Rapid Characterization and Modeling of Pattern-Dependent Variation in Chemical-Mechanical Polishing

Brian E. Stine, Dennis O. Ouma, *Member, IEEE,* Rajesh R. Divecha, *Member, IEEE,* Duane S. Boning, *Member, IEEE,* James E. Chung, *Member, IEEE,* Dale L. Hetherington, *Member, IEEE,* C. Randy Harwood, O. Samuel Nakagawa, and Soo-Young Oh, *Member, IEEE*

***Abstract*— Pattern-dependent effects are a key concern in chemical-mechanical polishing (CMP) processes. In oxide CMP, variation in the interlevel dielectric (ILD) thickness across each die and across the wafer can impact circuit performance and reduce yield. In this work, we present new test mask designs and associated measurement and analysis methods to efficiently characterize and model polishing behavior as a function of layout pattern factors—specifically area, pattern density, pitch, and perimeter/area effects. An important goal of this approach is rapid learning which requires rapid data collection. While the masks are applicable to a variety of CMP applications including back-end, shallow-trench, or damascene processes, in this study we focus on a typical interconnect oxide planarization process, and compare the pattern-dependent variation models for two different polishing pads. For the process and pads considered, we find that pattern density is a strongly dominant factor, while structure area, pitch, and perimeter/area (aspect ratio) play only a minor role.**

***Index Terms*— Chemical-mechanical polishing, interconnect, pattern dependencies, spatial variation.**

I. Introduction

IN RECENT years, chemical-mechanical polishing (CMP) has emerged as the primary technique for planarizing dielectrics [1], [2]. CMP is very effective at reducing feature-level or local step height and achieves a measure of global planarization not possible with spin-on and resist etch back techniques [1]; however, CMP processes are hampered by pattern sensitivities which cause regions on a chip to have thicker dielectric layers than other regions due to differences in underlying topography [2]–[4]. This problem has become especially acute as performance requirements have increased and dimensions have scaled. Also, CMP has found wider application in the entire VLSI development and production cycle serving as an enabling tool for shallow trench isolation [5]–[7], damescene technologies [8], and other novel process techniques.

In this paper, we present four masks for characterizing and modeling pattern dependent variation in CMP processes, consumables, and tools. Using these masks, we present methods for the rapid characterization, empirical modeling, and comparison of pattern dependencies as a function of processes, consumables, or equipment options. This is achieved in two ways. First, each mask is targeted toward an individual source of pattern dependent variation. To this end, four separate single-layer masks have been designed to probe structure area, pattern density, line pitch, and structure aspect ratio effects, respectively. Second, the masks support simplified metrology tools and techniques including optical film thickness and profilometry measurements.

Several masks have been developed for analyzing pattern dependent variation. Warnock [9] and Hayashide *et al*. [10] used a simple mask composed of a set of lines and spaces varied across the die. Renteln *et al*. [11] also used a set of gratings but over very large feature scales (1–10 mm). SRAM or gate array cells are sometimes used [12], [13]. Burke [3] used a set of two test masks but does not describe their design or construction. With a few exceptions, these masks and accompanying analysis techniques are not discussed fully or at all, and appear primarily targeted at testing specific models rather than to aid in characterizing or generating models.

Chang *et al*. [14] and Stine *et al*. [15] have demonstrated the use and analysis of test masks for screening experiments designed to investigate and identify the potential factors contributing to interlevel dielectric (ILD) thickness nonuniformity. These masks, however, were designed for electrical probing; as a result, their use requires several masking steps and is often restricted by the availability and throughput of probing equipment. Substantial effort is also required to analyze electrical probe data, including the conversion of electrical data to ILD thickness information. Electrical test masks do provide valuable complementary information to the simpler test masks described in this paper: small linewidth and spacing

Manuscript received October 14, 1996; revised July 12, 1997. This work was supported by DARPA under Contract DABT-63-95-C-0088 and AASERT Grant DAAHA04-95-I-0459, the NSF/SRC Engineering Research Center for Environmentally Benign Semiconductor Manufacturing, and by an Intel Foundation Fellowship.

B. E. Stine, D. O. Ouma, D. S. Boning, and J. E. Chung are with the Microsystems Technology Laboratories, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge MA 02139 USA (e-mail: boning@mtl.mit.edu).

R. R. Divecha was with the Microsystems Technology Laboratories, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge MA 02139 USA. He is now with Rockwell Semiconductor Systems, Newport Beach, CA 92660 USA.

D. L. Hetherington is with Sandia National Laboratories, Albuquerque, NM 87185 USA.

C. R. Harwood is with IPEC/Planar, Phoenix, AZ 85034 USA.

O. S. Nakagawa and S.-Y. Oh are with the Hewlett Packard Co., Palo Alto, CA 94304 USA.

Publisher Item Identifier S 0894-6507(98)00332-7.

features can be used and combinations of layout factors that more closely mimic realistic circuits can be constructed.

At a broader level, the characterization mask sets described in this paper represent an application of statistical metrology [16], [17], which is concerned with assessing and modeling both random and systematic variation and determining the impact of this variation on circuit performance and manufacturability. The primary purpose of the mask sets presented in this paper is characterizing and modeling CMP induced variation. Also, the analytical techniques developed and utilized in this paper draw heavily on previous statistical metrology research efforts [18]–[20].

These characterization masks can be utilized to investigate pattern dependencies in a variety of CMP process applications, including traditional back-end [21], trench isolation [5]–[7], [21], and damascene or inlaid metal processes [8], [22]. Pattern-dependent issues include both "dishing" of features being polished and "erosion" or regions with lower or higher density of features. In this paper, we consider pattern dependencies in a traditional back-end interconnect process.

In this paper, the CMP characterization masks are presented in Section II, and the metrology tools and techniques for which the masks were designed are described in Section III. An experimental application of the test masks is presented in Section IV, in which the polishing performance of two different pads are evaluated. Simplified analysis and modeling methods of pattern dependencies are presented and illustrated for this pad experiment in Section V. Finally, a summary and future work are offered in Section VI.

II. Mask Descriptions

The CMP Characterization Mask Set (Fig. 1) is designed for rapid CMP consumable, process, and tool characterization and evaluation. Each mask is designed to produce a 1.2 cm $\times$ 1.2 cm die, but a larger die can be generated by scaling the layout. The first mask, the *area mask*, [Fig. 1(a)] has patterned structures with areas ranging from $10 \times 10\ \mu\text{m}^2$ to $3 \times 3\ \text{mm}^2$ across a variety of pattern densities achieved by altering the fill pattern inside each structure. In addition to the area structures, there are also structures to test the role of geometric orientation (horizontal lines versus vertical lines).

The *pitch mask* is the second mask [Fig. 1(b)]. The density of each structure is fixed at 50% (equal linewidth and linespace), and the pitch is varied from 2 μm to 1000 μm for a total of 36 structures for each die. With the exception of the structures with a pitch less than 20 μm, each structure is 2 mm $\times$ 2 mm in size. Features with a pitch of 20 μm or less are lumped into a single 2 mm $\times$ 2 mm structure. There are also spatial replicates for many of the structures so that pitch effects can be separated from spatial location effects.

In the *density mask*, the third mask [Fig. 1(c)], the pattern density (the ratio of raised metal area in each structure to the total area of each structure) is varied systematically from 4% to 100% from lowest in the lower left corner to greatest in the upper right corner while the pitch of each structure is fixed at 250 μm. A total of 25 structures, each 2 mm $\times$ 2 mm, are arranged in a 5 $\times$ 5 grid, in addition to a border region which extends 1 mm from the edge of the 25 structures to act as a buffer. Without the border, structures with the lowest density would contact structures with the highest density (on a neighboring die) and interact strongly.

The fourth mask, or the *aspect ratio (perimeter/area)* mask [Fig. 1(d)], is designed to explore the role of aspect ratio (the ratio of the length of the structure to the width of the structure) and the ratio of perimeter to area. This mask targets any systematic edge/corner effects which may be present. A total of eight structures replicated twice are designed. The area of each structure is fixed at approximately 1 mm^2 and the ratio of width to length is fixed at 1:1, 1:4, 1:14, 1:16, 1:50, 1:62, 1:82, and 1:100. Each set of 16 structures is replicated six times across the die but with different spacings between structures ranging linearly from 10 to 60 μm for a total of 96 structures.

