EVOLUTION OF TOPOGRAPHY DURING 1ST STEP CMP OF CU-PLATED DAMASCENE STRUCTURES

Steve Hymes, Tom Brown, Paul LeFevre, Bob Mikkola, Ishmail Emesh,
***Rajeev Bajaj, *Konstantin Smekalin, *Yutao Ma, *Fritz Redeker,**
****Tae Park, **Tamba Tugbawa, **Duane Boning, ***John Nguyen**

2706 Montopolis Drive, SEMATECH, Austin, TX, 78741,
***3111 Coronado Drive, M/S 1512, Applied Materials, Santa, Clara, CA, 95054,**
****MIT, EECS, Room 39-567 Cambridge, MA 02139,**
*****SpeedFam-IPEC, 4717 E. Hilton Ave. Phoenix AZ 85034.**

ABSTRACT

Electroplate has emerged as a viable deposition technique for advanced copper-based metallization. Wafer-scale, array-scale and feature-scale metal thickness nonuniformity occurs during the plating process and contributes to post-CMP metal and dielectric film thickness variation. The resulting topography from this electroplated metal thickness variation is highly pattern dependent. Successful process integration requires the subsequent CMP process to planarize this as-plated topography prior to reaching the barrier film. CMP has its own length-scale-specific response function. Consequently, the level of topography which can be accommodated by the CMP process is a complex function of pattern. Understanding of the interaction of such length-scale dependencies of the plating process with the corresponding length-scale dependencies of the CMP process is critical to successful process integration. Modeling of this interaction is a highly desirable undertaking and provides a means by which a more in-depth understanding of the electroplating-CMP interaction can be derived as well as providing an efficient predictive tool for use in optimization of a manufacturable flow.

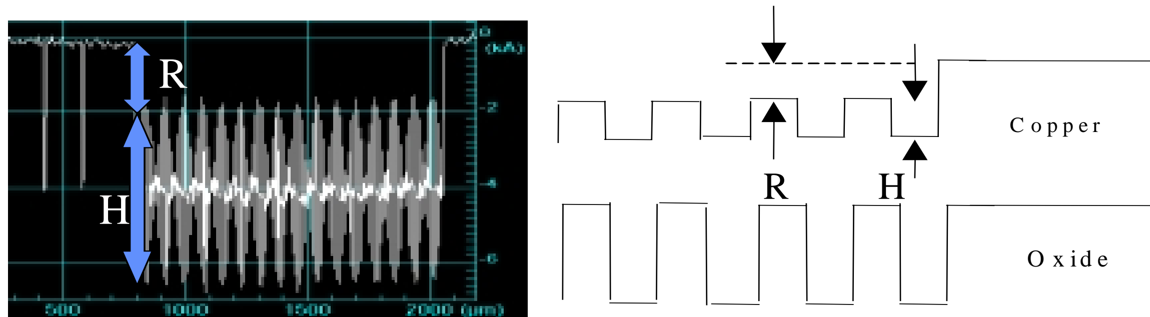
INTRODUCTION

In direct analogy to the "dishing" and "erosion" which is used to characterize the topography of post-CMP processing, the topography which occurs in the as-plated condition can be categorized similarly as as-plated feature-scale "dishing" (H) and as-plated array-scale "erosion" (R) as shown in Figure 1. Both modes of as-plated topography depend heavily on layout pattern and plating process parameters, platform and chemistry and must be planarized prior to reaching the barrier material. For the design rules of advanced microelectronics, the largest single features are bond pads (≈ 100 micron) and bus lines (≈ 5-20 micron) associated with upper levels of metallization. From a CMP perspective, the most aggressive combined structures are the low pitch (submicron), high metal density arrays (mm dimension) of the lower metal layers. Thus, it is important to understand planarization performance throughout the submicron

through millimeter range length-scale. The initial part of the polish process can thus be viewed as acting to remove preexisting topography while the latter clearing stage aims to prevent the generation of topography. A strong drive exists towards processes that yield little as-plated topography on both the individual feature and associated array-scale dimensions in addition to requisite high across-wafer uniformity and excellent fill capability.

