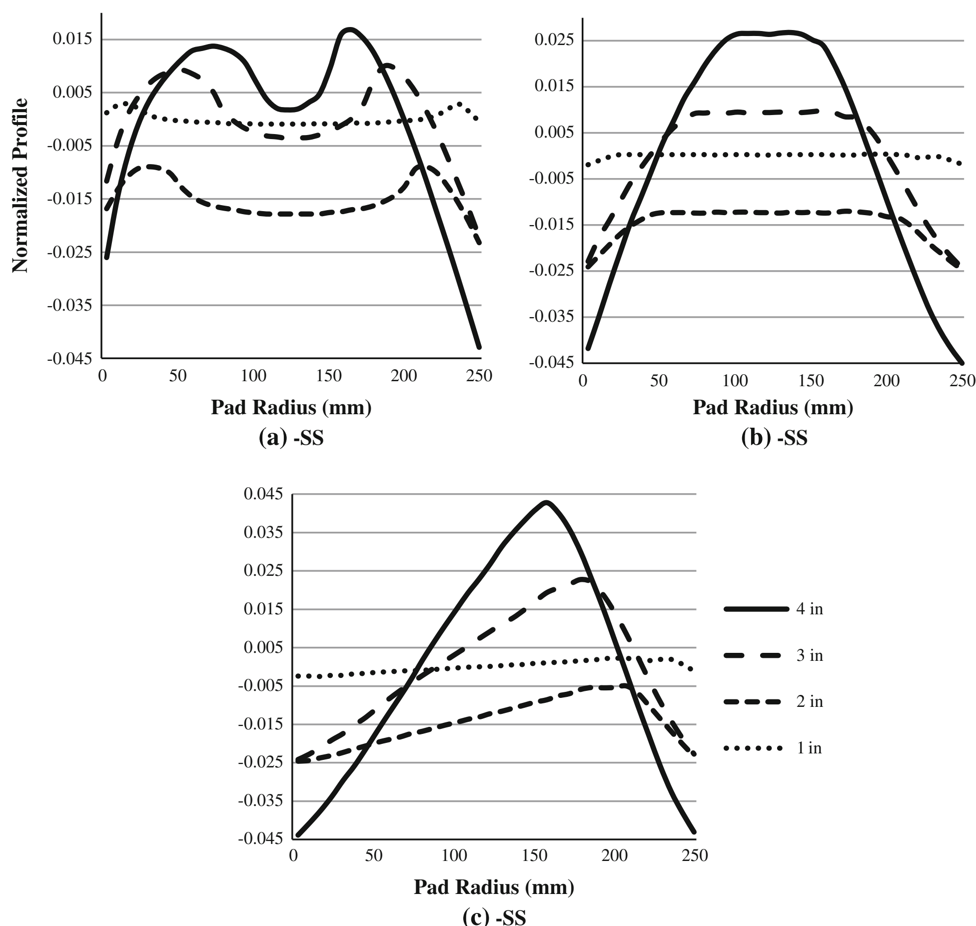

Fig. 9. Impact of conditioner radius on the pad wear profile for the different sweep profiles: (a) SS, (b) CS1, and (c) CS2. All profiles are normalized to the mean.

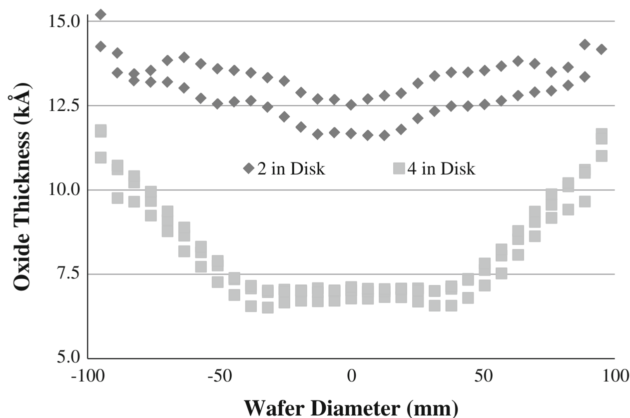

Fig. 10. Impact of conditioner radius on post-CMP wafer profiles for SS with 2-inch and 4-inch disks. The improvement in wafer profile with a reduced disk dimension is clearly illustrated.

