Nanotopography Effects on Chemical Mechanical Polishing for Shallow Trench Isolation

Brian Lee, Terence Gan,
Duane Boning
Mass. Institute of Technology
Cambridge, MA 02139
e-mail: lee21@mtl.mit.edu
e-mail: boning@mtl.mit.edu

Pat Hester, Noel Poduje
ADE Corporation
Westwood, MA 02090
e-mail: pat_hester@ade.com
e-mail: noelp@ade.com

Winthrop Baylies
BayTech Group
Weston, MA 02493
e-mail: winb@mediaone.net

Abstract

Nanotopography is the nanometer-scale height variation that occurs over lateral millimeter length scales on unpatterned silicon wafers [1][2]. This height variation can result in excess thinning of surface films during chemical mechanical polishing (CMP) of shallow trench isolation (STI) structures. The development of an accurate nanotopography CMP modeling and characterization procedure will allow for the proper diagnosis of potential problems due to wafer nanotopography in a given STI CMP process. In this work, a nanotopography modeling methodology is proposed which relates the length scale of nanotopography features to the length scale of the CMP process. A combined density/step-height polishing model indicates that when the nanotopography features occur over a range comparable to or shorter than the planarization length, appreciable thinning is predicted. A contact wear CMP model similarly shows that as the pad stiffness increases, film thinning also increases. These simulation results indicate that the effect of nanotopography on STI CMP may be a substantial concern.

Keywords

Nanotopography, CMP, STI

1. Introduction

Height variation of 20 to 80 nanometers, or nanotopography, is known to exist over lateral distances of several millimeters on unpatterned silicon wafers. Figure 1 shows a topography SQM map of a wafer surface that illustrates these nanotopographical features on an 8" wafer. Nanotopography is becoming a serious concern in IC fabrication [1]. A major concern, which is the focus of this work, is the interaction between wafer nanotopography and chemical mechanical polishing processes. This interaction can cause undesired thinning of surface films that results in yield concerns in shallow trench isolation processing. Another concern, which will not be discussed here, is the effect of nanotopography on lithography, particularly with respect to polysilicon critical dimension printability.

Chemical mechanical polishing (CMP) is the planarization technique used in current IC fabrication facilities. It can be crudely thought of as "wet-sanding" the film to be planarized, using a polyurethane pad with a chemical abrasive slurry to polish the wafer. CMP is used in both front-end and back-end processing to selectively remove raised films, and to reduce step heights around patterned features, in order to meet lithography and other planarization criteria.

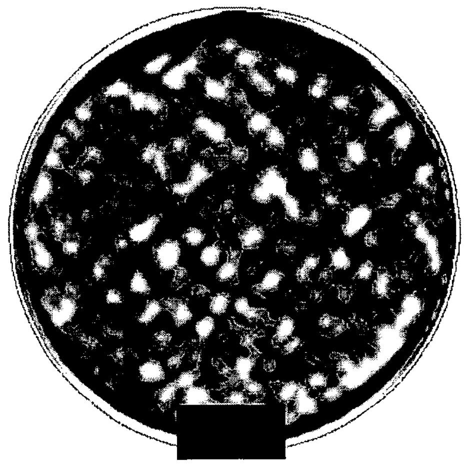

Figure 1. SQM topography map of raw silicon; shading indicates 0 to 100 nm height variations.

Shallow trench isolation is the isolation technique of choice in advanced IC fabrication. A typical STI process involves using silicon nitride to mask off active areas on a wafer, etching isolation trenches into the silicon, depositing an oxide film to sufficiently fill the trenches, and performing CMP to remove oxide over active areas, using the nitride as a polish stop. It is critical to clear oxide over all active areas, while also ensuring that the nitride is not excessively polished over any active area.

In this work, the thinning of surface films due to the interaction of CMP and nanotopography is modeled, and the implications of such variation for STI considered. Film thinning during CMP due to planarization of topography is discussed in Section 2. Some relevant CMP models are presented as background in Section 3. A nanotopography CMP simulation approach is proposed in Section 4, and simulation results presented in Section 5. The implications for STI processes are discussed in Section 6, and Section 7 draws conclusions and identifies areas for future work.

