PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Copper interconnect topography simulation in 3D NAND designs

Li, Yang, Li, Rick, Jiang, Peng, Fan, Luming, Zheng, Aman, et al.

Yang Li, Rick Li, Peng Jiang, Luming Fan, Aman Zheng, Sicong Wang, Guangyi Wang, Zhengfang Liu, Chunshan Du, Ruben Ghulghazaryan, Qijian Wan, Xinyi Hu, "Copper interconnect topography simulation in 3D NAND designs," Proc. SPIE 10962, Design-Process-Technology Co-optimization for Manufacturability XIII, 1096210 (20 March 2019); doi: 10.1117/12.2515161

SPIE.

Event: SPIE Advanced Lithography, 2019, San Jose, California, United States

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022 Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Copper interconnect topography simulation in 3D NAND designs

Yang Li[a], Rick Li[a], Peng Jiang[a], Luming Fan[a], Aman Zheng[a], Sicong Wang[a], Guangyi Wang[a] , Zhengfang Liu[b], Chunshan Du[b], Ruben Ghulghazaryan[b], Qijiang Wan[b], Xinyi Hu[b]
[a]Yangtze Memory Technologies Co., Ltd., 18 Gaoxin 4TH Road, East Lake High-Tech Development Zone, Wuhan, Hubei, China
[b]Mentor Graphics Corporation, 5F, Building 2, Lujiazui Century Financial Plaza, No. 759 South Yang Gao Road Pudong New Area, Shanghai 200127, China

ABSTRACT

Vertical NAND (3D NAND) designs provide unprecedented improvements in input/output (I/O) performance and storage density, but require additional analysis to ensure manufacturing and market success. While 3D stacked architectures greatly reduce chip area at advanced technology nodes, greater topology uniformity is essential, not only for inter-layers stacking, but also for the chip bonding process. As the link between design and manufacturing, design for manufacturing (DFM) predicts potential manufacturing issues during the design stage, enabling design teams to modify the layout and mitigate the risk. The copper interconnect process can be modeled through multiple process steps, from film stacking, etch, and copper deposition to polishing. The simulated topology of a given design predicts potential risky areas that may be fixed by changing designs or inserting dummy fill prior to manufacturing. This simulation is a useful technique during yield ramp-up, and can shorten the cycle from design to manufacturing. This paper presents a solution for BEOL CMP modeling and analysis on BEOL copper interconnect of a 3D NAND flow.

Keywords: DFM, CMP, Modeling, Topology, 3D NAND, BEOL

1. INTRODUCTION

Critical dimensions of integrated circuits (ICs) continuously shrink for each new technology node, imposing more demanding requirements on wafer planarity (surface evenness) and depth of focus (DOF) limits. Achieving high-quality planarization during the chemical-mechanical polishing (CMP) step in manufacturing is critical for surface uniformity and the improvement of DOF in the manufacturing process. A rough surface topography impacts successive layers as they are manufactured, and can result in reduced yield or chip failure. Incorporating CMP modeling during design for manufacturing (DFM) optimization allows design and manufacturing teams to find and mitigate potential planarization issues before manufacturing. Solving these issues during the design flow reduces manufacturing time and costs, and can lead to faster yield ramp and/or improved yield.

Compared to traditional 2D memory design, vertical NAND (3D NAND) memory designs stack memory cells vertically and use a charge trap flash architecture (Figure 1). These vertical layers allow larger areal bit densities without requiring smaller individual cells. While most 3D NAND designs are manufactured on one wafer using a single process technology, the Xtacking™ architecture [1] enables the periphery circuits that handle data I/O and memory cell operations to be processed on a separate wafer from the memory arrays, using the logic technology node that best enables the desired I/O speed and functions. Once the logic and array wafers are complete, the two wafers are connected electrically through millions of metal vias that are formed simultaneously between and across the wafers in one process step, using the innovative Xtacking technology, with a limited increase in total cost. This dual-wafer approach provides unprecedented NAND I/O performance, higher bit density, and faster time-to-market.

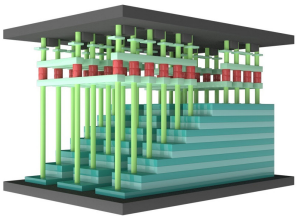

Figure 1. 3D NAND created using the Xtacking architecture.

