Die-Scale Nanotopography Characterization: New Insight

Viorel Balan[1], Yorrick Exbrayat[1], Sebastien Mermoz[2]
[1] Univ. Grenoble Alpes, CEA, LETi, F-38000 Grenoble, viorel.balan@cea.fr
[1] Univ. Grenoble Alpes, CEA, LETI, F-38000 Grenoble, yorrick.exbrayat@cea.fr
[2] STMicroelectronics, Crolles, France, sebastien.mermoz@st.com

INTRODUCTION

Building in the third dimension requires perfect flatness, whether in the macroscopic world or in microelectronics. 3D stacked architectures (FinFET transistors, 3D CoolCube™ stacked transistors, 3D NAND, Direct Hybrid Bonding) allows to reduce chip area at advanced technology nodes but require excellent topography uniformity both for inter-layers stacking, and chip bonding process. Chasing every nanometer of residual topography is mandatory in order to avoid defocus lithography issues or bonding defects [1]: this is the main goal of the Chemical Mechanical Planarization steps in semiconductor manufacturing. This topography chase need appropriate techniques in order to be able to control and improve the process performance: "You can't control what you can't measure" [2]. CMP became an enabler for advanced nodes lithography but also for direct hybrid bonding [3]: therefore, controlling surface topography is crucial in semiconductor fabrication process.

Surface topography analysis detects surface height variation z as function of x,y coordinates. Optical profilometry by interferometry allows acquiring high-resolution images at die-scale and, therefore, global die topography evaluation and hotspots detection [4].

BACKGROUND

Chemical Mechanical Polishing process was introduced in semiconductor manufacturing in order to planarize surfaces, to remove any surface deviation from its mean plane. Both trends in device fabrication, scaling down, following Moore's law, or stacking up, in a More than Moore paradigm, asks for more and more aggressive planarization specifications. From this perspective, any process control requires appropriate measuring and characterization metrics. In CEA Leti, in addition to profile characterization tool as mechanical profilometry and Atomic Force Microscopy, we developed methodology for full-die topography characterization by optical interferometry [5-7]. This technique allowed us to image die-level topography with a nanometric z-resolution and micrometric (x,y) resolution and, by his means, to demonstrate topographical amplitude increase over technological steps, as shown in **Erreur ! Source du renvoi introuvable.Erreur ! Source du renvoi introuvable.**.

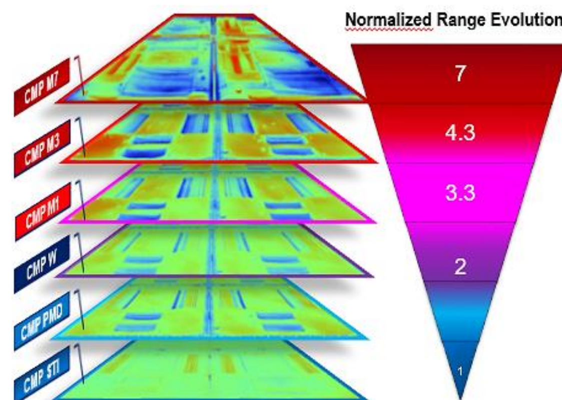


Fig. 1 - Die topography evolution after different CMP steps (from [1])

Surface topography can be described by a sum of different geometrical structures with different lateral wavelengths. Extraction of surface texture parameters requires a step of form removal in order to separate the nominal shape of a measured surface from its texture:

$$\text{Raw Data} = \text{Form} + \text{Local Topography} \quad (1)$$

For multiscale structured surfaces [8] containing deterministic pattern, as semiconductor devices, *"a priori"* knowledge is required to determine the nominal form present in the surface measurement data in order to perform effective surface texture analysis that is fit for purpose and the natural question which arises is: are we getting the right answer? In ISO 21 920-2 [9], polynomial shapes are indicated as being one of the most used form removal technique. Polynomial fit (see Fig. 2) aims to determine the

coefficients a_0, a_1, a_2... of an equation that characterizes a geometric shape geometric shape in such a way that this shape is as close as possible to the topographic data:

$$f(x,y) = a_0 + a_1x + a_2y + a_3xy + \cdots + a_nx^ny^n \quad (2)$$

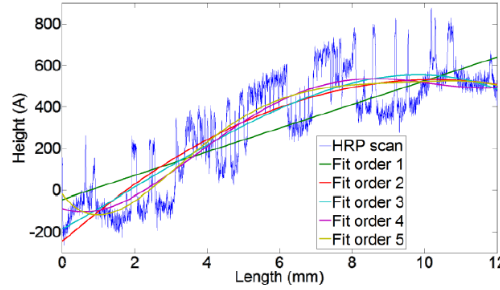


Fig. 2 - Polynomial fit for form removal from raw 2D topographical profile (from [6])

In linear regression we assume that there is a linear relationship between the dependent and the independent variables (see Fig.3).

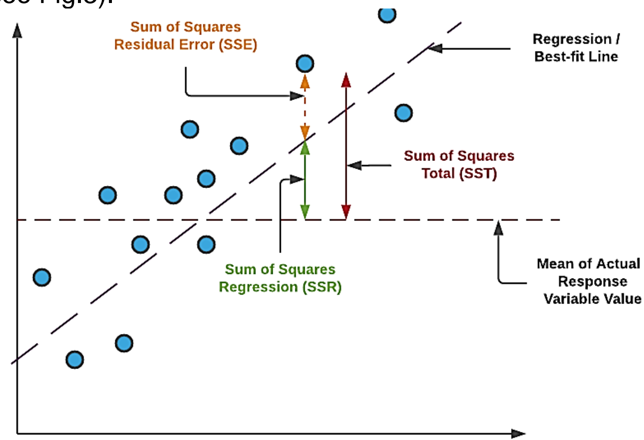


Fig. 3 - SSR, SSE and SST representation in relation to linear regression (*from https://vitalflux.com/linear-regression-explained-python-sklearn-examples*)

The goal is to minimize the sum of the squares of the distances between the data points and the model generated points, in order increase the model explained variability of data (Sum of Squares due to Regression – SSR) and reduce the remaining unexplained one (Sum of Squares Errors – SSE).

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2) + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2) \quad (3)$$

Total Variability (SST) = Model Explained Variability SSR + Unexplained Variability SSE

with y_i: *data point*; $\bar{y}$: *mean value of n* y_i *points*; $\hat{y}_i$: *model generated point*

In the case of polynomial linear regression, increasing the polynomial degree adds more independent variables to the model, leading to more complex form, therefore allowing to better explain the data points variability while reducing the unexplained variability. The risk, bellow inherent issues with polynomial linear regression, is to overwrite the local topography of the patterns on the chip. Furthermore, one can assume that, since the topographical complexity should increase with the technological level, the degree of the polynomial to be applied should be different. Therefore, we need to study the impact of polynomial degree on form removal and remaining topography. Correlating Equation (1) and (3), we can use SSR and SSE as metrics for this study.

EXPERIMENTAL

In this work, we used ContourGT-X 3D Optical Profiler from Bruker in order to perform non-contact die-level surface measurements at different CMP steps: STI, PMD, W, Cu M1, M3, M7 and Hybrid Bonding Metal level. This tool is equipped with a Michelson interferometer in a Phase Shifting mode, with a lateral resolution of 3.6 µm and a sub-nanometer vertical resolution. This configuration provides a field of view of 2.4 mm by 1.8 mm but image-stitching capability allows extending the acquisition area up to several square centimeters, covering full-die measurements.

In order to estimate post-CMP remaining topography, background polynomial fitting was done on full-die images in order to separate local nanotopography, due to pattern features, from longer-range

nominal shape[6]. Polynomial form extraction was done in Python, with StatsModels library, using Huber M-estimator in a Robust Linear Model regression [10]. Several polynomial degrees were tested in order to better explain variability through regression model. For background extraction, windowing function influence was studied: calculation was done by skipping points in X, and Y direction, with windows of (8,8), (16,16), and (32,32) points. Polynomial degree impact on Sum of Squares Error (SSE) and Sum of Squares due to Regression (SSR) was studied in order to understand topography evolution but also die-scale deformations along the technological flow. Finally, Machine Learning classification was applied in order to predict CMP step topography based on SSR and SSE values.