The current models governing the evolution of film thickness and surface topography have yet to mature. Various models of the planarization performance of the CMP process have been proposed. The pattern dependency of the polish rate for the single material system of oxide CMP has been investigated.(1) Perfect planarization efficiency was assumed. No polishing of the 'down' regions existed and for the etched pattern in blanket oxide the initial topography was feature-scale only, in the absence of an array-scale contribution. In contrast, for oxide deposited over patterned metal, array-scale

Figure 1. Profilometer scans of the surface of an isolated line and associated array in the as-plated condition.



topography exists. However, the model does not take into account the influence of either preexisting or induced array-scale topography on the effective local pressure and associated local rate. Modeling of topography evolution for the case of imperfect planarization efficiency during the latter stages of CMP has also been undertaken.(2) In this case, after a finite feature-specific 'contact' time, the pad influences the recessed region and polishing of both raised and recessed regions occurred simultaneously. The effect of array-scale topography was not incorporated into this model either. A more recent publication focused upon large feature-scale planarization efficiency in the single system material of copper.(3) For very wide trenches, the planarization efficiency tended to zero as the local pressure in the recessed region approaches that in the field. The key elements of this latter paper can be combined with that of the former two and applied to planarization of the copper overburden during copper CMP. In this work, such an extension is made to the existing published models so as to capture a more complete description of the plating-polish interaction. The polish time and layout pattern dependence of the thickness and topography of the copper metal overburden is presented for select plate-polish processing.

CHARACTERIZATION OF AS-PLATED TOPOGRAPHY

To more fully understand the as-plated surface, the topography for a standard patterned wafer using the SEMATECH 954AZ 1st level metal mask was characterized.

The reticle layout consists of a number of modules where each module contains two isolated lines and a corresponding 1.25mm x 1.25mm array. Modules with 25, 50 and 75% metal density and pitch from 1 to 200 µm are represented. Figure 2 displays the corresponding topography as a function of module type for a wafer processed using a conventional fill chemistry.

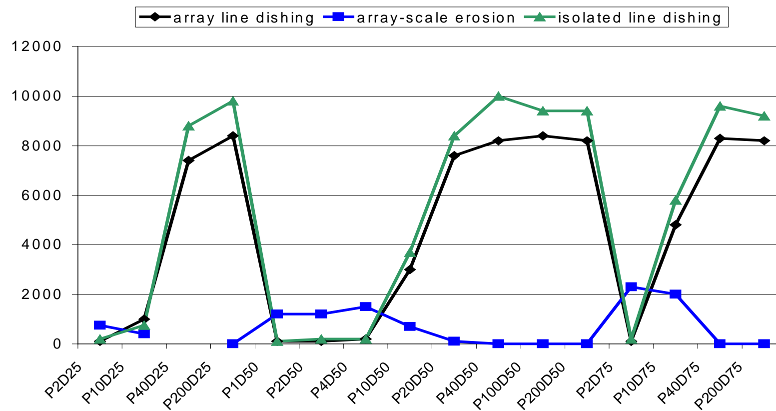


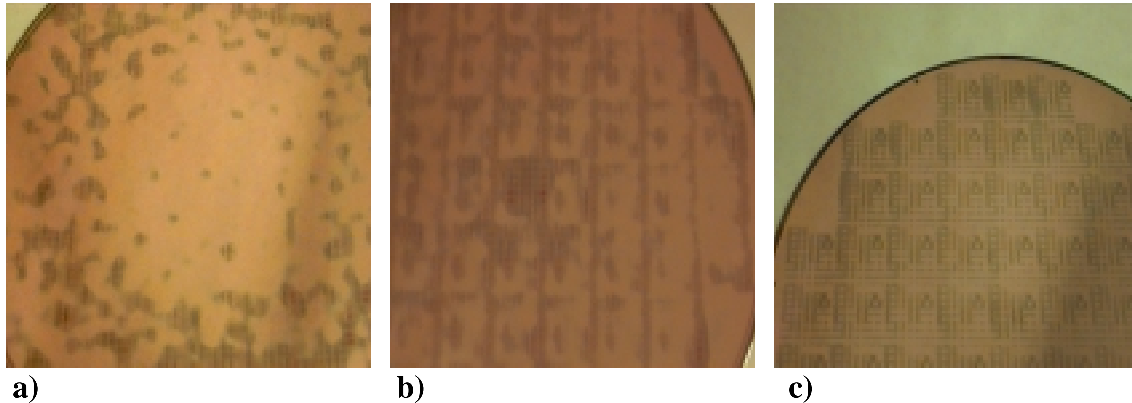
Figure 2. As-plated isolated line "dishing", array line "dishing" and array "erosion" for a typical as-plated wafer using a conventional fill chemistry (i.e. P2D25 = 2µm pitch, 25% metal density with 0.5µ=2.0 x 25% isolated lines).

