

Optimization of Within-Die Planarity and Defectivity for Chemical-mechanical Planarization

Wei-Tsu Tseng,* Raghuveer Patlolla, Injo Ok, Donald Canaperi,
Todd Buley,# and Mark Hucks#

IBM Semiconductor Research, 257 Fuller Road, Albany, NY 12203, USA

#DuPont Electronic and Imaging CMP Technologies, 451 Bellevue Road, Newark, DE 19713, USA

INTRODUCTION

Within-die non-uniformity (WiDNU) resulting from topography differences between features of varying length scales is among the most critical challenges in CMP. Quite often, the tolerance for WiDNU and WiWNU in topography and line height for advanced semiconductor chips such as embedded memories and stacked-FET is on the order of a few nm for some critical layers. Naturally, such nm-scale uniformity requirements along planar (x-y) as well as vertical (z- axis) directions would impose equally stringent defect tolerance because any surface discontinuity caused by defects such as scratches could translate to loss in yield and uniformity. The study here will demonstrate how CMP pads alone can reduce both WiD topography and defectivity. The underlying mechanisms involved will be discussed.

BACKGROUND

In semiconductor industry, the continuing drive towards shrinking geometry in advanced technology nodes imposes pressing challenges to all process technologies. For CMP, one of the most stringent performance requirements is the need to meet within-die uniformity in term of planarity and thickness across patterns of varying critical dimension.

Within a die area of a few cm^2 , long range (i.e., $100\mu\text{m} \sim 1000\mu\text{m}$) planarity is required to ensure uniformity across chiplets and repeating macros. Meanwhile, short range planarity (i.e., $0.1\mu\text{m} \sim 1\mu\text{m}$) is critically needed to meet the electrical specs of macros and devices.

Materials chemistry, surface morphology, and mechanical characteristics of CMP pads play crucial roles in removal rates, defectivity, and planarization efficiency. The usual practice in the industry is to apply “hard” pads for planarization. Besides the hardness of top pad itself, surface textures, the type of sub-pad (e.g., with different thickness and compliance) can modulate long range planarity.

On the other hand, dishing and erosion within a macro leads to loss in short range planarity. In this case, the amount of topography difference within the macro is often smaller (\sim a few nm) compared with that required for long range planarity (i.e., 100s of nm) in advanced tech node. However, their impact on device performance can be significant, as illustrated in Fig. 1 in the case of Cu interconnects.

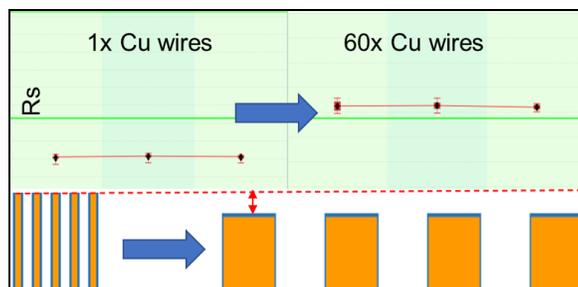


Fig. 1: R_s delta between 1x and 60x Cu wires resulting from deficiency in short-range planarity.

For FEOL oxide and STI CMP application, “hard pad” combined with ceria slurry is widely adopted to meet the planarity and selectivity requirements. In general, hard pads provide superior planarity but tend to generate more scratch defects. Additionally, ceria slurries are known to cause high polish scratches (PS) or even mechanical damage (MD) defects due to their morphology. Meanwhile, for BEOL Cu CMP, the use of

Corresponding Author:

#Wei-Tsu Tseng

Tel: +1 914-299-7479

E-mail: wei-tsu.tseng1@ibm.com

257 Fuller Road, Albany NY 12203, USA

poromeric “soft pads” with high pore volume is a standard practice for purpose of defect reduction. However, by nature, “soft pads” are more conformal to wafer surface topography, leading to loss in both long- and short-range planarity.

In this study, results from CMP pad evaluation and selection for within-die planarity improvement as well as defect reduction will be presented. The balance and trade-off between planarity and defectivity with traditional PU pads will be discussed in the context of their physical and mechanical properties. Based on the result, the needs for new pads and the trend for CMP pad technology development will be deliberated and projected.

