A CLOSED-FORM ANALYTIC MODEL FOR ILD THICKNESS VARIATION IN CMP PROCESSES

B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung
Massachusetts Institute of Technology, Cambridge, MA 02139

D. L. Hetherington, Sandia National Laboratories, Albuquerque, NM 87123

I. Ali, G. Shinn, J. Clark, Texas Instruments, Dallas, TX 75243

O.S. Nakagawa, S.-Y. Oh, Hewlett Packard Co., Palo Alto, CA 94304

ABSTRACT

CMP planarization of oxide results in excellent long-range uniformity compared to other planarization techniques but remains hampered by systematic pattern sensitivities. In the recent literature, several semi-empirical or physically-based models have been proposed to explain ILD thickness pattern sensitivities in CMP, but all of these models either fail to predict key empirical results, are not described fully, or do not present tractable closed form models. In this paper, we develop and derive a closed form model for ILD thickness variation and verify this model on datasets obtained over different polishing tools, consumable sets, and process conditions, and as a function of polishing time.

I. INTRODUCTION

In recent years, chemical-mechanical polishing (CMP) has emerged as the primary technique for planarizing dielectrics [1,2]. Although 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], 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, 3, 4]. This problem has become especially acute as performance requirements have increased and dimensions have scaled. CMP has also found wider application in the entire VLSI development and production cycle serving as an enabling tool for shallow trench isolation [5, 6, 7], damescene technologies [8], and other novel process techniques.

In this paper, we develop a closed form model for intra-die ILD thickness variation – specifically pattern density dependent – and verify this model on datasets obtained over different polishing tools, consumable sets, and process conditions, and as a function of polishing time. In Sections II and III, important background literature and the interrelationships of pattern-density and ILD thickness in CMP will be reviewed and discussed. Also, Section III attempts to more precisely define the concept of pattern density as this has been the source of much confusion in previous papers and discussions. Section IV begins with Preston's equation, a well known model for removal rate on blanket wafers, and couples the concepts and definitions discussed in Section III to form the closed form model for ILD thickness variation. The experimental methodology and model validation is discussed in Section V. Finally, Section VI summarizes the results of this paper and discusses several important limitations of the model.

II. BACKGROUND

CMP process modeling has received increased attention in recent years but the reported models do not provide concise ILD thickness prediction for a given die layout. Models by Burke [3], Hayashide et al. [9] and Renteln [10] are an exception but these are either not fully described or closed form solutions are not provided. Burke proposes closed form models for "up" and "down" areas of a layout structure. The polishing characteristics of the "down" areas are partitioned into linear and logarithmic regimes, and the two regimes are then empirically modeled as functions of blanket polish rate and step height. The "up" areas are modeled by introducing a density enhancement factor into the "down" area models. A planarization (or step height removal) rate is developed as:

$$-\frac{dS}{dt} = \frac{(1 - D_0)}{S_0} \cdot S \cdot U \tag{1}$$

where D_0 is the ratio of the polish rate of "down" areas to "up" areas, S is the step height, S_0 is the initial step height, and U is the polish rate of "up" areas. This equation states that the polish rate is proportional to the step height, but the physical motivation or mechanism for this assumption is not clear.

Hayashide et al. have proposed a model where the polish characteristic of the whole chip is evaluated by partitioning the chip into cells and determining the removal rate of a cell as functions of cell density, height and an enhancement factor. The enhancement factor is obtained by an FEM analysis of the bending characteristics of the pad. This results in possible prediction of edge rounding and polish characteristics of down areas. The density within a cell changes with time but the model does not specify how this is evaluated. The numerical thickness prediction also minimizes the utility of the model for quick evaluation of the relative removal rates of various layout patterns.

Renteln has presented a program which simulates the polishing characteristics of a die given the topography scan of the surface prior to polishing. The details of the implementation of the program are not presented and the utility of the program is limited by its availability.