III. Metrology Techniques

In this section, the application of several metrology tools and techniques are discussed. Specifically, particular attention is directed toward optical interferometry and profilometry techniques as well as atomic force microscopy (AFM) and scanning electron microscopy (SEM).

Optical interferometry is a commonly used technique [23] for directly measuring film thicknesses. It offers reasonably high throughput as well as an absolute thickness measurement assuming the tool is properly calibrated. For the characterization masks, either five-die, nine-die, or full-wafer sampling schemes are useful. In a five die strategy, five-die near the center of the wafer (which is typically more uniform than elsewhere on the wafer) are measured. This is useful for quick analysis and initial model development, but does not provide any information about edge or spatial effects across the wafer. In a nine-die sampling scheme, one die at each edge of the wafer (top, left, right, and bottom) and at the same radius are sampled in addition to five die from the center of the wafer. This technique can provide some information about edge effects. Finally, a full-wafer scenario can be used in which every die is sampled on the wafer. While the throughput of this technique is quite low, more powerful analytic techniques such as those described in [18] can then be used. Automated optical metrology is limited to structures with linewidths greater than 10 μm. For smaller structures (down to about 4 μm), optical metrology can still be used but only in manual model and with less reliable results.

The center of each structure above the metal (or above the field in the case of a trench process) is measured for the specified die, resulting in 20–30 measurements per die. The aspect ratio mask is an exception; in this case each structure from only one of the six spacings regions is sampled (16 measurements), and then one structure from each spacing region is also sampled across all spacings (five additional measurements) for a total of 21 structures. If desired, the thickness in the field next to the metal lines can also be measured to provide planarization information, and knowledge about the polishing time can be used to compute removal and planarization rates.

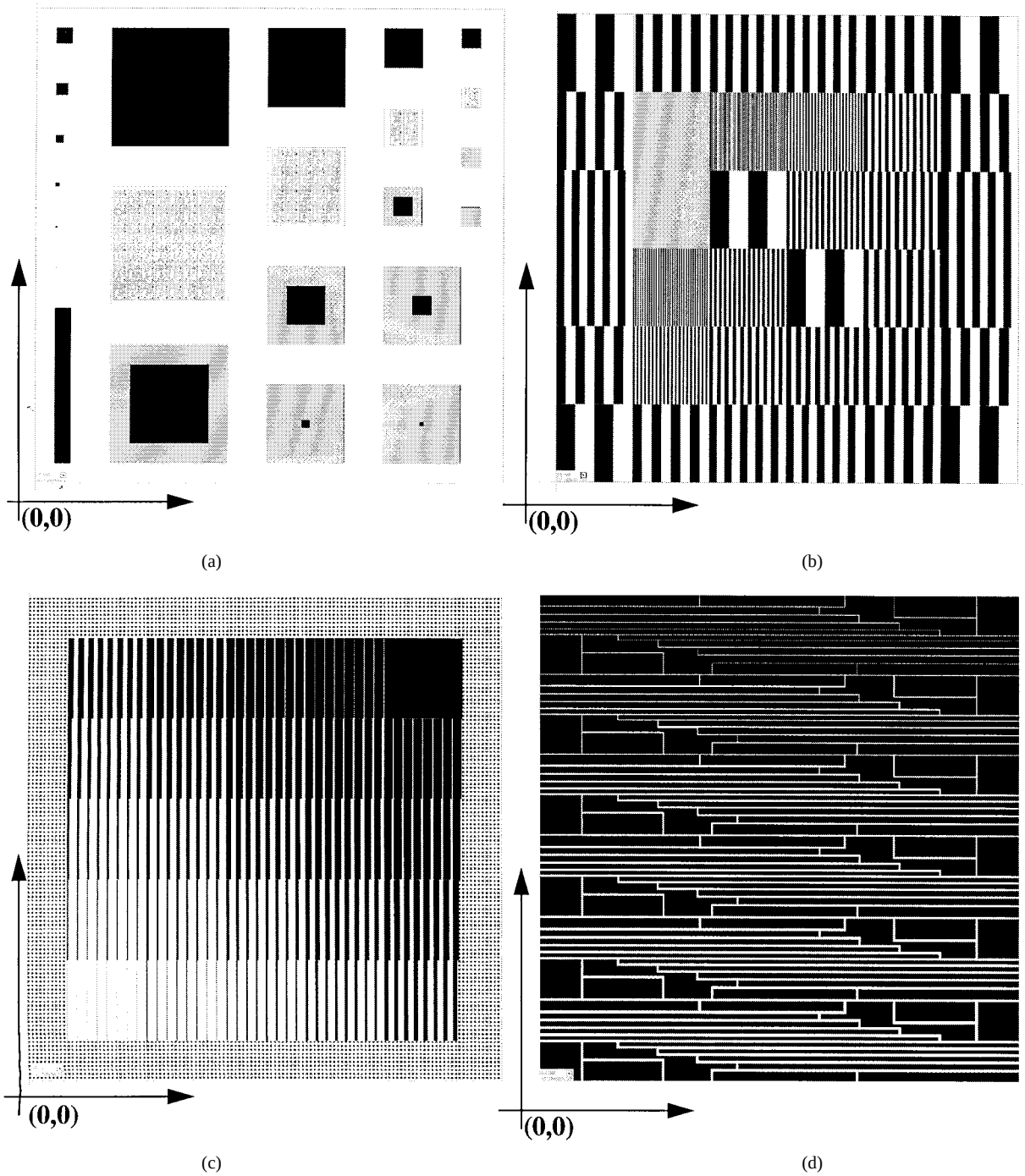

Fig. 1. The CMP characterization masks: the area, pitch, density, and aspect ratio masks are shown in Fig. 1(a)–(d), respectively. Each die is 12 mm × 12 mm.

In profilometry, a sharp stylus is dragged across the surface of interest and deviations in the stylus are measured [23]. Profilometry can be used in three-dimensional (3-D) mode to generate an entire die map with high throughput or in two-dimensional (2-D) mode to generate planarization information over centimeter scale distances. In profilometry, measurements are susceptible to stage tilt and bias as well as wafer bow and warp. This problem can be compensated by using a combined optical/profilometry technique: several points are selected on the die which are measured using both optical interferometry and profilometry measurements. The surface formed by the measurements from profilometry is forced to match the surface formed by the optically measured values; the difference between these two surfaces is called the correction surface. The effectiveness of this technique is dictated by the number of measurements which make up the correction surface: using only five measurement sites to form the correction surface enables correction for linear deviations.

AFM and SEM [23] are also useful metrology tools for use with the characterization masks. In each case, detailed information about planarization can be obtained, and small structures which cannot be measured using optical interferometry can also be examined. Since the effective throughput of these techniques is extremely low, the application of these

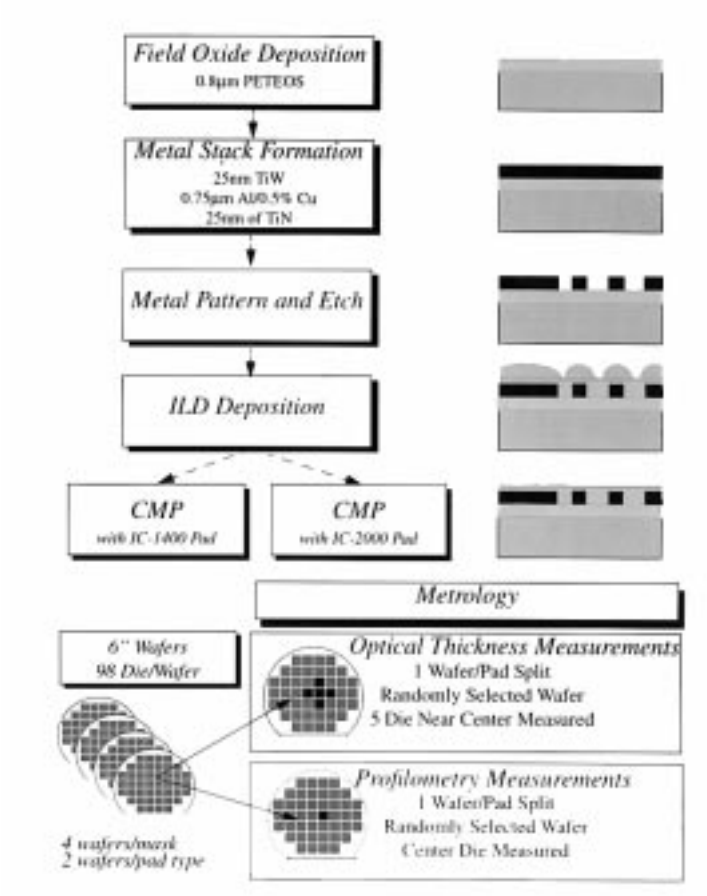

Fig. 2. The short-flow process flow used in the pad comparison study.

