Spectral interferometry for fully integrated device metrology

Daniel Schmidt,[a,*] Manasa Medikonda,[a] Michael Rizzolo,[a] Claire Silvestre,[a] Julien Frougier,[a] Andrew Greene,[a] Mary Breton,[a] Aron Cepler,[b] Jacob Ofek,[c] Itzik Kaplan,[c] Roy Koret,[c] and Igor Turovets[c]

[a]IBM Research, Albany, New York, United States
[b]Nova Measuring Instruments Inc., Fremont, California, United States
[c]Nova Ltd., Rehovot, Israel

Abstract. A spectral interferometry technique called vertical travelling scatterometry (VTS) is introduced, demonstrated, and discussed. VTS utilizes unique information from spectral interferometry and enables solutions for applications that are infeasible with traditional scatterometry approaches. The technique allows for data filtering related to spectral information from buried layers, which can then be ignored in the optical model. Therefore, using VTS, selective analyses of the topmost part of an arbitrarily complex stack are possible within a single metrology step. This methodology helps to overcome geometrical complexities and allows for focusing on parameters of interest through dramatically simplified optical modeling. Such model simplifications are specifically desired for back-end-of-line applications. Three examples are monitored discussed: (i) the critical dimensions (CDs) of a first metal level on top of nanosheet gate-all-around transistor structures, (ii) the thickness of an interlayer dielectric above embedded memory in the active area, and (iii) the CDs of trenches on top of tall stacks in the micrometer range comprising many layered dielectrics. It was found that, in all three cases, data filtering through VTS allowed for a simple optical model capable of delivering parameters of interest. The validity and accuracy of the VTS solution results were confirmed by extensive reference metrology obtained by traditional scatterometry, scanning electron microscopy, and transmission electron microscopy. Furthermore, it was shown that machine learning models trained with VTS filtered data can converge to a robust solution with a smaller dataset compared with models training with traditional scatterometry data.
© *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JMM.22.3.031203]

Keywords: spectral interferometry; vertical travelling scatterometry; scatterometry; optical critical dimension; spectral reflectometry.

Paper 22053SS received Aug. 26, 2022; accepted for publication Feb. 1, 2023; published online Feb. 21, 2023.

1 Introduction

Due to the ever-shrinking dimensions of semiconductor devices, there is a constant drive for in-die and fully integrated device metrology to capture subtle differences where they matter rather than going through a dedicated metrology target in the scribe line. More importantly, scribe line targets usually comprise simplified stacks and designs to focus on a specific process step of interest and to not be impacted by prior process variations where possible. These simplifications may result in minor structural differences compared with fully integrated device stacks, and they are starting to matter more with shrinking dimensions. Furthermore, the recent developments of embedding memory elements in the early back-end-of-line (BEOL) of the manufacturing process demand advanced materials and dimensional metrology very late in the process that was previously not required.[1] In particular, there is now a significant need for optical critical dimension (OCD) metrology during these later stages of device manufacturing.[2] The most advanced

*Address all correspondence to Daniel Schmidt, schmidt@ibm.com

back-end memory devices are embedded between the first and second metal levels on top of front-end complementary metal-oxide-semiconductor (CMOS) logic. The integration scheme requires a dielectric chemical-mechanical polishing (CMP) planarization that necessitates in-line monitoring on the active device for tight process control. This results in substantial modeling challenges due to the large number of layers and buried three-dimensional features and front-end architecture. The cumulative statistical process variations would all need to be considered for an accurate optical model, leading to a large number of floating parameters. Depending on the stack complexity, traditional OCD approaches may no longer be capable of solving these tasks with the precision required for the parameters of interest.

Recently, OCD was enhanced with spectral interferometry technology by providing absolute phase information to improve the sensitivity to weak target parameters and to reduce parameter correlations.[3] Here a spectral interferometry method that enables OCD to overcome geometrical complexities and allows for focusing on the parameters of interest is introduced. The new spectral interferometry technique, called vertical travelling scatterometry (VTS), allows for a selective measurement of the top part of a stack separately from the bottom part within a single metrology step only. Thus VTS solutions can exclude complex or unknown underlayers, which allows for solving applications that are not feasible with traditional scatterometry approaches.[4] In this paper, scatterometry challenges are discussed, and one approach that helps to overcome some of the shortcomings of the model-based optical technique is presented in more detail. VTS capabilities for three different BEOL examples are demonstrated: (i) measuring critical dimensions (CDs) of a first metal level on top of nanosheet gate-all-around (NS-GAA) transistor structures,[5,6] (ii) measuring the thickness of an interlayer dielectric (ILD) above embedded memory in the active area,[1] and (iii) measuring CDs of trenches on top of tall stacks in the micrometer range comprising many layered dielectrics. In addition, the significantly improved prediction capabilities of VTS trained machine learning models over traditional approaches are shown. All VTS results are validated with traditional OCD, scanning electron microscopy (SEM), or transmission electron microscopy (TEM).

2 Scatterometry Challenges and Strategies

With the continued pitch scaling, enabled through the insertion of extreme ultraviolet (EUV) lithography, and the increasing device complexity, nondestructive, inline metrology continues to gain importance. Semiconductor development or production lines could not run without dimensional process monitoring by OCD scatterometry. One of the major advantages of the nondestructive optical spectroscopy technique is the fast measurement speed, which allows for sufficient sampling to gain within-wafer and wafer-to-wafer process uniformity information. Traditionally, OCD is a model-based technique, and dedicated metrology targets with a homogeneous periodic pattern are measured. Smaller targets are preferred because they allow for a more active chip area. Ultimately, because of decreasing device dimensions, in-die measurements are desired because seemingly small differences between metrology proxy targets (typically found in the scribe line) and active areas may play an important role for overall device performance. However, stacks with buried complex device architectures are very challenging for any model-based metrology technique such as scatterometry.