EXPERIMENTAL

All CMP experiments are conducted in a 300mm commercial polisher with 3 platens. Post cleaning process contains two-stage PVA brush scrubbing steps with a proprietary alkaline clean chemical, followed by IPA drying. For inter-level planarization (ILP), a 1-step timed polish process with silica slurry of pH ~ 9.5 is utilized to planarize the low-*k* dielectric over device macros with various pitches and pattern density, as shown in Fig. 2.

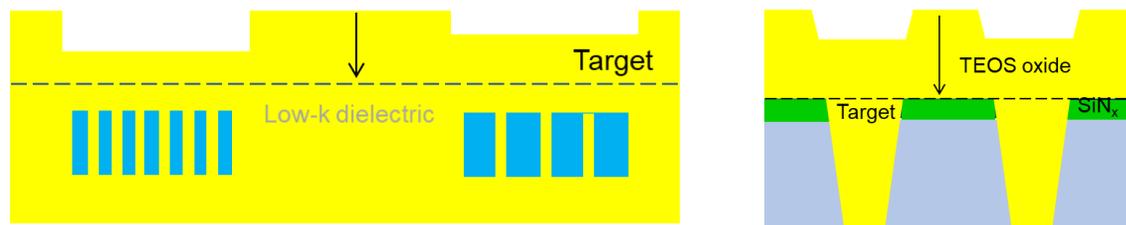


Fig. 2: ILP (*left*) and polish oxide stop on nitride (SON, *right*) CMP process schemes in the current study.

For defectivity study, a ceria slurry of pH 5.0 is applied to polish oxide and stop on nitride (SON) for an advanced memory device. For blanket wafers, post SP3 defect scan is carried out at 40nm resolution. For patterned wafers, defect scan is operated at dark field mode, followed by SEM review and classification.

RESULTS

Two different pads, POR and Pad C were tested for ILP process to planarize the low-*k* dielectric over large macros with different pitches and pattern density for a particular memory device. After polishing to the target oxide thickness, AFM scan was performed on top of 2 different such macros and their traces are shown in Fig. 3. The scan length (x-axis) is on the order of hundreds of μm while the topography (y-axis) is on the order of tens of nm.

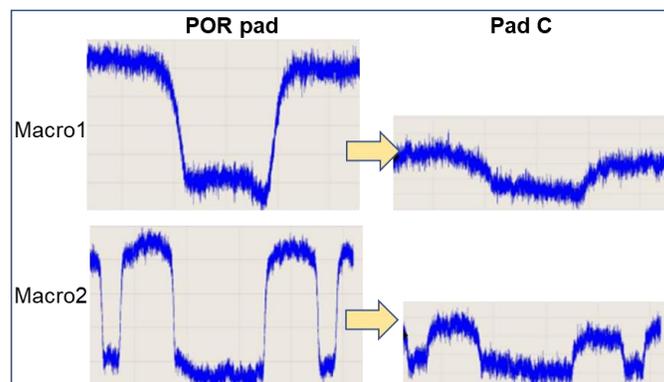


Fig. 3: AFM traces across Macros 1 & 2 on POR vs. Pad C. The topo on y-axis is adjusted to the same scale for comparison.

It is apparent from Fig. 3 that Pad C leads to > 50% lower topography over 2 different macros, relative to POR. In an extended study, a variant of Pad C (i.e., Pad D) plus an additional macro (Macro 3) were

included for further evaluation of planarization efficiency. The pattern density of Macro 3 is slightly higher than that of Macro 2 while the scan length remains the same. For Pads C and D, 3 timed polishes with different duration were conducted. Their resulting post thickness and topography from AFM traces are summarized in Fig. 4. Topography from POR pad polished to target thickness is included for comparison.