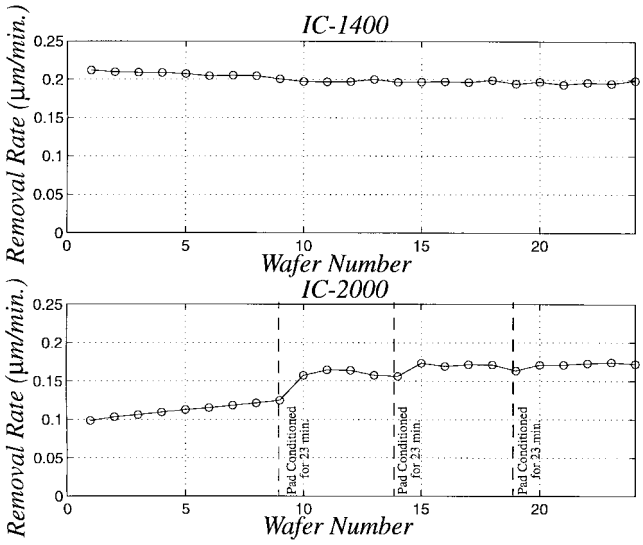

Fig. 3. Average removal rate versus wafer number for the first few wafers on each pad type showing the break-in characteristics of each pad.

tools is limited to only a few key measurements. An additional drawback of SEM techniques is that they are destructive.

Finally, a simple but effective approach should not be neglected: visual inspection and interpretation of fabricated wafers. In certain circumstances, differences between structures and wafers can be discerned visually by looking for subtle color variations between each case, and a color chart

TABLE I
POLISHING PARAMETERS USED IN THE PAD COMPARISON STUDY

Tool Settings	
Down Force	7 PSI
Table Speed	32 RPM
Quill Speed	28 RPM
Back Pressure	0.2 PSI
Slurry Flow Rate	150 mL/min
Consumable Settings	
Slurry	SS-12
Carrier	Flat
Carrier Insert	R200 T3
Conditioning Technique	in-situ concurrent conditioning with a flat profile

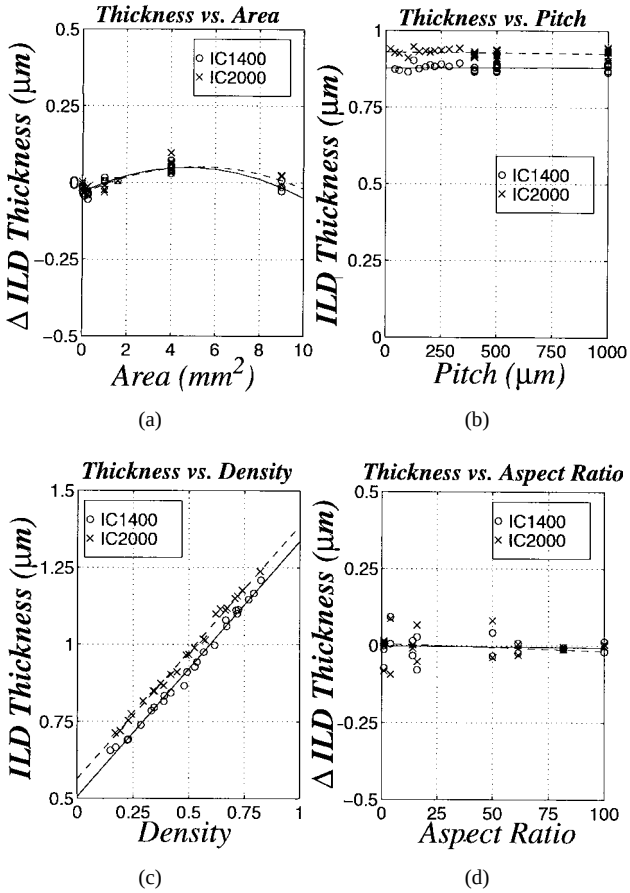

Fig. 4. ILD thickness versus structure area, pitch, pattern density, and aspect ratio for each pad type. Note that for the area and aspect ratio data, ΔILD thickness values are listed since the effect of pattern density has been removed from the data.

[23] can be used to roughly quantify thickness differences. Although simple, this technique can be quite useful in identifying gross differences associated with particular consumable, process, and tool choices during rapid evaluation or optimization.

IV. EXPERIMENTAL RESULTS

In this section, a pad comparison experiment and raw data are described to illustrate the use and application of the

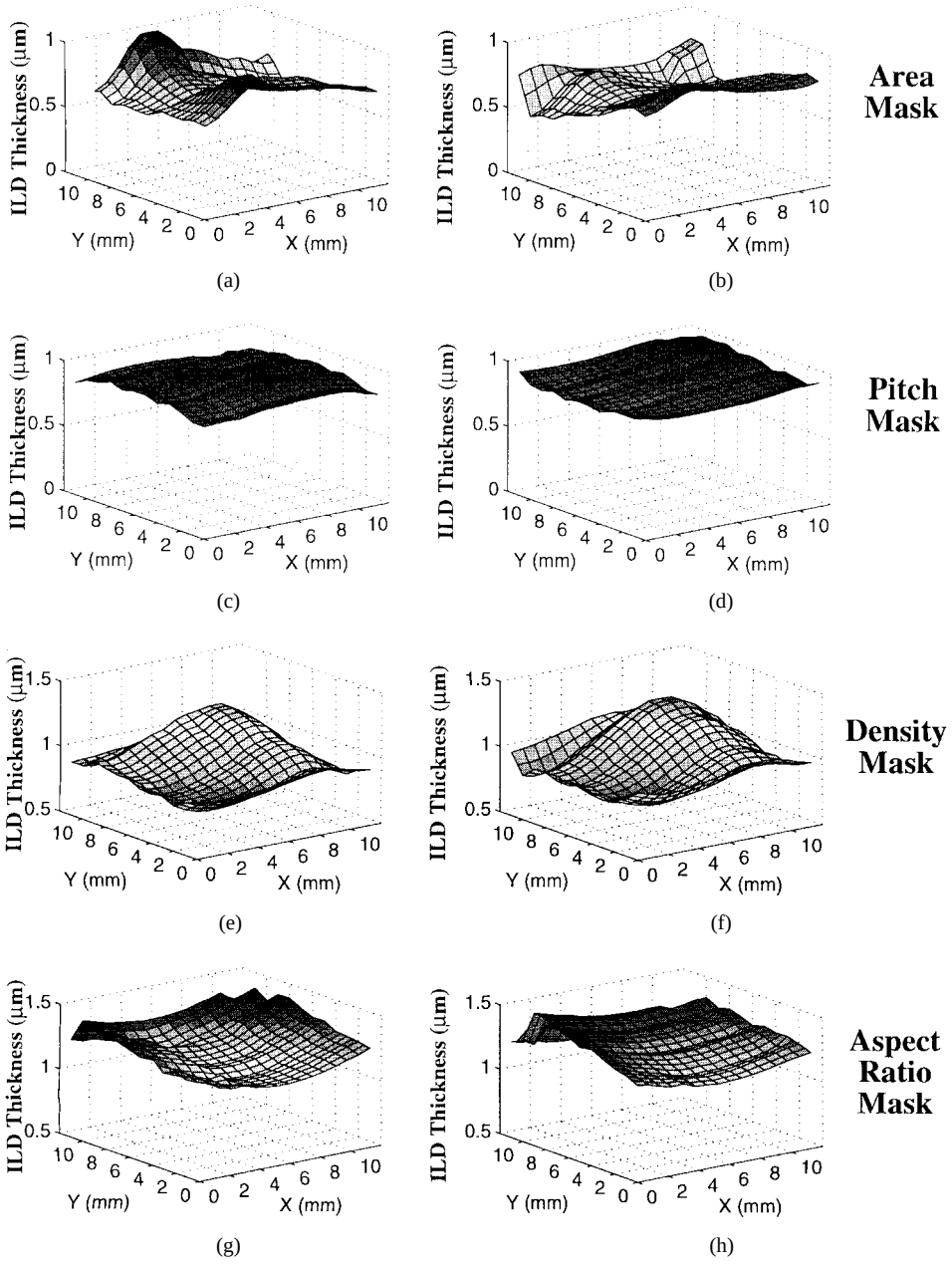

Fig. 5. The 3-D profilometry measurements for each mask across each pad type. The profilometry measurements have been leveled using the combined optical/profilometry techniques.

characterization mask set. The analysis and modeling of this data is discussed in the following section. For this experiment, a back-end process is used in which field oxide is initially deposited followed by metal deposition, pattern, and etch. After depositing the ILD (inter layer dielectric), the wafers are polished to achieve uniform surface heights. A detailed process flow used in this study can be found in Fig. 2.