2. Film Thinning during CMP

Due to the planarization capability of the CMP process, film thinning over certain raised areas of material will occur. Thinning is defined here as greater removal of material in the raised areas of the pre-CMP film than in the low areas of the film. The degree of thinning that occurs depends on the relative length scale of the CMP process (planarization length) and the size of the structures being polished. For example, Figure 2 shows the thinning of small surface heights over long-range (8mm) nanotopography. The thinning of silicon nitride can lead to problems in an STI process, as discussed in Section 6.

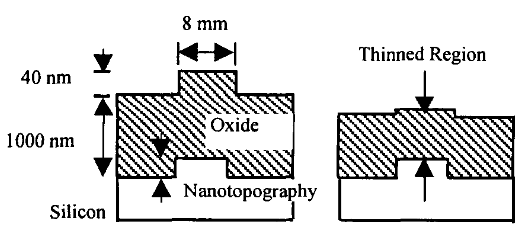

Figure 2. CMP of oxide film over nanotopography on the silicon layer results in thinning of the oxide film.

In conventional CMP (e.g. in the polishing of interlevel dielectric films between metal layers in back-end or interconnect processes), thinning is desirable because it means that planarization of the surface is occurring. Indeed, CMP is very effective at removing the step-height resulting from oxide deposition over very short-range lateral features, such as micron-scale patterned metal lines. However, CMP is known to have a limit to the lateral range over which step height can be removed. As illustrated in Figure 3, during polishing a point is reached at which the CMP pad and process no longer "see" the step height, and after which little additional planarization or step height reduction will occur. This planarization length, typically in the 2-10 mm range, serves as one important metric for a CMP consumable set and process.

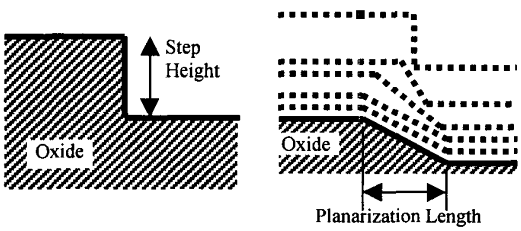

Figure 3. Evolution of step height across a large distance during CMP illustrating the planarization length limit.

The fundamental hypothesis of this paper is that the interaction of nanotopography and CMP depends on the relative scale or extent of the nanotopography and of the planarization length. As illustrated in Figure 4 the thinning effect may or may not be appreciable, depending on the wafer topography and on the process. For structures much smaller than the planarization length, significant thinning will occur; for structures larger than the planarization length, little or no thinning will occur. In this work, we will apply CMP models based on the planarization length and effective density, or on the bending of an elastic pad material around surface structures, in order to predict the evolution of nanotopographical features during CMP.

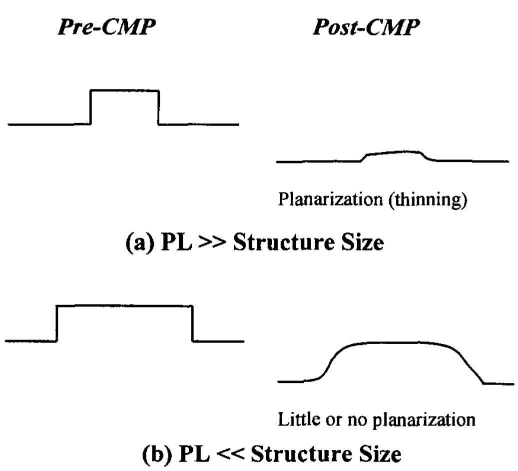

Figure 4. Planarization length and thinning implications in the CMP of films

3. CMP Modeling

There are several potential methods for modeling the CMP of nanotopography; in this work we consider two approaches. The first method is to use a density and step height based CMP model (referred to as the Smith model here) to predict the polish. The second method is to use a polish model based on contact mechanics between the wafer and pad.

Density/Step Height CMP Model

The first simulation method is implemented using the density based model described by Stine [4], together with the modification added by Smith [5] to more accurately model the polish of features with small step heights.

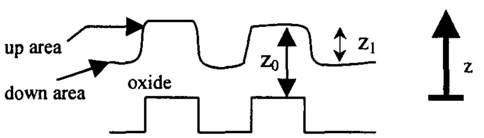

Figure 5. Definition of terms in the Stine model for the planarization of surface step features.