III. PATTERN DENSITY DEFINITIONS

As reported in the literature [3, 4, 11, 12, 13] and as apparent from simple visual inspection of patterned post-CMP wafers, underlying pattern density is a key factor affecting polish in CMP processes. A major obstacle to modeling pattern density dependencies in CMP rests with finding a suitable and compact definition for a density metric. In this section, we give a specific definition of pattern density and of interaction distance. We then examine the relationship between interaction distance and planarization length often discussed in the literature.

An example helps to define and illustrate subtleties in the definition of pattern density. Figure 1 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. We note that In many situations the deposition is conformal and not as shown in Figure 1, and the oxide profile 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.
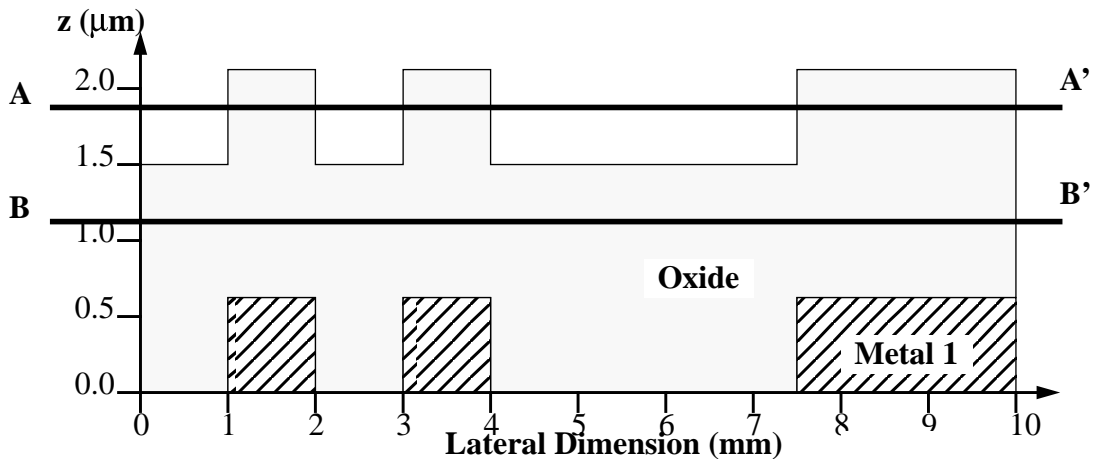

Figure 1. A simple example to aid in defining pattern density.

Pattern density, for the purpose of CMP modeling, can be defined as the v olume fraction of oxide within an infinitesimally thin surface. For example, the surface formed by A-A' in Figure 1 is dz thick and the total volume of oxide inside this surface is $1 + 1 + 2.5 = 4.5$ mm x dz while the maximum volume possible within A-A' is 10mm x dz where dz is an infinitesimal. In this way, the pattern density is $4.5/10 = 0.45 = 45\%$. As Figure 1 shows, however, pattern density is also a function of z; thus, the pattern density at B-B' is $10.0/10.0 = 1 = 100\%$ which is significantly different than the 45% pattern density at surface A-A'. Note that in CMP since material is always being removed at a given rate, the pattern density that the pad "sees" (i.e., the pattern density near the pad-oxide interface) is a function of time.