For the polishing process, the wafers are split on polishing pad type. Half of the wafers are polished with an IC-1400 polishing pad while the other half are polished with an IC-2000 polishing pad, both from Rodel, on an IPEC/Planar 472 polishing tool. The polishing parameters for the pad experiment can be found in Table I. Several dummy wafers, or wafers which contain only a thick blanket oxide layer, are polished on each pad type to "break-in" the pad, i.e., to stabilize the removal rate of the pad. As Fig. 3 shows, the IC-2000 pad has a much longer break-in period than the IC-1400 pad. Longer break-in periods are undesirable and costly in that they lead to excessive usage of monitor wafers and slurry. Note, however, that the polishing process used in this experiment was not optimized for use with the IC-2000 pad and only slightly optimized for use with the IC-1400 pad.

Fig. 4(a)–(d) show ILD thickness, measured on a Prometrix FT-650—an optical thin film measurement tool, versus area, pitch, density, and aspect ratio, respectively. Each data point in Fig. 4 represents the average of five die near the center of the wafer. Note that Fig. 4(a) and (d) are shown as ΔILD thickness versus area and aspect ratio. ΔILD represents the component of ILD thickness variation that can be explained by area or aspect ratio alone. The component of ILD thickness

variation which could be explained by differences in pattern density was removed prior to determination of the area or aspect ratio contribution; this procedure is described in detail in Section V.

Fig. 5(a)–(h) show the measured profilometry traces for each mask measured using a Tencor P-22 profilometry for a die near the center. The leveling techniques discussed in the previous section were used to remove wafer bow and warp and stage bias. A linear correction surface was used as discussed in the previous section. In all cases, the origin of the profilometry measurements was chosen as close as possible to the lower left corner of each pattern. The differences between the two pads, especially for the area mask, are discussed and explained in the next section.

In addition to measuring the post-CMP ILD thickness, we also selectively measured the pre-CMP (i.e., immediately after deposition) ILD thickness. From these measurements and based on the fact that we measured five die near the center of the wafer where incoming deposition variation is smallest, we concluded that deposition variation is small (typically ± 10 nm); thus, post-CMP measured thicknesses did not have to be corrected for incoming film deposition variation.

V. Pattern-Dependent Variation Model Generation

In this section, the use of the characterization masks for pattern dependent variation modeling in CMP processes is presented. Several modeling methodologies, spanning a range of applicability and sophistication, are presented for each mask and applied to data obtained from the pad experiment described in Section IV. The resulting models are especially useful for verifying physical models, for process optimization, or for studying the impact of variation on circuit performance or manufacturability.

In order to generate semi-empirical models, two wafers were randomly selected for each mask—one from each pad split. Five die from the center of each wafer were averaged together to partially suppress within-wafer effects. Models of ILD thickness variation were formed as a function of area, pitch, density, and perimeter/area (aspect ratio).

A. Pattern Density Modeling

As reported in the literature [3], [4], [14], [15], [24] and apparent from a visual inspection of the data (Figs. 4 and 5), the ILD thickness is quite sensitive to pattern density. A major obstacle to modeling pattern density dependencies in CMP rests with finding a suitable and compact definition for a density metric which not only yields a good fit to the available data, but is also physically intuitive [25], [26].

An example helps to illustrate the definition of pattern density. Fig. 6 shows a simple cross section through a fictitious test structure composed of two 1 mm wide metal lines separated by 1 mm and a 5 mm line which is separated from the 1 mm lines by 3.5 mm. Since the lines are very wide, we can assume that the deposition profile can be approximated by the metal profile. In this example, a 1.5-μm layer of oxide was deposited. In many situations the deposition is conformal and not as shown in Fig. 6, and the oxide profile

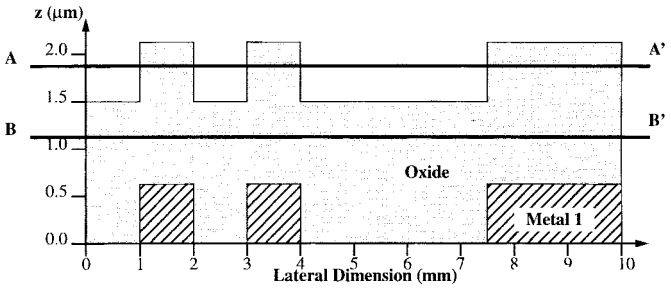

Fig. 6. A simplified example to aid in the definition of pattern-density.

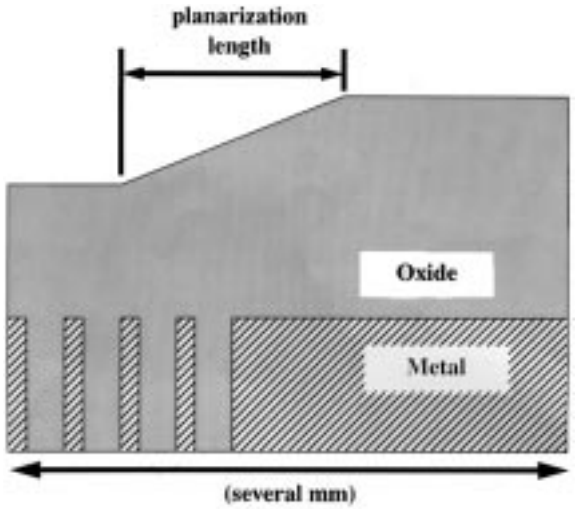

Fig. 7. The definition of planarization length. Typical planarization lengths are on a scale of several millimeters.

cannot always be approximated by the metal profile; this is most evident in tight pitches or small spaces. For this reason, computations of pattern density also depend upon accurate deposition profiles or models, and deposition parameters, tools, and materials are an important integration/modeling issue in CMP [26]. Fortunately, the linewidth and space (except for the pitch mask) are all greater than 10 μm, and approximating the deposition profile with the metal profile is a reasonable approximation.

According to the model proposed by Stine *et al*. [25], the relationship between ILD thickness and pattern density can be expressed as

$$z = z_0 - z_1 - Kt + \rho z_1 \quad \text{for } t > \frac{z_1}{K} \tag{1}$$

where z is the ILD thickness referenced from the top of metal regions, z_0 is the amount of dielectric deposited before CMP, z_1 is the as-deposited step height, K is the removal rate of blanket or unpatterned wafers, t is time, and ρ is pattern density.

In the literature, much discussion has been generated over what is the planarization length for any particular CMP process. Fig. 7 illustrates a definition of planarization distance. If one seeks to planarize "vertical" oxide profiles (as shown in Fig. 6) over two regions with a step change in pattern density (as shown in Fig. 7), the low-density region will polish faster than the high-density region with a transition ramp in between

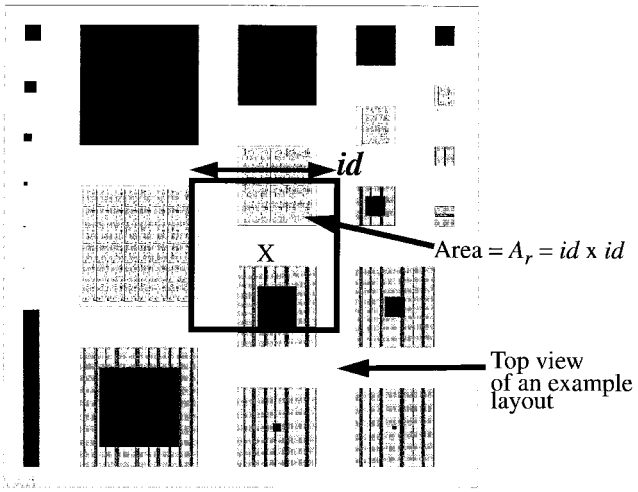

Fig. 8. The definition of interaction distance shown on an example layout.

the two regions. The final oxide profile will be similar to the ramp shape shown in Fig. 7, and the planarization length is defined to be the width of this transition ramp.