The Stine model is summarized in Figure 5 and described by Equation 1:

$$RR_u = K/\rho, \quad RR_d = 0, \qquad z > z_0 - z_1 \qquad (1a)$$
$$RR_u = K, \quad RR_d = K, \qquad z < z_0 - z_1 \qquad (1b)$$

where RR_u is the instantaneous up area film removal rate, RR_d is the instantaneous down area film removal rate, K is the blanket (100% density) removal rate, ρ is the effective pattern density, z is the instantaneous up area film thickness, z_0 is the initial film thickness, and z_1 is the initial step height of the film. The basic premise of the model is that before local step height removal, the up area polishes at the blanket polish rate weighted inversely by feature density, and there is no down area polish. After local step height removal, both up and down areas polish at the blanket removal rate.

Coupled with the notion of the density-based model is the planarization length parameter; this parameter is the length scale over which the effective pattern density at a given point is calculated. The planarization length determines the distance over which neighboring topography around a point on the die affects the polishing at that particular point. CMP processes with longer planarization lengths are usually considered to be desirable because they provide better die feature density averaging, which leads to better within-die uniformity in the planarization of pattern features across the chip.

Smith [5] added a step-height and time-dependent modification (based on Grillaert's work in [8]) to the Stine pattern density model that considered the polish of down areas even before the local step height is completely removed. The "contact" with down areas occurs only after a certain step height (defined as the contact height) is reached during the polish. The time at which the contact height is reached is called the contact time. The model states that the removal rates of the up and down areas exponentially change in time after the contact height is reached. The contact height varies from site to site on the die and depends on the density of features at a particular site.

The Stine model is considered for nanotopography modeling, as it accounts for the impact of raised material (the effective density) across millimeter scale distances. The Smith modification is hypothesized to be important in the nanotopography modeling problem, because small step heights (comparable to or smaller than the contact height) are being polished. The modified equations (referred to as the density/step height or Smith model) are:

$$RR_u = K/\rho, \quad RR_d = 0 \qquad t < t_c \qquad (2a)$$
$$RR_u = K + (1-\rho)*(h_1/\tau)*\exp(-(t-t_c)/\tau), \quad t > t_c \qquad (2b)$$
$$RR_d = K - \rho*(h_1/\tau)*\exp(-(t-t_c)/\tau)$$

where h_1 is the contact height, t_c is the contact time, and τ is an exponential time constant. The effective density ρ is calculated using the planarization length parameter and pattern layout or topography information with moving average or Fourier transform methods described in [3] and [4].

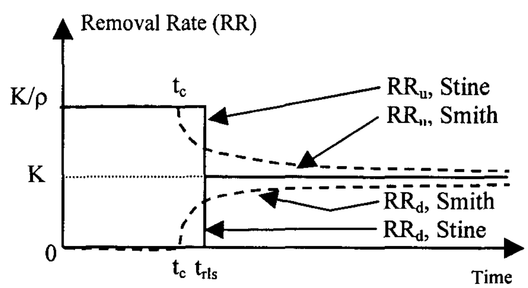

Figure 6. Removal Rate diagrams for Stine and Smith models. Note that t_c, the contact time, occurs before t_{rls}, the time to remove the local step height.

Figure 6 illustrates the difference in removal rates for the Stine vs. Smith models. The removal rate equations can be integrated directly to find the amount removed for a given polish time, or can be used in an iterative time-stepping algorithm to similarly calculate the final post-CMP film profile.

Contact Wear CMP Model

The modeling of CMP over a surface film with underlying nanotopography can also be done through consideration of wafer-pad contact mechanics. Contact wear CMP models have been developed by Chekina [6] and Yoshida [7], and have been re-implemented for use in this work.

The concept behind the wafer-pad contact mechanics model is to relate the local pressures on the pad and wafer with the displacements of the pad, assuming the pad to behave as an elastic material. Figure 7 depicts the contact mechanics problem. The pressure-displacement equation can be discretized into matrix format, and used in an iterative scheme to calculate the amount of film removed over time.