The isolated "dishing" is similar to that for the array "dishing", being slightly greater in magnitude. For a given metal density, the topography can be categorized as primarily array-scale "erosion" for low pitch values, both array-scale "erosion" and feature-scale "dishing" for moderate pitch values and pure "dishing" for large pitch values. The transition point is a strong function of pitch and a weaker function of density. One also notes that the magnitude of the "dishing" for large pitches saturates to a value approximately that of the etched trench-depth, as would be expected. In contrast, the magnitude of topography for array-scale "erosion" is only a fraction of the trench depth. In all cases, the topography follows the conventional positive recess relative to field for this conventional-fill plate chemistry. The magnitude of "dishing" can be much higher than that for "erosion". However, this does not mean that "dishing" is more important, because for CMP, the planarization efficiency is a function of the length-scale and the array structure length-scale is much more slowly planarizing than the "dishing" structure length-scale. When the across-wafer nonuniformity of a CMP process is sufficiently low, the within-die thickness nonuniformity becomes the primary limitation. In general, the array-scale topography limits the within die thickness range. Thus, particular attention needs to be paid to minimizing array-scale topography.

THE PLATING-POLISH INTERACTION

Since the array-scale "erosion" topography is a less efficiently planarizing feature than that associated with a single shorter length-scale feature, the critical feature is not necessarily that with the largest as-plated topography magnitude. For example, Figures 3a-c show photographs of 3 wafers that were polished under different conditions to the point of barrier exposure. The three conditions differ by the planarization performance of the process. As seen in Figure 3a, the wafer has just broken through to barrier uniformly

across the wafer. There is no die-scale repetition to the clearing, nor does there appear to be a significant wafer-scale (radial) dependence to the clearing. The lack of die-scale pattern is an indication that the wafer has planarized both the short length-scale structures as well as the arrays. As the planarization efficiency decreases, a die-scale pattern to the clear begins to emerge as shown in Figure 3b.



a) **b)** **c)**

Figures 3a,b,c Photographs of the surface of three separate wafers which depict decreasing levels of planarization in presentation order**.**

Note that one is only able to evidence these die-scale nonuniformity provided the wafer-scale variation is sufficiently low. Figure 3c shows an even more extreme case of this die-scale pattern in which both field and specific repeatable smaller areas of the die contain residual overlaying copper while other regions of the die have been fully cleared of residual copper. Table I maps out the as-plated topography in terms of "dishing" and "erosion" for each structure type and indicates whether or not the associated structures have cleared. Residual metal exists for all structure above the dotted line whereas those structures below the dotted line in the table were cleared of copper. For this particular conventional-fill plating process, there is a monotonic increase in the extent of as-plated "dishing" for increasing pitch at fixed density. For fixed pitch, there in an increase in the extent of both "dishing" AND "erosion" as the density increases. For this particular plating process, which exhibited positive (recessed) topography for both "dishing" and "erosion", it is not surprising that the moderate and high-pitch, high-density structures are earlier to clear.

Table I. As-plated topography (Å) for a wafer polished to the copper/barrier transition as a function of structure type off the SEMATECH 954AZ M1 mask. Copper residual was found in the field region as well.

Initial Topography						
	D75		D50		D25	
Pitch	Erosion	Dishing	Erosion	Dishing	Erosion	Dishing
P1	X	X	-1100	0	X	X
P2	-1800	0	-1100	0	-550	0
P4	X	X	-1230	0	X	X
P10	-4200	-1500	-2000	-1000	-550	-400
P40	-500	-8400	X	X	0	-8800
P100	X	X	0	-8800	X	X
P200	0	-8800	0	-8800	0	-8800

As-plated within DIE topography and metal thickness range contributes to overpolish.
-ve topography refers to conventional recess

Residual

Cleared

MODELING OF COPPER OVERBURDEN CMP

Modeling of the evolution of the copper overburden during CMP was conducted. The approach aims to account for initial topography (array-scale recess and feature-scale "dishing") in determination of polishing behaviors. The methodology includes extraction of the planarization length using the MIT pattern density oxide model [1] extended to the copper system incorporating the idea of local removal rate change on step height reduction [2] which is also extended to array-scale recess [3]. Grillaert's paper [2] focused only on local step height reduction and how that affects local removal rate. We have extended this idea to global step height reduction for array-scale recess and understanding how array-scale recess influences polish rates and behaviors on a macroscopic level (e.g. die-level).

