Modeling of Planarization in Chemical-Mechanical Polishing

Leonard J. Borucki[1]

1 Introduction

Chemical-mechanical polishing (CMP) is a commonly used procedure for planarizing material layers that have been deposited on silicon wafer surfaces. Planarization considerably simplifies the task of building multilayer integrated circuit structures while staying within the depth-of-focus limitations of photolithography tools. First applied to microelectronics by IBM in the early 1980s [1], CMP is one of the fastest growing and most essential processing techniques in the electronics industry.

In one large class of single wafer CMP tools, the wafer is held upside down by a rotating wafer carrier while being pressed against a soft rotating polishing pad as shown in Figure 1. During polishing, a chemically-reactive slurry containing a small weight fraction of fine (0.1 micron, where 1 micron =1e-6 m) abrasive particles is sprayed on the pad ahead of the wafer. Usually, a diamond-covered rotating disk called a conditioner is also swept back and forth radially across the pad. The purpose of the conditioner is to refresh and maintain the roughness of the pad surface. Pad surface asperities (explained below) that are tall enough to touch the wafer trap abrasive particles and drag them across the surface. This abrasive action, combined with chemical attack of the exposed wafer surface by the slurry, is thought to be responsible for polishing. Although the broad outlines of the mechanisms underlying the CMP process are qualitatively understood, many aspects of the process are not understood in detail and have not been modeled.

A magnified cross section of a commonly used type of void-filled polyurethane polishing pad is shown in Figure 2. Rather than being ideally smooth, the pad surface is covered with a large number of voids and bumps, or asperities. The voids average 30 microns in diameter and occupy about half of the volume of the pad. The surface can be statistically characterized by dragging a fine needle across the surface with a profilometer and sampling the height of the needle tip. A typical profilometer measurement is shown in Figure 3. It may be seen that the height of the pad surface has variations of tens of microns. The pad surface height pdf is often not Gaussian but can usually be described by one of the Pearson family of distributions, which are solutions of

$$\frac{\mathrm{d}\phi}{\mathrm{d}z} = \frac{z-a}{b_0 + b_1 z + b_2 z^2}\phi \tag{1}$$

for suitable values of a, b_0, b_1 and b_2.

---

[1] * Digital DNA Laboratory, Motorola, Inc., Mesa, Arizona 85202.

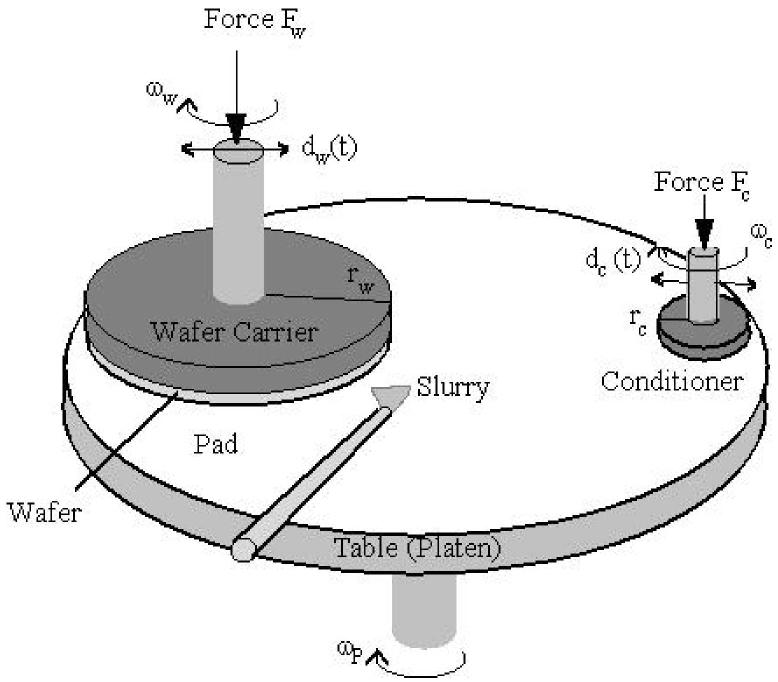

Figure 1: Schematic of a typical single-wafer CMP tool.

Since the pad surface is very rough, the contact mechanics between the pad and the wafer is not easily described using linear elasticity theory. However, Greenwood and Williamson [2] developed a statistical theory of contact based on a solution derived by Hertz [4] for the contact between a sphere and a planar surface. In the Greenwood and Williamson theory, the relationship between the average contact pressure p and the separation distance d between a rough surface and a smooth one is described by

$$p(d) = 4/3\eta E^*/\kappa_s^{1/2} \int_d^\infty (z-d)^{3/2}\phi(z)dz \tag{2}$$

where η is the asperity areal density and κ is the asperity tip mean curvature. The parameter E^* is the Young's modulus divided by $1-\nu^2$ where ν is the Poisson ratio. One implication of the Greenwood and Williamson theory is the actual area of contact between a rough surface and a smooth one is much smaller than the nominal area and that the actual asperity contact pressure is much higher than the nominal pressure.

