Cutting-edge CMP modeling for front-end-of-line (FEOL) and full stack hotspot detection for advanced technologies

Ushasree Katakamsetty, Jiansheng Jansen Chee, Yongfu Li, Chiu Wing Hui, Yaodong Huang, et al.
Cutting-edge CMP modeling for front-end-of-line (FEOL) and full stack hotspot detection for advanced technologies

Ushasree Katakamsetty, Jiansheng Jansen Chee, Yongfu Li, Hui Chiu Wing Colin, Yaodong Huang, Ernesto Gene de la Garza

1 GLOBALFOUNDRIES, 60 Woodlands Ind. Park D Street 2, 738406, Singapore
2 GLOBALFOUNDRIES, 400 Stone Break Road extension Malta, NY 12020, United States

ABSTRACT

As process technology scales down, the number of Chemical Mechanical Polishing (CMP) processes and steps used in chip manufacturing are increasing exponentially. Shrinking process margins increase the risk of excessive metal or oxide thickness or topography variations, causing potential yield problems such as dishing, erosion, resist lifting or printability issues.

Present DFM CMP modeling and applications mainly focus on the hotspot detection and fixing methodology for the Back-End-Of-Line (BEOL) layers [1]. Today, the present methodology is no longer sufficient to eliminate all the CMP related manufacturing defects. There is a strong demand for STI, poly and contact silicon calibrated CMP models to predict and fix the related CMP hotspots.

Shallow Trench Isolation (STI) and Poly CMP planarity is very critical in advanced technologies with Diffusion layer FIN structures and Replacement Metal Gate Process flow [2]. Gate uniformity after CMP will improve device performance, reduce CMP defects and increases the yield. Contact (Tungsten) CMP polishing is another important step that defines contact planarity, which will influence metal layer CMP planarization [3].

This paper will discuss design dependent CMP variations for STI, Poly and Contact CMP steps and showcase the importance of FEOL CMP modeling. We present the methodology for Silicon calibrated STI CMP, Poly and Contact CMP models and the applications of FEOL CMP models in CMP dishing and erosion hotspot analysis. We also present FEOL plus BEOL multi stack CMP simulations applications and provide design guidelines to fix CMP hotspots.

Keywords: DFM, CMP, FEOL, BEOL, STI, FinFET, CMP model

1. INTRODUCTION

Highly integrated circuits, like logic and memory chips, consist of many material layers. In the course of manufacturing, the surfaces have to be planarized over and over again to obtain sufficient process windows for critical processes like lithography and etching. Furthermore, planarization ensures defined sizes of the structures, thus ensuring reliable functioning of the electronic elements. The chemical-mechanical planarization (CMP) is the state of the art to reach the necessary planarization.

CMOS scaling has led to transistor devices with critical structures approaching a few atoms in size. New materials and new structures of existing materials have been introduced to optimize device size and performance. Many of these changes in IC device structures and materials have driven new planarization requirements. Dielectric material planarization is widely used in front-end-of-line (FEOL) processing for device isolation and in back-end-of-line (BEOL) processing for interconnection.

2. NEED FOR FEOL CMP MODELING

Conventional CMP model development and applications are mainly for BEOL interconnect layers. Advanced technologies with FinFET and gate-last HKMG processes increase the need for FEOL CMP modeling. Topography variations from Contact CMP process will influence Metal1 and above BEOL layers surface height variations. As FEOL features are very small in advanced technologies, pattern density rules are not sufficient to avoid CMP hotspots. CMP
Model based hotspot analysis is required to capture perimeter density variation along with line width and other design related parameters.

Shallow trench isolation (STI), is the mainstream CMOS isolation technology. STI CMP step removes excess deposited oxide and attains a planar surface for successive process steps. Despite advances in STI CMP technology, pattern dependencies cause large post-CMP topography variation that can result in functional and parametric yield loss. Figure[1] shows a pattern density dependent STI CMP hotspots. CMP defects can be fixed by process optimization as shown in Figure[2], but early detection of CMP hotspots will help chip designers and foundry to achieve first-time-right products. STI planarity variation may impact poly and contact layers topography.