In this paper, we define pattern density at a particular location (x, y) on a die as the area of all polygons inside a square region called the *density window* (see Fig. 8) divided by the area of the density window. We call the length of a side of the density window the *interaction distance (id)*. An intuitive physical interpretation of the *interaction distance* is the macroscopic distance over which the pad bends and conforms to the wafer surface and is typically several mm. In the pad experiment described in Section IV, the IC-2000 is a stiffer pad compared to the IC-1400 and also has a longer interaction distance as revealed below. It can also be shown for pattern density computed using a square density window that the interaction distance is identical to the planarization length.

Regardless of the techniques used to compute pattern density, a procedure is needed for optimizing the choice of the interaction distance parameter. Three ways are available for determining the interaction distance:

1) via direct measurement of the planarization length;
2) via a slope regression method; or
3) via a maximum R^2 method.

Direct measurement of the planarization length is by far the quickest and most direct method of determining the interaction distance. Unfortunately, none of the structures on any of the characterization masks are large enough or contain a large step change in density; thus, this technique was not used in this experiment.

In the slope regression method, the slope of a line fit of ILD thickness to pattern density for a particular mask is computed over a wide range of interaction distances. Thus, if the pattern density is computed at the proper interaction distance, the slope of a regressed line of ILD thickness to pattern density should equal the as-deposited step height. If the regressed slope does not equal the measured as-deposited step height, the procedure continues on a different choice of interaction distance until convergence is obtained.

Fig. 9(a) and (c) shows a plot of the regressed slope versus interaction distance for data from the density mask and the

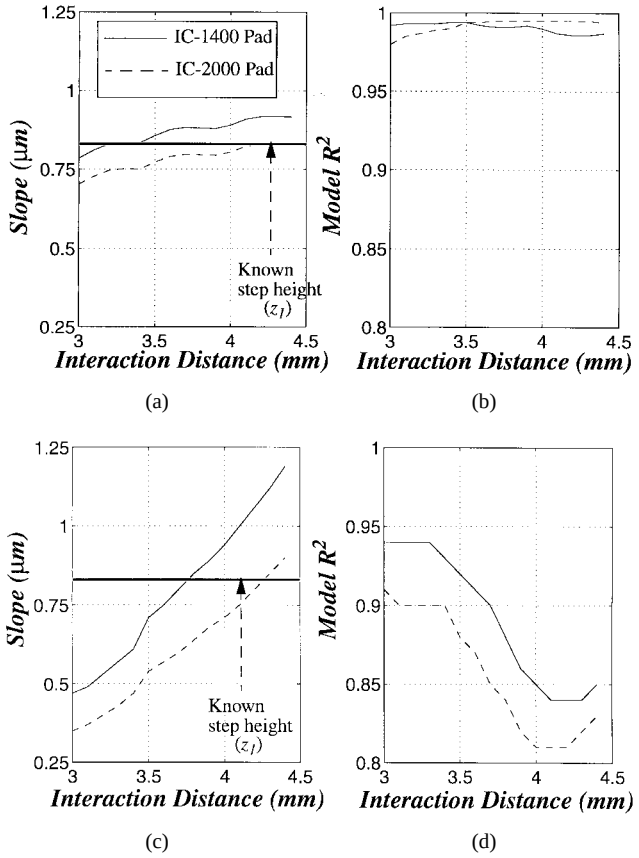

Fig. 9. Estimated slope of ILD thickness versus pattern density computed at different interaction distances from 3 to 4.5 mm for (a) the density mask and (c) for the area mask. For comparison the maximum R^2 method is shown in (b) and (d) for the density mask and for the area mask, respectively.

area mask. The amount of dielectric deposited (z_0) was known with considerable certainty, but the as-deposited step height, z_1, could only be estimated. Fortunately, data from a similar process was available and the amount of pattern sensitivity in z_1 is small. Based on this information, z_1 was estimated at 0.83 μm} $\pm$ 0.01 μm. As Fig. 9 shows, the estimated interaction distance is 3.2–3.4 mm for the IC-1400 pad and 4.2–4.4 mm for the IC-2000 pad as measured from the density mask data. The area mask data indicates that the interaction distance is 3.6 mm for the IC-1400 pad and 4.2 mm for the IC-2000 pad. Note that the area mask is more accurate in detecting the interaction distance compared to the density mask. Since the density mask was designed with gradual gradations in pattern density from one region to the next, this result is not surprising. The area mask has density values spread more randomly across the mask. This unplanned design feature results in greater discrimination of the interaction distance.

In the maximum R^2 method [26], [27], the interaction distance is determined by regressing a linear function of pattern density to ILD thickness for several different values of the interaction distance. The R^2 value, a measure of the quality of the model, is computed for each value of interaction distance. The interaction distance which maximizes R^2 is selected. We have found that the maximum R^2 method is unreliable and imprecise and can yield misleading results. Fig. 9(b) and (d) shows a plot of R^2 versus interaction distance

for data from the density mask and the area mask. We note that the slope regression method is considerably more precise in detecting the interaction distance, and for the area mask, the interaction distance extracted using the R^2 method is significantly different than the slope regression method. This is due to the presence of an underlying design factor, in this case area, confounding with density. In the density mask, there are no significant confounding design factors available to degrade the determination of the interaction distance; thus, the R^2 method can still be used albeit with diminished precision compared to the slope regression method.

Table II and Fig. 4 show the fitted model of ILD thickness versus density. The slopes of the model for each pad are essentially identical by construction. The intercept for the IC-1400 pad data is less than the intercept for the IC-2000 pad data since the blanket removal rate of the IC-1400 pad is greater than the blanket removal rate for the IC-2000 pad and all wafers were polished for approximately the same amount of time. The interaction distance, however, is significantly different for each pad type. The larger interaction distance for the IC-2000 pad compared to the IC-1400 pad correlates well with the physical stiffness/hardness of the IC-2000 pad versus the IC-1400 pad.

The difference in interaction distance between the two pad types explains the differences in the data for the area mask shown in Fig. 5. For an interaction distance near 3 mm (IC-1400), the pattern density near the center of the larger features in Fig. 1(a) is nearly 100% while for an interaction distance near 4 mm (IC-2000) and for the same structure, a substantial amount of low density regions surrounding the large features are averaged together. Thus, transitions between high and low densities with the IC-2000 pad are substantially smoother and a smaller range in pattern density and ILD thickness variation can be expected compared to the IC-1400 pad.

Since the area mask set also has structures which span different ranges of pattern density, the area mask set data can also be used to check the predictive power of models generated from the masks. Fig. 10 shows the predicted ILD thickness variation and observed ILD thickness variation measured using the combined optical/profilometry technique on the area mask. The ILD thickness variation model in (1) was used with an interaction distance of 3.4 mm (extracted from the density mask). As the results show, the agreement is good. Discrepancies between the prediction and the measured values is due to lack of precise knowledge about the shape of the density window, uncompensated metrology errors, and second-order effects.