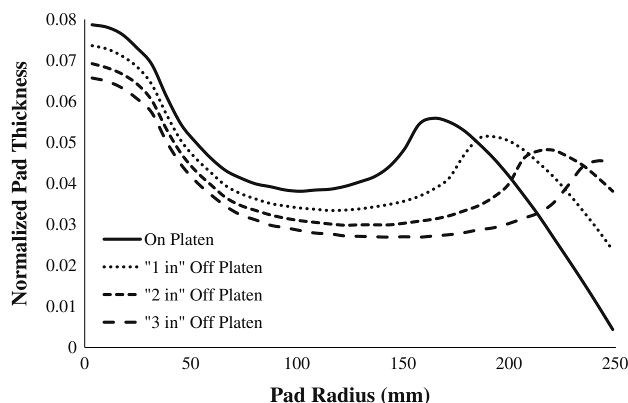

Fig. 11. Impact of conditioner sweep extension on the pad thickness–velocity profile using a sinusoidal conditioning sweep; the more the extension is increased, the more planar the obtained profile.

pad surface of postpolish residues and agglomerated slurry particles; this is accomplished by combing the pad fibers and allowing fresh slurry to be introduced to the pad–wafer interface.[1,21] However, it is also common to observe macroscratches on the wafer surface originating from a diamond abrasive that came loose from the conditioner. The probability of the conditioner disk losing a diamond abrasive is higher when the sweep is extended outside the pad radius. Additionally, the protruding diamond

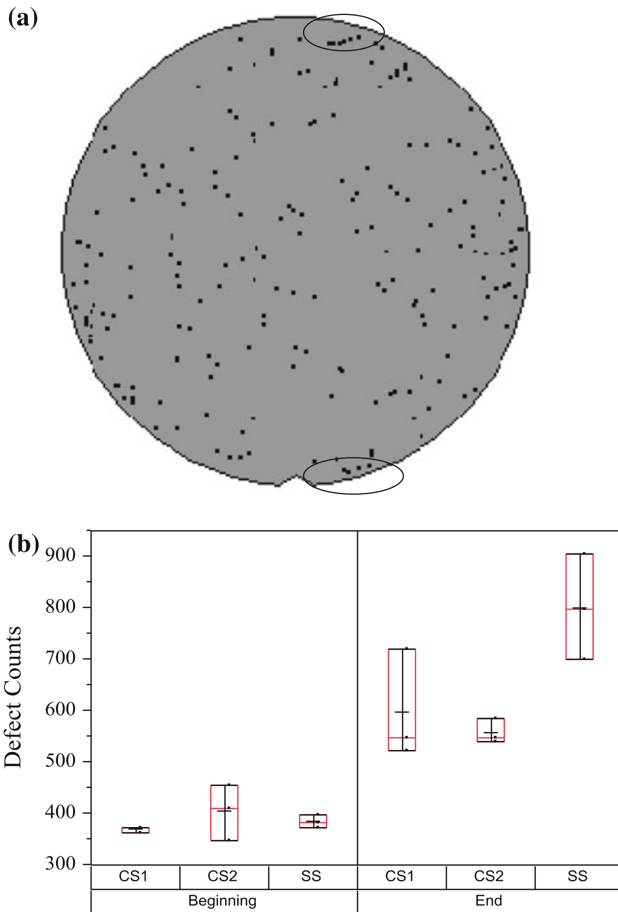

Fig. 12. (a) Postpolish defect map on the blanket TEOS wafer polished with a sinusoidal conditioning profile extending 2 inches out of the pad radius. The scratches formed on the wafer edges are circled on the wafer map. (b) Defect count comparison on blanket TEOS wafers polished with the SS, CS1, and CS2 conditioning profiles confined within the wafer radius in the beginning and at the end of the pad life. The increase in the defect counts clearly shows the impact of pad aging on defect elevation.

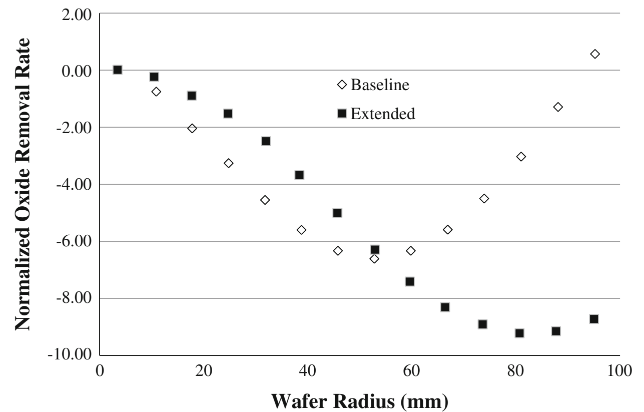

Fig. 13. Normalized wafer removal rate predicted by the model using SS confined within the pad radius and extended out of the pad radius.

