Nanotopography Issues in Shallow Trench Isolation CMP

Duane Boning and Brian Lee

Abstract

As advancing technologies increase the demand for planarity in integrated circuits, nanotopography has emerged as an important concern in shallow trench isolation (STI) on wafers polished by means of chemical–mechanical planarization (CMP). Previous work has shown that nanotopography—small surface-height variations of 10–100 nm in amplitude extending across millimeter-scale lateral distances on virgin wafers—can result in CMP-induced localized thinning of surface films such as the oxides or nitrides used in STI. A contact-wear CMP model can be employed to produce maps of regions on a given starting wafer that are prone to particular STI failures, such as the lack of complete clearing of the oxide in low spots and excessive erosion of nitride layers in high spots on the wafer. Stiffer CMP pads result in increased nitride thinning. A chip-scale pattern-dependent CMP simulation shows that substantial additional dishing and erosion occur because of the overpolishing time required due to nanotopography. Projections indicate that nanotopography height specifications will likely need to decrease in order to scale with smaller feature sizes in future IC technologies.

Keywords: *chemical–mechanical planarization, chemical–mechanical polishing, CMP, contact-wear simulation, nanotopography, shallow trench isolation (STI), silicon wafers.*

Introduction

Nanotopography is a term used to describe relatively gentle (10–100-nm) surface-height variations occurring over lateral distances of 1–10 mm on unpatterned silicon wafers,[1] as shown in Figure 1. Conformal thin films deposited on wafers with nanotopography have been shown to exhibit localized deviations in film thickness after undergoing chemical–mechanical polishing, or planarization (CMP), due to the planarizing action of the CMP pad and process (Figure 2).[2–4] These deviations are of particular concern in modern integrated-circuit shallow trench isolation (STI) structures fabricated using the CMP process.

Figure 3a depicts a typical trench structure used to isolate "active" regions where devices will be built. The nitride layer has been patterned and a shallow trench etched into the silicon. An oxide has then been deposited into the trench, which also results in overburden oxide above the nitride active areas. In the ideal CMP process, the oxide is removed completely in all active regions, leaving oxide only in the trench regions (Figure 3b). Three key failure mechanisms may arise, however, as shown in Figures 3c–3e. First, the CMP process may fail to completely clear the overburden oxide. Second, the process may succeed in clearing the oxide, but result in an excessive removal (erosion) of the nitride film. Finally, excessive removal of the oxide (dishing) within the trench may occur.

In this review, we focus on how nanotopography may exacerbate these STI CMP failure mechanisms. We first summarize a simulation methodology that can be used to understand how a wafer surface with underlying nanotopography will evolve during CMP. A blanket wafer simulation, using an unpatterned wafer with a conformal ("blanket") film layer, is then used to study the first two failure mechanisms. Finally, additional dishing and erosion in patterned STI wafers due to nanotopography are simulated, using a pattern-dependent STI CMP model.

Contact-Wear Model for Nanotopography/CMP Interaction

Our first goal is to understand how nanotopography across large regions of the wafer (here, the center 100-mm region of a wafer is simulated) affects the polishing of a blanket oxide/nitride film stack, as illustrated in Figure 4. A variety of modeling approaches have been proposed to convert initial wafer nanotopography to final post-CMP film-thickness information, including statistical measures,[5] simple scal-

