UC Berkeley
Precision Manufacturing Group

Title
Pad Surface Roughness and Slurry Particle Size Distribution Effects on Material Removal Rate in Chemical Mechanical Planarization

Permalink

Authors
Wang, C
Chandra, A.
Dornfeld, David

Publication Date
2005

Peer reviewed

Pad Surface Roughness and Slurry Particle Size Distribution Effects on Material Removal Rate in Chemical Mechanical Planarization

C. Wang[1], P. Sherman[1], A. Chandra[2], D. Dornfeld[3](1)
[1]Dept. of Aerospace Engineering and Dept. of Statistics, Iowa State University, Ames, IA, U.S.A
[2]Dept. of Mechanical Engineering, Iowa State University, Ames, IA, U.S.A
[3]Laboratory for Manufacturing Automation, Dept. of Mechanical Engineering, University of California, Berkeley, CA, U.S.A

Abstract
The ability to predict material removal rates in chemical mechanical planarization (CMP) is an essential ingredient for low cost, high quality IC chips. Recently, models that address the slurry particles have been proposed. We address three such models. The first two differ only in how the number of active particles is computed. Both assume that pad asperities are identical and nonrandom. The third is dynamic in accommodating changing pad properties. For larger mean particle size (diameter), the role of the standard deviation of particle size distribution is uncertain. The dynamic behavior of the third model is compared with experimental observations.

Keywords
Modeling; Polishing; Statistical Distribution

1 INTRODUCTION

Chemical mechanical planarization (CMP) has emerged as an enabling technology for the next generation of integrated chip manufacturing, and has become the second fastest growing area of semiconductor equipment manufacturing. The CMP process entails sliding a wafer surface over a relatively soft polymeric porous pad, flooded with chemically active slurry containing abrasive particles of sub-micron diameter. Many models have been proposed to investigate the mechanisms of material removal in the CMP process. Features considered, include statistics of pad asperity height and spatial distribution [1], local deformation of individual cells [2], elastic contact between the wafer and the pad [3], and multi-level contact evolution at particle and asperity scales, leading to several domains of wafer/particle/pad contact [4]. Fu and Chandra [5] investigated the effects of viscoelastic pad properties on the material removal rate (MRR). Guo et al. [6] investigated the effects of time-dependent viscoelastic pad properties on dishing and step height reduction. Evans et al. [7] reviewed the fundamental mechanisms of material removal in lapping and polishing processes and identified key areas where further work is required. Komanduri [8] discussed material removal mechanisms in finishing of advanced ceramics and glasses.

In our previous MRR decay model [9], it was found that, by assuming inelastic contact between wafer and pad as well as the pad asperity topography evolution and suitable initial pad asperity height distribution, prediction of MRR was notably improved, as measured relative to the experimental results of Stein et al. [10].

All of the works referenced above considered MRR in relation to direct contact between the pad and the wafer. The influence of the slurry was not explicitly considered. However, the slurry abrasive particles entrapped in the wafer-pad contact region are directly responsible for wafer material removal. As a result of pad pressure, they are indented into the wafer surface and cut material from the wafer surface. Several investigators have explored the mechanisms of MRR due to the pad-slurry particle-wafer contact. Examples include Luo and Dornfeld [3, 11], and Zhao and Chang [12], Fu et al. [2].

A key distinction between the models in [3, 11] and the models in [2, 12], is the method of estimating the number of slurry particles actively participating in the material removal process. In [3, 11], as a pad asperity is brought into contact with the wafer, it is assumed that all of the slurry particles within the free volume occupied by the asperities become entrapped between the wafer and the pad asperity. It is then assumed that only a very small fraction of these particles are actively involved in the material removal. In [2, 12], it is assumed that most of the particles are squeezed away as the pad asperity approaches the wafer surface. Only a monolayer of the slurry particles is assumed to be trapped. However, all of these particles are assumed to be active.

In this paper, we first compare the MRR predictions accompanying these two different assumptions. Elements of these two models are then combined with the time-varying MRR model first proposed by Borucki [13], and subsequently extended in Wang et al. [9, 14].