2 Trench Planarization

One important quality of a CMP process is the rate at which it can planarize an irregular surface. This rate can be characterized using isolated rectangular

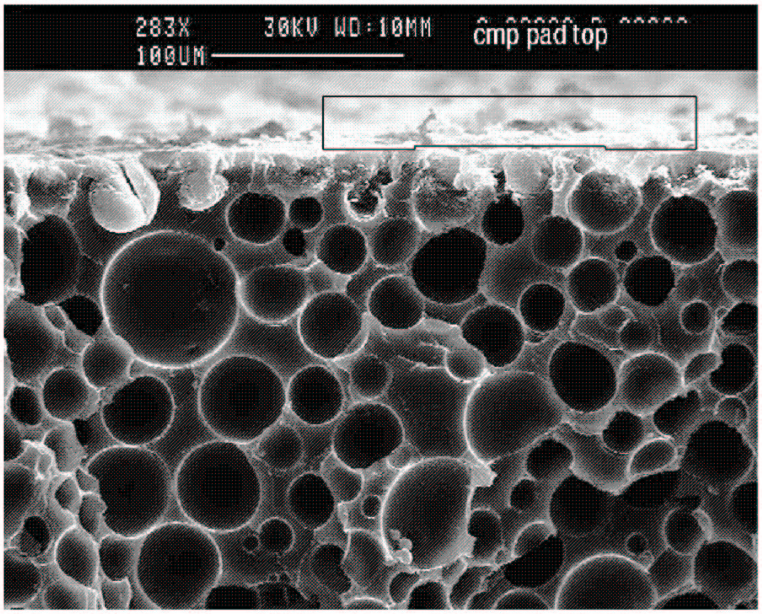

Figure 2: Scanning Electron Micrograph (SEM) cross section of a used, conditioned void-filled polyurethane polishing pad. Surface asperities can be seen at the top of the image. The scale bar at the top center is 100 microns (0.1 mm) long. Voids average about 30 microns in diameter and occupy about 50% of the material volume. Shown for comparison is a sketch of a trench 100 microns wide and 2 microns deep. (Data by Letitia Malina, Motorola)

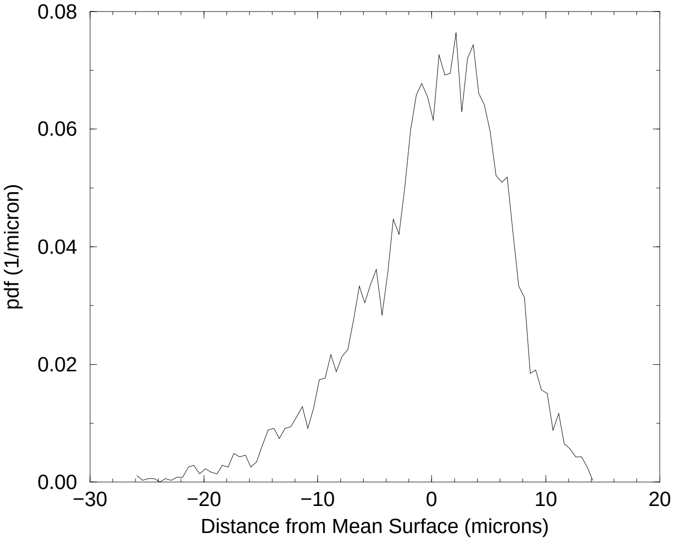

Figure 3: Surface height probability distribution obtained from a profilometry scan of a conditioned pad surface. Heights are referenced to the mean value, which is taken to be zero. (Data from Georgia Tech)

trenches of various widths and depths that have been etched into a thin film (Figure 4). The rate of removal of the film surface around the trench is always higher than the rate of removal at the bottom of the trench. This effect leads to gradual reduction of the trench depth and increased overall planarity. Experimentally, the removal rate can often be described by Preston's law [3], which states that the removal rate is proportional to the product of the nominal contact pressure between the pad and the wafer and the relative sliding velocity.

Figure 5 shows a typical set of measurements of the amount removed at both the trench bottom and the surrounding area during a 90 second polish of silicon dioxide at an average applied nominal contact pressure of 28,000 Pa. Typically, at a fixed depth, the amount removed at the bottom increases as the trench width increases, becoming nearly equal to the far field removal amount when the trench is very wide. Figure 6 shows a log plot of the same data as a function of trench depth for various trench widths. It can be seen that the amount removed at the trench bottom appears to decrease exponentially with the trench depth.

The objective of the group working on this problem will be to try to explain several of the features of this data:

- Why does the removal rate at a trench bottom decrease exponentially with depth?
- Why does the removal rate become nearly independent of trench width for small widths?
- The bottom removal curves in Figure 5 all appear to have a maximum at around 6 microns. Is this real?
- The data ends at 0.1 mm, yet most trenches in real circuits have widths in the 0.1 to 10 micron range. What would be expected to happen for such narrow trenches?

References

[1] K. Beyer, "The Inception of Chemical-Mechanical Polishing for Device Applications at IBM", MicroNews, Vol. 5, No. 4. http://www-3.ibm.com/chips/micronews/vol5_no4/beyer.html.