A critical parameter in this definition of pattern density is the range o ver which pattern density is computed. In the previous example, this value was 10mm. If we define the range over which pattern density is computed as being a surface of area A_r and infinitesimally thin, then a key parameter becomes A_r. Because of the nature of CMP, we would expect A_r not to be as large as the entire chip area nor to be as small as indi-vidual lines. Also, since A_r is formed by a tw o-dimensional surface, the shape (e.g. rectangular, circular, square) of this surface becomes an important parameter. As shown in Figure 2, in this paper we will assume that the volume used in computing pattern density is formed by a square of area A_r and infinitesimally thin. The determination and use of non-square surfaces will be presented elsewhere. We define the width of the square as the *interaction distance*, or *id* and the square of area A_r as the *density window*. Since *id* is generally not equal to the length of the chip side, the pattern density will vary with position (i.e. the pattern density in the upper right hand corner of the layout in Figure 2 is v ery different from the pattern density in the lo wer left hand corner of the layout). An intuiti ve physical interpretation of A_r is as the macroscopic re gion over which the pad bends and conforms to the w afer surface and is typically several mm^2. A procedure to deter-mine the precise value to use for the interaction distance parameter is discussed in Section V.
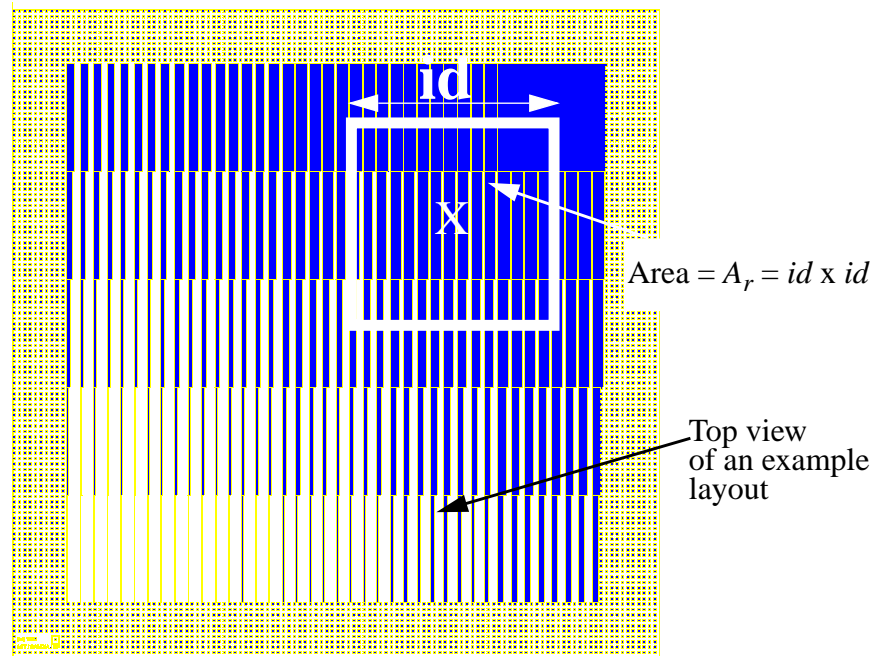

Figure 2. The definition of interaction distance.

In the literature, much discussion has been generated over what is the planarization length for a par-ticular CMP process. Figure 3 illustrates a definition of planarization distance. If one seeks to planarize "ver-tical" oxide profiles (as shown in Figure 1) over two regions with a step change in pattern density (as shown in Figure 3), the low density region will polish faster than the high density region with a transition ramp in between the two regions. The final oxide profile will be similar to the ramp shape shown in Figure 3, and the planarization length is defined to be the width of this transition ramp. It can be sho wn, for pattern density

computed using a square window as described above, that the interaction distance is identical to the planarization length. The equality of planarization distance to interaction distance lends physical intuition to the concept of interaction distance.
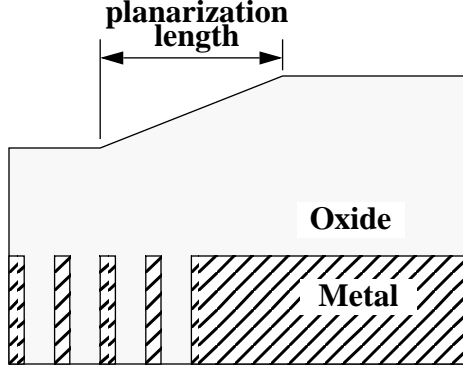

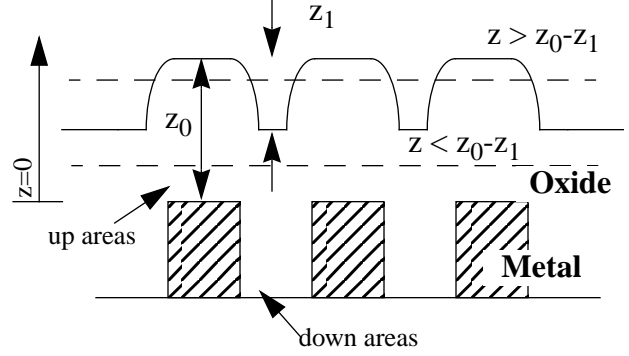**Figure 3. Definition of Planarization Distance**

Figure 4. Important definitions used to develop the closed form ILD thickness variation model.