DISCUSSION

In Fig.4, we have represented the dependence of the sum of Square Error (left), the Sum of Square Regression (middle) and Sum of Square Total (right) against the CMP level. Here, a polynomial of degree 6 has been applied to extract the nominal form of the background. One can note that the total variability of the data is mainly driven by the SSR, so by the extracted form, while SSE values (the remaining topography after form removal) are negligible at this scale. It is interesting to note that total and explained variability (SSR) for STI, PMD and M7 are equivalent, while M1 and M3 show lower variability of data. It's PMD level and, especially HBM which show high data variability.

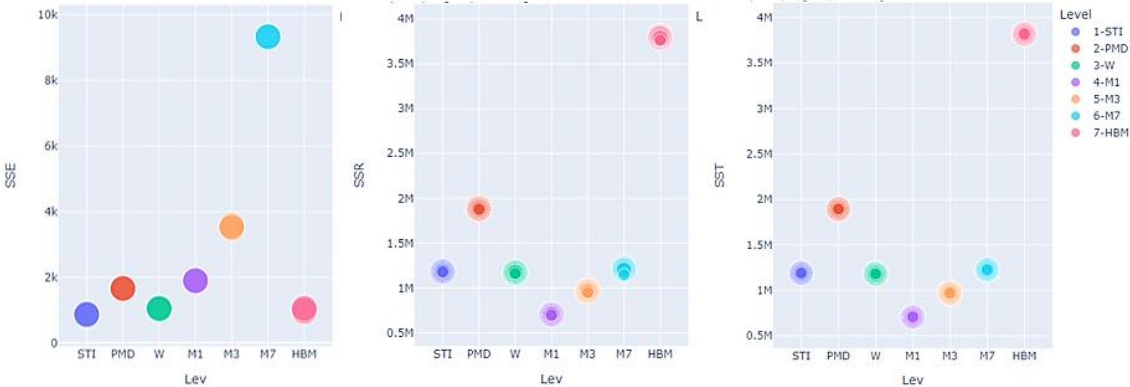


Fig. 4 - SSE (left), SRR (middle) and SST (right) evolution with CMP step

The SSE increase over the CMP steps and to dramatically decrease for the Hybrid Bonding level; where the CMP process is specially designed to address critical topographical specifications of the bonding. Note, also, the "bump" for the PMD level, knowing that corresponding CMP is the only one to be a direct planarization step, compared to the others, which are in a damascene configuration. This increase of SSE over CMP steps is in agreement with topographical range increase observed in Fig. 1. This can be easily explained by the fact that SSE is directly correlated with standard deviation (STD) of the n topographical data points:

$$STD = \sqrt{\frac{SSE}{n-1}} \qquad (4)$$

Therefore, the unexplained variation SSE could be a metric to express the short range topographic variation of the patterns.

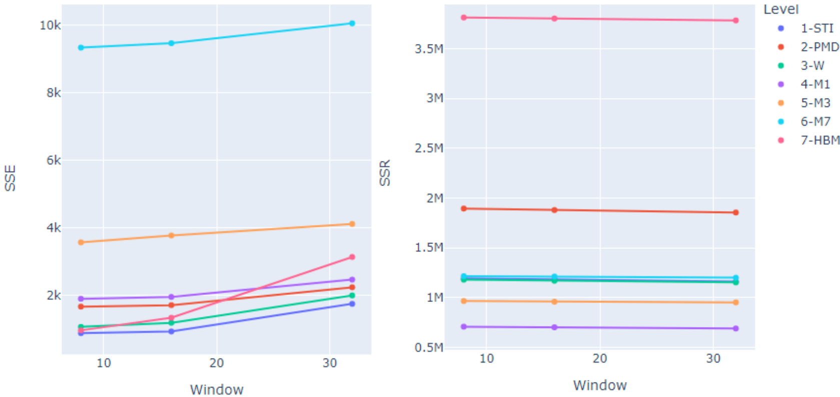


Fig. 5 - SSE and SSR dependence against the skipped point's window

The influence of the skipped points window on the explained and unexplained variability is shown in Fig. 4. Increasing this window lead to SSE augmentation, especially for HBM level, and very slight SSR decrease. While working with wider windows reduce the computing cost of the polynomial fit, attention must be payed to this factor in order to optimize form removal for different CMP steps.