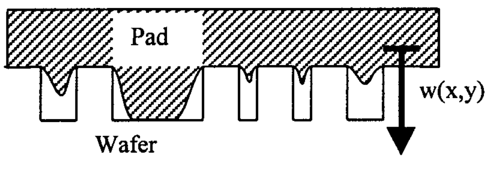

Figure 7. Wafer-pad contact diagram. Initially, either pressure or displacement is known at every point. Unknown pressures can be calculated by solving the pressure-displacement equation.

The basic pressure-displacement equation is given by Equation 3:

$$w(x,y)=[(1-v^2)/(\pi E)]\int_{\omega} p(\xi,\eta)*[(x-\xi)^2+(y-\eta)^2]^{-1/2}d\xi d\eta \qquad (3)$$

where w is the displacement of the pad at lateral position (x,y), p is the pressure at a point (ξ,η), v is Poisson's ratio, E is Young's modulus, and ω is the contact area of the pad to the wafer.

Yoshida [7] describes how boundary element methods may be used to discretize this formula. For a given initial wafer surface profile, displacement is known for some set of locations (where the pad is in contact with the wafer), and pressure is known for a complementary set of locations (zero where the pad is not in contact with the wafer). Yoshida also describes an algorithm to solve the initial matrix equation that specifies the known pressures and displacements for the pressures at every location on the wafer. Once the local pressure is known everywhere, the removal rate can be calculated using Preston's equation:

$$RR = K_p\, p\, v \qquad (4)$$

where K_p is Preston's coefficient, p is the local pressure of the pad on the wafer, and v is the relative velocity of the pad and wafer. Once the removal rate is known, it is possible to increment the time step, calculate the amount removed, calculate the new wafer surface profile, and iterate using Equations 3 and 4 to calculate the new pressure profile until the desired polish time has been reached.

4. Simulation Methodology

Nanotopography on a wafer is simulated by creating a millimeter-scale grid of cylindrical structures of random millimeter-scale diameters, as illustrated in Figure 8. The height of the structures is assumed to be 400 Angstroms (as was illustrated in Figure 2). The nanotopography length, defined as the grid spacing of the cylindrical structures, mimics the lateral length scale of typical wafer nanotopography. A 1 μm conformal oxide film is assumed deposited on the simulated topography.

CMP of the film is simulated using the Smith model described earlier for various combinations of nanotopography length and planarization length. Nanotopography lengths of 2 mm, 5 mm, and 8 mm are used, along with planarization lengths of 2 mm, 5 mm, and 8 mm. A blanket removal rate of 1900 Angstroms is assumed. Grillaert [8] notes that exponential time constant parameter values extracted from data are in the range of 45 to 70 seconds. For this work, a τ of 50 seconds is assumed. A polish time of 60 seconds is used for all simulations.

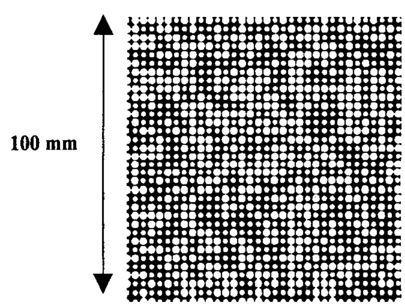

Figure 8. Example nanotopography simulation scenario: an array of random millimeter-scale features on a specified grid.

Simulation of the polish is also performed using the contact wear model, with various combinations of nanotopography length and

Young's modulus. Adjustment of the Young's modulus of the pad will vary the planarization length of the CMP process. While a direct correlation between Young's modulus and planarization length cannot be done, a general comparison (i.e., stiffer pads should result in longer planarization lengths, as shown in [9]) is considered here.

Yoshida [7] specifies Young's modulus parameters of 147 MPa and Preston's coefficient of 1.32 x 10^{-7} 1/MPa for his simulation of silicon dioxide polish; these numbers have been used for the simulation here. In addition, Young's moduli of 72 and 294 MPa are simulated. These three Young's moduli (different pad stiffness) cases are simulated for nanotopography lengths of 2 mm, 5mm, and 8 mm. A nominal pressure of 34.3 kPa and velocity of 700 mm/sec is used here (also from [7]). A Poisson's ratio of 0.2 has also been used. A polish time of 60 seconds is used for all simulations. For these values, the result is a blanket removal rate of 1900 Angstroms per minute.