We define an array-scale recess as R and feature-scale "dishing" (step height) as H as presented in Figure 1. MIT's pattern density model states that the removal rate is inversely proportional to pattern density as in equation [1] where K is blanket removal rate. The original work and model has been developed and used for oxide, but this model can be applied to copper polishing during the bulk copper removal phase before reaching barrier layer. Such application of the model can be made since we are dealing with a single material polishing and since we attribute the main polishing mechanism to chemical-mechanical factors even though copper CMP may have a greater component of a pure chemical factor than exists for oxide polishing. The principal approach consists of the following steps:

1) Determination of the effective average density over a spatial region centered about the point of interest as a function of characteristic dimension using an appropriate weighting function. For this work, a square profile of uniform spatial weighting was employed.
2) Determination of the array-scale modification to the effective local pressure through a linear term α.
3) Determination of the feature-scale modification to the raised and recessed feature-scale regions through a linear term β.
3) Determinations of the corresponding instantaneous polish rate and film thickness across the range of available structure types.
4) Iteratively apply the sequence and predict the film thickness across a range of test structures.
5) Determine the best fit across this range of experimental structures and subsequently extract the planarization length of the process as well as copper overburden film thickness profile. From the basic oxide model is:

$$\text{Removal Rate} = K/P_{ox}, \qquad [1]$$

where P_{ox} is oxide density of the etched dielectric layer and K is the blanket removal rate. After Cu plating, the effective density of the copper burden approximates that of the underlying oxide density. The extent to which this is valid is debatable and represents a potential area of refinement. To account for initial or induced topography, we make a modification to the above expression through the term "α" as follows:

$$\text{Removal Rate} = (K/P_{ox})*\alpha \quad [2]$$

"α" is a linear expression that accounts for effective pressure on the recessed array relative to the field and has a value between 0 and 1. Here, $\alpha = 1$ when $R = 0$ which is the case absent of any array-scale recess relative to field and $\alpha = 0$ when $R >= R_{crit}$, where R_{crit} is the global step-height (array-scale recess) at which the pad loses contact with the recessed surface. This could be thought of as a maximum flexing limit of a pad and is clearly a function of the CMP process and recessed area length-scale. The value of α and R_{crit} for a given array width would be analogous to array-scale planarization efficiency and zero planarization efficiency length values, respectively. We assume a linear relationship between α and R and we can relate them as follows:

$$\alpha = 1 - R/R_{crit} \quad \text{for R less than } R_{crit} \quad [3]$$
$$\alpha = 0 \quad \text{for } R >= R_{crit} \quad [4]$$

Now we make a second addition to the model as follows:

$$\text{Removal Rate} = (K/P_{ox})*\alpha*\beta \quad [5]$$

In a similar analogy to the global recess and term "α", the "β" is a term to account for the local step height influence on the removal rate and has a value between P_{ox} and 1. The value of β can be interpreted as the portion of the local pressure going to the *original* 'up' area of the structure. As it was the case before, we assume a liner relationship between β and H as follows.

$$\beta = P_{ox} + [(1- P_{ox})/H_{crit}]*H \quad \text{for } H < H_{crit} \quad [6]$$
$$\beta = 1 \quad \text{for } H >= H_{crit} \quad [7]$$

When H (local step height) is greater than H_{crit}, all the pad pressure is exerted on the 'up' areas of local steps, and thus β is 1 and perfect feature-scale planarization efficiency is achieved as the low regions are not contacted by the pad and thus, are not polished. Now, H_{crit}, analogous to R_{crit}, is the critical step height below which a pad starts to "see" and polish the 'down' area. When H is zero, the pad pressure is now evenly distributed between the 'up' areas and 'down' areas. Thus $\beta = P_{ox}$ as the original 'up' area receives an area ratio of the local pressure with this ratio being the oxide density. With $H=0$, $\beta = P_{ox}$, which cancels P_{ox} in equation [5] and thus the removal rate is equal to the blanket rate adjusted by the global recess factor α, as expected. In direct analogy to R_{crit}, H_{crit} must be determined for each process and pattern factor (e.g. trench width).