particles can result in chunks of pad material being torn apart at the pad edges, which can also result in surface scratches. Figure 12a shows the defectivity analyses conducted on the blanket oxide wafers polished with a 4-inch conditioner disk with SS extending out of the pad radius by 2 inches on a pad broken in for 10 min, as is typical in conventional production. Scratches were formed on the wafer edges, as shown by the KLA-SP1 scans of the polished wafers.

In addition to the conditioner sweep extension, pad aging also affects the probability of defect formation during CMP. Figure 12b shows a comparison of the defect counts on the blanket TEOS wafers for the three conditioning sweeps that swept the pad radius after 10 min versus after 1 h of pad conditioning (before polishing was conducted). As expected, as the pad aged, the defect count increased. Inevitably, the extended pad conditioning leads to more pad wear and enhances the probability of slurry particles agglomerating and embedding in the pad surface, which will elevate the surface defectivity.

HVM Implementation Example

The findings in the previous sections were implemented on a production line for a 90-nm STI CMP process. The SS sweep range of the baseline process was extended by 22.5%, resulting in a part of the 4-inch conditioner drifting beyond the edge of the pad at the maximum extension. As was demonstrated by the simulation results in Fig. 11, the extension of the conditioning pad sweep resulted in increased pad wear towards the edge of the pad. The sweep extension is expected to reduce the pressure applied at the edge of the wafer and therefore cause a reduction in the oxide removal rate at the wafer edge, as shown in Fig. 13. This removal rate profile is indeed in agreement with the STI oxide deposition profiles and should result in a reduction in the thickness range of the post-CMP oxide. In agreement with these expectations, the extended conditioner sweep profile resulted in a significant improvement of the post-CMP product oxide thickness range; this is demonstrated in Fig. 14, which illustrates the oxide thickness range for the tested 90-nm STI process over a 12-month timeframe, compared with data generated by the baseline process. The improvement in the range was mainly driven by the measurement sites that were located near the edge of the wafer (beyond a radius of 75 mm, where the most drastic change is predicted by the model). The wafer edge consistently polished faster despite the increased pad aging throughout the testing. Along with the implementation of the extended sweep, the pad life was increased by 20%, indicating that the new, optimized sweep condition was able to maintain a better uniformity throughout the pad usage. The dataset and simulations outlined above are evidence of the fact that the pad wear profile does not necessarily need to be uniform but

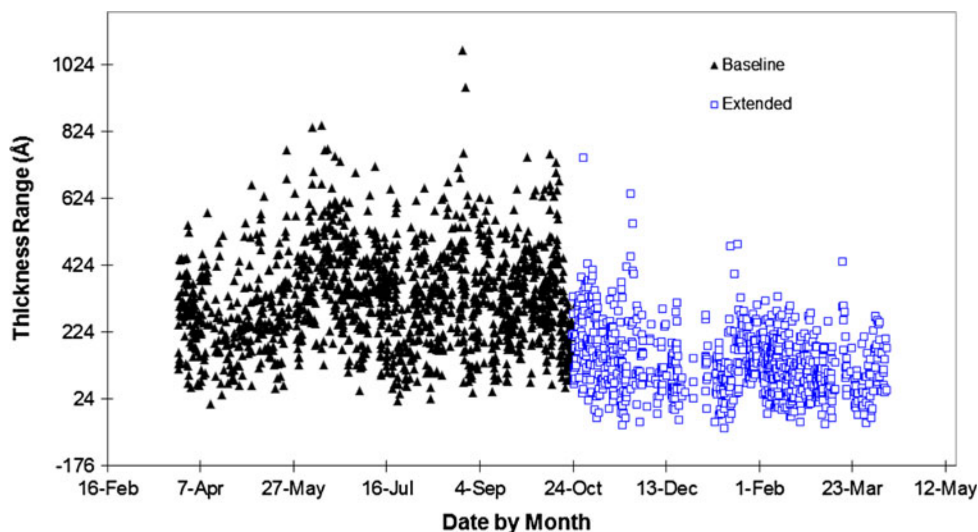

Fig. 14. Oxide thickness range change as a function of time of the 90-nm STI process, when SS was confined to the pad radius and extended out of the pad radius. Recorded range values were normalized by taking the baseline process mean range as 0.