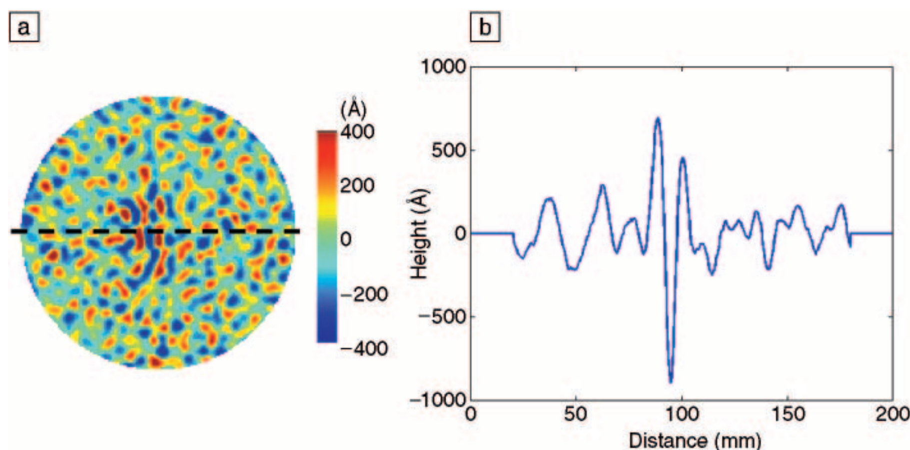

Figure 1. (a) Top view and (b) cross-section graph of wafer nanotopography. Dotted line in (a) shows path of scan. The x axis in (b) indicates the distance along the scan path in (a), moving from left to right.

ing models,[6] and filter-based approaches.[4,6] While these simplified methods can predict the results of CMP on a single-material oxide over nanotopography, it is not clear how to extend these to the polishing of STI stacked films, where both oxide and nitride are polished simultaneously with different removal rates. Instead of these approaches, we have found it best to use a contact-wear model for nanotopography that can naturally account for film stacks.[6]

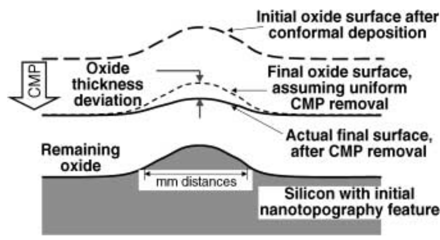

Figure 2. Film thinning resulting from the chemical–mechanical planarization (CMP) of conformal films over the wafer's nanotopography.

We begin with a contact-mechanics model, as described by Chekina and Keer[7] and Yoshida.[8] The concept is to relate local pressures on the wafer surface to the polishing-pad displacement and use the local pressures to derive local removal rates, assuming that the polish rate is linearly proportional to the pressure. As the film surface evolves through time, the pad displacement changes, and thus local pressures and removal rates are modified. The simulation surface is sectioned into discrete elements, and a time-stepped algorithm is used to determine the final post-CMP film surface.[6]

Treating the pad as a linear elastic material, the relationship between pad displacement and pad pressure is given by

$$w(x,y) = \frac{(1-\nu^2)}{\pi E}\iint \frac{p(\xi,\eta)}{\sqrt{(x-\xi)^2+(y-\eta)^2}}\,d\xi\,d\eta, \quad (1)$$

where $w(x,y)$ is the displacement of the pad at point (x,y) on the wafer surface, $p(x,y)$ is the local pressure of the pad, ν is the Poisson ratio of the pad, and E is the Young's modulus of the pad. Yoshida[8] offers a boundary element method to solve the matrix form of this equation. This is useful for the general case where the pad displacement and pad pressures are not known completely. While the method is effective in accurately computing pressures and displacements, it suffers from a long computation time for large simulation domains.

Here, we use a special case that takes advantage of fast Fourier transform (FFT) analysis to efficiently compute the pad pressure given the pad displacement. We assume full pad–wafer contact everywhere, so that $w(x,y)$ is fully specified initially and is known at all subsequent points in time just prior to computation of the corresponding $p(x,y)$. At each time step, Equation 1 is then seen to be a convolution of the pressure and an inverse distance kernel (i.e., impulse response); this convolution becomes a multiplication after Fourier transform, enabling rapid calculation of the pressure distribution at each time step.