![Figure 1: RX pattern density dependant CMP polishing defect](image1)

![Figure 2: RX pattern density dependent CMP polishing defect](image2)

The gate is essentially the heart of the transistor. Extreme control is needed for overall gate processing steps, to ensure proper device function. The gate-last process has few additional CMP steps compared to the gate-first process. Figure[3] shows the major steps in replacement metal gate (RMG) process flow, utilizing poly opening CMP (POC) and RMG polish steps. Because of the small dimensions and consequent small dimensional tolerance of the gate structure, pattern density and design related CMP hotspots has to be addressed early.
The Contact layer connects the FEOL to the BEOL layers. CMP variations from the Contact layer will directly influence Metal topography variations. Contact CMP process is mainly influenced by pattern density and perimeter density variations. Accounting for Contact layer CMP variations will improve the accuracy of Metal1 and above BEOL layers CMP predictions and hotspot detection. Figure[4] shows Metal1 Cu thickness difference with and without Contact layer simulations.

3. SILICON CALIBRATED FEOL CMP MODELING FLOW

Silicon calibrated CMP modeling flow starts with CMP testchip planning. The CMP testchip should cover all the line width and density design rule space defined for the technology node. The testchip should contain CMP test macros on modeling layers and dummy fill or uniform patterns on the available below layers. This is to maintain a flat profile on the below layers, thus eliminating the noise from the below layer topography on the modeling layer. Metrology methods and requirements also need to be planned during testchip planning stage.
3.1 FEOL CMP testchip plan for model calibration

There are 2 approaches for FEOL CMP testchip planning and model calibration. First approach is to stack RX, PC and CA CMP macros with density and linewidth variation on top of each other and calibrate RX, PC and CA sequentially using multistack CMP modeling flow as shown in Figure[5]. The advantage of the method is smaller testchip size, but this flow has lot of limitations. The developer has to spend lot of time studying multilayer metrology data and eliminate the noise arising at different layers. Modeling has to be done sequentially, so developer has to hold off till the completion of the previous model even though the metrology data is available. This approach also limits CMP modeling based on Fab priorities, for example CA modeling can only be done after completing RX, POC and RMG.

![Figure[5]: Sequential FEOL CMP modeling flow](image)

The second approach for FEOL CMP testchip planning and model calibration is to place RX, PC and CA CMP structures into separate macros. The density and linewidth for the layer under study is varied while maintaining uniform density on the remaining two layers. For example, PC and CA density will be uniform for the RX CMP modeling macros. This approach enables parallel development of all FEOL layers CMP models figure[6], thus reducing the overall CMP modeling time. Metrology data will be clean as the testchip contains only single layer topography variation. This flow also allows developer to prioritize CMP modeling based on Fab requirements. This approach will provide faster CMP models with a little bit overhead on the testchip area.

![Figure[6]: Agile FEOL CMP modeling flow](image)
3.2 FEOL CMP modeling flow

BEOL layer Cu deposition is carried out by Electroplating while FEOL layers deposition techniques are different for each layer and material. Common FEOL layers deposition techniques are CVD, FCVD or HDPCVD etc. Some layers can have mix of two different deposition techniques. Collecting measurements for all the deposition steps can be time and resource intensive. From the FEOL layers’ post-deposition metrology data, we concluded that total surface height variation is very minimal as FEOL features are very small and the process is highly controlled. This led to the assumption of flat profile after deposition, and we calibrated only the final CMP step for FEOL modeling, even though CMP modeling tools are capable of supporting all FEOL deposition techniques. This assumption led to faster closure of CMP modeling effort without compromising the accuracy. This flow also helped to reduce metrology time, optimize metrology resources and reduce wafer cost by reducing the spare wafers for metrology from 12 to 5, as shown in Figure[7].

![Figure 7](https://www.spiedigitallibrary.org/conference-proceedings-of-spie)

Figure[7]: Original vs optimized wafer split plan for FEOL metrology

3.3 FEOL multi-stack CMP model validation

After calibrating each FEOL layer separately, the FEOL layers CMP models combined into a single model file and multistack validation is carried out. Take note of the negative mask settings as per the process steps, during modeling and validation stage. Figure[8] shows the FEOL layers CMP model validation against silicon data.