IV. DERIVATION OF MODEL

The derivation of a closed form expression for ILD thickness variation begins with the well known Preston equation which states that the removal rate on blanket wafers is proportional to the product of pressure and velocity:

$$= \;—\; = \; \kappa P v \tag{2}$$

where κ is a proportionality constant. If the pressure term is represented as F/A where A is the oxide area contacted by the pad then Preston's equation can be rewritten as:

$$— \; = \; \frac{\kappa F v}{(id)^2 \rho(x, y, z)}. \tag{3}$$

In (3), $\rho(x,y,z)$ is pattern density and is a function of x,y since it varies across the chip, and is a function of z since as oxide is removed the pattern density changes (as in Figure 1). Also, note that the removal rate, RR, has been rewritten as a differential and that A has been replaced by $(id)^2\rho(x,y,z)$ – which is the oxide area contacted by the pad at a particular z. In (3), we can lump constants together and rewrite the equation as:

$$— \; = \; \frac{}{\rho(x, y, z)} \qquad K \; = \; \frac{\kappa F v}{(id)^2} \tag{4}$$

and K can be interpreted as the removal rate of a blanket wafer (or 100% density region).

Realistically, $\rho(x,y,z)$ can be expected to also be a function of deposition conditions and local line width and space. For an initial oxide step height of z_1 and an initial oxide thickness of z_0 (see Figure 4), and assuming the deposition profile can be approximated vertically using the metal profile, the pattern density can be approximated as:

$$\rho(x, y, z) \; = \; \begin{cases} \rho_0(x, y) & z > z_0 - z_1 \\ 1 & z < z_0 - z_1 \end{cases}. \tag{5}$$

In addition, we will also assume that "down" areas (or regions between the steps in Figure 4) polish at a negligible rate compared to the "up" regions (or regions near the top of the steps in Figure 4). Substituting (5) into (4) yields a separable ordinary differential equation. This differential equation can be solved for z:

$$z = z_0 - \left(\frac{Kt}{\rho_0(x, y)}\right) \qquad Kt < \rho_0(x, y)z_1$$
$$z = z_0 - z_1 - Kt + \rho_0(x, y)z_1 \qquad Kt > \rho_0(x, y)z_1 \tag{6}$$