The broadband optical beam of currently available OCD tools can be focused to an area of less than $40 \times 40\ \mu\text{m}^2$ without sacrificing metrology performance. It is important to note that, although the nonimaging technique may allow for subnanometer precision, the measured quantity is always an average value from across the entire probed area, which is identical to the focused beam size. It is interesting to note that CD uniformity values, can be extracted from within the measured area using machine learning algorithms.[7] Nonetheless, there is a desire to significantly reduce the spot size area by at least 100×, which would benefit measurements in active areas as well as enable mapping to understand local process variabilities between the center and edge of memory arrays. Because reducing the measurement area of a broadband instrument is not straightforward, the interesting concept of microsphere-assisted spectroscopic reflectometry was recently demonstrated as one option to dramatically reduce the spot size even below the diffraction limit.[8] Although smaller beam sizes are desired to measure in-die active

areas, they would also allow for significantly smaller dedicated metrology targets (with simplified designs more suitable for optical modeling) that can be placed with much more flexibility across the chip and even in or near device areas.

Workhorse OCD tools used in logic manufacturing today employ high-intensity broadband light sources (within the wavelength range from ∼200 to 1000 nm), which allow for sensitive and versatile measurements of semiconductor device architectures, such as gate-all-around nanosheet transistors.[9] In certain scenarios, shorter wavelengths toward EUV and x-rays would be desired for increased sensitivity to weak parameters of interest, roughness, precise shape analyses, or measurements of optically opaque materials.[10–12] In contrast, longer wavelengths in the infrared or terahertz would be beneficial for contactless electrical transport property characterization of patterned materials.[13,14] Although x-ray scattering[15] and mid-infrared ellipsometry[16] in-line tools have been demonstrated for CD metrology in 3D NAND channel holes, further tool developments are required to efficiently deploy them for metrology on patterned wafers in advanced logic manufacturing. For example, increased brightness of x-ray sources, decreased acquisition times, and decreased spot size allow for measurements on a sufficiently small area.

To obtain any quantitative geometrical information from a scatterometry experiment, typically a parameterized optical model is required, which describes the smallest unit cell of the periodic pattern within the measured area. It defines dimensional, shape, and profile information of the measured features and requires the dielectric function of each material within the stack. Rigorous coupled-wave analysis (RCWA) is then used to solve Maxwell's equations and to calculate the optical response corresponding to the measurement conditions.[17] Even though RCWA calculations are used because of their efficiency, depending on the complexity of the stack and the size of the unit cell, it can become computationally intensive. Minimization algorithms are utilized to vary the floating parameters in the optical model until the difference between calculated and measured spectra reaches a minimum.[18]

There are multiple challenges related to optical modeling that may render the accurate extraction of desired parameters not feasible. First, the unit cell optical model is always an idealized replica of the measured features and therefore an approximation. It has to be built based on process assumptions and available structural information usually obtained from cross-sectional TEM images that may not contain or reveal all details. Structural nonidealities (known or unknown), such as sidewall/edge and interface roughness, concentration gradients, or compositional variations, are often ignored because there may be limited sensitivity in otherwise complicated stacks. Corner rounding or feature footings may be represented by a simple slant rather than a radius. Such approximations are also accepted because of lower computational efforts, especially because some trade-offs regarding RCWA calculations may have to be made for stacks with intricate architectures in terms of the number of slices and harmonics.[17] Additionally, with increasing device complexity, the number of floating parameters required to account for statistical process variations can quickly reach double digits and increasing parameter correlations will negatively impact the analysis.[9] Furthermore, the optical constants of nanometer thin and patterned materials can change with thickness and dimensions due to quantum confinement effects and electron–phonon interactions, so neither bulk nor properties of very thin but continuous films may be appropriate. In addition, compositional variations as well as strain have an effect on the dielectric function and must be considered appropriately for accurate analyses.[18]

Although each individual approximation of the above-mentioned challenges may only have a minor influence on the final geometric parameters of interest in a complex architecture, compounding effects can lead to unacceptable inaccuracies. It should be noted that, in state-of-the-art nanosheet transistor architectures, small volume changes need to be monitored with high accuracy. For example, in the inner spacer module, such volumes may be comparable to some rounding or footing that would have been previously approximated with a slant.[9]

With this in mind, it is important to work also on solutions that reduce the modeling complexity. One opportunity, which is applied to some extend already, is a methodology that may be called design metrology co-optimization (DMCO). The idea is to develop metrology targets that mimic the device architecture as closely as possible but with design simplifications for optimized sensitivity to the desired parameters. Continuously increasing device complexity requires an increased focus on DMCO, which is relevant not only for scatterometry but also for other inline metrology techniques, such as ellipsometry, Raman spectroscopy,[19] and overlay.

Another path may be material-specific metrology techniques, such as grazing exit x-ray fluorescence, in which it is possible to reconstruct an element-specific geometry based on a two-dimensional fluorescence map.[20] Resonant soft x-ray scattering has been shown to significantly improve contrast between components compared with typical hard x-ray scattering experiments.[21] It may be possible to reduce the required geometrical model to layers comprising specific elements and ignore other features altogether, which would significantly simplify the efforts. However, both techniques are still far from being in-line capable.