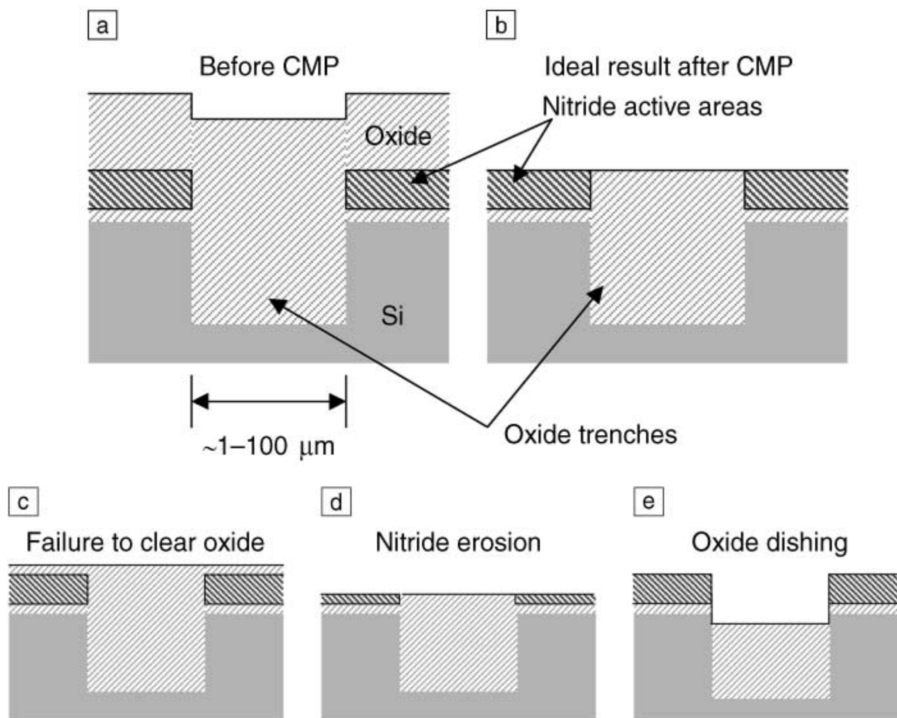

Figure 3. (a) Initial structure before CMP: typical trench isolation structure used to isolate "active" regions on a wafer where devices will be built. The nitride layer has been patterned and a shallow trench etched into the silicon. An oxide has then been deposited into the trench, which also results in overburden oxide above the nitride active areas. (b) Ideal result after CMP: the oxide is removed completely in all active regions, leaving oxide only in the trench regions. Three key failure mechanisms may arise: (c) failure to clear oxide from nitride active areas, (d) excessive removal (erosion) of nitride in active areas, and (e) excess removal of oxide (dishing) within the trench.

Experimental Description and Simulation Results

Experimental data are taken from previously reported studies on nanotopography.[3] The experiment used sets of 200-mm epitaxial silicon wafers, each set containing wafers with five different nanotopography signatures. Starting wafers have 1 μm of thermal oxide grown on them. Each set of wafers is polished using a different CMP process consisting of a specific combination of tool, consumable set, and process conditions.

For the purposes of tuning the model, data from three different CMP processes are used. The planarization lengths (as defined by Ouma et al.[9]) of these three CMP processes are 3.4 mm (a "soft" pad process), 6.4 mm (a "medium" pad process), and 9.7 mm (a "stiff" pad process). Initial nanotopography height (NH) data for a given wafer are measured on a NanoMapper (an automated surface-mapping system) and filtered with a 20-mm double Gaussian filter to remove long-range wafer-geometry effects. Oxide thickness deviation (OTD) data are obtained by computations from pre- and post-CMP measurements performed using a full-wafer mapping tool called AcuMap. Nanotopography and OTD data are interpolated onto a common grid with spacings of 0.2 mm. The OTD data are then also filtered using a 20-mm double Gaussian filter to remove wafer-scale CMP trends.