2 MRR MODEL WITH IDENTICAL HEIGHT PAD ASPERITIES (BASED ON MODELS IN [3, 11])

In this model, the number of active slurry particles is taken to be a fraction of the number of particles assumed to be entrapped beneath a pad asperity that comes into contact with the wafer. This number is assumed to equal the number of slurry particles in a slurry volume equal to the asperity volume. The fraction of this number is determined by a chosen "gap" parameter g. For example, in [3, 11] the value of this parameter is chosen to be the mean particle size (diameter) plus three standard deviations. We note that the randomness of the overall MRR comes exclusively from the randomness of slurry particle size distribution. The pad asperities are assumed to be equal in height and nonrandom. It is also assumed that slurry particles are spaced in a nonrandom, uniform fashion throughout the slurry. Thus, for a specified slurry particle density, the total number of particles will be inversely related to the

assumed mean particle size, and, as we shall see, to a lesser degree on the assumed standard deviation of particle size distribution.

The exact expression for the average (mean) overall MRR in thickness (given as equation (4.18) of [14]) is rewritten here as the following equation (1). It is rather complex, in that it depends on a large number of variables and parameters.

$$\overline{MRR} = C_1 \frac{(E[X_{active}(g)])^2}{(E[X])^3}(1 - \Phi_X(g)) \tag{1}$$

where

$$E[X_{aactive}(g)] = E[X] + \sigma_X^2 \frac{p_X(g)}{1 - \Phi_X(g)} \tag{2}$$

$$g = \frac{H_p - P_{cont}/4}{H_p} E[X_{active}(g)] \tag{3}$$

And C_1 is a constant coefficient which depends on pad surface properties (asperity density, asperity tip curvature, asperity volume, Young's modulus), slurry particle volume concentration, nominal contact area, wafer hardness, applied down pressure and sliding velocity between pad and wafer surface (see equation (4.18) of [14]). In the above equations, E[ ] is the expectation operation, X represents random particle size, σ_X^2 is the variance of particle size distribution, $\Phi_X(g)$ is the cumulative distribution function (CDF) percentile of normal distribution at position g, $p_X(g)$ is the probability density function (PDF) value of normal distribution at position g, H_p is pad hardness, P_{cont} is the local contact pressure between pad asperity and wafer surface which is a constant depending on applied down pressure and pad asperity density, asperity tip curvature and pad Young's modulus (see equation (4.8) of [14]) . The effect of particle size distribution on the average (mean) MRR is totally reflected by the term without the constant C_1 in the right side of (1) and is illustrated in Figure 1 (named as normalized mean MRR), for a specified applied down pressure 1924 Pa, pad hardness 50 MPa, and a pad surface roughness parameter (ratio of actual contact area over nominal contact area) 1/1000.
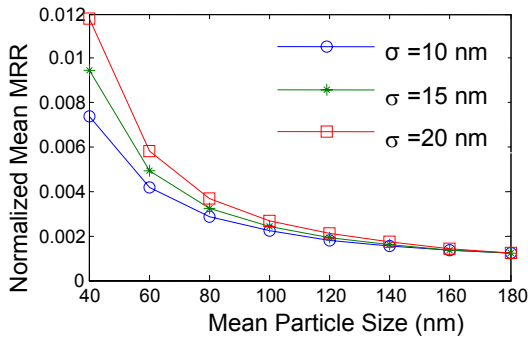

Figure 1: Normalized Mean MRR.

This figure shows that the average (mean) MRR decreases with increasing mean particle size, and that it increases with the increasing standard deviation σ of particle size distribution. However, when the particle diameter standard deviation is much smaller than its mean, MRR becomes essentially independent of the particle standard deviation. The MRR trend shown in Figure 1 matches many experimental results when the particle size is small, typically below 1 µm (e.g. [2], [11] and [15]). We note that in [14], as well as [3, 11] and the above related references it is assumed that the slurry particle diameter is normally distributed. Clearly, this assumption puts into question the validity of the leftmost portion of Figure 1. For example, the left tail of the normal density function corresponding to a mean of 40 and a standard deviation of 20 will extend well into the region of meaningless (i.e. negative) particle diameters.

The general behavior illustrated in Figure 1 follows from the fact that the total number of particles entrapped beneath a pad asperity is inversely proportional to the mean particle size (diameter) to the third power [4]. Thus, even though larger mean particles remove more material individually, there are far fewer of them.

3 MRR MODEL WITH IDENTICAL HEIGHT PAD ASEPRITIES (BASED ON MODELS IN [2, 12])

In these works it is assumed that only a mono-layer of particles is entrapped beneath a contacting pad asperity, but that every particle is actively involved in the MRR. In this case, the expression for the mean overall MRR (equation (4.24) of [14]) shows that the mean MRR is determined by the largest diameter particle entrapped, and by the second and third moments of the particle diameter, which is rewritten here as the following equation (4).

$$\overline{MRR} = C_2 \frac{X_{max}E[X^2]}{E[X^3]} \tag{4}$$

where

$$C_2 = 6\chi u * \left(\frac{1}{2\pi H_w}\right)^{3/2} \left(\frac{4E^*\kappa_s}{3}\right)^{1/3} \eta_s^{-1/6}(\overline{p}) \tag{5}$$

Here χ is slurry particle volume concentration, u is the sliding velocity between wafer and pad surface, H_w is wafer hardness, E^* is pad equivalent Young's modulus, $\overline{p}$ is the applied down pressure, and κ_s, η_s are pad asperity density and asperity tip curvature respectively.