5. Simulation Results

The results from the Smith model analysis, shown in Figure 9, illustrate that oxide thinning occurs more significantly when the planarization length is much greater than the nanotopography length. Specifically, it can be seen that for planarization lengths much smaller than the nanotopography length, no thinning of the surface film is predicted. Also, it can be noted that for planarization lengths much larger than the nanotopography length, more than half of the initial step height of 400 Angstroms is thinned away.

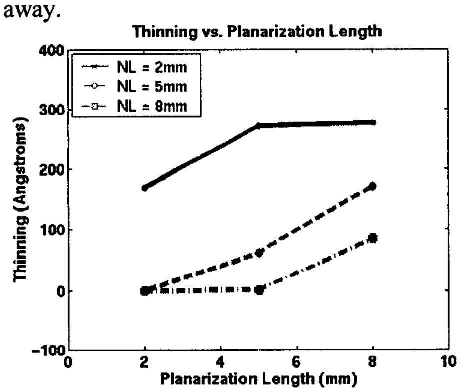

Figure 9. Thinning vs. Planarization Length. Model results are using the Smith model with a nominal τ value of 50 seconds.

Note that the increase in thinning is clear from the simulated results. Also note the significant increase in thinning, as the planarization length becomes comparable to and larger than the nanotopography length. The model predicts that the longer the planarization length, the larger the amount of thinning that occurs.

The results from the contact wear analysis, given in Figure 10, show that for stiffer pads (larger Young's modulus), thinning occurs to a greater degree than with lower Young's moduli. This agrees well with our initial hypothesis that the results using a stiffer pad and a longer planarization length should be comparable based on previous studies [9]. Here, the argument can be made that a stiffer pad does not deform to contact low areas and so preferentially removes the raised nanotopography areas, resulting in the thinning of films in these areas.

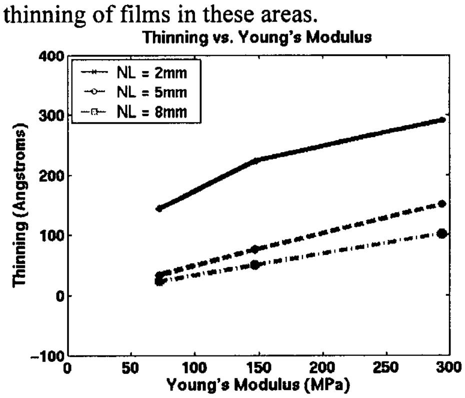

Figure 10. Thinning vs. Young's modulus, using a contact wear model with nominal Preston's coefficient value of 1.32×10^{-7} and Poisson's ratio of 0.2.

Simulation of a longer polish (120 seconds) was performed using a contact wear model for the 5 mm nanotopography length case. The final amount of thinning is shown in Figure 11.

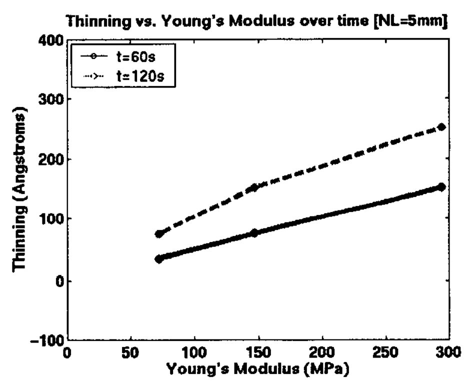

Figure 11. Thinning over time, using contact wear modeling.

The simulation results show that thinning increases significantly with time. This is to be expected since a CMP process nominally planarizes over time until step heights are removed. This increased thinning may be an issue for STI processes with large initial oxide thicknesses that require more polish time.

These results indicate a possibly conflicting set of requirements. Nominally, industry has been moving toward the stiffer CMP pad types, which result in processes with long planarization lengths, and thus better within-die uniformity. However, the results of this simulation indicate that using such processes can lead to potential problems with yielding of STI devices, as excessive thinning of nitride can result in device degradation or failure. The amount of thinning that results is a function of the initial step height (i.e. height of nanotopography), the planarization length of the CMP process (or similarly, the material properties of the CMP pad used), and the lateral length scale of the nanotopography on the initial wafers.