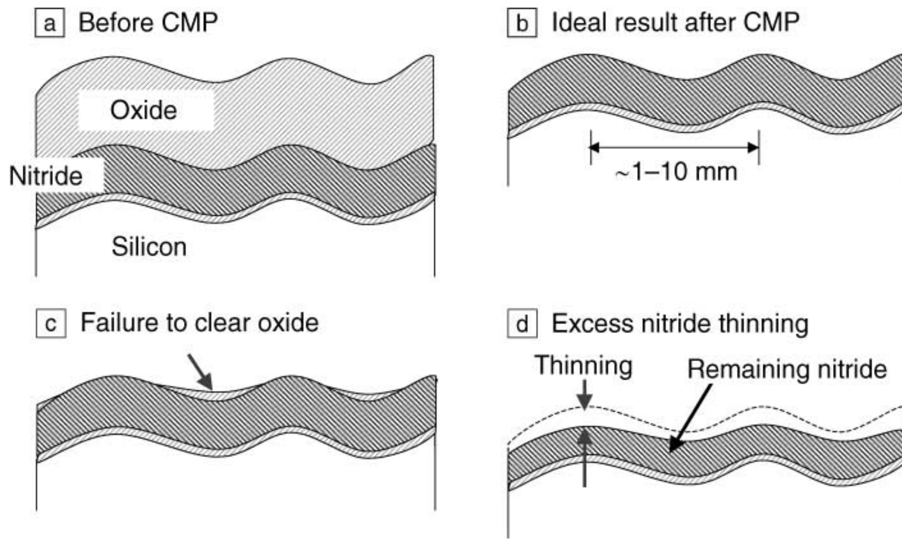

Figure 4. Key concerns in STI CMP over nanotopography. (a) Before CMP, the oxide film is uniform in thickness, but varies in surface height. (b) The ideal result after CMP is the complete removal of oxide, with no removal of nitride. (c) Failure to clear the oxide occurs when CMP is unable to remove the oxide layer completely in low regions created by nanotopography. (d) Overpolishing to clear the oxide results in excessive localized nitride thinning.

In addition, the OTD data are shifted and rotated into alignment with the nanotopography data. In the simulation analyses to follow, only the central 100 mm × 100 mm region of the wafer data set is used, in order to focus on the effect of nanotopography and exclude any possible effects due to wafer edge proximity.

Contact-Wear Model Results

The contact-wear model is fit by minimizing the square error between model prediction and measurement data for the polishing of a measured nanotopography wafer. The fitting extracts an effective pad modulus E, as well as an overall mean centering or removal-rate parameter. Depending on the pad and process used, typical extracted pad moduli range from 20 MPa to 900 MPa, with rms fitting errors of 1–2.8 nm.[6] An example set of cross sections is shown in Figure 5 with an extracted E of 147 MPa, illustrating both the "transmission" of nanotopography into oxide removal and the good accuracy achieved with the fitted model.

Once tuned, the contact-wear model can be used to predict both oxide and nitride film thicknesses as a function of time for other measured nanotopography wafer maps.[10] In the STI stack simulations reviewed here,[10] we assume a 300-nm oxide film over a 200-nm nitride film. A 5:1 oxide:nitride removal-rate selectivity ratio is assumed.

Problem 1: Failure to Clear Oxide

The first application of the contact-wear model is to understand the failure to clear the oxide fully in some regions on the wafer. This results in immediate device failure because the nitride will not be stripped as required in non-cleared regions, preventing transistor formation in the corresponding active areas. A simulation is first performed, assuming a fixed CMP overpolishing time of 14 s beyond the time at which the oxide just begins to clear somewhere on the wafer. As seen in Figure 6b, this time is too short to completely clear the oxide everywhere. Careful examination of the failure map shows that the clearing failure spots correspond to low regions on the initial nanotopography map.

Problem 2: Excess Nitride Thinning

Rather than risk device failure due to incomplete clearing of the oxide, CMP processes employ enough overpolishing time, beyond an end-point signal triggered when the oxide film just begins to clear, to ensure completion of clearing across the entire wafer. Here, we run the simulation until the oxide has just cleared everywhere (corresponding to "perfect" knowledge of the necessary overpolishing time). Unfortunately, this overpolishing results in localized thinning of the nitride film. Excessive thinning of the nitride can affect the shape of the trench and active area corners and the depth of the oxide recess, leading to device degradation.[11] We will assume a budget of 10% of the initial 200-nm nitride film that can be polished away before degradation occurs. Shown in Figure 6c are those regions where nitride in excess of this budget has been