Design-Process-Technology Co-optimization for Manufacturability XIII, edited by Jason P. Cain, Proc. of SPIE Vol. 10962, 1096210 · © 2019 SPIE · CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2515161

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

However, the topology quality of each wafer before the bonding process directly impacts the yield and reliability of the full chip. Accurate CMP simulation and process-related hotspot detection are required for process improvement and design optimization.

This paper presents a method for modeling the interconnect layers of a 3D NAND design. We build and validate a back-end-of-line (BEOL) CMP model using Calibre® CMP ModelBuilder and CMPAnalyzer tools.

2. MODELING AND EXPERIMENT PLAN

In our approach, CMP model building and simulation are separated into two steps (Figure 2). First, experiments are planned for data collection, taking into consideration the selected process and design rules. With this data, a CMP model is calibrated against the physical measurements on a test wafer. The final model provides accurate simulation of CMP process results, enabling designers to confidently perform dummy fill optimization and hotspots fixing in the design layout.

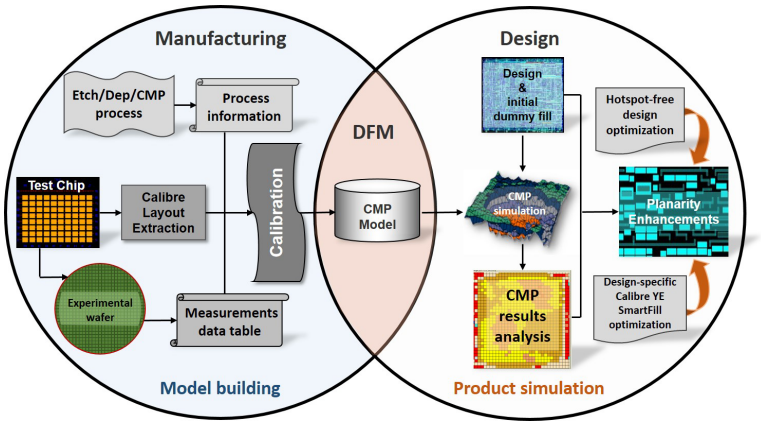

Figure 2. CMP modeling solution workflow.

Accurate topology simulation for a given layer depends on both the layer design and the topography of underlying layers. To accurately predict wafer bonding hotspots in a 3D NAND design, multi-layer CMP simulations of BEOL layers must account for topography variation accumulation from layer to layer. We accomplish this by building separate CMP models for three BEOL layers (M1, M2, and M3), then combining them to create a full-stack CMP model for multi-layer simulations. The model building flow is shown in Figure 3.

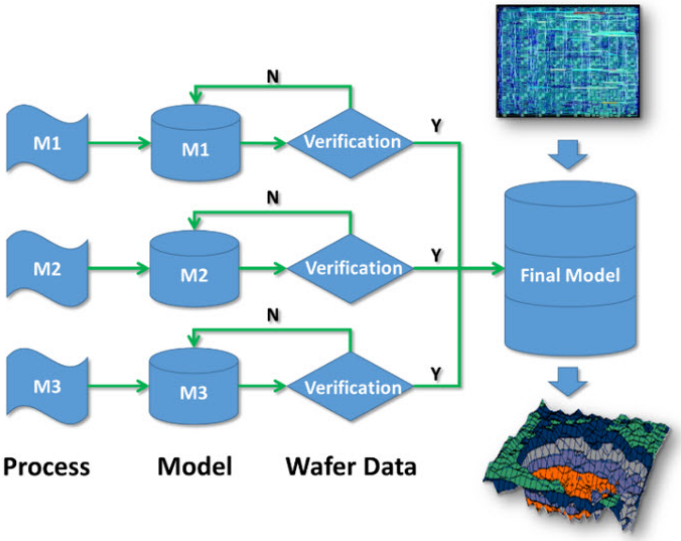

Figure 3. Model creation flow for multiple layer simulation.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

CMP test chips play a critical role in CMP model building. The test chip design must provide good coverage of possible width/space/density possibilities for a given technology node (Figure 4). Test keys should be designed as matrices consisting of parallel trenches with different combinations of width, space, and density [3, 4]. The diversity of test patterns should comply with the design rules for polygon width, space, and density of the modeling layer. However, to improve model predictability and flexibility, non-compliant test patterns may be included, as well as other structures used to capture the process behavior for model validation.