Equation (6) implies that if features are planarized for a long enough time (to complete the removal of the local step), a linear relationship between pattern density and ILD thickness results as seen frequently in the literature [14, 15]. We call this region of operation of the model the linear regime, and we term the other case the locally non-planar regime. Note that a transition time can be identified, equal to $\rho z_1/K$, which defines the time for a given pattern density at which local planarization is achieved. Also, (6) implies that the polishing time to guarantee local planarization over all features is $t_t = z_1/K$. After this time, no further planarization can occur (see Figure 5).
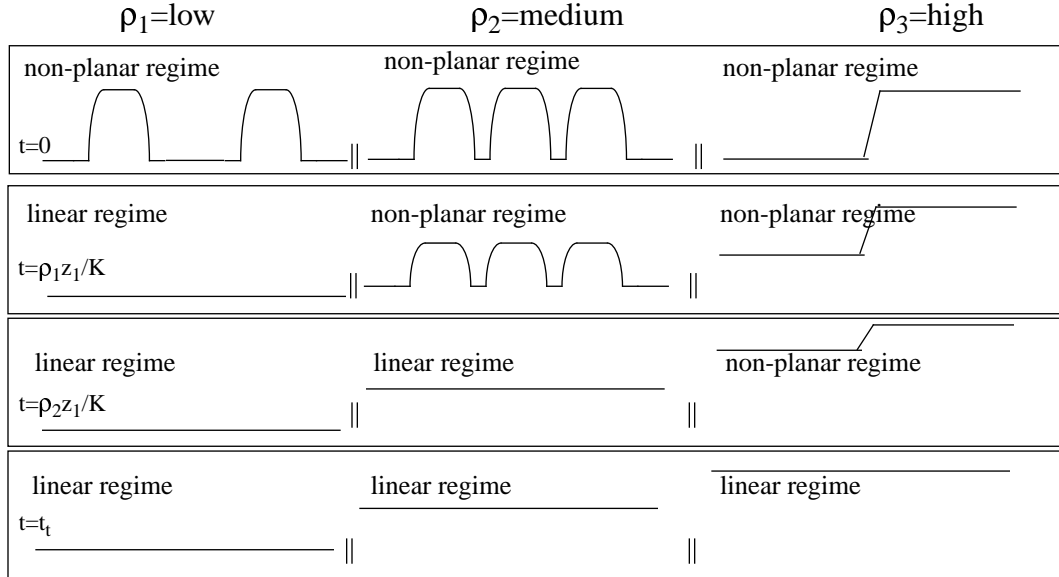

Figure 5. Illustration of the time dependent nature of the model. Initially, all of the features are at the same height. As time progresses, lower density features erode quickly and achieve local planarization. As the polishing process proceeds, features with higher density begin to achieve local planarization. Finally at the transition time, t_t, all of the features have achieved local planarization and no further global planarization is possible.

A physical explanation of this model is basically a polish volume statement. The CMP process removes a certain volume of oxide per unit time regardless of the pattern density (or contact area of the pad). Over high density regions there is much more oxide volume to be removed per vertical increment compared to regions of low density where there is less oxide volume to be removed per vertical dz increment. Thus, oxide over high density regions will ultimately remain thicker than over less dense regions.

V. EXPERIMENTAL VALIDATION OF MODEL

In order to validate the model, the density mask from the CMP characterization mask set [15] was used. This mask, shown in Figure 2, is composed of 25 structures, each 2 mm x 2 mm in size, of a fixed pattern density, at least within each structure, ranging from 4% (lower left corner) to 100% (upper right corner). In this mask, the pitch was fixed at approximately 250 μm and linewidths were kept above 20 μm to permit optical thickness metrology. Since the linewidths and spaces were relatively large, approximating the deposition profile with the metal profile is valid.

The masks were fabricated in a short-flow back-end-of-line manufacturing process beginning with a field oxide deposition for isolation followed by metal deposition and patterning. Finally, a thick layer of TEOS was deposited and planarized. For this experiment, we varied the polishing pad, IC-1000/Suba-IV versus IC-1400, the polishing tool, as well as the process conditions, specifically the down force and the table speed, as shown in Figure 6. Also, wafers were polished at 1/3 and 2/3 the total polishing time in addition to the full polishing time. For each process, pad, and tool, the final polishing time was adjusted so that the final ILD thickness in all cases was approximately the same. For each wafer, nine measurements were taken on each die using optical film thickness metrology tools, and all thickness measurements from the same structure were averaged together to take into account die to die variation. The averaged thickness values for a representative experiment is shown in Table I.