6. Implications of nanotopography interaction with STI CMP

In the CMP of dielectric films, it has been shown that variations in pattern densities over a die can cause areas to polish at different rates, leading to post-CMP height variation [4]. For STI processes, this can lead to uneven clearing, excessive nitride thinning (also known as nitride erosion), and trench oxide dishing (see Figure 12). This density effect problem can be approached by engineering the nitride/oxide slurry selectivity during the CMP process. This creates a low nitride removal rate while the remaining oxide overburden is removed from above active areas.

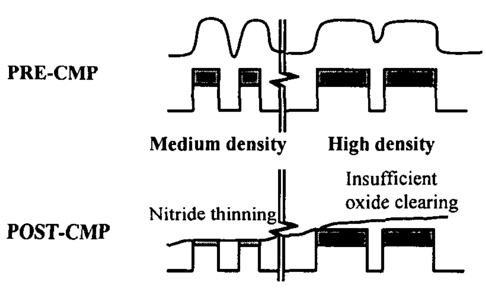

Figure 12. Density effect in STI CMP

The CMP process itself can achieve good local planarity but generates global non-planarity due to varying pattern densities over a die. In other words, local step heights are nominally planarized, but the density mismatch causes non-planarity over longer length scales across the die and wafer. Figure 13 illustrates this concept, pictured for the case of topography over top of patterned metal features. After local step height removal, the higher density feature area becomes thicker than the lower density feature region.

The length scale over which this global non-planarity exists is the planarization length [3], as originally illustrated in Figure 3. This is the same parameter used in the Smith model to calculate effective feature density. Any feature smaller than the planarization length will be planarized, given sufficient polish time; features larger than the planarization length will result in little or no planarization.

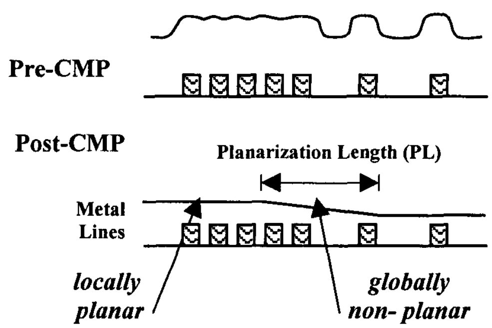

Figure 13. Local vs. global planarity in a CMP process.

It is conjectured that the comparative length scales of a CMP process (planarization length) and the nanotopography of wafers to be used in that CMP process determine whether a particular wafer film will planarize (thin) under that CMP process. The thinning that results from this interaction can potentially have serious implications in STI CMP, and may result in device failure in extreme cases.

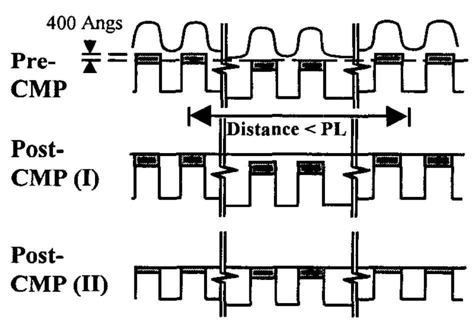

Figure 14. Nanotopography can lead to incomplete oxide clearing (I), or nitride overpolish (II)

Figure 14 illustrates two potential problems that the thinning effect can cause in STI CMP. The figure shows two raised nanotopographical areas surrounding a low nanotopographical area.

If the nanotopography length is much less than the planarization length of the CMP process, the thinning that results can be problematic in two ways.

First, if the polish is stopped when one just begins to clear the oxide over the nitride in the raised nanotopographical areas, then this may result in incomplete oxide clearing over the low areas. Second, if one polishes long enough to ensure clearing through to the nitride in the low areas, then this results in overpolish and thinning of the nitride in the raised areas.

In both cases, device yield may be a concern; if the oxide is incompletely cleared, then the nitride mask layer may not be completely removed during subsequent processing. Conversely, if the nitride is thinned by a significant amount (or removed completely), then device degradation, or in the worse case, device failure, may result.