Optical modeling approaches including machine learning solutions (in which no geometric optical model is required) are discussed under consideration of specific BEOL examples. Figure 1 illustrates three typical scenarios that pose significant challenges for traditional scatterometry and may render a conventional optical model infeasible. Measuring the line height and CDs of the first metal layer (M1) on top of an NS-GAA device architecture using OCD is very challenging [Fig. 1(a)].[5,6] Even though, in this example, the metrology target does not comprise middle-of-line (MOL) contacts, an optical model would require many degrees of freedom to account for statistical process variations associated with the nanosheets including the gate stack and source/drain region. Furthermore, in the case of any front-end-of-line (FEOL) process changes, a model would likely require a significant update and reoptimization, leading to a long time-to-solution. Figure 1(b) depicts an example of MRAM embedded between M1 and M2. In addition to CD measurements of the magnetic tunnel junction (MTJ) and its electrodes, the ILD thickness also must be monitored in the active region (on device). The integration scheme requires planarization before M2 patterning, and the CMP process homogeneity across the chip is critical to preventing subsequent processing issues.[1] A traditional optical model for the active area is infeasible because the OCD spectra carry information from the entire device architecture including FEOL and wiring all the way to the memory pillar and ILD. The last example of trenches at the top of a tall stack comprising many layers of optically transparent dielectrics may appear like a simple one [Fig. 1(c)]. However, once the total stack thickness reaches several micrometers, the measured OCD spectra are dominated by strong oscillations due to the thin film interference effects related to the tall stack and the many layer interfaces. A traditional OCD solution would be extremely cumbersome, and sensitivity to the critical parameters of interest will likely be low.

Typically, BEOL OCD developments are reported for wafers with short-loop processing only, i.e., no FEOL or MOL content is present.[22–26] Transferring these learnings to manufacturing is not straightforward because of the aforementioned challenges. Timoney et al. discussed the substantial modeling complexities for M2 on fully integrated targets at the 7- and 14-nm nodes.[27] The incoming process variations can impact the measurement uncertainty and may lead to a significant measurement error of the features of interest.

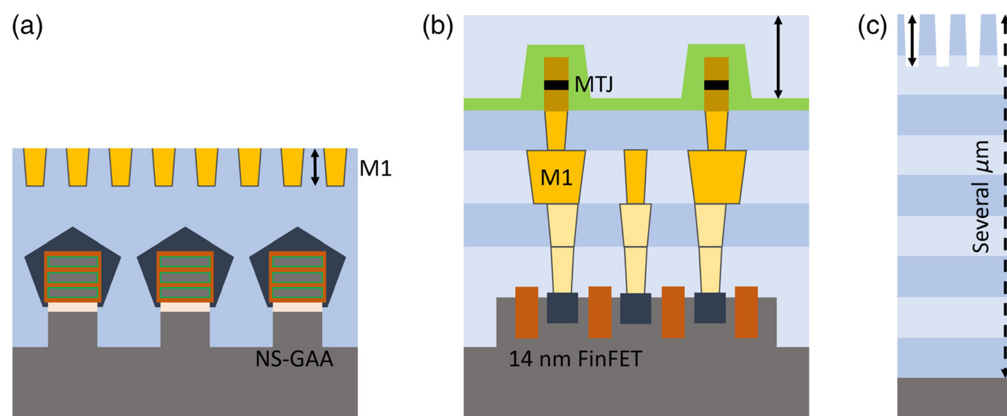

Fig. 1 Schematics of challenging BEOL scenarios for traditional scatterometry. (a) CDs of the M1 metal level on top of NS-GAA transistor structures; (b) the thickness of the topmost ILD above embedded memory in the active area; and (c) CDs of trenches on top of tall stacks in the micrometer range comprising many layered dielectrics.

There are several pathways to a solution for the discussed scenarios, but they are coming with numerous disadvantages. Although a feedforward or hybrid model reduces the number of degrees of freedom, it still requires a complex geometric optical model that includes data from either previous or other measurement steps.[24,25,27,28] Another variation of hybrid metrology is related to "differential modeling," which requires measurements before and after the process step of interest along with appropriate algorithms.[29] There are certain constraints, and the methodology does not necessary simplify the solution.

Machine learning can enable scatterometry solutions without the need for a geometrical model and may even allow for the prediction of parameters that are inaccessible with traditional model solutions.[7,9,25,26,30] However, the required reference data are not always easy to obtain. It may also be very time-consuming and expensive if, for example, only cross-sectional TEM reference metrology is an option. Interestingly, in the case of tall stacks [Fig. 1(c)], a very large reference dataset would be required for a successful machine learning model because the measured optical spectra contain the convolution of parameters of interest related to the trenches and information of each buried layer.

If measurements can be performed on specifically designed targets in the scribe line, design modifications can help to engineer the optical response such that the solution may be simplified. In the case of tall stacks with many transparent dielectric layers, it may be helpful to introduce buried patterned metal layers if the process flow allows. These can help to suppress contributions in the optical response from the layers below, thereby reducing interference fringes and limiting sensitivity to buried layers. The materials, CDs, volume, number of patterned layers, and orientation must be considered during the design phase to achieve desired simplifications. Of course, the patterned layers must then be considered in the final optical model.

In all described scenarios, the root cause of the BEOL modeling complexities is the fact that the measured optical spectra contain information from the entire stack. Hence, the irrelevant information of all buried layers and architectures is convoluted with the optical response from the topmost features of interest. A deconvolution is desired to separate relevant from irrelevant information and thus to substantially increase sensitivity to the parameters of interest and to dramatically simplify the optical solution.

3 Vertical Travelling Scatterometry

Over the last decades, OCD metrology has proven itself as an enabler for advanced node semiconductor device fabrication. The current workhorses in this field are the well-established normal and oblique incidence spectral reflectometry and ellipsometry techniques. The constant progress in semiconductor technology requires design and fabrication of ever smaller devices and architectures with increased complexities. This development must be accompanied by concurrent improvement in metrology capabilities to monitor and control the fabrication process. The spectral interferometry technology, which is implemented in the NOVA PRISM platform, is an example of such an improvement. Spectral interferometry technology enables access to the absolute spectral phase information, which is not available otherwise. In addition to providing improved sensitivity to structural parameters and decorrelation capabilities, spectral interferometry also enables an entirely new methodology for solving complex applications called VTS.