The size (diameter) of the largest of the entrapped particles is a random variable, and it determines the thickness of the monolayer. If a large number of particles are entrapped, then one can approximate it as equal to the mean plus three standard deviations. In this case, the influence of particle size mean and standard deviation on MRR is illustrated in Figure 2 (the right side of (4) without the constant C₂ normalized by scaling to be comparable in magnitude to that shown in Figure 1), The behavior in Figure 2 is similar to that in Figure 1, except that for a large mean particle size, the standard deviation plays a significant role. This difference is due, in part, to the fact that in this model the number of active particles, being a monolayer, as opposed to a fraction of a free volume, is inversely proportional to the particle diameter to the first power.

In summary, both models for the number of active particles give similar MRR trends as a function of particle size (diameter) mean and standard deviation. Furthermore both conform reasonably well to experimental results [11, 15] for moderate mean particle sizes. No published results are available for larger mean particle sizes, so that it is not reasonable to speculate as to which model is more accurate in that region.
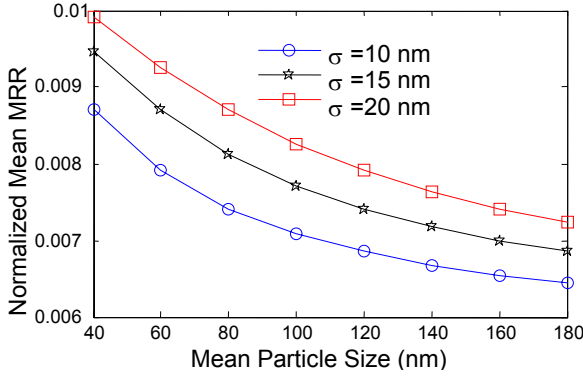
Figure 2: Normalized Mean MRR.

4 MRR MODEL WITH VARYING HEIGHT PAD ASPERITIES (BASED ON MODELS IN [2, 12])

In this section, we offer an extension to the MRR model of the last section to reflect a more realistic situation; namely that pad asperities have random heights and that the CMP process is dynamic; that is, time-varying. The model in this section may be viewed as an extension of the dynamic MRR decay models [9, 13, 14] which considered the MRR in relation to the direct contact between wafer and pad, ignoring the explicit influence of the slurry particles. Even though the model in this section is dynamic, if the pad asperity height and slurry particle distributions are given at any time, then it can be compared to the above models at that time.

Since in this model, pad asperity height is randomly distributed, so are the local contact area, the local contact load, and the local contact pressure between the wafer and pad asperities. In the previous sections these three quantities were non-random. Hence, in this sense, the MRR model in this section is the most general of the dynamic MRR models proposed to date.

The material removed by a single active particle depends on the particle size and the local contact pressure, which is linearly related to the difference between pad asperity height and the mean separation distance. The total number of active particles depends on both the pad asperity height and particle diameter distributions. It follows that the material removal per particle, and consequently, the total MRR is a random variable. Here, we restrict our attention to the mean of the overall MRR, the derivation of the mean overall MRR is (given as equation (4.48) in [14]) rewritten here as the following equation (6).

$$\overline{MRR} = C_3 \frac{X_{max}E[X^2]}{E[X^3]} \int_{d(t)}^{\infty} [z - d(t)]^{7/4} \phi(z,t)dz \quad (6)$$

where

$$C_3 = 6 * \chi * u * \eta_s * \frac{1}{\kappa_s^{1/4}} * \left(\frac{2E^*}{3\pi H_w}\right)^{3/2} \quad (7)$$

In (6), d(t) is the separation distance between the wafer surface and the mean plane of pad asperity heights, and $\phi(z,t)$ is the asperity height time-dependent probability density function (PDF). As shown in the MRR decay model [9, 14], the pad asperity height PDF will evolve according to

$$\frac{\partial}{\partial t}\phi(z,t) = \frac{C_a 4E^*\kappa_s^{1/2}}{3\pi} \frac{\partial}{\partial z}\{[z - d(t)]^{1/2}\phi(z,t)\} \quad (8)$$

for z>d(t). The PDF evolution rate is zero for $z \le d(t)$. Equation (6) is a static MRR model at any fixed time t, while equations (6) and (8) form the dynamic MRR model with initial condition PDF $\phi(z,0)$.