Moving forward, we studied the impact of polynomial degree (from 2 to 9) on the explained and unexplained variability. Concerning the explained variability (right-side on Fig. 4), increasing complexity of removed form by rising the polynomial degree shows visible impact for M7 and HBM: one can explain this by the increased die topography complexity at elevated CMP levels produced by accumulated erosion of copper arrays. On the other hand, looking at the shorter range topography, expressed by SSE (left-side on Fig.5), we can see that at lower CMP SSE values hit a threshold for low polynomial degree. On the contrary, M7 and HBM levels need high polynomial degree ~6 in order hit a threshold value, indicating a good compromise for polynomial form removal, in agreement with SSR behavior we emphasize before.

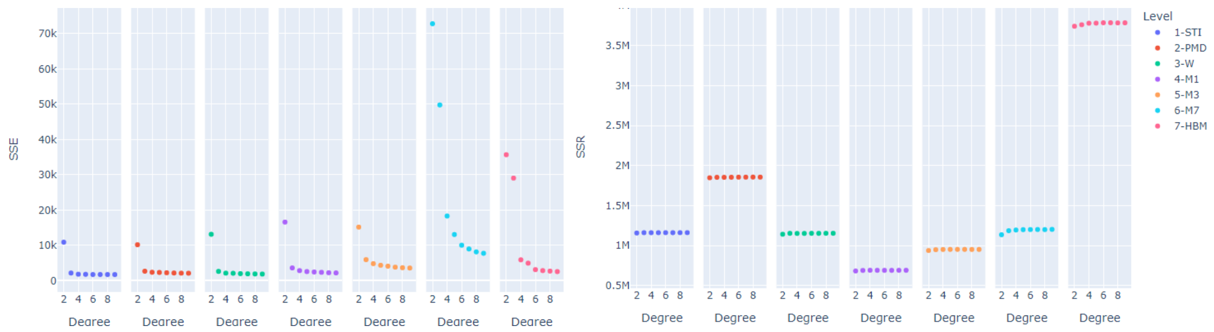


Fig. 62 - SSE and SSR dependence against polynomial degree

Finally we used Machine Learning in JASP [11] models in order to classify CMP levels based on SSR and SSE values, obtained using polynomial degrees from 3 to 7 and skipped points windows (8,8), (16,16) and (32,32). Table 1 shows obtained test accuracies for different models we used. The models predicted well the CMP corresponding level based on SSE and SSR values.

Table 1: Test accuracies for different Machine Learning classification models

Classification Model	Test Accuracy
Decision Tree	0.868
Support Vector Machine	0.897
Boosting Classification	0,908
Random Forest	0.948

Best result was obtained for Random Forest model, as detailed in Appendix 1. One can see that the model prediction gets some erroneous prediction for lower damascene CMP levels, STI and W, but gets very accurate predictions for higher CMP levels.

If previously we focused on studying CMP hotspots on global die topography after polynomial form subtraction, in this work we addressed, in addition, the extracted background polynomial shape. The SSR metric, corresponding to the explained variability of the polynomial fit give us information on the long-range topography amplitude, which is very valuable information when stacking levels, dies and wafers. Coupling both SSR and SSE information, prediction based on machine-learning models should allow us to optimize CMP process in order to reduce globally the die topography at both long and short range.