VTS uses unique information from spectral interferometry along with novel algorithms to enable a deconvolution of spectral information relative to the depth from which they originate. Hence, it allows for a separation of relevant and irrelevant optical information, thereby enabling selective measurements of the topmost layers of interest, for example, within a single metrology step. The algorithm allows for filtering of data obtained through spectral interferometry channels such that the dominant part of the information from below a user-defined cut-off can be selectively removed from the measured optical response (Fig. 2). Hence, underlayer contributions may be ignored in the optical model, which can dramatically reduce the geometrical complexities and significantly increase the sensitivity to parameters of interest. Additionally, the simplified optical model improves the time-to-solution. The utilization of VTS allows for solving applications that are otherwise not feasible with traditional scatterometry approaches. Because of the variable filter cut-off and the ability to ignore some measured information, the same optical

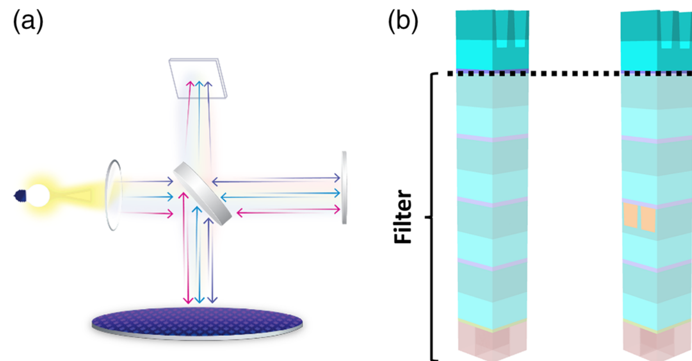

Fig. 2 (a) Schematic of a spectral interferometer: white light is split between the sample and a reference arm and then recombined and measured with a spectrometer. (b) Illustration of VTS filtering in which spectral information from below the cut-off (dashed line) can be identified and removed from the optical response.

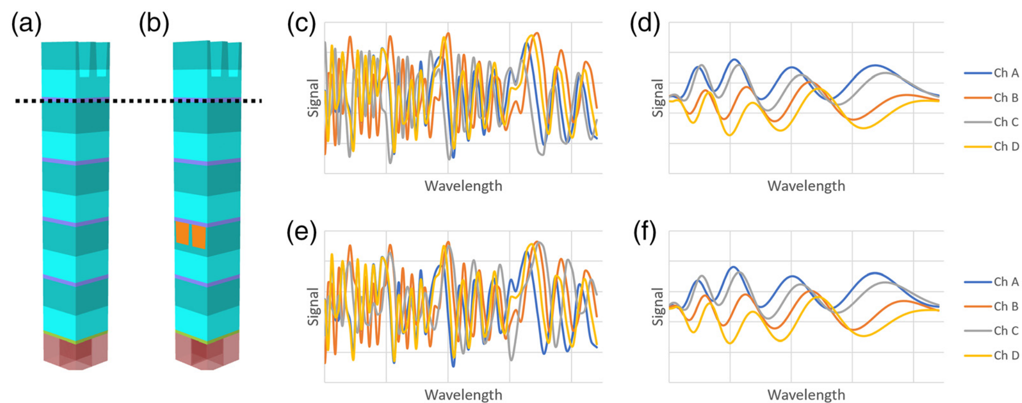

Fig. 3 Comparison of tall stacks (4.5-μm total thickness) with (a) only solid dielectric and (b) solid dielectric and a patterned Cu underlayers. (c), (e) Original spectral interferometry channels and (d), (f) VTS filtered spectra with (c), (d) top and (e), (f) bottom rows showing simulated graphs for stacks (a) and (b), respectively. The dashed lined indicates the VTS filter cut-off position. Each graph shows a total of four different polarization-dependent channels (Ch A through Ch D). Both axes are identical for each of the four graphs.

model can be used regardless of the type of underlayers (solid or patterned) and its process variations. Figure 3 depicts an example of two tall stacks with a total thickness of 4.5 μm that have a difference in one of the underlayers (patterned Cu lines introduced in one of the stacks) but otherwise have identical trench patterning (the top of the stack). The simulated spectral interferometry data for four different polarization-dependent channels within the UV to NIR range are dominated by strong oscillations related to thin-film interference effects. It is evident that the original spectra are different because of the variation in one of the buried layers [Figs. 3(c) and 3(e)]. Applying a VTS filter removes the dominant part of the underlayer signals and hence substantially reduces the oscillations [Figs. 3(d) and 3(f)]. When using the filtered VTS spectra, a simplified optical model that does not need to account for underlayer contributions and allows for focusing on the desired trench characterization can be employed.

4 Application Examples

4.1 *BEOL Trenches on Tall Stacks*

To verify VTS for tall stacks, the selected application requires monitoring the trench height and CD for the fifth metal layer of a short-loop BEOL flow. For the specific metrology target under

investigation, all layers below the trenches of interest are not patterned and comprise only thin films of various BEOL dielectrics, which are optically transparent within the measured spectrum. The total stack thickness measured from the substrate interface is ∼4.5 μm. As discussed above, the collected OCD spectra are overwhelmingly dominated by oscillations, and a traditional OCD model attempt did not yield any reasonable results. However, through VTS filtering, the optical model can be reduced to the topmost part of the stack (above the filter) comprising the trenches, which allows for a metrology solution. Four wafers with dose variations were exposed to intentionally change the trench CD: three wafers with fixed dose values and one wafer with a dose stripe (variable dose across the wafer). After collection of the optical spectra, top-down CDSEM reference data were obtained from 12 dies per wafer. On a different set of five wafers, a trench height variation was achieved through etch time differences. For wafers 1 and 3, an increased and decreased etch time was used, respectively. The other three wafers were etched at nominal conditions. To verify the trench depth, cross-sectional SEM reference data were obtained for two dies from each of the five wafers. A schematic of the stack and the VTS results in comparison to reference metrology are shown in Fig. 4. The VTS solution results show very good correlation to reference metrology, particularly for the trench height. These results confirm the basic working principle of VTS and verify that it is possible to filter irrelevant spectral information from buried layers.