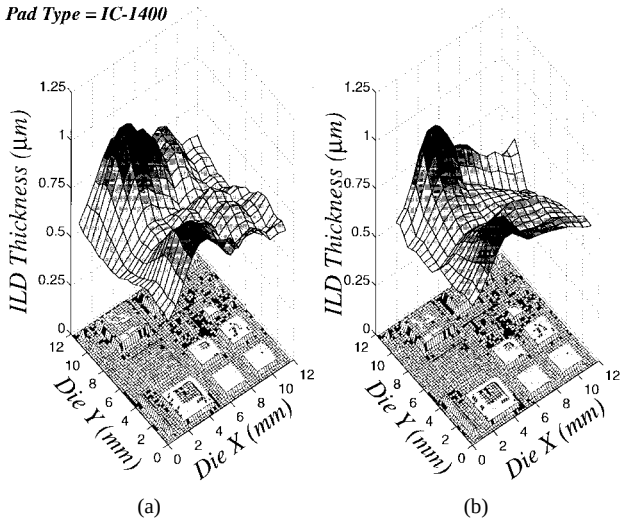

Fig. 10. Profilometry measurements for a center die from (b) the area mask versus (a) the predicted ILD thickness based on computed pattern density. Discrepancies between the model and profilometry measurements are due to lack of knowledge about the window shape, uncorrected stage bias and wafer warping, and second-order effects.

B. Pitch Modeling

Unlike pattern density, pitch is a locally defined parameter; thus, relatively simple procedures can be used to develop a model for the dependence of ILD thickness variation on pitch. A simple regression procedure, assuming a linear or polynomial model, to the available data can be performed. A slightly more accurate model can be generated by removing any spatial dependence (such spatial dependence may arise, for example due to any within-wafer nonuniformity localized within the measured die). The spatial dependence of the data can be estimated by generating a surface formed by replicated structures with the same pitch, such as for the 1000-μm pitch structures which were replicated six times across the die. This surface, as expressed as a function of x and y coordinates, can then be used to estimate the contribution of spatial position to each observation. Table II and Fig. 4 shows the model and regression line and data taken from a wafer from each pad split averaged over five interior die and with spatial dependence removed. A key observation is that the dependence of ILD thickness on pitch, at least for pitches greater than 40 μm, is very slight. Also, the slope of the model does not appear to vary across pad type as indicated by the confidence intervals for the slope coefficient in Table II. As before, the intercept values are different because of unequal removal rates across similar polishing times for each pad split. To compute the confidence intervals, a simultaneous confidence interval at a 95% level of confidence was used [28].

C. Minor Effects

Since structure area and aspect ratio are confounded with pattern density, a simple regression based approach, as in the pitch mask, cannot be used. For data from these masks, the portion of variation which can be attributed to pattern density must be removed first by modeling the ILD thickness variation for all structures as a function of density. The interaction distance for this density model should also be determined via one of the methods outlined in the previous subsection. Alternatively, the density model and interaction distance obtained from the density mask can be used. Once a density model has been assembled, the contribution of area or aspect ratio to ILD thickness variation can be estimated by subtracting the component which can be attributed to density from the observed data. Finally, the remaining variation can be fitted to a linear or polynomial function of area or aspect ratio.

TABLE II
ILD THICKNESS VARIATION MODELS FOR THE PAD EXPERIMENT

Model

Density Mask

IC-1 400

ILD = 0.83ρ + 0.51 [μm] (R^2 = 0.993) **(id = 3.2-3.4mm)**

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	0.51	0.0078	64.7042	0.0000	+/- 0.0205
ρ	0.83	0.0145	57.4407	0.0000	+/- 0.0382

IC-2000

ILD = 0.82ρ + 0.56 [μm] (R^2 = 0.995) **(id = 4.2-4.4mm)**

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	0.56	0.0066	84.9305	0.0000	+/- 0.0174
ρ	0.82	0.0128	64.4719	0.0000	+/- 0.0337

Pitch Mask

IC-1400

ILD = 2.6932×10^{-6}pitch + 0.86 [μm] (R^2 = 0.0088)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	0.86	0.0027	316.4	0.0000	+/- 0.0069
pitch [μm]	0.00	0.0000	0.5373	0.5975	+/- 1.3×10^{-5}

IC-2000

ILD = -5.3658×10^{-6}pitch + 0.91 [μm] (R^2 = 0.0097)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	0.91	0.0028	320.5640	0.0000	+/- 0.0072
pitch [μm]	0.00	0.0000	-0.5602	0.5793	+/- 1.4×10^{-5}

Area Mask

IC-1400

ILD = 0.8ρ + 0.54 + ΔILD [μm] (R^2 = 0.89) **(id = 3.6mm)**

ΔILD = $-0.003582\text{Area}^2 + 0.0339\text{Area} - 0.0312$ [μm] (R^2 = 0.84)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	-0.0349	0.0051	-6.9016	0.0000	+/- 0.0156
Area [mm^2]	0.0367	0.0036	10.0861	0.0000	+/- 0.0110
Area^2 [mm^4]	-0.0038	0.0004	-9.3123	0.0000	+/- 0.0012

IC-2000

ILD = 0.8ρ + 0.60 + ΔILD [μm] (R^2 = 0.82) **(id = 4.2mm)**

ΔILD = $-0.003182\text{Area}^2 + 0.0335\text{Area} - 0.0365$ [μm] (R^2 = 0.75)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	-0.0365	0.0065	-4.5949	0.0002	+/- 0.0199
Area [mm^2]	0.0335	0.0047	7.1488	0.0000	+/- 0.0144
Area^2 [mm^4]	-0.0032	0.0005	-6.0541	0.0000	+/- 0.0015

Aspect Ratio Mask

IC-1400

ILD = 0.80ρ + 0.63 + ΔILD [μm] **(id = 3.6mm)**

ΔILD = -4.62056×10^{-5}AR - 0.00193 [μm] (R^2 = 0.002)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	-0.0019	0.0109	-0.1764	0.8619	+/- 0.0289
AR	-4.6×10^{-5}	0.0002	-0.1994	0.8441	+/- 0.0005

IC-2000

ILD = 0.80ρ + 0.63 + ΔILD [μm] **(id = 4.2mm)**

ΔILD = -2.38×10^{-4}AR + 0.00576 [μm] (R^2 = 0.037)

Coefficient	Value	Std. Err.	t value	Pr(>|t|)	Conf. Int.
(intercept)	0.0058	0.0132	0.4370	0.6670	+/- 0.0350
AR	-2.4×10^{-4}	0.0003	-0.8555	0.4029	+/- 0.0008

Computing a model of density and area or density and aspect ratio (perimeter/area) simultaneously is inappropriate since the area, and aspect ratio are correlated with pattern density and the ILD thickness variation model (1) does not consider these effects simultaneously; thus, the model must be constructed hierarchically (i.e., the density dependence removed prior to any model development based on area or aspect ratio).

Table II and Fig. 4 shows the resulting area model for a wafer from each pad split. The density model used to remove any underlying density dependence for the area mask is also shown in Table II. In order to determine the appropriate model form and order, a leave-one-out/add-one-term stepwise regression technique based on Mallow's Cp statistic was used [28], [29]. From the stepwise procedure, a second-order polynomial model emerged as the best choice for each pad split as shown in Table II. From the 95% simultaneous confidence intervals and visual inspection of the data, the effect due to area for the two pad types is small.

The aspect ratio model for a representative wafer from each pad split is shown in Table II and Fig. 4. Again, any underlying density dependence had to be removed; however, because the range of pattern density spanned by the aspect

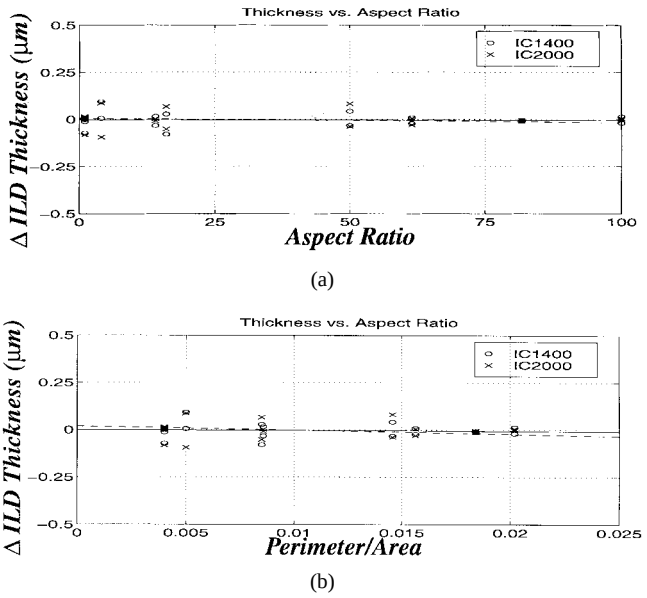

Fig. 11. A comparison of the dependence of ILD thickness on (b) perimeter/area versus (a) aspect ratio for sample wafers from each pad type. The effect of pattern density has been removed to show only the perimeter/area or aspect ratio dependence.

ratio mask is small (on the order of 20%), the estimation of the interaction distance directly from the data for the aspect ratio mask via the slope regression technique was not used. Instead, the density dependence for the area mask was assumed to apply to the aspect ratio mask data as well since all wafers were polished with the same polishing conditions and at the same time except for the split on pad type. As the data shows, an aspect ratio effect is essentially nonexistent as measured by the 95% simultaneous confidence intervals. In fact, an examination of the t-statistic for each coefficient and the overall model R^2 value indicates the best model should be $\Delta \mathrm{ILD} = 0$.