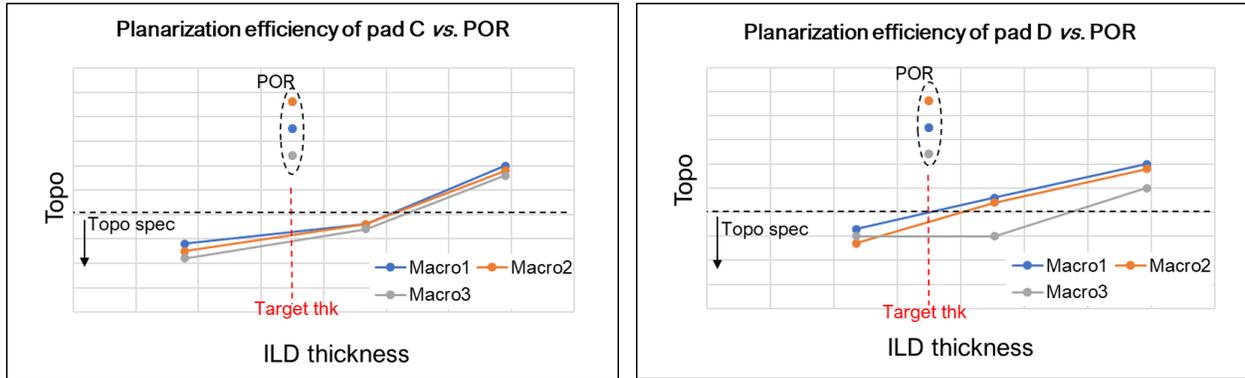


Fig. 4: Topography vs. remaining ILD thickness for Pads C (*left*) and D (*right*). POR pad at target ILD thickness is included for comparison.

The topography vs. ILD thickness plots in Fig. 4 indicate that, with POR pad, when ILD is already polished to target thickness, the measured topography across all 3 macros is still higher than the spec. On the other hand, both pads C and D afford lower than spec topography when remaining ILD thickness meets the target. Consequently, one can conclude that both pads C and D exhibit superior planarization efficiency than POR pad. A closer look at the plots reveals that, provided the same ILD thickness, pad C results in the least amount of topography difference among the 3 macros, followed by pad D and then POR.

For defectivity study, a ceria slurry of pH 5.0 is applied to polish blanket oxide wafers and patterned SON wafers on two different pads, Pad X (POR) and Pad Y of about the same hardness and % porosity. For blanket wafers, post SP3 defect scan is carried out at 40nm resolution. For patterned wafers, post CMP defect scan is operated at dark field mode, followed by SEM review and defect classification. Wafer maps and defect density from both pads are presented in Fig. 5 for both blanket and patterned wafers.

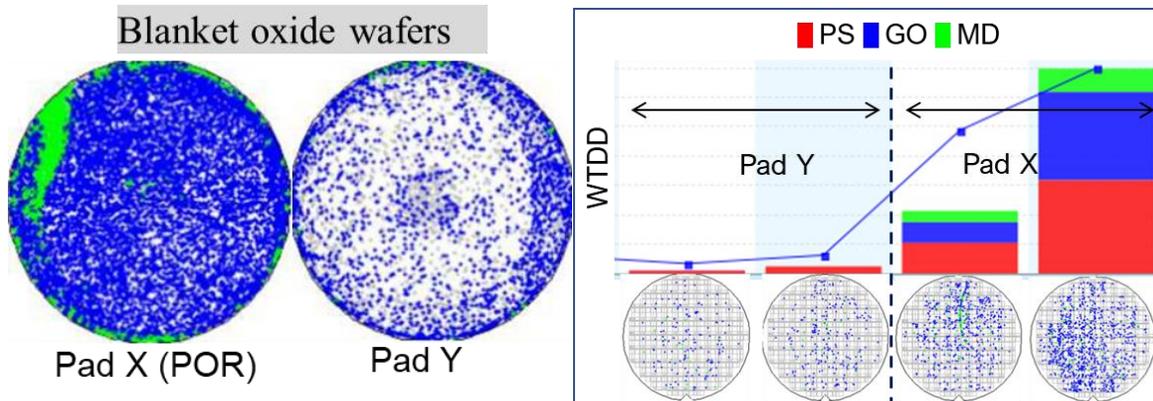


Fig. 5: (*Left*) Blanket oxide defect wafer maps; and (*right*) weighted total defect density (WTDD) of polish scratches (PS), gouge (GO), and mechanical damage (MD) on SON patterned wafers from pads X and Y.