The VTS methodology can also be very beneficial for machine learning-based solutions. Training a model with VTS-filtered data may require only a reduced reference dataset with variations of the parameters of interest. In contrast, training a comparable machine learning model with unfiltered, original scatterometry or spectral interferometry data may require a much larger reference set, which requires additionally process variations of all underlayers. To support these statements, three different machine learning solutions were trained, and test results are depicted in Fig. 5. In all cases, the machine learning models were trained with a small, labeled training dataset comprising optical spectra and reference values from five wafers with nominal processing. The prediction capability of the machine learning model was then tested on two additional wafers not included in the training dataset. The supervised learning is done with CDSEM reference data and differs based on the optical input data. As expected, the machine learning model trained with VTS filtered data yields very good results when tested on two independent wafers despite the small training dataset with R^2 and slope of 0.94 and 0.95, respectively [Fig. 5(a)]. Here the data filter was applied such that spectral information pertaining to layers below the trenches was not considered. Figure 5(b) shows the same dataset, but the input data comprised unfiltered VTS data; hence spectral information from the entire stack is considered. This results in a degraded test correlation between machine learning prediction and CDSEM reference because of the influence of the buried dielectric film variation. The training set is not large enough for a robust machine learning solution. However, this approach is still preferred over using unfiltered spectral interferometry spectra for training [Fig. 5(c)]. Any changes in the dielectric film stack will shift the spectral oscillations [see Fig. 3(c)] such that the algorithm cannot find patterns within the small training set, leading to low prediction confidence.

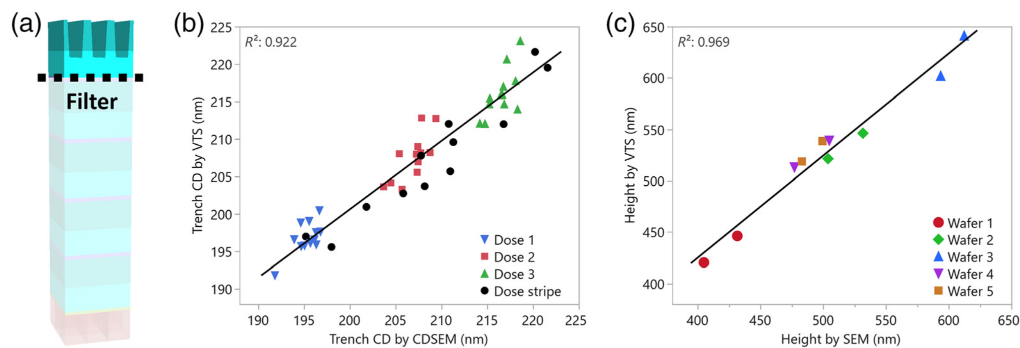

Fig. 4 (a) Schematic of the stack including VTS filter cut-off and comparison between reference metrology and VTS results for (b) trench CD and (c) trench height. The trench CD reference was obtained by CDSEM and the trench height by cross-sectional SEM.

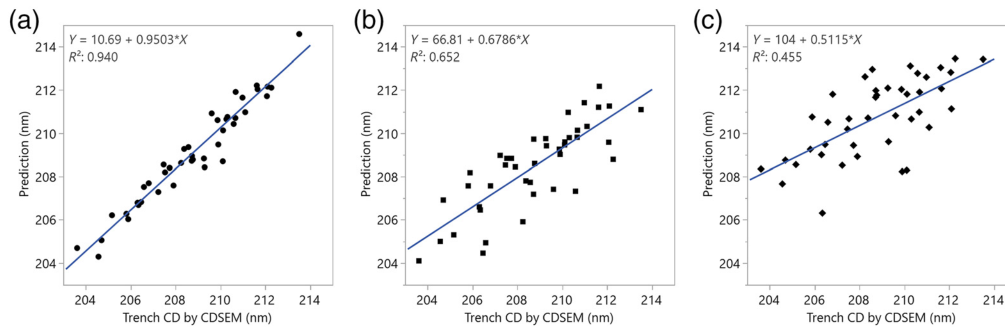

Fig. 5 Trench CD predicted by different machine learning-based solutions as a function of CDSEM reference data. The difference is related to the scatterometry input data: (a) VTS spectra with a filter cut-off just below the trenches; (b) VTS spectra with no cut-off filter; and (c) unfiltered spectral interferometry spectra. Note that the machine learning models were trained on five nominal wafers and then tested on two different nominal wafers.

A significantly larger training set with sufficient process variation in each dielectric film would be required for a robust solution capable of predicting the trench CD. The main issue is that each wavelength within the spectral data is a convolution of optical information from the entire transparent stack. Applying VTS and using deconvoluted spectral information relative to the depth from which they originate enables a machine learning algorithm to identify pattern variations that match the reference data. Therefore, in situations in which complex stacks are involved, large process variations unrelated to the features of interest present, or only small reference datasets are available, machine learning solutions trained with VTS spectra are advantageous over traditional approaches.

4.2 *M2 ILD CMP on Fully Integrated Embedded MRAM Devices*

Embedding memory elements in the early BEOL requires a dielectric planarization before the subsequent metal patterning because the MRAM pillars introduce topography that is otherwise not present. The ILD thickness is location- and design-dependent, and process variations may lead to metallization defects. In the particular case discussed here, measuring a proxy target in the scribe line is not sufficient for detecting process excursion, and thickness monitoring on fully integrated devices post-CMP is required. The MRAM elements are embedded between M1 and M2 on top of a fully built CMOS architecture [Ede20]. A traditional OCD approach would be extremely challenging, if at all possible, due to the optical model complexities. The many degrees of freedom required to account for statistical process variations very likely would lead to a large measurement uncertainty that is not sufficient to control the process.