TABLE I. Measured final ILD Thickness (Averaged over all Measured Die on Wafer)

Structure (Designed Density)	0.08	0.16	0.24	0.44	0.52	0.60	0.80	0.88	1.00
Thickness (μm)	1.088	1.148	1.192	1.279	1.383	1.352	1.489	1.563	1.554

For the model shown in (6), an important issue is determining the interaction distance parameter, *id*. One technique for obtaining this parameter is to experimentally measure the planarization distance as discussed in Section III, and use this value for the interaction distance. Another method, which is the one used in the rest of this paper, works by first computing the pattern density from the layout for different interaction distances starting from 2 mm up to some suitably large value. Table II shows the computed pattern density as a function of the interaction distance up to 6 mm for the nine sites which were measured in this experiment. These values are calculated using CAD tools and under the assumption that the deposition profile can be

TABLE II. Pattern Density vs. Interaction Distance for each Measured Structure

Structure (Designed Density)	Interaction Distance (mm)								
	2.0	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**	**5.5**	**6.0**
0.08	0.078	0.136	0.174	0.200	0.220	0.236	0.248	0.258	0.267
0.16	0.157	0.210	0.245	0.270	0.289	0.304	0.315	0.315	0.334
0.24	0.236	0.277	0.304	0.324	0.338	0.350	0.359	0.366	0.373
0.44	0.433	0.450	0.460	0.466	0.471	0.474	0.477	0.479	0.480
0.52	0.501	0.511	0.513	0.515	0.516	0.517	0.517	0.518	0.517
0.60	0.580	0.563	0.553	0.546	0.541	0.538	0.535	0.533	0.531
0.80	0.801	0.755	0.725	0.704	0.688	0.676	0.666	0.658	0.648
0.88	0.882	0.824	0.786	0.759	0.739	0.724	0.711	0.701	0.689
1.00	0.999	0.894	0.830	0.787	0.757	0.734	0.717	0.703	0.688

approximated by the metal profile. According to (6), if all local features have been eroded away ($t > t_t$) then the ILD thickness versus pattern density relationship should be a straight line with a slope of z_1 with respect to density. Thus, if a straight line is regressed on the observed ILD thickness as a function of pattern density, for final polishing times only and at a specified interaction distance, the interaction distance which yields a fit of a line with a slope closest to z_1 is chosen as the optimal interaction distance. Note that in order to minimize any error in this approach the value of z_1 is measured on pre-CMP wafers. The interaction distances for each process and pad ranged in values from 3.2 mm to 3.6 mm. Table III lists the computed slope of the fitted lines versus different interaction distances for one of the process conditions with the optimal choice appearing to be at about 3.5 mm. For this example, the measured value of z_1 was 0.829 μm. Figures 6 and 7 show plots of ILD thickness versus pattern density (evaluated over the optimal interaction distance) for all of

the process conditions and pads examined. Both the model and the experimentally observed values are visible. From these figures, we see that the model explains the observed data quite well.

TABLE III. Slope of ILD Thickness vs. Pattern Density Across Different Interaction Distances.

Interaction Distance (mm)	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
slope	0.528	0.636	0.730	0.812	0.885	0.951	1.009	1.062	1.123

VI. CONCLUSION

In this paper, we have proposed a closed-form model for ILD thickness variation which assumes a density dependent modification to Preston's equation. The density dependence to the removal rate appears to be universal; other deposition profiles in addition to the "vertical" shape presented here can also be accommodated. The appropriate interaction distance or window for computation of density, on the other hand, is pad and process dependent and at present requires empirical determination.

The global planarization model explains a wide spectrum of observed pattern dependent behavior in CMP. At the same time, however, there are instances where the model does not explain observed behavior. First, the model only predicts the polish evolution of the oxide thickness above the metal lines ("up" features) while regions in the spaces ("down" features) are not considered. An assumption in the model is that "up" features polish much faster than "down" features; such polish rate ratios are often on the order of 5-10 [3]. Also, as Figure 8 shows, the model does not explain pattern dependent variation in CMP for extremely low pattern density (typically less than 15%). Almost certainly in this regime other physical effects (e.g. dishing, stress-related acceleration) are playing a large role. Finally, this model only aims at predicting intra-die or pattern dependent variation and not within-wafer variation where different effects such as equipment parameters and wafer stress play a role. These topics and others are the subject of considerable research and will be reported elsewhere as they mature.