An example of the normalized time-varying mean MRR decay due to pad surface topography evolution predicted by (6) and (8) is illustrated in Figure 3 (the integral part in the right side of (6) with asperity height PDF evolution according to (8)). With continued polishing without pad conditioning, the pad surface topography evolves due to the wear process and becomes glazed (flattened). This causes the MRR decay in Figure 3.
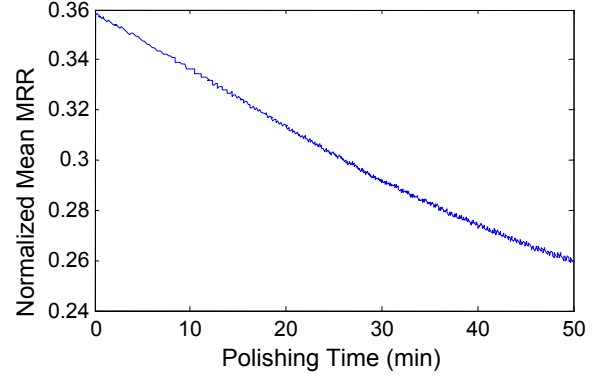

Figure 3: Normalized mean MRR decay due to pad surface topography evolution (Model prediction).

Notice that in (6) the particle properties control the scale, but do not influence the dynamical behavior of MRR (MRR decays with polishing time). For this reason, since Figure 3 includes normalized MRR values, it is not necessary to specify particle statistics. To assess the reasonability of the trend shown in Figure 3, we include Figure 4, which is an experimental measurement of MRR decay over time given by Stein [10]. A comparison of Figures 3 and 4 shows that the overall decay trends are similar. Since no particle information is provided in [10] and the MRR in Figure 3 is normalized, a direct comparison of these figures is not appropriate. However, one can observe that the relative rates of decay are similar. Figure 3 shows an MRR decay of ~ 30% over the polishing interval. The Lot A curve in Figure 4 shows ~ 40% decay.
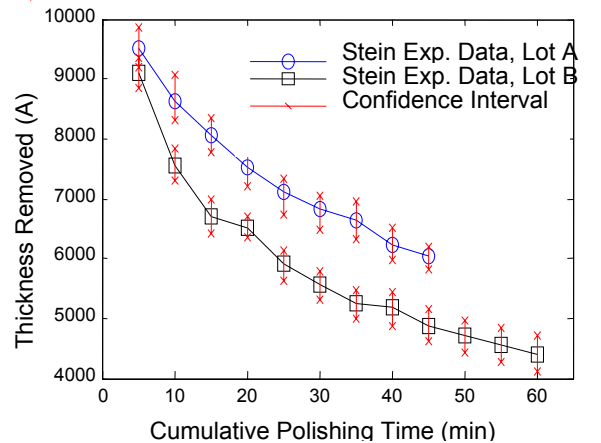

Figure 4: Experimental results of MRR decay (from [10])

Finally, we offer a discussion of the model defined by equations (6) and (8) and the static model in the last section. As mentioned above, for any fixed time in the polishing interval the model defined by equation (6) is a static model. A key distinction between this model and those in the previous sections is that it takes the pad behavior into account in a very clear way. Even though

pad manufacturers often provide asperity statistics for new pads, this information becomes less and less useful as the polishing process progresses. The equation (8) allows one to predict the evolution of the pad asperity height PDF at any time. Neither the models in the two previous sections, nor any other static models we are aware of (e.g. [3, 11, 12]) are equipped to account for changing asperity properties in conjunction with particle information.

Notice that the middle fractional term in equation (6) controls the influence of the slurry particle size distribution on the mean overall MRR. This term is exactly the fraction term in equation (4) (or equation (4.24) of [14]), which resulted in Figure 2. Hence, the normalized MRR at any given time will behave in the manner illustrated in that figure. A quantitative comparison of equation (6) with other models and/or experimental data would require information related to the pad, slurry particle and wafer mechanical properties (hardness, Young's modulus, Poisson's ratio), pad surface roughness parameters (asperity density, asperity spherical tip curvature), as well as CMP operating condition (sliding velocity or rotation speed), applied down pressure, slurry particle volume concentration. Since this information is not available, our discussion must be limited to normalized MRR behavior.