Four wafers with different CMP split conditions but otherwise identical processing were manufactured. Spectra for both dedicated metrology targets in the scribe line (MTJ only) and fully integrated devices in the active area (MTJ + CMOS) were collected post-ILD CMP on all wafers (Fig. 6). Both targets comprise MRAM pillars with identical dimensions, and the difference between the scribe line targets and the active area was merely the patterning below the memory elements. Specifically, the MTJ-only targets do not comprise any patterned features but only several thin films below the memory pillar [Fig. 6(a)]. Figure 6(b) shows the distinct differences in ILD height after CMP between the two target areas. Moreover, different CMP process conditions do not lead to a constant offset between the two measured targets, which is the main reason why in-die measurements are key for successful in-line process monitoring.

A geometric optical model was built to analyze the dedicated MTJ only metrology target. Using RCWA algorithms, it is possible to solve for the total ILD thickness and the remaining ILD thickness above the pillar. Due to the above discussed structural complexities, a geometric optical model was not attempted for the measured in-die active area. Rather, VTS was used to filter the measured spectra from the fully integrated area such that the irrelevant spectral

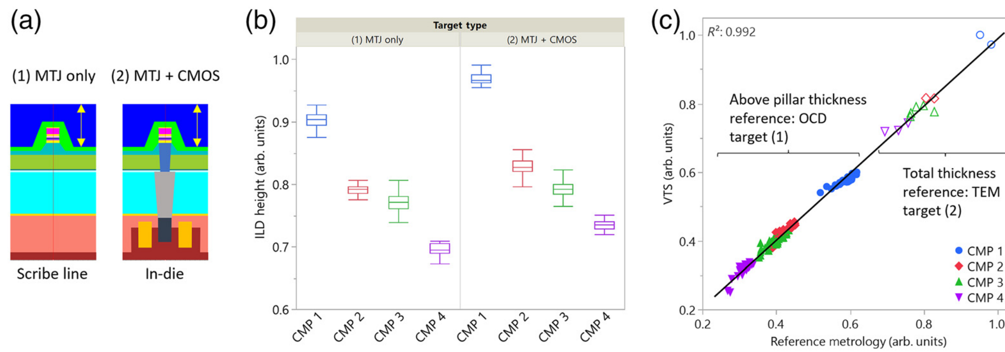

Fig. 6 (a) Schematics of dedicated MRAM metrology target in the scribe line and active device (in-die) comprising MRAM pillars on top of a 14-nm CMOS node. (b) ILD height box plot with across-wafer data for both targets as a function of four different CMP split conditions. (c) Correlation between VTS results and reference measurements from all four split conditions. The open symbols correspond to TEM and the solid symbols to OCD reference data.

information from the buried CMOS does not need to be considered. Consequently, the optical model can be dramatically simplified. This VTS solution was then used to analyze measured spectra from both targets: MTJ only and MTJ + CMOS. The excellent match of VTS results to reference metrology is shown in Fig. 6(c). For the MTJ only target, the ILD height as a result of the traditional OCD model serves as reference. Note that here the remaining ILD thickness above the pillar rather than the height in the open areas is plotted. A total of 12 cross-section TEM images from the MTJ + CMOS targets were analyzed to obtain the total ILD thickness in the area between the pillars for all four CMP conditions. These results confirm that the VTS algorithm can be used to analyze fully integrated targets to obtain device information when a traditional model is infeasible. Moreover, it is possible to use the same VTS solution for analyzing different target types.

4.3 *M1 CMP on Integrated Nanosheet Gate-All-Around Devices*

For early BEOL applications such as M1, there is a desire to measure critical dimensional parameters using OCD on targets that are device-like or with little simplifications to match device processing closely rather than on targets that skip FEOL patterning altogether. For example, subtle metal line height differences that may be observed after CMP related to microtopography do not have to be considered and characterized when measuring and monitoring on device-like targets. Often, this is not done because of the inherent model complexities and the increased measurement uncertainty that then significantly outweigh the benefits. Or often cumbersome hybrid model techniques have to be developed to lower measurement uncertainties, but they still require a detailed and complicated geometric optical model. Specifically in the process development stage, an OCD solution comprising a full optical model is usually not a preferred path. Continuous process improvements in the FEOL and MOL would require frequent updates of the optical model, which is not economical due to the long time-to-solution. VTS enables monitoring of M1 CDs on top of a full FEOL build comprising NS-GAA architectures without the need to consider or model any of the FEOL features. Figure 7(a) depicts a cross-sectional TEM image of a dedicated metrology target comprising nanosheets including gate and source drain (MOL contacts are omitted here). VTS can be used to filter the spectral information related to the FEOL and allows for focusing on the M1 characteristics only. The results of the VTS solution show a very good correlation to OCD and cross-sectional TEM reference metrology (Fig. 7). Note that a traditional OCD model was not developed for the device-like target containing M1 and NS-GAA FEOL due to the architectural complexity. For simplicity and to avoid a time-consuming model solution, the depicted VTS to OCD comparison is using results from a short-loop BEOL wafer with only M1 patterning. The TEM results are obtained from full-flow wafers and the targets comprising M1 on top of a full FEOL build with NS-GAA features [Fig. 7(a)] as well as targets comprising patterned M1 only [Fig. 8(b)], respectively. For all three