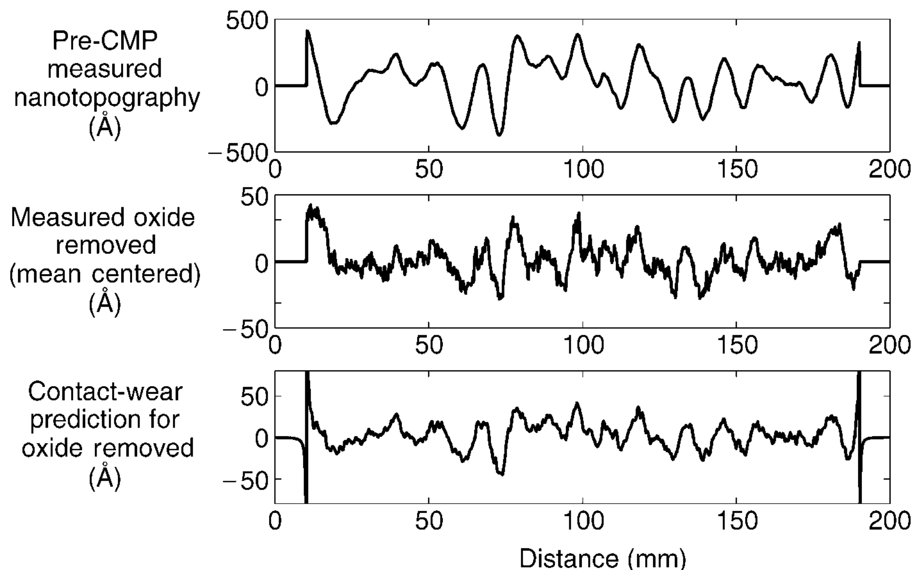

Figure 5. Cross sections through example wafer, showing pre-CMP nanotopography, post-CMP oxide thickness removed, and the contact-wear prediction, based on the simulation for the amount of oxide removed. The x axis is the position along a selected diameter scan across the wafer.

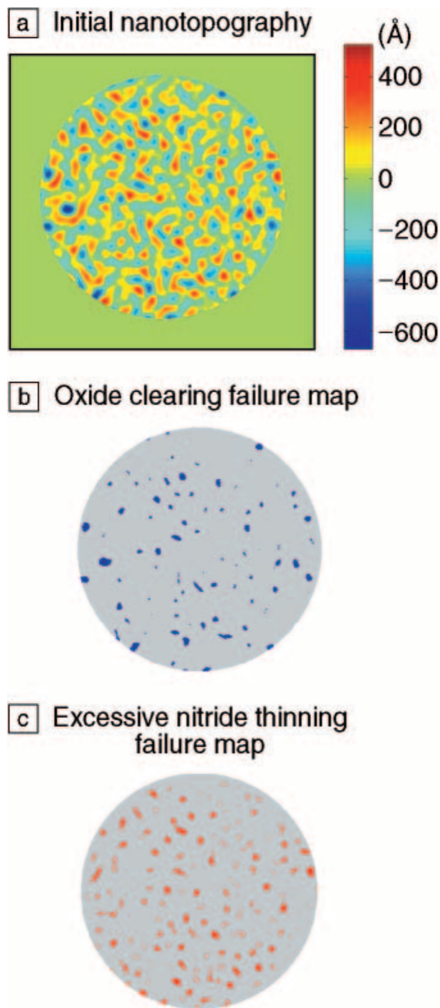

Figure 6. Simulation of potential device-failure areas. (a) Nanotopography height map for unpolished wafer. (b) Oxide-clearing failure map; blue regions indicate where oxide remains for a specified polishing time. (c) Nitride-thinning failure map; red regions indicate where >10% of the original nitride thickness has been eroded for a given overpolishing time; those regions fail because too much of the nitride has been removed.