5 SUMMARY AND CONCLUSION

The purpose of this work was to present CMP models for material removal that explicitly take into account the slurry particles, which are the primary mechanism for material removal. Three models were presented. The first two models differ only in how the number of active particles is arrived at. They also assume asperities are nonrandom and equal. The third model was an extension of the second, in that it relaxed the assumption of equal nonrandom pad asperities. In fact, it was dynamic, in the sense that it included a time-evolution model that can be solved to estimate asperity height distribution at any given time throughout the CMP operation.

In the static case where time is fixed in the third model, it reverts to the second model; but with the advantage of having pad asperity height distribution information. In this case, both the first and second models show similar MRR decay trends, in relation to particle mean and standard deviation values. A key difference is that even for relatively large mean particle sizes, the second model suggests that the particle diameter standard deviation plays a significant role in MRR decay.

Because the third model is dynamic, we presented a comparison of its dynamic behavior with the experimentally observed MRR decay behavior observed by Stein [10]. The model behavior was shown to be consistent with the experimental results.

Finally, it should be mentioned again that because the various models, as well as the experimental results from [10] are lacking in specific information, it was possible here to address only normalized MRR behavior. Hence, the value of this work is related to the ability to better understand how particle properties, in conjunction with various amounts of pad information, influence MRR trends.

7 REFERENCE

[1] Yu, T.K., Yu, C.C., Orlowski, M., 1993, A statistical polishing pad model for chemical-mechanical polishing, IEEE IEDM Washington DC, 865-868.

[2] Fu, G., Chandra, A., Guha, S., Subhash, G., 2001, A Plasticity-Based Model of Material Removal in Chemical-Mechanical Polishing (CMP), IEEE Trans. Semiconductor Manufacturing, 14(4):406-417.

[3] Luo, J., Dornfeld, D., 2001, Material removal mechanism in chemical-mechanical polishing: theory and modeling, IEEE Trans. Semiconductor Manufacturing, 14 (2):112-133.

[4] Bastawros, A.F., Chandra, A., Guo, Y., Yan, B., 2002, Pad Effects on Material Removal Rate in Chemical Mechanical Planarization, J. Electronic Materials, 31(10):1-10.

[5] Fu, G., Chandra, A., 2002, A Model for Wafer Scale Variation of Material Removal Rate in Chemical Mechanical Polishing Based on Viscoelastic Pad Deformation, J. Electronic Materials, 31(10):1-8.

[6] Guo, Y., Chandra, A., Bastawros, A.F., 2004, Analytical Dishing and Step Height Model for CMP with a Viscoelastic Pad, J. Electrochem. Soc., to appear.

[7] Evans, C.J., Paul , E., Dornfeld, D., Lucca, D.A., Byrne, G., Tricard, M., Klocke, F., Dambon, O., Mullany, B.A, 2003, Material removal mechanisms in lapping and polishing, CIRP keynotes paper, STC G, 52(2):611-634.

[8] Komanduri R, 1996, On material removal mechanisms in finishing of advanced ceramics and glasses. CIRP Annals, 45(1):596-60.

[9] Wang, C., Sherman, P., Chandra, A., 2004, A Stochastic Model for the Effects of Pad Surface Topography Evolution on Material Removal Rate Decay in Chemical Mechanical Polishing (CMP), CMP-MIC, California. Revised detail version has been submitted to IEEE Trans. Semiconductor Manufacturing and is now in review.

[10] Stein, D., Hetherington, D., Dugger, M., Stout, T., 1996, Optical interferometry for surface measurement of CMP pads, J. Electronic Materials 25(10):1623-1627.

[11] Luo, J., Dornfeld, D., 2003, Effects of abrasive size distribution in chemical mechanical planarization: modeling and verification, IEEE Trans. Semiconductor Manufacturing, 16(3):469 - 476.

[12] Zhao, Y., Chang, L., 2002, A micro-contact and wear model for chemical-mechanical polishing of silicon wafers, Wear, 252:220-226.

[13] Borucki, L., 2002, Mathematical modeling of polish-rate decay in chemical-mechanical polishing, J. Engineering Mathematics, 43:105-114.

[14] Wang, C., 2004, Stochastic models for material removal rate (MRR) in chemical mechanical planarization (CMP) process, Ph.D dissertation, Iowa State University.

[15] Bielman, M., Mahajan, U., Singh, R.K., 1999, Effect of particle size during tungsten chemical mechanical polishing, Electrochem. And Solid-State Lett., Vol.2, pp. 401-403.