This analysis is worst-case since it neglects low area polish that may result due to the compressibility of the pad (see [8]). Thus it may be possible to achieve the desired low nanotopographical area polish depending on the properties of the consumable set used. This low area polish is similar to the dishing phenomenon described in dual material polishes, where height variations are introduced due to the difference in polish rates of two materials that are polished simultaneously.

Nanotopography effects are topographical in nature, and thus may not be best approached by engineering CMP slurry selectivity. Indeed, high slurry selectivity may be detrimental in this case since it will lead to a longer polish time to clear the oxide in low nanotopographical areas. Infinite selectivity is worse because the oxide in the low areas may be even more difficulty to clear. In the ideal case, infinitely selectivity means that the oxide in low areas will never clear because the nitride in the raised areas will always support the pad. Again, this assumes negligible low area polish due to pad compressibility.

7. Conclusion

It has been proposed that nanotopography introduces a new topographical effect in STI CMP that must be considered in addition to pattern density, dishing, and erosion effects, and may not necessarily benefit from approaches typically applied to these problems.

In addition, it has been proposed that the relative length scales of the CMP process and the nanotopography on a raw wafer may affect the degree of yield for STI devices on that wafer.

Simulation of the CMP polish of silicon dioxide using a density/step height model for various combinations of planarization length and nanotopography length agrees with this hypothesis. Simulation using a contact wear model for various combinations of Young's moduli and nanotopography lengths shows that stiffer pads result in more significant film thinning, as expected. In addition, thinning seems to increase with time, this may have implications on STI processes with large initial thicknesses, since such processes may require longer polish times.

The thinning that results can have implications in the STI process in the form of device degradation and potential failure. Analysis of the relative length scales of the wafer nanotopography and potential CMP processes may help in CMP process consumable and parameter selection.

There is much future work that will be explored in this area. First, there is a need for experimental work on the simultaneous characterization of planarization length and nanotopography evolution during polish. These experiments can serve to validate the simulation results presented in this work.

In addition, the simulation results shown here are for thinning of a single material; since STI CMP involves a dual-material polish (both oxide and nitride), the model will need to be extended to include the accurate prediction of nanotopography impact on nitride thinning during STI CMP.

Acknowledgments

We would like to thank Dale Hetherington at Sandia National Laboratories for helpful discussions on this topic.

References

[1] S. Xu, "Effects of Silicon Front Surface Topography on Silicon Oxide Chemical Mechanical Planarization," *Electrochem. and Sol. State Lett.*, vol. 1, no. 4, pp. 181-182, 1998.

[2] K.V. Ravi, *Future Fab International*, no. 7, pp. 207.

[3] D. Ouma, Modeling of Chemical Mechanical Polishing for Dielectric Planarization, *Ph.D. Thesis*, MIT, Nov. 1998.

[4] B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung, D.L. Hetherington, I. Ali, G.

Shinn, J. Clark, O.S. Nakagawa, S.-Y. Oh, "A closed form analytic model for ILD thickness variation in CMP processes," *Proc. CMP-MIC Conf.*, pp. 266-273, Feb. 1997.

[5] T. Smith, S. Fang, D. Boning, G. Shinn, J. Stefani, "A CMP model combining density and time dependencies," *Proc. CMP-MIC Conf.*, pp. 97-104, Feb. 1999.

[6] O.G. Chekina, L.M. Keer, "Wear-contact problems and modeling of chemical mechanical polishing," *J. Elec. Soc.*, vol. 145, no. 6, pp. 2100-2106, June 1998.

[7] T. Yoshida, "Three-dimensional chemical mechanical polishing process model by BEM," *ECS Conf.*, Oct. 1999.

[8] J. Grillaert, M. Meuris, N. Heylen, K. Devriendt, E. Vrancken, M. Heyns, "Modelling step height reduction and local removal rates based on pad-substrate interactions," *Proc. CMP-MIC Conf.*, pp. 79-86, Feb. 1998.

[9] R. Jin, J. David, B. Abbassi, T. Osterheld, F. Redeker, "A production-proven shallow trench isolation (STI) solution using novel CMP concepts," *Proc. CMP-MIC Conf.*, pp. 314-321, Feb.1999.