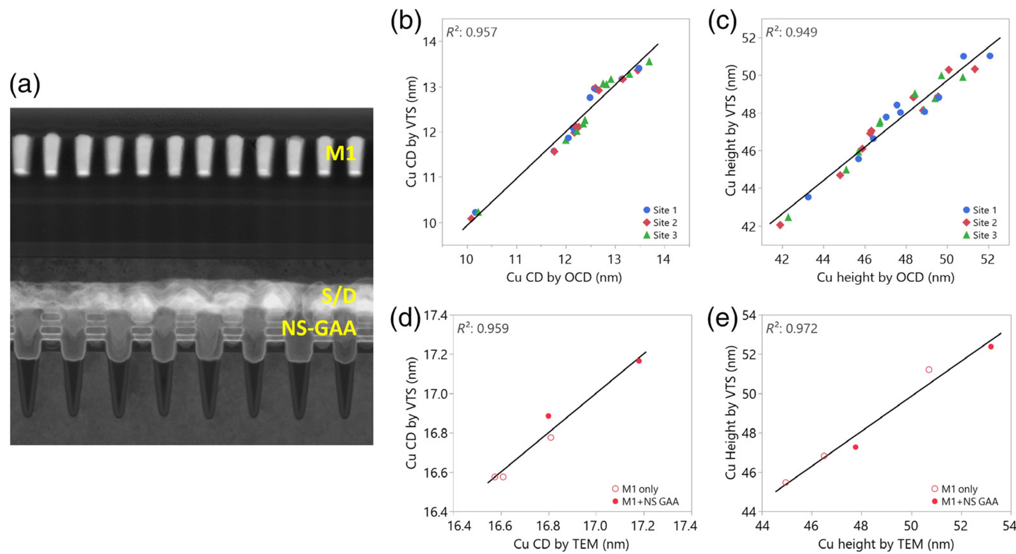

Fig. 7 (a) TEM cross section of a device-like metrology target comprising M1 on top of a full FEOL build with NS-GAA features including gate stack and source/drain; MOL contacts are omitted. (b) M1 Cu CD and (c) M1 Cu height determined by VTS in comparison to OCD reference metrology for three sites and ten measurements each. (d) M1 Cu CD and (e) M1 Cu height determined by VTS in comparison to TEM reference metrology. Open symbols are referring to data from a target having only M1 patterning, and solid symbols depict values obtained from a target depicted in (a).

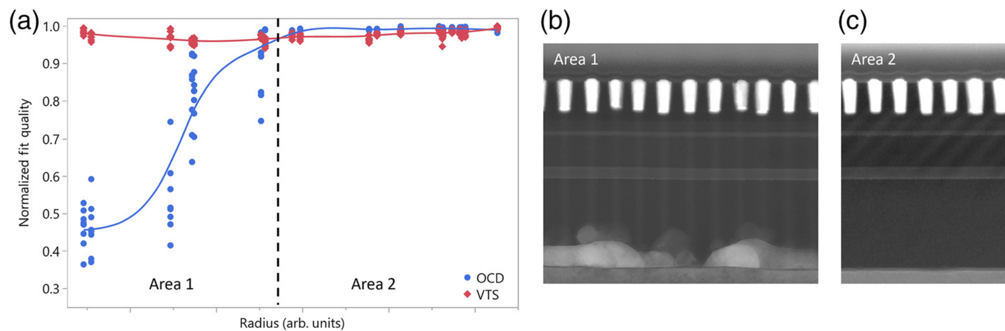

Fig. 8 (a) Normalized fit quality obtained from OCD and VTS solutions for a dedicated metrology target comprising M1 patterning only. The cross-section images correspond to areas (b) toward to the center of the wafer (area 1) and (c) toward the edge of the wafer (area 2).

measured scenarios (short-loop M1 only, full flow M1 only, and M1 + NS-GAA), the same VTS model was used.

While analyzing the dedicated metrology targets with M1 patterning only on some experimental wafers, an interesting observation was made. As expected, the VTS solution delivered consistently good results for all sampled target locations across the wafer. However, the OCD solution fit quality degraded significantly toward the center of the wafer (Fig. 8). Cross-sectional TEM images from two representative locations revealed that, in area 1 (toward the center of the wafer), defects are present at the substrate interface. The traditional OCD model, utilizing the full spectral information, starts to fail as soon as defects are present in the probed region. The VTS solution, however, uses filtered spectral content and is therefore only sensitive to the topmost layers including the M1 layer of interest. Hence, VTS is agnostic to the defects and can deliver equally good results across the entire wafer. This example confirms once again that VTS spectral filtering works as intended, and irrelevant spectral information from buried layers can be ignored for a simplified optical solution.

5 Conclusions

In this paper, the spectral interferometry technique VTS is introduced and demonstrated for three metrology challenges related to state-of-the-art device manufacturing. VTS utilizes information from spectral interferometry and allows for data filtering of irrelevant spectral information from buried layers and features. Therefore, VTS enables selective measurements of the topmost part of a stack and is agnostic to patterned or blanket underlayer variations. As shown, this can dramatically simplify optical modeling, increase sensitivity to parameters of interest, and significantly decrease time-to-solution. Due to the spectral filtering, applications can be solved with VTS when traditional OCD modeling is not feasible.

It was shown that VTS can filter out reflections from buried layers and successfully deliver critical dimensional parameters of trenches on top of a tall stack with many dielectric films. In addition, it was discussed and confirmed that a machine learning model trained with filtered VTS data exhibits superior performance over similar solutions trained with either unfiltered or traditionally measured spectra. When using standard spectra, in the presented case, a much larger training dataset is required to approach VTS correlation levels due to the convolution of depth information across the entire measured spectrum.

Furthermore, M2 ILD CMP monitoring on fully integrated embedded MRAM devices was demonstrated. VTS was used to characterize the remaining ILD thickness in the active area, a task not feasible with traditional OCD modeling that would require a complex geometric model and many floating parameters to account for any process variations. Finally, an example of M1 monitoring on top of a NS-GAA FEOL device architecture was presented. The simplified VTS model is capable of delivering M1 dimensional parameters, and it was shown that the relevant metrics can be reported regardless of underlayer variations. The VTS solution is agnostic to the presence of buried defects and thus to FEOL process variations, meaning the VTS solution does not need to be updated in the case of any process changes affecting buried layers or features.