SEM pictures of PS, GO, and MD defects as referenced in Fig. 5 are displayed in Fig. 6. While all 3 types of defects fall into the general category of scratches, PS refers to short and shallow scratches, GO signifies an area of structure that is “gouged” out, and MD represents deep and heavy scratches spanning across multiple patterns.

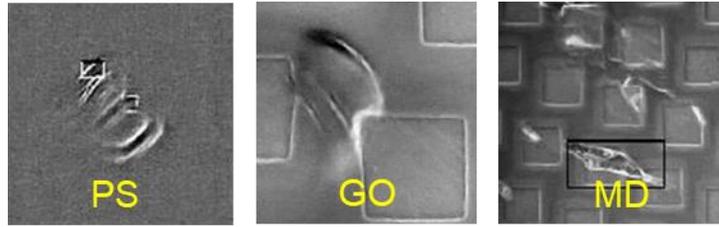


Fig. 6: SEM pictures of PS, GO, and MD defects on patterned wafers from Fig. 5.

The defect results in Figs 5 and 6 clearly demonstrates significant scratch defect reduction by Pad Y compared with Pad X (POR) even though both pads share the same hardness.

The above topography and defectivity studies illustrate that, without resorting to any new tooling, hardware, software algorithm, or even altering the incoming deposition process, CMP pads alone can modulate the topography and within-die uniformity to a large extent. Likewise, without changing slurries or clean chemical, CMP pads can reduce scratch defects remarkably. Pad surface characteristics are keys to the planarization and defectivity performance as will be discussed in the next session.

DISCUSSION

Fig. 7 captures how planarization efficiency (PE) is commonly defined as applicable to any length scale, feature/ die. At any length scale, it is a measure of how well the topography is reduced related to material removal due to CMP. It is related to material removal on the higher areas relative to lower area in any given feature/ die scale. For a similar set of features neighboring the feature of interest, pad properties determine the level of interaction of the pad with the high and low areas in any feature. Depending on the length scale (λ) of the features of interest, the relevant properties of the pad that determine the mechanical action of the pad in the low areas towards RR_{Low} can be expected to vary.

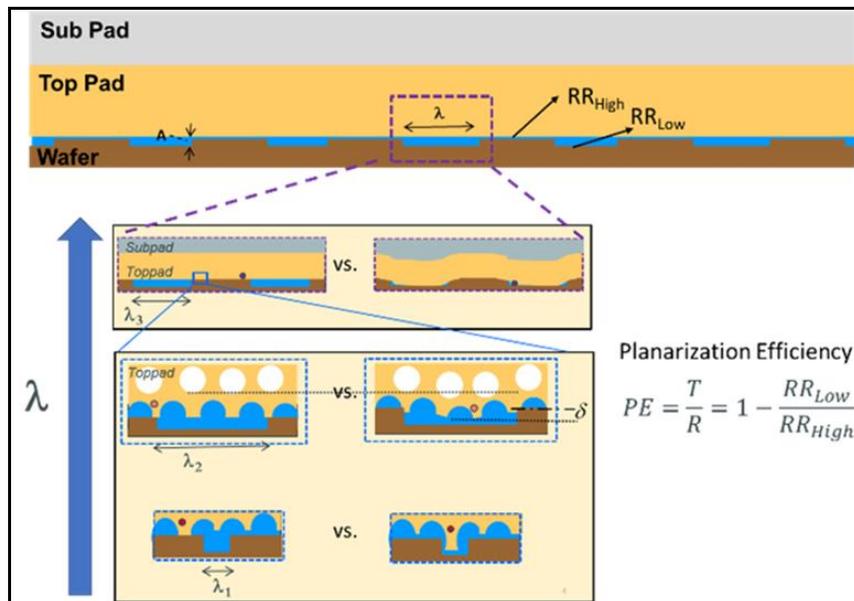


Fig. 7: Planarization efficiency and different length scales.

Table I captures the key contributors at appropriate feature length scales, and the relevant DuPont's proprietary property parameters (α_{LAA} , β_{Bend} , and γ_{Stack}). Understanding these dependencies is critical for pad selection for the process need.