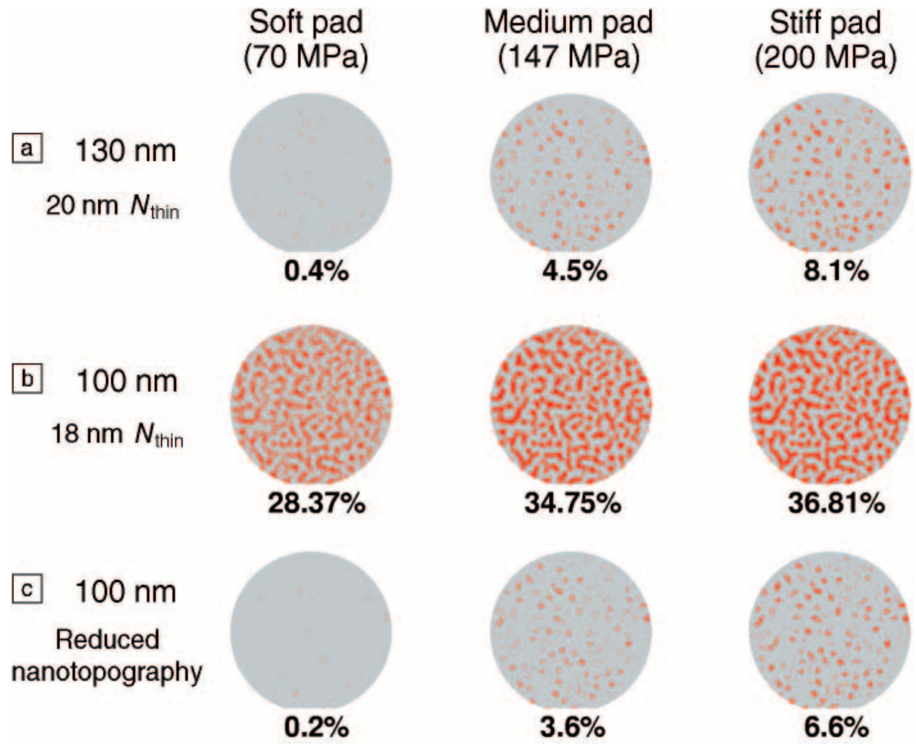

Figure 7. Nitride-erosion failure maps for three different pad stiffnesses (70 MPa, 147 MPa, and 200 MPa). (a) Failure areas on 130-nm technology wafers (130 nm is the minimum patterned feature size) corresponding to erosion (N_{thin}, the allowed amount of nitride thinning) greater than 20 nm. (b) Failure areas in scaled 100-nm technology corresponding to >18-nm nitride loss, assuming the same starting wafer as in (a). (c) Failure areas if the starting wafer's nanotopography height is also scaled by a 0.9 factor.

removed; thus, those regions fail. Comparison with the initial nanotopography map shows that these regions correspond to starting high spots, as expected.

In Figure 7, we repeat the simulation for three different extracted effective pad moduli (for three different CMP processes). We see that the nitride thinning and failure areas increase with pad stiffness. The current trend in STI CMP toward stiffer pads to deal with chip-scale pattern dependencies will increase the sensitivity of the process to nanotopography. This is seen in Figure 7, where we also consider a hypothetical STI scaling from 130-nm to 100-nm technology, where the nitride film thickness reduces by a factor of 0.9, from 200 nm to 180 nm.[10] With a 10% thinning budget, the excess nitride failure region increases dramatically. If we also scale the initial allowed nanotopography by a 0.9 factor, smaller failure areas are again observed. This suggests that as the nitride film scales with future technology generations, the nanotopography specification must also scale. A move toward stiffer pads, however, may require substantially tighter nanotopography specifications.

Problem 3: Additional Dishing and Erosion Due to Nanotopography

The final question we consider cannot be answered by the wafer-scale nanotopography model alone. It is well known that on patterned STI wafers, dishing into the trench and erosion of the nitride result from differences in layout pattern densities across the chip.[12] Substantial work has been devoted to the development of chip-scale, pattern-dependent models for the STI CMP process.[13] Here, we use such a model to investigate the additional dishing and erosion that occur in a hypothetical chip due to the overpolishing time required to clear the entire wafer based on its nanotopography.[13] An example chip-scale simulation of dishing and erosion is shown in Figure 8. The predicted dishing and erosion for a 45-s polish are shown in Figure 8b; with the additional 13 s of overpolishing required to completely clear a given wafer due to its nanotopography, the chip suffers the additional dishing and erosion shown in Figure 8c, corresponding to a 32% increase in dishing and a 130% increase in nitride erosion.