This is of high value not only for process development but also for high-volume manufacturing due to the increased sensitivity to the parameters of interest and the significantly simplified optical model, which leads to a fast time-to-solution. The VTS technique is not only relevant for solving BEOL logic applications but also highly valuable for nonvolatile flash memory when tall stacks with many layers have to be measured.

Acknowledgments

The authors would like to thank Marjorie Cheng, Haibo Liu, Daphna Peimer, Ilya Osherov, Michael Sendler (all Nova), Stefan Schoeche, Chih-Chao Yang, Daniel Edelstein, and Nelson Felix (all IBM) for support of this work. Nova acknowledges funding for VTS software development from the Electronic Components and Systems for European Leadership Joint Undertaking (Grant No. 826422).

References

1. D. Edelstein et al., "A 14 nm embedded STT-MRAM CMOS technology," in *IEEE Int. Electron. Devices Meeting (IEDM)*, San Francisco, CA, USA (2020).
2. M. Medikonda et al., "Scatterometry based methodologies for characterization of MRAM technology," *Proc. SPIE* **12053**, 120530J (2022).
3. D. Schmidt et al., "OCD enhanced: implementation and validation of spectral interferometry for nanosheet inner spacer indentation," *Proc. SPIE* **11611**, 116111U (2021).
4. D. Schmidt et al., "Vertical travelling scatterometry for metrology on fully integrated devices," *Proc. SPIE* **12053**, 120530S (2022).
5. N. Loubet et al., "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," in *Symp. VLSI Tech.*, pp. 230–231 (2017).
6. J. Zhang et al., "Full bottom dielectric isolation to enable stacked nanosheet transistor for low power and high performance applications," in *IEEE Int. Electron. Devices Meeting (IEDM)*, San Francisco, CA, USA (2019).

7. D. Kong et al., "Measuring local CD uniformity in EUV vias with scatterometry and machine learning," *Proc. SPIE* **11325**, 113251I (2020).
8. S. Kwon et al., "Microsphere-assisted, nanospot, non-destructive metrology for semiconductor devices," *Light Sci. Appl.* **11**, 32 (2022).
9. D. Schmidt et al., "Development of SiGe indentation process control for gate-all-around FET technology enablement," *IEEE Trans. Semicond. Manuf.* **35**, 412 (2022).
10. N. G. Orji et al., "Metrology for the next generation of semiconductor devices," *Nat. Electron.* **1**, 532 (2018).
11. A. Fernández Herrero et al., "Analysis of line-edge roughness using EUV scatterometry," *Nanomanuf. Metrol.* **5**, 149 (2022).
12. C. Porter et al., "Soft X-ray scatterometry: at-resolution, 3D metrology for the EUV era," in *Int. Conf. Frontiers of Characterization and Metrol. for Nanoelectron.* (2022).
13. A. Boosalis et al., "Free-charge carrier profile of iso- and aniso-type Si homojunctions determined by terahertz and mid-infrared ellipsometry," *Thin Solid Films* **519**, 2604 (2011).
14. T. Hofmann, D. Schmidt, and M. Schubert, "THz generalized ellipsometry characterization of highly-ordered three-dimensional nanostructures," in *Ellipsometry at the Nanoscale*, M. Losurdo and K. Hingerl, Eds., Springer (2013).
15. M. Wormington et al., "X-ray critical dimension metrology solution for high aspect ratio semiconductor structures," *Proc. SPIE* **11611**, 116110W (2021).
16. G. A. Antonelli et al., "Ellipsometric critical dimension metrology employing mid-infrared wavelengths for high-aspect-ratio channel hole module etch processes," *Proc. SPIE* **11611**, 116111O (2021).
17. M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," *J. Opt. Soc. Am.* **72**, 1385 (1982).
18. S. J. O'Mullane, D. Dixit, and A. C. Diebold, "Advanced ellipsometry measurements for semiconductor structures," in *Characterization and Metrology for Nanoelectronics and Nanostructures*, Pan Stanford (2016).
19. D. Schmidt et al., "In-line Raman spectroscopy for gate-all-around nanosheet device manufacturing," *J. Micro/Nanopattern. Mater. Metrol.* **21**, 021203 (2022).
20. P. Hönicke et al., "Simultaneous dimensional and analytical characterization of ordered nanostructures," *Small* **18**, 2105776 (2022).
21. D. F. Sunday et al., "Characterizing patterned block copolymer thin films with soft x-rays," *ACS Appl. Mater. Interfaces* **9**, 31325 (2017).
22. L. Towidjaja et al., "Back end of line metrology control applications using scatterometry," *Proc. SPIE* **6152**, 61521X (2006).
23. S. Zangooie et al., "Characterization of 32nm node BEOL grating structures using scatterometry," *Proc. SPIE* **6922**, 69220S (2008).
24. B. L'Herron et al., "Hybridization of XRF/XPS and scatterometry for Cu CMP process control," *Proc. SPIE* **9424**, 94241M (2015).
25. M. Breton et al., "Electrical test prediction using hybrid metrology and machine learning," *Proc. SPIE* **10145**, 1014504 (2017).
26. S. Das et al., "Scatterometry solutions for 14nm half-pitch BEOL layers patterned by EUV single exposure," *Proc. SPIE* **11611**, 116112A (2021).
27. P. Timoney et al., "Hybrid scatterometry measurement for BEOL process control," *Proc. SPIE* **10145**, 1014506 (2017).
28. A. Vaid et al., "Implementation of hybrid metrology at HVM fab for 20nm and beyond," *Proc. SPIE* **8681**, 868103 (2013).
29. I. Osherov et al., "Application of advanced hybrid metrology method to nanoimprint lithography," *Proc. SPIE* **10145**, 101451X (2017).
30. D. Schmidt et al., "Advanced EUV resist characterization using scatterometry and machine learning," in *Annu. SEMI Adv. Semicond. Manuf. Conf. (ASMC)*, Milpitas, CA, USA (2021).

Biographies of the authors are not available.