ACKNOWLEDGEMENTS

REFERENCES

[1] I. Ali, S. Roy, and G. Shinn, *Solid State Technology*, vol. 37, no. 10, pp. 63-70, October 1994.
[2] M. Fury, *Solid State Technology*, vol. 38, no. 5, pp. 47-54, April 1995.
[3] P. Burke, *VLSI Multilevel Interconnect Conference*, pp. 379-384, Santa Clara, CA, June 1991.
[4] S. Sivaram, et al, *Solid State Technology*, vol. 35, no. 5, May 1992.
[5] J. Pierce, et al., *Proc. of the 3rd Intl. Symp. on ULSI Science and Tech of the Electrochemical Soc.*, vol 91-11, p. 650-656, 1991.
[6] A. Perera, et al., *IEDM Technical Digest*, pp. 679-682, Dec. 1995.
[7] A. Bryant, W. Hansch, T. Mii, *IEDM Technical Digest*, pp. 671-674, Dec. 1994.
[8] C. Kaanta, et al., *VLSI Multilevel Interconnect Conference*, pp. 144-152, Santa Clara, CA, June 1991.
[9] Y. Hayashide, et al., *VLSI Multilevel Interconnect Conference*, pp. 464-470, Santa Clara, CA, June 1995.
[10] P. Renteln, et al., *VLSI Multilevel Interconnect Conference*, pp. 57-63, Santa Clara, CA, June 1990.
[11] E. Chang, et al., *IEDM Tech. Digest*, Dec. 1995.
[12] B. Stine, et al., *VLSI Multilevel Interconnect Conference*, pp. 421-423, Santa Clara, CA, June 1996.
[13] R. Divecha, et al., *First Intl. Workshop on Statistical Metrology*, Honolulu, HI, June 1996.
[14] W. Leipold, et al., *VLSI Multilevel Interconnect Conference*, pp. 473-475, 1995.
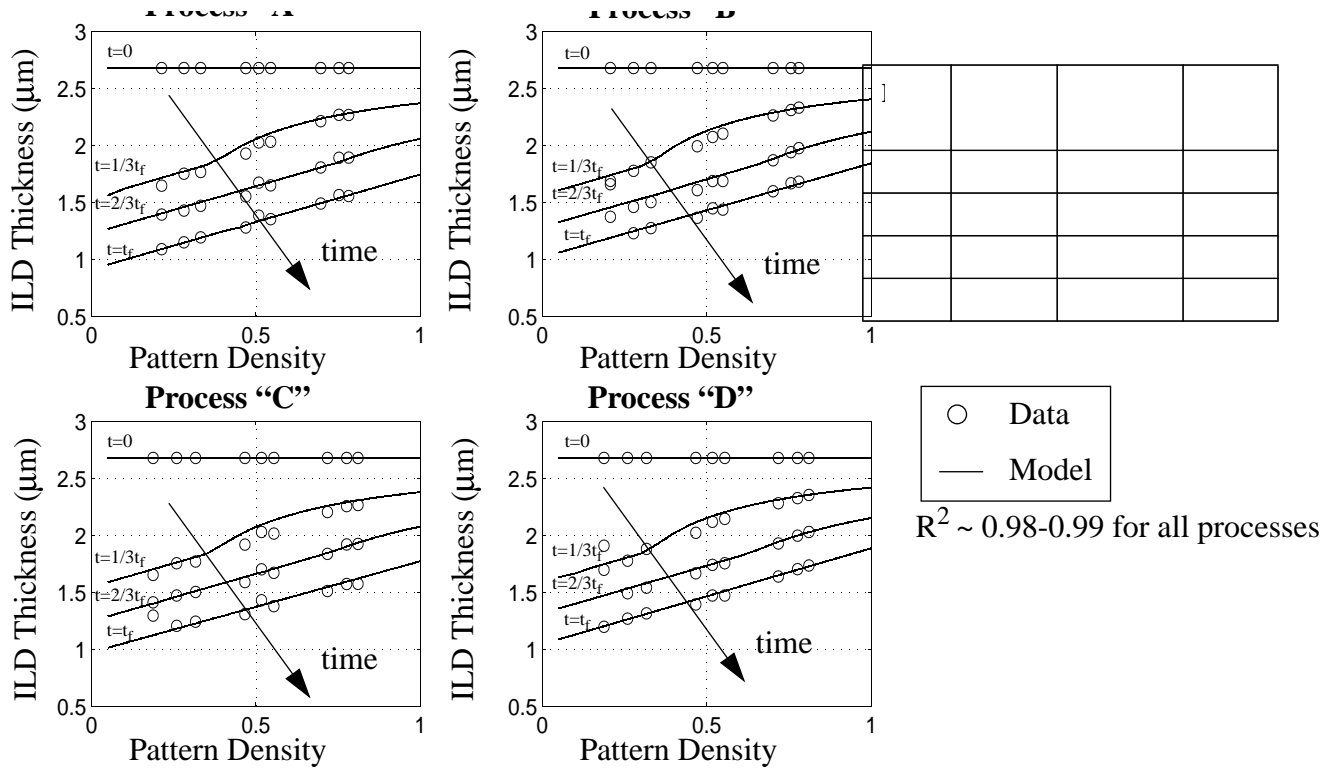[15] B. Stine, et al., Submitted to *IEEE Trans. on Semi. Manuf.*, Oct. 1996.