Conclusions

Nanotopography is an emerging concern in chemical–mechanical planarization processes for shallow trench isolation of integrated-circuit wafers. Not only must careful consideration be given to wafer-level CMP uniformity and chip-scale pattern-dependent uniformity, but the impact of millimeter-length nanotopography height variations must also be examined. Contact-wear and pattern-dependent CMP simulations indicate that this concern is

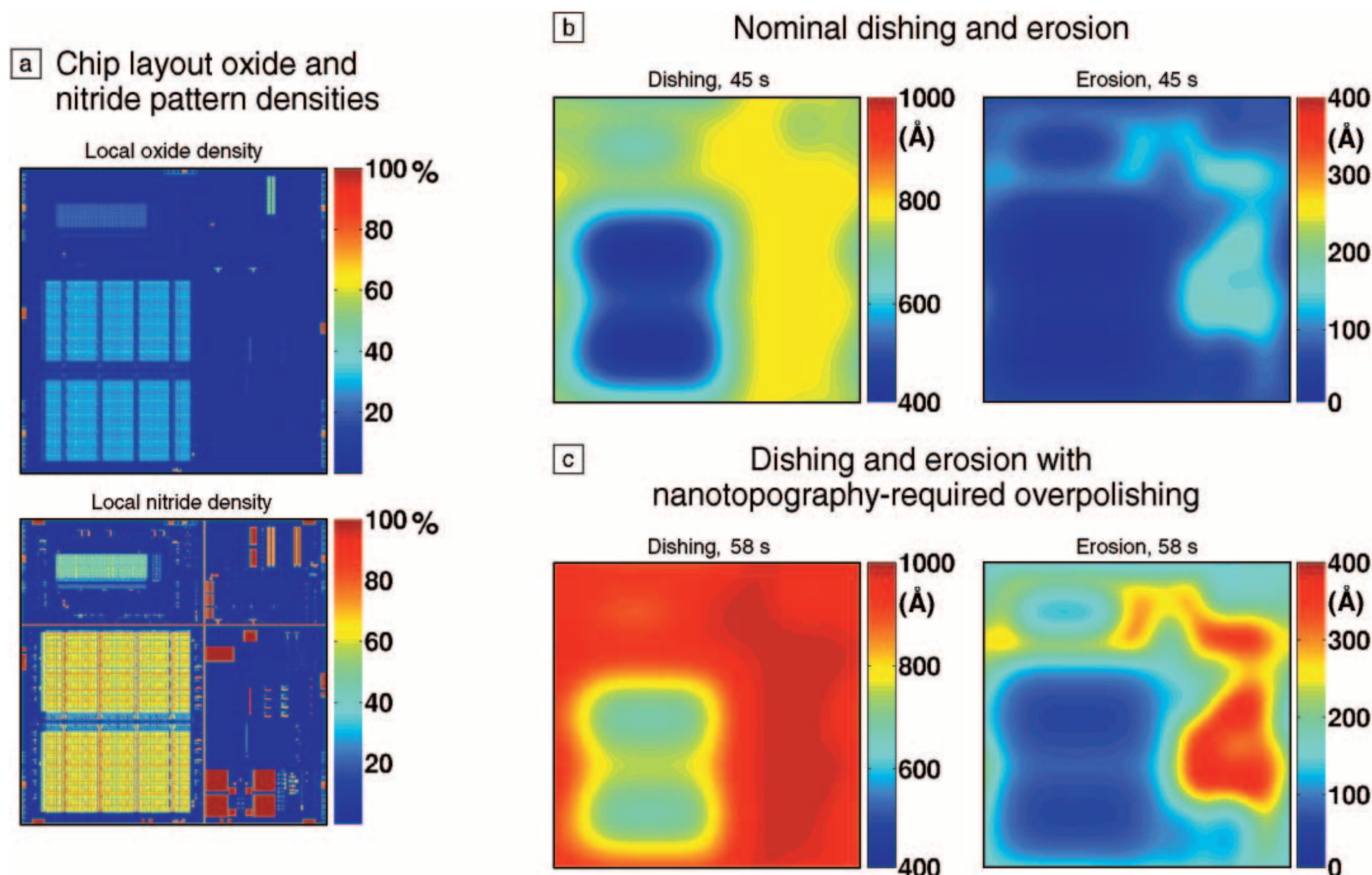

Figure 8. Chip-scale simulations of pattern-dependent dishing and erosion. (a) Test chip layout pattern densities with static random-access memory and other features. (b) Predicted dishing and erosion due to layout pattern density effects for a 45-s polish. (c) Predicted dishing and erosion after an additional 13-s overpolish required to clear nanotopography across the wafer.

most likely to impact STI processes that use stiffer pads. More work is needed to define wafer specifications for nanotopography and to more fully integrate nanotopography and STI pattern simulations.

Acknowledgments

The work at MIT on nanotopography has been supported by ADE Corp. We appreciate the long-term collaboration in this research with Noel Poduje, Yong Xia, Pat Hester, and others at ADE Corp.; Win Baylies of BayTech Group; and John Valley of ADE Phase Shift.

References

1. K.V. Ravi, *Future Fab Int.* **7** (1999) p. 207.
2. C.S. Xu, E. Zhao, R. Jairath, and W. Krussell, *Electrochem. Solid-State Lett.* **1** (4) (1998) p. 181.
3. B. Lee, D.S. Boning, W. Baylies, N. Poduje, P. Hester, Y. Xia, J. Valley, C. Koliopoulos, D. Hetherington, H. Sun, and M. Lacy, in *Chemical–Mechanical Polishing 2001—Advances and Future Challenges*, edited by S.V. Babu, K.C. Cadien, and H. Yano (Mater. Res. Soc. Symp. Proc. **671**, Warrendale, PA, 2001) p. M4.9.1.