We have so far only referenced and related β in terms of the remaining step height H. The question now remains as to how we relate the local step 'up' removal rate versus the local 'down' area removal rate to find out how H evolves with polish time. When the step height H is smaller than $H_{crit,}$ as presented in the first stage of the mathematical copper CMP model development [4] we assume a linear decrease and increase for 'up' and 'down' area removal rates, respectively. The original MIT's oxide pattern density model assumes perfect planarization efficiency where there is no 'down' area polishing until all the 'up' areas get polished, and this assumption was used only as an approximation to simplify the initial modeling work on oxide. In this work, we do not make this

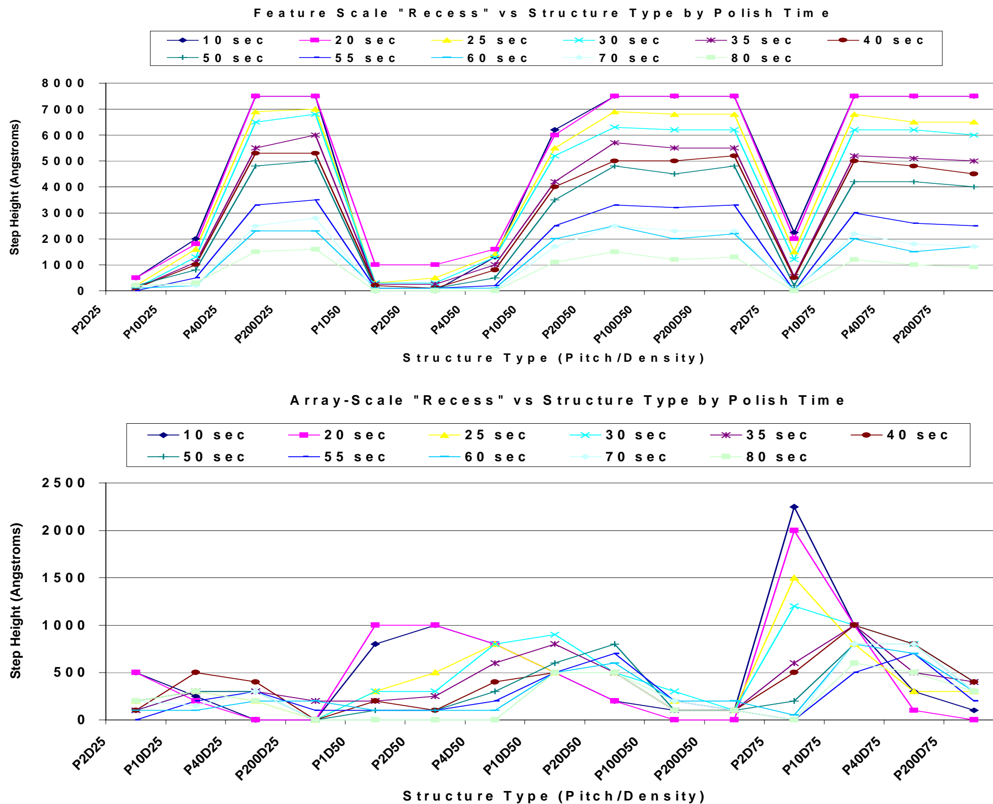
assumption of perfect planarization efficiency because it is physically non-intuitive and it has been shown that the 'down' area does indeed get polished before the local step height becomes zero [2]. Thus we have the following relationship for H and the 'up' and 'down' area removal rates are obtained from equation (5) substituting in equation (6) and neglecting the array-scale recess (α=1).

$$H = \text{'up' Area Removed} - \text{'down' Area Removed} \qquad [8]$$

$$\text{'up' Area Removal Rate} = K\{1 + H/H_{crit}\,(1/P_{ox}-1)\} \qquad \text{For } H <= H_{crit} \qquad [9]$$
$$\text{'up' Area Removal Rate} = K/P_{ox} \qquad \text{For } H > H_{crit} \qquad [10]$$

$$\text{'down' Area Removal Rate} = K\{1 - H / H_{crit}\} \qquad \text{For } H <= H_{crit} \qquad [11]$$
$$\text{'down' Area Removal Rate} = 0 \qquad \text{For } H > H_{crit} \qquad [12]$$

In this new formulation of a copper CMP model, we have captured the removal rate dependency on pattern density as well as the global step height (and the associated pad flexing limit) and local step height (and associated feature-scale length). The new model presented here does not have a closed form solution. Both of the terms R and H change with polish time. Thus, the terms α and β change accordingly and interactively with R and H. As a result, the simulation must be performed in a small time increment, where to compute the remaining copper thickness at time T, the values of the model parameters (e.g. α and β) at time T-1 are needed.