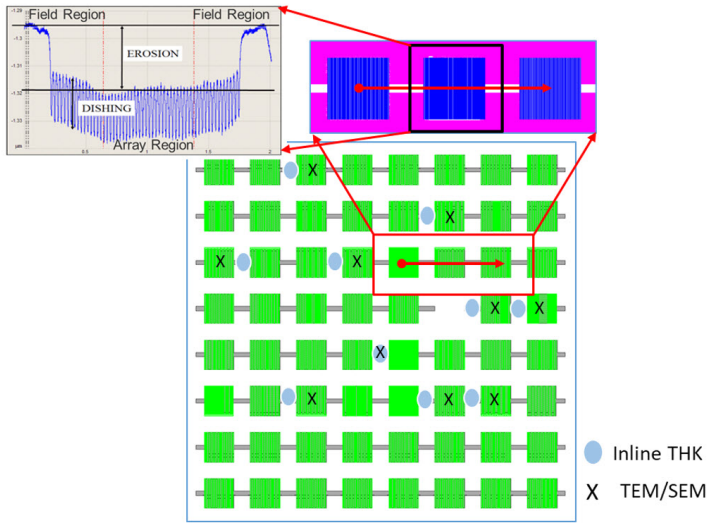

Figure 4. Test chip and data collection positioning.

Planarity measurement data collection and cleanup is another crucial step in building accurate models. In general, thickness data from physical failure analysis (PFA) cross-sections at selected design points and atomic force microscopy (AFM) profile scans are the methods used to characterize post-CMP topography. Typically, several wafers must be processed to provide sufficient measurement data for accurate model building (Figure 5).

Layer	Wafer NO.	Split Run						Data Collection		
		Non-Trench Deposition	Etch	Trench Deposition	CMP-P1	CMP-P2	CMP-P3	Inline Thickness	AFM Profile	PFA
Mx	1	√	√	-	-	-	-	-	√	√
	2	√	√	√	-	-	-	√	√	√
	3	√	√	√	√	√	-	√	√	√
	4	√	√	√	√	√	√	√	√	√

Figure 5. Wafer split run plan.

3. MODELING CALIBRATION AND VERIFICATION

Deposition results provide the initial silicon topography for the CMP process. In our experiment, tungsten is used as the interconnect material for the Metal1 (M1) layer. It is deposited using the chemical vapor deposition (CVD) technology. CVD provides an almost flat topology for patterns with small width trenches, and an isotropic or conformal deposition profile with a step for large width trenches (Figure 6).

Figure 6. Tungsten CVD deposition.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Copper is used for the interconnect material for the Metal2 (M2) and Metal3 (M3) layers. It is deposited using the electro-chemical deposition (ECD) process. The copper profile after an ECD process is significantly different from the tungsten CVD profile, since ECD generates bumps for patterns with small-width trenches, while it follows an isotropic deposition trend with steps for large-width trenches (Figure 7).

Figure 7. Copper ECD deposition.

Separate CMP models were built for the three BEOL layers (M1, M2, M3). First, measured data was collected from test chips using AFM line scans, transmission electron microscopy (TEM), and inline optical measurements tools. This data was then analyzed and any outliers removed. Next, a measurement data table with erosion, dishing, and thickness data was prepared for each deposition and CMP step. Recipe files containing process conditions and layer stack information were constructed for each layer. After completing these preparation steps, we used the Calibre CMP ModelBuilder tool to construct deposition models for tungsten CVD for the M1 layer, and copper ECD for the M2 and M3 layers. Then we constructed polishing models for the corresponding Platen 1, Platen 2, and Platen 3 steps for each layer. Measurement data fittings for the best CMP models are shown in Figures 8-10, demonstrating that all the models fit the measured data and closely follow trends.

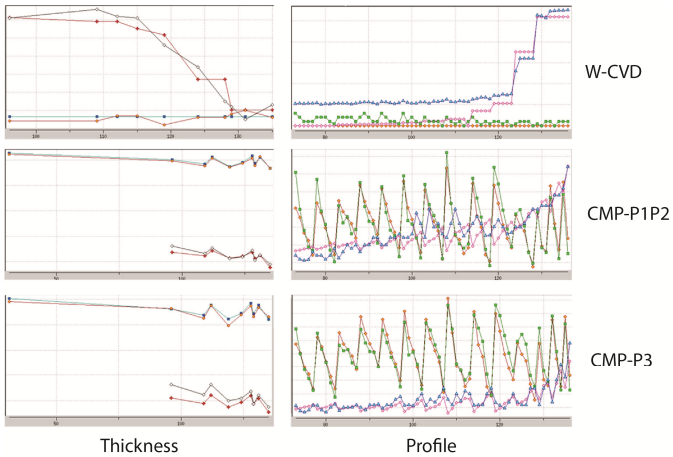

Figure 8. Measured erosion and dishing, and tungsten and oxide thickness data fitting for M1 layer.