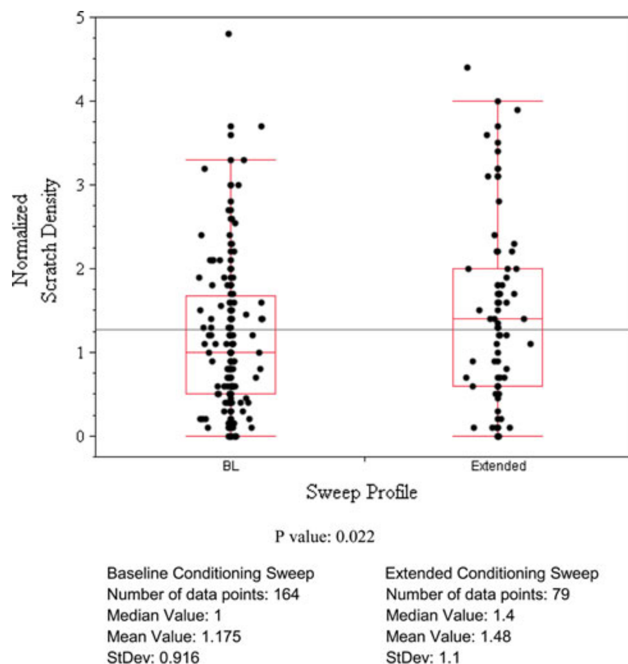

Fig. 15. Box plot of the defect density for the extended and nonextended pad conditioning sweeps for the 90-nm STI CMP process. It can be observed that the extended sweep profile results in an increase in the defectivity of the polished wafers.

that it needs to match the incoming wafer thickness profile within the limits of the pad life that are specified by the process requirements.

The main concern of extending the conditioner sweep beyond the edge of the pad is the increasing number of defects. The box plot of the normalized scratch density for the HVM testing demonstrated that extended sweeps increased the scratch density, as shown in Fig. 15. There was a slight increase in the defect density with the extended sweep profile, as expected and verified by the p value, reported as 0.022, based on a statistical t-test evaluation. The impact of defectivity versus profile improvements needs to be evaluated on a case-by-case basis, because some CMP layers can tolerate additional defects, and the improvement in thickness range is more important for these layers. It must be noted that nonuniform pad wear rates will result in a dynamic removal rate profile, which is a function of pad life. If the cut-rate for a certain process is high, this may result in a narrow operating window in pad life. One approach could be to initially use a shaping conditioner sweep to customize the pad profile as desired and then switch to a uniform sweep profile to preserve the achieved pad profile for the remainder of the pad life.

CONCLUSIONS

A comprehensive model relating pad thickness profiles shaped by conditioner sweep parameters was developed and validated using experimental data on three different types of sweep profiles. The changes in the pad thickness driven by the conditioning wear modulated the slurry flow characteristics and the local pressure applied on the wafer surface, resulting in a direct correlation between the pad thickness profile and the postpolish wafer thickness profile. The model was experimentally validated on blanket TEOS wafers, on patterned test wafers, and on a production line for STI CMP; all three types of applications demonstrated good agreement between the simulated and experimental values. The main conclusion is that the radial pad profiles can be tuned through conditioning wear to achieve the most complementary shape to the incoming wafer profile to improve planarization. Therefore, the first requirement to achieve

improved planarization is an analysis of the incoming wafer profile; one type of pad conditioning cannot compensate for every incoming profile. However, the results obtained in this study give us a strong tool that can optimize the shape of the pad wear through conditioning to improve the material removal rate profiles and achieve the most planar postpolish thicknesses. It can also be recommended that the pad conditioning be set up to be different from platen to platen, to enable better planarization of the bulk film material on the initial platen polish while keeping the pad more planar on the subsequent platens once the interface is reached. This also indicates that the pad break in procedures for different platens needs to be different for this multistep CMP procedure.