In addition to aspect ratio, the aspect ratio mask can also be used to model the effect of the ratio of perimeter to area on ILD thickness variation using the same techniques mentioned above in connection to aspect ratio modeling. The resulting perimeter to area model and data as well as the aspect ratio model and data can be seen in Fig. 11 for representative wafers from each pad split. Again, no statistically significant difference was seen across the pad split nor was any sizable perimeter/area dependence detected.

From the models and confidence intervals displayed in Table II, the significance of differences between pads across different models can be identified. In this experiment we find that the pattern dependent behavior of the two pads is very nearly the same except for a difference in pad interaction distance (3.4 mm for the IC-1400 compared to 4.2 mm for the IC-2000).

VI. Summary

The masks and associated metrology and analysis methods presented in this paper represent important tools for rapidly evaluating the role of CMP process, consumable, and tool options on polishing performance, and should have application to other hybrid or novel planarization schemes. The data and models generated using the test masks can serve as an important base for evaluating and verifying physically based models and in investigating manufacturability/yield issues associated with CMP—especially for process integration, development, or implementation issues. Finally, the resulting characterization models for ILD polishing as a function of key layout pattern dependencies have the potential for use in layout-specific evaluation of circuit performance degradation due to thickness variation.

In addition to contributing new mask designs for exploration of pattern-dependencies, several metrology tools and techniques have been presented, as well as analysis tools and methods, for generating ILD thickness models using relatively simple procedures. A simple pad comparison experiment has been described. For the process and pads studied in this experiment, we find that density is the dominant layout factor contributing to pattern-dependent variation, and that area, pitch, and aspect ratio represent a second-order or no measurable effect.

Several extensions and applications of this work can be identified. The masks and analysis procedures detailed in this paper are being applied to extensive pad comparison studies and experiments to understand at both an empirical and physical level the role of conditioning and pad properties on ILD thickness variation. Investigation is also underway to evaluate novel process techniques such as silicon nitride capping layers for reducing intra-die ILD thickness variation using the masks. Finally, models extracted from characterization mask data are in use to study the impact of intra-die ILD thickness variation on circuit performance for a number of representative circuit architectures (e.g. SRAM arrays, standard logic cells) through modified layout parasitic extraction procedures.

Acknowledgment

The authors would like to thank J. Schneir, D. Keller, and J. See at Tencor for profilometry measurements and assistance and C. Zenner at Tencor for programming assistance on the Prometrix UV-1050, and M. Fury from Rodel for additional impetus in pursuing research in this area. They would also like to thank T. Hill and K. Achuthan at Sandia National Laboratories, P. Burke and K. Yang from AMD, L. Camilletti and T. Equi from Digital Equipment Corporation, H. Landis from IBM, D. Shy and T. Bibby from IPEC/Planar, M. Berman, D. Towery, S. Prasad, and A. Kapoor from LSI Logic, K. Robinson and S. Miekle from Micron Technology, P. Renteln from National Semiconductor, J. Curry from Strasbaugh, and G. Shinn, I. Ali, C. Appel, H. Hosack, and others at Texas Instruments for feedback in the design of the characterization masks. Finally, the numerous contributions of the members of the staff of Sandia National Laboratories Microsystems Development Laboratory are acknowledged. They would also like to thank the anonymous reviewers for improving the quality and clarity of this paper.

References

[1] I. Ali, S. Roy, and G. Shinn, "Chemical mechanical polishing of interlayer dielectric: A review," *Solid State Technol.*, vol. 37, no. 10, pp. 63–70, Oct. 1994.

[2] M. Fury, "Emerging developments in CMP for semiconductor planarization," *Solid State Technol.*, vol. 38, no. 5, pp. 47–54, Apr. 1995.
[3] P. Burke, "Semi-empirical modeling of SiO_2 chemical mechanical polishing planarization," in *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1991, pp. 379–384.
[4] S. Sivaram, H. Bath, R. Leggett, A. Maury, K. Monnig, and R. Tolles, "Planarizing interlevel dielectrics by chemical-mechanical polishing," *Solid State Technol.*, vol. 35, no. 5, May 1992.
[5] J. Pierce, P. Renteln, W. Burger, and S. Ahn, "Oxide-filled trench isolation planarized using chemical/mechanical polishing," in *Proc. 3rd Int. Symp. ULSI Science Tech. Electrochemical Soc.*, 1991, vol. 91–111, pp. 650–656.
[6] A. Perera, J. Lui, Y. Ku, M. Arzak, B. Taylor, J. Hayden, M. Thompson, and M. Blackwell, "Trench isolation for 0.45 μm active pitch and below," in *IEDM Tech. Dig.*, Dec. 1995, pp. 679–682.
[7] A. Bryant, W. Hansch, and T. Mii, "Characteristics of CMOS device isolation for the ULSI age," in *IEDM Tech. Dig.*, Dec. 1994, pp. 671–674.
[8] C. Kaanta, S. Bombardier, W. Cote, W. Hill, G. Kerszykowski, H. Landis, D. Poindexter, C. Pollard, G. Ross, J. Ryan, S. Wolff, and J. Cronin, *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1991, pp. 144–152.
[9] J. Warnock, "A two-dimensional process model for chemimechanical polish planarization," *J. Electrochem. Soc.*, vol. 138, no. 8, Aug. 1991.
[10] Y. Hayashide, M. Matsuura, and M. Hirayam, "A novel optimization of chemical mechanical polishing (CMP)," in *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1995, pp. 464–470.
[11] P. Renteln, M. Thomas, and J. Pierce, "Characterization of mechanical planarization processes," in *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1990, pp. 57–63.
[12] D. Hetherington, A. Farino, and Y. Strausser, "Characterizing variations in ILD CMP planarization rates using atomic force microscopy," in *Chemical Mechanical Polishing of VLSI/ULSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1996, pp. 74–81.
[13] M. Thomas, P. Renteln, S. Sekigahama, and J. Pierce, "Mechanical planarization process characterization," in *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1990, pp. 438–441.
[14] E. Chang, B. Stine, T. Maung, R. Divecha, D. Boning, J.Chung, K. Chang, G. Ray, D. Bradbury, O. S. Nakagawa, S. Oh, and D. Bartelink, "Using a statistical metrology framework to identify systematic and random sources of die- and wafer-level ILD thickness variation in CMP processes," in *IEDM Tech. Dig*,, Dec. 1995.
[15] B. Stine, D. Boning, J. Chung, L. Camilletti, E. Equi, S. Prasad, W. Loh, and A. Kapoor, "The role of dummy fill patterning practice on intra-die ild thickness variation in CMP processes," *VLSI Multilevel Interconnect Conf.*, Santa Clara, CA, June 1996, pp. 421–423.
[16] D. Bartelink, "Statistical metrology—At the root of manufacturing control," *J. Vac. Sci. Technol. B*, vol. 12, no. 4, pp. 2785–2794, July/Aug. 1994.
[17] D. Boning and J. Chung, "Statistical metrology: Understanding spatial variation in semiconductor manufacturing," Microelectronic Manufacturing Yield, Reliability, and Failure Analysis II, *SPIE 1996 Symp. Microelectronic Manufacturing*, Austin, TX, Oct. 1996.
[18] B. Stine, D. Boning, and J. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," *IEEE Trans. Semiconduct Manufact.*, vol. 10, pp. 24–41, Feb. 1997.
[19] C. Yu, T. Maung, C. Spanos, D. Boning, J. Chung, H. Liu, K. Chang, and D. Bartelink, "Use of short-loop electrical measurements for yield improvement," *IEEE Trans. Semiconduct. Manufact.*, vol. 8, pp. 150–159, May 1994.
[20] B. Stine, D. Boning, and J. Chung, "Inter- and intra-die polysilicon critical dimension variation," Microelectronic Manufacturing Yield, Reliability, and Failure Analysis II, *SPIE 1996 Symp. Microelectronic Manufacturing*, Austin, TX, Oct. 1996.
[21] S. Wolf, *Silicon Processing for the VLSI Era: Volume 2—Process Integration*. Sunset Beach, CA: Lattice, 1990.
[22] J. Steigerwald, R. Zirpoli, S. Muraka, D. Price, and R. Gutmann, "Pattern geometry effects in the chemical-mechanical polishing of inlaid copper structures," *J. Electrochem. Soc.*, vol. 141, no. 10, Oct. 1994.
[23] S. Wolf, *Silicon Processing for the VLSI Era: Volume 1—Process Technology*. Sunset Beach, CA: Lattice, 1986.
[24] R. Divecha, B. Stine, E. Chang, D. Ouma, D. Boning, J. Chung, O. Nakagawa, S. Oh, S. Prasad, W. Loh, and A. Kapoor, "Assessing and characterizing inter- and intra-die variation using a statistical metrology framework: A CMP case study," *First Int. Workshop on Statistical Metrology*, Honolulu, HI, June 1996.
[25] B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung, D. Hetherington, I. Ali, G. Shinn, J. Clark, O. S. Nakagawa, and S.-Y. Oh, "A closed-form analytic model for ILD thickness variation in CMP processes," in *Proc. 2nd Int. Conf. Chemical Mechanical Polishing for ULSI Multilevel Interconnect Conf.*, Santa Clara, CA, Feb. 1997, pp. 267–273.
[26] R. Divecha, B. Stine, D. Ouma, J. Yoon, D. Boning, J. Chung, O. Nakagawa, and S.-Y. Oh, "Effect of fine-line density and pitch on interconnect ILD thickness variation in oxide CMP processes," in *Proc. 2nd Int. Conf. Chemical Mechanical Polishing for ULSI Multilevel Interconnect Conf.*, Santa Clara, CA, Feb. 1997, pp. 29–36.
[27] O. Nakagawa, K. Rahmat, N. Chang, S.-Y. Oh, P. Nikkel, and D. Crook, "Impact of CMP ILD thickness variation in interconnect capacitance and circuit performance," in *Proc. 2nd Int. Conf. Chemical Mechanical Polishing for ULSI Multilevel Interconnect Conf.*, Santa Clara, CA, Feb. 1997, pp. 251–257.
[28] R. Johnson and D. Wichern, *Applied Multivariate Statistical Analysis*. Englewood Cliffs, NJ: Prentice-Hall 1992.
[29] *S-Plus Guide to Statistical & Mathematical Analysis*, Mathsoft, Inc., 1995.