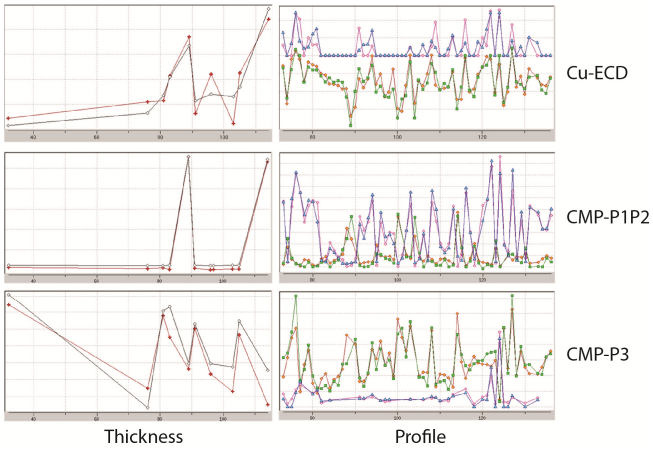

Figure 9. Measured erosion and dishing, and copper and oxide thickness fitting data for M2 layer.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

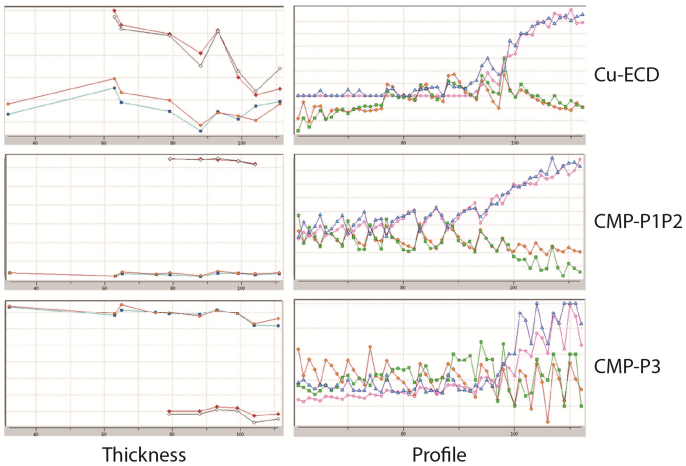

Figure 10. Measured erosion and dishing, and copper and oxide thickness data fitting for M3 layer.

One method for validating the accuracy of the model predictions is to compare simulated topography profile line scans with measured AFM line scans. In Figure 11, comparative plots of simulated line scans versus measured line scans are shown for the M1-M3 layers. The models' predictions track closely with the measurements, and the simulated topography matches well with silicon data.

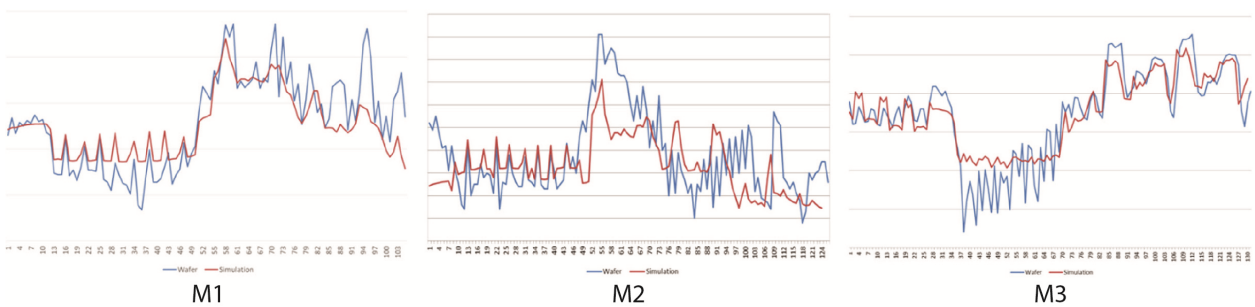

Figure 11. Simulated versus measured topology verification for M1-M3 layers.