[2] Greenwood, J. A. and Williamson, J. B. P. Contact of nominally flat surfaces. *Proceedings, Royal Society*, A295, 300-319, 1966.

[3] Preston, F. W. The Theory and Design of Plate Glass Polishing Machines. *Journal of the Society of Glass Technology*, 11, 214, 1927.

[4] Johnson, K. L. *Contact Mechanics*, pages 92-93, Cambridege University Press (1985).

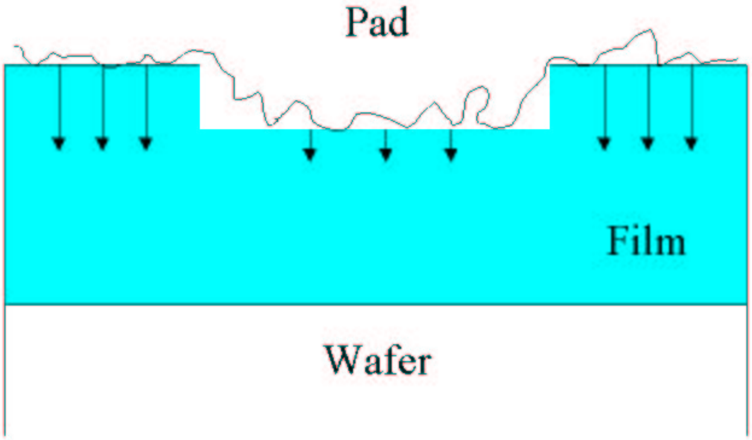

Figure 4: Cross section of a trench in a film deposited on a wafer surface. During CMP, the rate of removal of material around the trench is higher than the rate at the bottom, as indicated by the arrows. As illustrated in the previous figure, the scale of roughness of the pad surface can be larger than the trench depth.

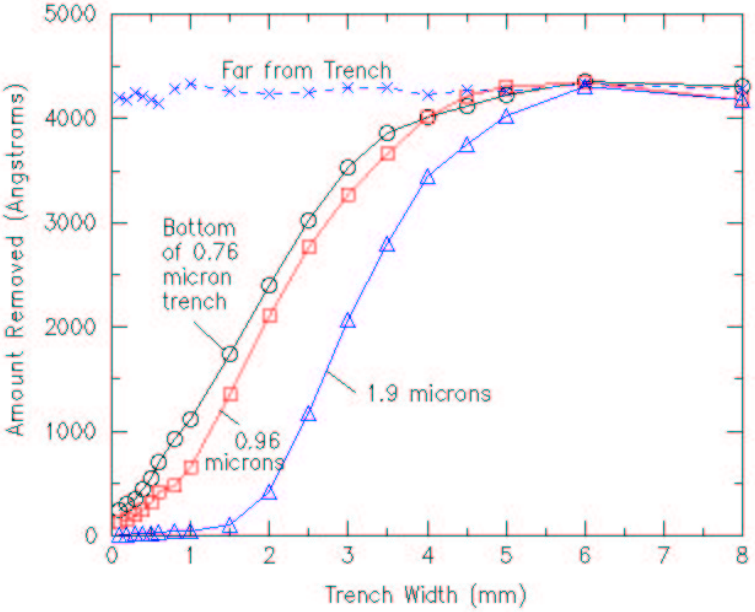

Figure 5: Experimental data showing the amount of material removed from trench bottoms or from the surrounding area as a function of trench width and depth at the end of 90 seconds of polishing at 28,000 Pascals (4 psi). The trenches were etched into silicon dioxide and ranged in from 0.1 mm to 8 mm in width and 0.76 microns to 1.9 microns in depth.

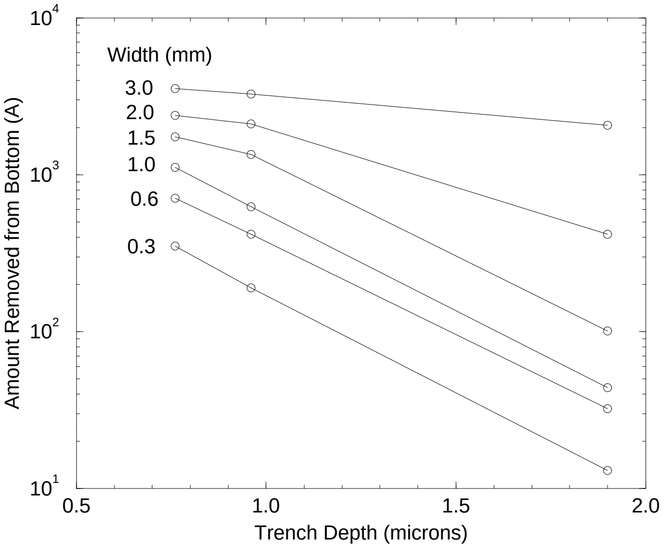

Figure 6: The data from Figure 5 shown on a log plot as a function of trench depth for various trench widths. The variation in the amount removed at the trench bottom appears to be approximately an exponential function of the depth. When the trench width is narrow, the slope appears to be nearly independent of the width.