Figure 6. Experimental verification of the model for different process conditions (on tool "A")
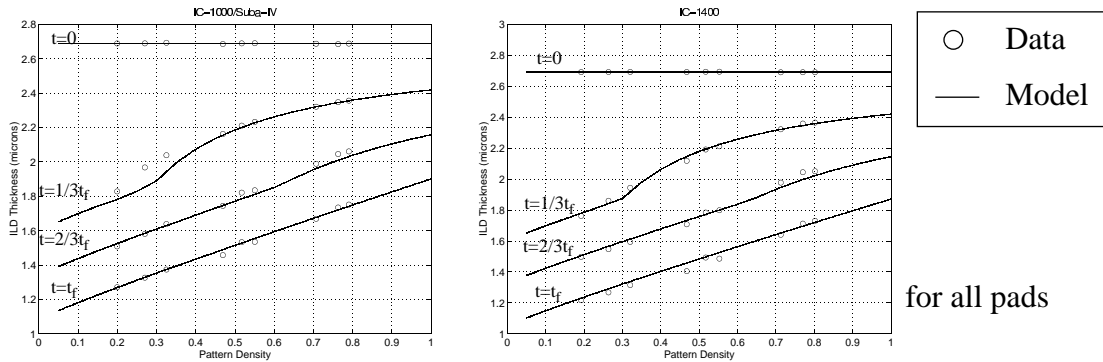

Figure 7. Experimental verification of the model for different polishing pads (on tool "B")
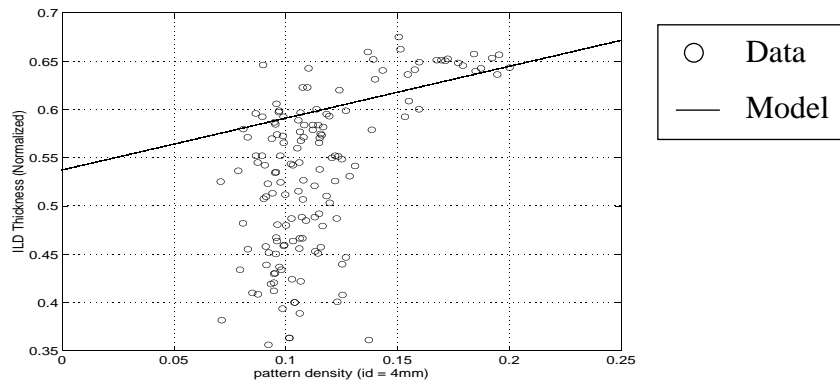

Figure 8. An example of low-density phenomenon not explained by the ILD thickness model presented