Based on material properties, compared to the POR pad, both Pad C and Pad D in Figs. 3 and 4 are expected to have a 50% lower tendency to interact with the low areas in the features of interest ($\lambda = 300\mu\text{m} \sim 600\mu\text{m}$). This explains the improved planarization efficiencies observed for the Pads C and D relative to the POR.

Feature length scale, λ	Key contributor towards RR_{Low}	Material property parameter of interest
λ_1 (10 μm)	Top pad polymer	α_{LAA}
λ_2 (100 μm)	Top pad composite	β_{Bend}
λ_3 (1000 μm)	Top pad+ subpad stack	γ_{Stack}

Table I: Expected key contributors and property parameters of interest for planarization for various feature scales.

The configurations of Pad C and Pad D use different subpads. The top-pad + subpad stack in Pad D is expected to offer 20% higher tendency to interact with low areas, and hence improved planarization than Pad C. However, the results in Fig. 4 indicate that in general Pad C seems to offer improved topography over all 3 Macros with tighter Macro to Macro delta (i.e., within-die topo uniformity). Among other factors, these observations point to the need for a more detailed understanding of any differences in pitches, and pattern density neighboring the features of interest across the 3 macros.

In any polish process, for a given set of consumables, the state of contact at wafer scale can be impacted by pad macro feature like grooves while that at texture scale can be impacted by micro features like pores, porosity, and the pad polymer properties. Table II compares the properties relevant for defectivity, of the pads tested along with typical expected contact areas. α_{force} is defined as the ability of full asperity to apply force back to wafer; α_{strength} signifies the resistance of the asperities to fracture; and α_{Rigidity} denotes the rigidity of asperity in polish environment. Compared with Pad X (POR), Pad Y with 20% lower α_{Force} and smaller pore sizes can be expected to offer more lubricated state increasing the pad-wafer gap which reduces the frequencies of particle-wafer collisions (Fig. 8) and hence the observed material removal rates.

	Pad X (POR)	Pad Y
SG	0.8	0.8
Avg pore size, μ	40	20
Porosity	0.3	0.31
α_{Force}	1	0.8
α_{Strength}	1	1.3
α_{Rigidity}	1	0.92
Static Contact Area at 2psi, %	0.05	0.15

Table II: Selected characteristics of Pad X and Pad Y used in defectivity study.

Different mechanisms lead to different types of defects. Here we focus on the polish scratches which can be explained based on the sliding action on the wafer of a harder indenting ceria particles sticking on a tall asperity (Fig. 9). The asperity's proximity to the wafer and its mechanical strength to take the wafer load play a key role. The less lubricated state of pad-wafer contacts with POR pad (from Fig. 8) coupled with its ability to take the wafer load (given its higher α_{Rigidity}) lead to more polish scratches compared to Pad Y. Depending on the mechanism at play, the knobs for controlling the defects can vary.