4. MODEL APPLICATION

After validation of the accuracy of the individual CMP models, we created a combined model for the M1-M3 layers to use for product design simulation for a multi-layer stacked full-loop process. We used this combined model in conjunction with the Calibre CMPAnalyzer tool to analyze a 3D NAND design. The Calibre platform provides a universal CMP modeling and simulation environment, including a GUI for visualizing simulated surface colormaps, histograms, profile cuts, simulated line scans, and predicted planarity hotspots. The advantage of CMP simulation is that not only can designers analyze the final topology results, but they can also analyze and save the topography of intermediate deposition, etch, and polishing steps.

For the 3D NAND process, all long-range, short-range, and local topology variations are critical, due to the wafer bonding process. Another primary benefit of CMP simulation is the prediction of planarity hotspots. The CMP simulation model provides information and details about the process flow that may be used for analyzing the cause of planarity hotspots and devising ways to eliminate them, using either dummy fill optimization or design modifications prior to manufacturing. There are three hotspot types in particular that identify key points of variation:

- Surface height hotspots identify points that are far from the topology average across the chip;
- Topology gradient hotspots identify points that have large height gradients with neighbors;
- Dishing hotspots identify points that have large step height or dishing for wide metal lines across the chip.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

In this work, we used CMP simulation to identify all potential planarity hotspots. We then used the Calibre YieldEnhancer SmartFill functionality to identify the optimal dummy fill solution specific to a given design and technology (Figure 12).

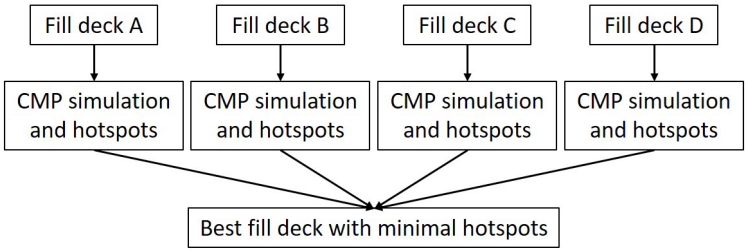

Figure 12. CMP model-driven best dummy fill selection.

After dummy fill optimization, most of the critical hotspots were eliminated. To remove the remaining hotspots, we optimized the design layout. After completing modifications of some patterns at the remaining hotspot sites, we used another CMP simulation to confirm that all hotspots were removed. Later analysis of silicon data of the final design confirmed the simulation results, and showed that the planarity of the final design matched well with the process requirements.

5. CONCLUSION

The ability to detect and mitigate potential planarity variations and hotspots is critical to yield success, particularly for designs like 3D NAND that employ multiple wafers in manufacturing. In this paper, we demonstrated a workflow for 3D NAND CMP modeling using the Calibre CMP ModelBuilder and CMP Analyzer tools. First we built separate models for the deposition and polishing steps of three metal layers from M1 to M3, then combined these models to create a full-stack CMP model for multi-layer simulations. Verification of the models against silicon data showed that the models provide a good fit, follow the process trend, and can accurately predict planarity hotspots on production designs. By using CMP simulations to predict planarity variation and hotspots, designers can modify design layouts and select optimal dummy fill solutions to eliminate hotspots and improve planarity results prior to manufacturing. Changing designs before physical production is faster and more cost-effective, enabling design companies to reduce both their time to market and yield ramp, and improve production yield.

REFERENCES

[1] Yangtze Memory Technologies, "Yangtze Memory Technologies Introduces New 3D NAND Architecture – Xtacking™," August 6, 2018. http://www.ymtc.com/index.php?s=/cms/172.html

[2] Mentor, a Siemens Business. Calibre® CMP ModelBuilder User and Reference Manuals.

[3] R. Ghulghazaryan, J. Wilson, N. Takeshita, "Building CMP Models for CMP Simulation and Hotspot Detection," mentor.com, July 2017, http://go.mentor.com/4TTwl.

[4] Ushasree Katakamsetty, Dinesh Koli, Sky Yeo, Colin Hui, Ruben G. Ghulghazaryan, et al. "20nm CMP model calibration with optimized metrology data and CMP model applications," Proc. SPIE 9427, Design-Process-Technology Co-optimization for Manufacturability IX, 94270U (18 March 2015).

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 05 Jan 2022
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use