![Figure 8](https://www.spiedigitallibrary.org/conference-proceedings-of-spie)
4. CHALLENGES WITH FEOL CMP MODELING AND SIMULATIONS

CMP modeling requires profile scan metrology data to derive erosion and dishing, and TEM cross-section data for absolute height measurements. The profile scan tool uses a stylus or laser tip to scan across the test structures. BEOL test structures can be drawn as long arrays so that the metrology tool scans across copper and dielectric material for erosion and dishing measurements, as shown in figure[9].

Advanced node FEOL structures can only be drawn as islands or short array lines due to complex design rules and process limitations. This poses a challenge as scanner may pass through only oxide material, instead of oxide and nitride as shown in figure[10]. To minimize this issue, maintain minimum S1 space and vary S2 space to achieve required density during the testchip planning stage. Multiple metrology scans are collected during the metrology stage and the results are scrutinized closely to retain only suitable and accurate data. All the FEOL layers profile scans must be taken with high resolution to accurately derive erosion and dishing[4].

Figure[8]: FEOL layers CMP model validation

Figure[9]: Parallel FEOL CMP modeling flow

Figure[10] Parallel FEOL CMP modeling flow
CMP simulations starts with design gds or oasis file extraction. Lot of layout Boolean operations involved in FinFET enabled designs and double or triple pattern designs. A simplified representation of the process of manufacturing FinFET structures is shown in Figure [11]. The key steps involved are definition of the active device areas, shown as the blue mandrels (temporary supporting structures). The fins (red) are formed by etching the mandrels(a). Then a cut mask is used to remove the unwanted parts of the structure (b) leaving the final pattern (c). RX mask generation involves OR, AND and NOT Boolean operations and CMP simulations need to support the required boolean operations. And number of Boolean layers increases for double or triple pattern designs. CMP simulators have been undergone enhancements to support layer Boolean operations, like AND, OR, NOT etc. within CMP tools, thus enabling full stack CMP simulations without runtime impact.

Figure[11] Paralle FEOL CMP modeling flow

5. FEOL + BEOL FULL STACK CMP SIMULATIONS AND APPLICATION

The focus of this paper is on FEOL + BEOL full stack CMP modeling and applications for physical hotspot detection/prediction and prevention. First, the validated CMP model is used to simulate the post CMP thicknesses and topography on a given design of interest, and then to predict the presence of hotspots on said design using custom written hotspot rules. If hotspots are present on the design, designer need to update the layout to fix the hotspots by modifying the floorplan or dummy fill in hotspot regions of the design. Third party IP may trigger CMP hotspots in some designs. It is not always possible to modify within the third party IP, but designers can modify the surrounding of the IP to prevent the CMP hotspots. Fixing CMP hotspots at early stages of the design will minimize the fixing efforts, leading to CMP friendly tapeout.

6. SUMMARY & CONCLUSION

This paper demonstrates the silicon calibrated full stack CMP modeling flow and methodology. We have shown that the use of the full stack model improves the CMP simulation prediction of higher layer hotspots. Two of the key features that define the CMP model in the advanced technology nodes is a) the need to carefully design the line-space array to maximize metrology data collection success and b) include more complex physical verification techniques into the modeling process. By making strategic assumptions on the CMP processes’ behavior, we were able to preclude the need to collect data for selected steps, thus reducing the CMP metrology cost by up to 2.4x. Meanwhile, an accuracy equal to or exceeding 90% of the silicon data is still consistently achieved. Novel hotspot detection and fixing flow was also implemented to enable designers to fix CMP-related hotspots while still in the design stage [3].
7. ACKNOWLEDGEMENTS

We would like to thank Dewei Xu, Sunit Mahajan and Wei Zhao for 14nm profile scan and TEM metrology support and Xiaochong Guan for 40nm metrology. We would also like to thank Ruben Ghulghazaryan from Mentor Graphics, for his excellent guidance and support on CMP modeling.

REFERENCES