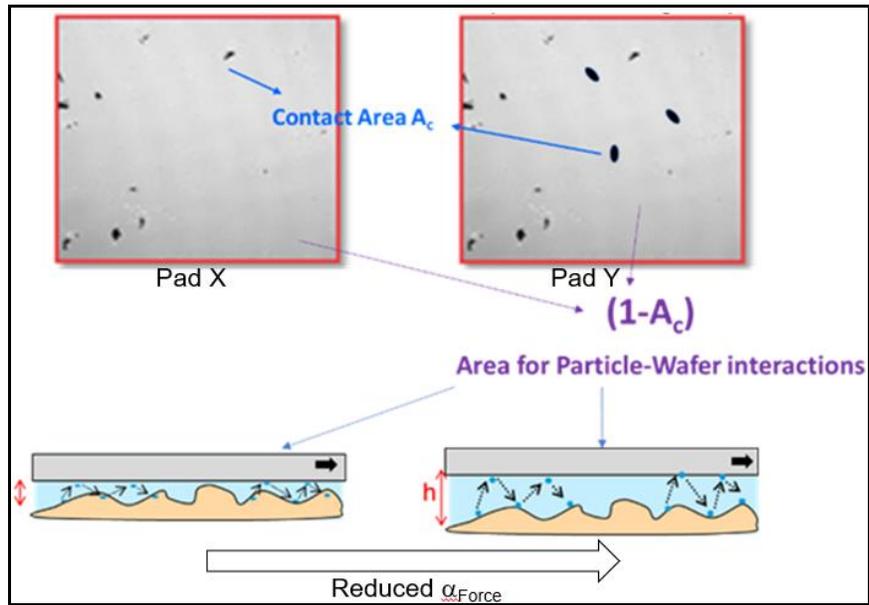


Fig. 8: Expectations of processes within Pad X and Pad Y based on typical contact areas and known α_{Force} parameters from Table II.

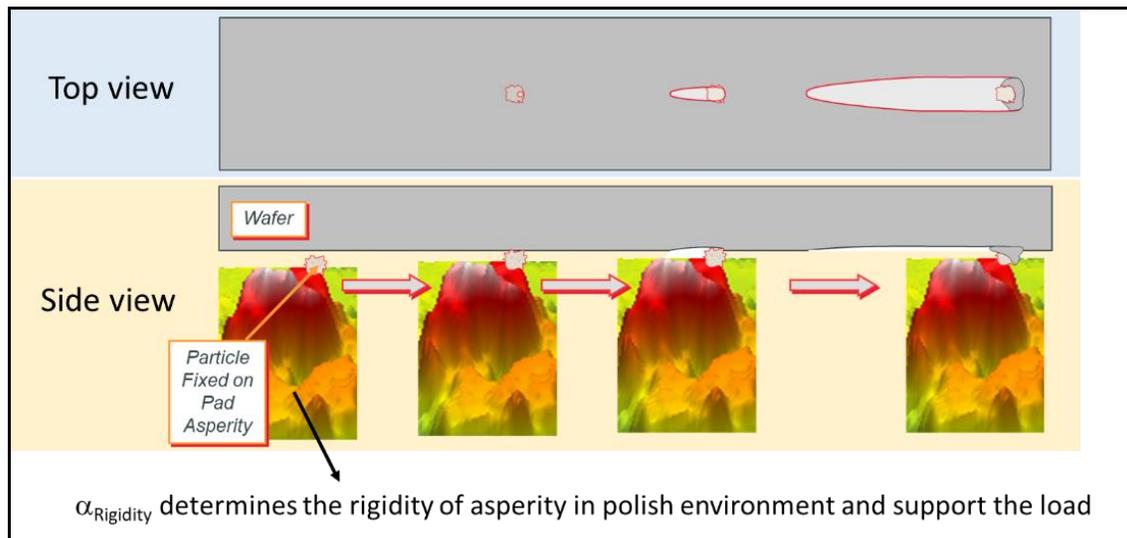


Figure 9: Mechanism for scratches in ceria polishing

It should be noted that the amount of scratch reduction from CMP pads can vary, depending upon the type of ceria slurry used. Ceria abrasives are known to be chemically reactive on oxide surface. The phase, morphology, and manufacturing process of ceria abrasives can all modulate the Ce^{3+}/Ce^{4+} ratio and hence chemical reactions on oxide surface. Similarly, additives in the slurry can play a role in the bonding strength between ceria and oxide surface. All these factors can interfere with α_{Force} parameter and modulate the extent of scratch reduction.

CONCLUSIONS

In this study, CMP pads of different designs and characteristics are tested to reduce within-die topography for ILP applications, and to reduce scratch defects with ceria slurry for SON CMP. Greatly

improved planarization efficiency is achieved with CMP pads of specific designs whereby more than 50% reduction in topography across feature length of hundreds of μm is demonstrated, compared with POR. Meanwhile, > 70% scratch defect reduction is accomplished by pads of the same hardness and porosity.

The keys to topography and defectivity reduction by CMP pads require multi-parameter optimization of pad characteristics. The mechanical properties of surface asperity as well as the bulk material properties can be optimized to design a pad that would not touch upon the low features wafer surface topography for better planarization efficiency. The properties of subpad can also play a role too. Smaller pore size combined with lower force that asperities can apply back to wafer can help reduce defects generated by ceria slurry polishing.