Brian E. Stine received the B.S. degree (summa cum laude) and the M.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, PA in 1993 and 1994 respectively. He is currently a Ph.D. candidate in the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, Cambridge, working on assessing and modeling spatial variation in semiconductor processes and on estimating the impact of spatial variation on circuit performance and product manufacturability. Most of this work has been focused on chemical mechanical polishing (CMP).

Mr. Stine is a member of Tau Beta Pi, Sigma Xi, and Eta Kappa Nu and is the recipient of an Intel Foundation Fellowship.

Dennis O. Ouma (S'95–M'96) received the S.B. and the M.Eng. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 1995. He is currently a Ph.D. candidate in the Department of Electrical Engineering and Computer Science at MIT. His research focuses on modeling of chemical mechanical polishing for dielectric planarization.

Rajesh R. Divecha (S'95–M'96) received the S.B. and M.Eng. degrees in electrical engineering and computer science in 1997 from Massachusetts Institute of Technology, Cambridge.

He is with Rockwell Semiconductor Systems, Newport Beach, CA, where he joined the TCAD group as a design automation engineer. He is involved in layout parasitic extraction, interconnect modeling and process/design interface related issues. Prior to joining Rockwell, he spent two summers as a summer intern at Hewlett-Packard ULSI Labs in Palo Alto, CA where he designed test structures and masks for characterizing CMP induced intra-die dielectric thickness variation. He has authored and coauthored several conference and journal papers in the CMP and statistical metrology field.

Mr. Divecha is a member of Sigma Xi.

Duane S. Boning (M'91) received the S.B. degrees in electrical engineering and in computer science in 1984, and the S.M. and Ph.D. degrees in 1986 and 1991, respectively, all from the Massachusetts Institute of Technology (MIT), Cambridge.

From 1991 to 1993, he was a Member of the Technical Staff, Texas Instruments Semiconductor Process and Device Center, Dallas, TX, where he worked on process/device simulation tool integration, semiconductor process representation, and statistical modeling and optimization. He is currently as Associate Professor in the Electrical Engineering and Computer Science Department at MIT. His research focuses on variation modeling and control in semiconductor processes, with special emphasis on chemical-mechanical polishing and plasma etch. Additional interests include tools and frameworks for process and device design, network technology for distributed design and fabrication, and computer integrated manufacturing.

Dr. Boning is an Associate Editor for the IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, and is a member of Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and the Association of Computing Machinery. He was an NSF Fellow from 1984 to 1989 and an Intel Graduate Fellow in 1990.

James E. Chung (S'87–M'90) received the B.S. degree in electrical engineering from the University of Illinois, Urbana-Champaign, in 1984 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1988 and 1990.

He is an Associate Professor in the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology (MIT), Cambridge, and a member of the MIT Microsystems Technology Laboratories. He has been a member of the MIT faculty since 1991. From 1991 to 1994, he held the Analog Devices Career Development Chair. He spent the fall of 1990 as a Visiting Researcher at the Motorola Advanced Products Research and Development Laboratory, Austin, TX. His research interests are in the areas of MOSFET device physics, VLSI technology and manufacturing, SOI materials and devices, and hot-electron and thin-dielectric reliability.

Dale L. Hetherington (S'87–M'92) received the B.S. degree in mining engineering from West Virginia University, Morgantown, in 1978, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Arizona, Tucson, in 1989 and 1992, respectively.

He is a Senior Member of the Technical Staff at Sandia National Laboratories, Albuquerque, NM, where he is primarily responsible for chemical-mechanical polishing (CMP) and post-CMP process development. He has written and coauthored over 20 technical papers on semiconductor devices and processing. He has been involved in a number of the industrial short courses dealing with CMP technology and is currently chairman of the SEMI standards CMP process metrics task force committee. His research interests include CMP, semiconductor cleaning technology, and multilevel interconnects.

C. Randy Harwood received the B.S. degree in chemistry from Michigan Technological University, Houghton, in 1976 and the M.S. degree in materials science from the University of Vermont, Burlington, in 1996.

His semiconductor experience includes 16 years with IBM, Burlington, as a mask manufacturing and process engineer and as a CMP process engineer in the 200-mm CMOS fab. He joined IPEC/Planar, Phoenix, AZ, in the Process Applications Laboratory in 1995 and is currently the Manager of Internal Process Applications.

O. Samuel Nakagawa received the B.S.E.E. degree from the University of Illinois, Urbana-Champaign, in 1985, and the M.S.E.E. and Ph.D.E.E. degrees from the Pennsylvania State University, University Park, in 1987 and 1994, respectively.

He is a member of technical staff at Hewlett-Packard Laboratory, Palo Alto, CA. His research interests include interconnect process and design modeling for deep-submicron ULSI.

Soo-Young Oh (M'88) received the B.S.E.E. degree from Seoul National University, Seoul, Korea, in 1972, and the M.S.E.E. and Ph.D. degrees in electrical engineering in 1977 and 1980, respectively, from Stanford University, Stanford, CA.

He then joined Hewlett-Packard, Palo Alto, CA, and is currently the Manager of the TCAD group in the ULSI Laboratory at HP Laboratories, working on process and device simulation, statistical metrology, and interconnect modeling and characterization. His work has resulted in over 40 publications, and he coauthored the book *Computer-Aided Design and VLSI Device Development* (Norwell, MA: Kluwer, 1986).