CONCLUSION

Using the 3rd dimension for building stacked devices asks for perfect flatness at different spatial wavelengths. Therefore, hiking the “way to zero planarity” [12] asks for metrics able to better describe die flatness. We showed that use of Sum of Squares due to Regression in addition to Sum of Squares Error of polynomial regression on full-die optical interferometry images improve nanotopography characterization. Finally, machine learning models allowed to classify technological levels based on SSR and SSE predictors, which could help to CMP process optimization.

Appendix 1

Random Forest Classification

Trees	Predictors per split	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy	OOB Accuracy
17	1	188	48	58	0.958	0.948	0.963

Note. The model is optimized with respect to the *out-of-bag accuracy* .

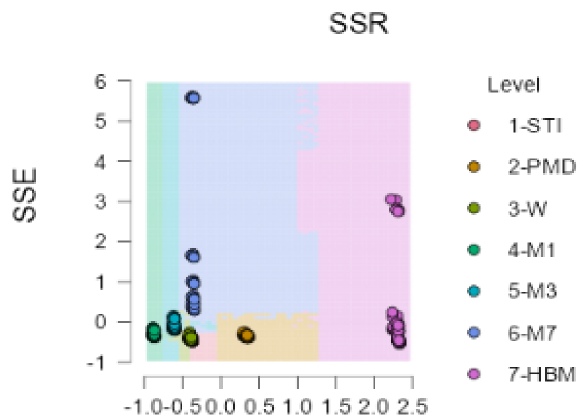
Data Split

Total: 294

Confusion Matrix

		Predicted						
		1-STI	2-PMD	3-W	4-M1	5-M3	6-M7	7-HBM
Observed	1-STI	4	0	0	0	0	0	0
	2-PMD	0	6	0	0	0	0	0
	3-W	2	0	6	0	0	0	0
	4-M1	0	0	0	15	0	0	0
	5-M3	0	0	0	0	7	0	0
	6-M7	0	0	0	0	0	9	0
	7-HBM	0	1	0	0	0	0	8

Decision Boundary Matrix



Preference: □ Oral
Topic: Consumables, equipment, and metrology

[1] "*Chase of Nanometer Topography in CMP for 3D Integration*", C. Euvrard, Y. Exbrayat, C Perrot, A. Seignard, S. Mermoz, V. Balan, ICPT 2017

[2] *"Controlling Software Projects"*, Tom de Marco, Management Measurement & Estimation", 1982

[3] "*CMP – 3D Enabler*", V. Balan C. Euvrard, A. Seignard, ICPT 2015

[4] "*Interferometry: a direct die level characterization technique",* F. Dettoni *et al.*, ICPT 2012;

[5] "*High resolution nanotopography characterization at die scale of 28 nm FDSOI CMOS front-end CMP processes",* F. Dettoni *et al*, Microelectronic Engineering,113 (2014): 105-108.

[6] "*Development of high resolution topographic characterization at die scale by interferometry*", F. Dettoni *et al*, FCMN 2013

[7] *"Challenges in Nanotopography measurements at die level",* C. Beitia, FCMN 2017

[8] "*Multiscale Analyses and Characterizations of Surface Topographies*", C.A. Brown *et al*, CIRP Annals - Manufacturing Technology, 67 (2) (2018), pp. 839-862

[8] "*Open questions in surface topography measurement: A roadmap",* R. Leach *et al,* Surface Topography Metrology and Properties 3(3):013001

[9] ISO/DIS 21 920-2 2020 “*Geometrical Product Specifications (GPS) - Surface Texture: Profile—Part 22: Terms, Definitions and Surface Texture Parameters”* (Geneva: International Organisation for Standardisation)

[10] *“Robust Linear Models”,* https://www.statsmodels.org/dev/rlm.html

[11] *JASP* (Version 0.16.3)[Computer software], JASP Team (2022).

[12] “*The way to zeros: The future of semiconductor device and chemical mechanical polishing technologies”,* M. Tsujimura, Jpn. J. Appl. Phys. 55 (2016) 06JA01

Corresponding Author:
Viorel Balan
+33 438 78 32 36
viorel.balan@cea.fr
Univ. Grenoble Alpes,
CEA, LETI,
F-38000 Grenoble
France