4. J.-G. Park, T. Katoh, H.C. Yoo, and J.-H. Park, *Jpn. J. Appl. Phys., Part 2: Lett.* **40** (2001) p. L857.
5. R. Schmolke, R. Deters, P. Thieme, R. Pech, H. Schwenk, and G. Diakourakis, *J. Electrochem. Soc.* **149** (4) (2002) p. G257.
6. B. Lee, D.S. Boning, W. Baylies, N. Poduje, and J. Valley, "Modeling and Mapping of Nanotopography Interactions with CMP," presented at Symposium I, Materials Research Society Meeting, San Francisco, April 2002, paper No. I1.5.
7. O.G. Chekina and L.M. Keer, *J. Electrochem. Soc.* **145** (6) (1998) p. 2100.
8. T. Yoshida, in *Proc. 3rd Int. Symp. on Chemical–Mechanical Planarization in IC Device Manufacturing*, Vol. 99-37 (The Electrochemical Society, Pennington, NJ, 1999) p. 593.
9. D. Ouma, D. Boning, J. Chung, G. Shinn, L. Olsen, and J. Clark, in *Proc. Int. Interconnect Technology Conf.* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1998) p. 67.
10. D. Boning, B. Lee, W. Baylies, N. Poduje, and J. Valley, "Impact of Nanotopography on STI CMP in Future Technologies," presented at the Workshop on Metrology for Silicon Wafers for 100-nm Technology Generations and Beyond, SEMI/Europa, Munich, April 2002.
11. D.P. Shum, J.M. Higman, M.G. Khazhinsky, K.Y. Wu, S. Kao, J.D. Burnett, and C.T. Swift, in *IEDM Tech. Dig.* (Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1997) p. 665.
12. J.T. Pan, D. Ouma, P. Li, D. Boning, F. Redecker, J. Chung, and J. Whitby, "Planarization and Integration of Shallow Trench Isolation," presented at the VLSI Multilevel Interconnect Conf., Santa Clara, CA, June 1998.
13. B. Lee, PhD thesis, Massachusetts Institute of Technology, May 2002. □

The Materials Gateway: www.mrs.org