Figures 4a,b "dishing" and "erosion" dependencies for the various structures of the 954AZ mask as a function of polish time.

EXPERIMENTAL VALIDATION OF THE MODEL

Experiments were conducted to check the accuracy of the prescribed model. A series of SEMATECH 954AZ patterned wafers were polished for small time increment differences using a hard-pad CMP process on a conventional rotary motion platform. Figures 4a,b display the associated "dishing" and "erosion" dependencies on polish time for the various structures of the 954AZ mask as obtained from high resolution profilometry. The "dishing" decreases monotonically in all cases, while the "erosion" does not. For some arrays, the "dishing" is rapidly converted to "erosion". It is only after this conversion that the "erosion" component then decreases monotonically. The critical feature is not necessarily the array with the greatest initial topography.

Pre and post sheet film resistance measurements were employed across a field region of the center die and the blanket polish rate of 3500 Å/min was calculated for the process using a global pattern density of 0.685 for this mask. An estimated 3500 Å was used for R_{crit} for these 1250 µm wide arrays based upon Figure 3 from Hymes et al.(3) The complete form of the model in equation [5] was used to account for initial topography recess and local step height. H_{crit} was extracted from the polish time increment data set we used in study. We examined step height reduction vs. polish time and extracted H_{crit} as a point when the reduction starts to deviate from a linear response to a sublinear, exponential-like dependence for the feature in question. The following figures show model fit for various polish times used. An extracted planarization length of 6mm was obtained based upon least error fitting to the characteristic length of the density region. The model fit overall is reasonably good and even though the density range in this study is not wide, the actual data and model fit show trends of faster removal of copper for low effective oxide density (i.e. high copper density). This is consistent with previous observations of faster clearing time for high copper density blocks.

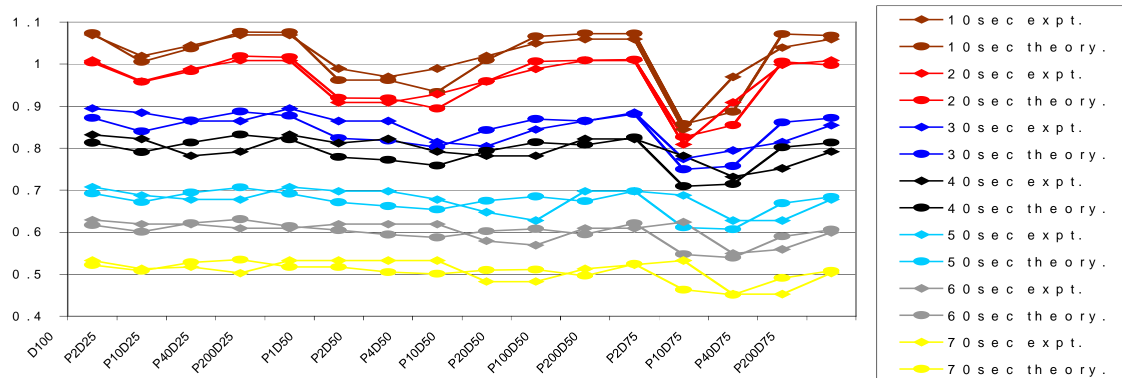


Figure 5 Comparison of the predicted thickness (um) values by structure type with the experimental values as a function of polish time.

CONCLUSION

As-plated topography for a conventional-fill plating process was characterized as a function of structure type using the SEMATECH/MIT 954AZ M1 mask. The analogs of conventional CMP "dishing" and "erosion" for the as-plated condition were presented and

their influence on metal thickness evolution modeled. The pattern density, feature-scale and array-scale contributions to the effective rate on the protruding and recessed structures was incorporated into the model of the CMP of the copper overburden. The model shows reasonable fit to the experimental data across a range of pitch and density.

REFERENCES

1. D. Ouma, et al., Proc. IEEE IITC, pp. 67-9, 1998.
2. J. Grillaert, et al., Proc. CMPMIC, pp.79-86, 1998.
3. S. Hymes, et al., Proc. MRS, April, 1999.
4. T. Tugbawa, et.al., ECS, Hawaii, Oct. 1999.