Using the model, it was also shown that smaller conditioner radii and extended sweep ranges improved the pad thickness uniformity. It is not possible to infinitely improve pad range by increasing the sweep zones and fine-tuning the residence time in each zone; as long as the conditioner has a radius greater than the zone width, part of the conditioner will still be in the adjacent zones. This impact can be remedied by using as small a conditioner diameter as possible (which has the disadvantages of possible underconditioning and/or a short conditioner life) or by extending the sweep of the conditioner beyond the edges of the pad (which may cause increased scratch levels, as the conditioner will introduce more large particles by sweeping onto the pad from the edge). The 90-nm process example from the HVM environment demonstrated that, by simply extending the conditioner sweep range, the post-CMP product oxide thickness range can be improved significantly with the added benefit of increased pad life. The increased pad life is due to matching the pad profile to the incoming thickness profile at the wafers, but this occurs at the expense of slightly increased scratch densities.

The presented study is not only important for highlighting how conventional conditioning practices can be improved through experimentation but also for introducing a simulation method to predict the process outcomes without a detailed experimental requirement. Additionally, to emphasize the optimization challenges of CMP, the defectivity performance of the alternative processes were presented, and the necessary trade-offs between perfect planarization and reduced defectivity were discussed.

ACKNOWLEDGEMENTS

The authors would like to acknowledge support from the Texas Instruments Incorporated DM5 Wafer Fab and Analog Technology Development, where the experimental work was conducted. Mr. G. Davis is acknowledged for suggesting one of the custom conditioning sweeps, and Dr. L. Olsen is acknowledged for supporting the HVM tests.

REFERENCES

1. M. Krishnan, J.W. Nalaskowski, and L.M. Cook, *Chem. Rev.* 110, 178 (2010).
2. Y. Gu, S. Chang, G. Zhang, K. Kirmse, D. Rogers, L. Olsen, and J. Lewellen, *Proc. SPIE* 6152, 1 (2006).
3. C.-P. Chen, B. Huang, W. Lee, W.-J. Chung, and H.T. Hou, *Proc. SPIE* 4344, 274 (2001).
4. G.-B. Wang, E.-H. Lin, H.-S. You, M.-W. Lee, F.-K. Hsiao, and C.-W. Lai, *IEEE* 8469-5/04, 178 (2004).
5. A. Chen, R.-.S. Guo, Y.L. Chou, C.L. Lin, J. Dun, and S.A. Wu, *IEEE* 5403-6/99, 229 (1999).
6. T.H. Smith, S.J. Fang, J.A. Stefani, G.B. Shinn, D.S. Boning, and S.W. Butler, *J. Vac. Sci. Technol. A* 17, 1384 (1999).
7. J. Luetzen, S. Pal, S. Gonzales, and B. Yuval, *Proc. SPIE* 3882, 36 (1999).
8. S.-J. Shiu, C.-C. Yu, and S.H. Shen, *J. Vac. Sci. Technol. B* 22, 1679 (2004).
9. C.E. Chemali, J. Moyne, K. Khan, R. Nadeau, P. Smith, J. Colt, J. Chapple-Sokol, and T. Parikh, *J. Vac. Sci. Technol. A* 18, 1287 (2000).
10. K. Achuthan, K.J. Curry, M. Lacy, D. Campbell, and S. Babu, *J. Electron. Mater.* 10, 117 (1996).
11. T. Manabu, *J. Electron.* 6, 96 (2001).
12. D. Ng, M. Kulkarni, G. Xu, P. Severs, R. Marvin, J. Xiao, and H. Liang, *J. ASTM Int.* 2, JAI12894 (2005).
13. O. Chang, K. Kim, K. Park, B. Park, H. Seo, and H. Jeong, *Microelectron. Eng.* 84, 577 (2007).
14. T. Feng, *IEEE Trans. Semicond. Manuf.* 20, 464 (2007).
15. S.R. Runnels, I. Kim, J. Schleuter, C. Karlsrud, and M. Desai, *IEEE Trans. Semicond. Manuf.* 11, 501 (1998).
16. D. Castillo-Meji and S. Beaudoin, *J. Electrochem. Soc.* 150, G96 (2003).
17. G.B. Basim, S. Kincal, and G. Davis, U.S. patent 7,899,571 (2011).
18. W.-T. Tseng, J.-H. Chin, and L.-C. Kang, *J. Electrochem. Soc.* 146, 1952 (1999).
19. F. Tyan, *IEEE* 3, 2052 (2005). 10.1109/ACC.2005.1470272.
20. C.-Y. Chen, C.-C. Yu, S.-H. Shen, and M. Ho, *J. Electrochem. Soc.* 147, 3922 (2000).
21. G.B. Basim, B.M. Moudgil, and J. Colloid, *Interface Sci.* 256, 137 (2002).