

Need for LWR metrology standardization: the imec roughness protocol

Gian Francesco Lorusso; Takumichi Sutani; Vito Rutigliani; Frieda van Roey; Alain Moussa; Anne-Laure Charley; Chris Mack; Patrick Naulleau; Chami Perera; Vassilios Constantoudis; Masami Ikota; Toru Ishimoto; Shunsuke Koshihara

Gian Francesco Lorusso

IMEC (Belgium)

Takumichi Sutani

Hitachi High-Technologies Corp. (Japan)

Vito Rutigliani , Frieda van Roey , Alain Moussa , Anne-Laure Charley

IMEC (Belgium)

Chris Mack

Fractilia, LLC (United States)

Patrick Naulleau

Lawrence Berkeley National Lab. (United States)

Chami Perera

EUV Technology (United States)

Vassilios Constantoudis

Institute of Nanoscience and Nanotechnology (Greece)

Masami Ikota

Hitachi High-Tech Science Corp. (Japan)

Toru Ishimoto , Shunsuke Koshihara

Hitachi High-Technologies Corp. (Japan)

Abstract

As semiconductor technology keeps moving forward, undeterred by the many challenges ahead, one specific deliverable is capturing the attention of many experts in the field: line width roughness (LWR) specifications are expected to be <2 nm in the near term, and to drop below 1 nm in just a few years. This is a daunting challenge and engineers throughout the industry are trying to meet these targets using every means at their disposal. However, although current efforts are surely admirable, we believe they are not enough. The fact is that a specification has a meaning only if there is an agreed methodology to verify if the criterion is met or not. Such standardization is critical in any field of science and technology and the question that we need to ask ourselves today is whether we have a standardized LWR metrology or not. In other words, if a single reference sample were provided, would everyone measuring it get reasonably comparable results? We came to realize that this is not the case and that the observed spread in the results throughout the industry is quite large. In our opinion, this makes the comparison of LWR data among institutions, or to a specification, very difficult. We report the spread of measured LWR data across the semiconductor industry. We investigate the impact of image acquisition, measurement algorithm, and frequency analysis parameters on LWR metrology. We review critically some of the International Technology Roadmap for Semiconductors (ITRS) metrology guidelines [such as measurement box length <2 μm and the need to correct for scanning electron microscope (SEM) noise]. We compare the SEM roughness results to AFM measurements. Finally, we propose a standardized LWR measurement protocol—the imec roughness protocol—intended to ensure that every time LWR measurements are compared (from various sources or to specifications), the comparison is sensible and sound. We deeply believe that the industry is at a point where it is imperative to guarantee

that when talking about a critical parameter such as LWR, everyone speaks the same language, which is not currently the case.

1.

Introduction

In the last few years, the importance of being able to correctly understand the impact of the stochastic nature of the various elements involved in the lithography process—from light to materials—has become more and more central for the advancement of the field of semiconductor technology. In this respect, the last International Technology Roadmap for Semiconductors (ITRS) forecast for line width roughness (LWR) specifications was quite daunting, requiring an LWR of 1.3 nm by 2020, down to 0.6 nm by 2028.¹ Meeting these goals will surely require a formidable effort from all parties involved, from material and equipment suppliers, to integrated circuit manufacturers (ICM).

Our focus here, however, is not how we are going to meet these specifications, but rather whether we are currently capable of properly assessing our performance in terms of roughness. The question we would like to answer in this paper is then the following: as we stand, are we able to judge if roughness requirements are met or not, and if not, what are the solutions? In other words, we are interested in making sure that we have credible roughness metrology, as we deeply believe that this is essential to achieve our goals.

In Sec. [2.1](#), we report the spread of measured LWR data across the semiconductor industry. In Sec. [2.2](#), we define the basic concepts of unbiased metrology. We investigate the impact of image acquisition averaging in Sec. [2.3](#). The consequences of digital filtering are analyzed in Sec. [2.4](#). In Sec. [2.5](#) we analyze the impact of pixel size and field of view (FOV) on roughness metrology. In Sec. [2.6](#) we re-evaluate critically the requirement of 2- μ m²- μ m measurement box length, while in Sec. [2.7](#) we compare the impact of biased and unbiased metrology on accuracy, using atomic force microscopy (AFM) as a reference metrology.

Finally, in Sec. [2.8](#), we propose a standardized LWR measurement protocol, the imec roughness protocol (iRP). Such a protocol is aimed to ensure LWR standardization across the industry, so that measurements from various sources can be compared. This paper is a joint effort of imec and four major providers of roughness metrology, as we all deeply believe that the industry is at a point where it is imperative to guarantee roughness metrology standardization.

2.

Results and Discussion

2.1.

ITRS Requirements and the Status of LWR/LER Metrology Today

In terms of metrology, the ITRS used to provide some guidelines.¹ The current ITRS definition, developed in the 2003 to 2005 timeframe, was a significant improvement at the time, replacing its predecessor which had become inadequate for 45/32 nodes. First, a 2- μ m²- μ m measurement box length longer, with 10 nm sampling distance in yy, was requested to avoid artificially low LWR measurements. In fact, the measured roughness does decrease when shortening the measurement box length. Specifically, this metrology requirement resulted in the implementation and adoption of the rectangular image acquisition mode across the industry, using different pixel sizes in XX and YY.

A second ITRS requirement demanded to correct for the scanning electron microscope (SEM) noise.^{2,3} In fact, the image noise does alter the roughness measurements results, as it interferes with the edge detection algorithms. Hence, it is necessary to remove this contribution to get a proper roughness estimate.

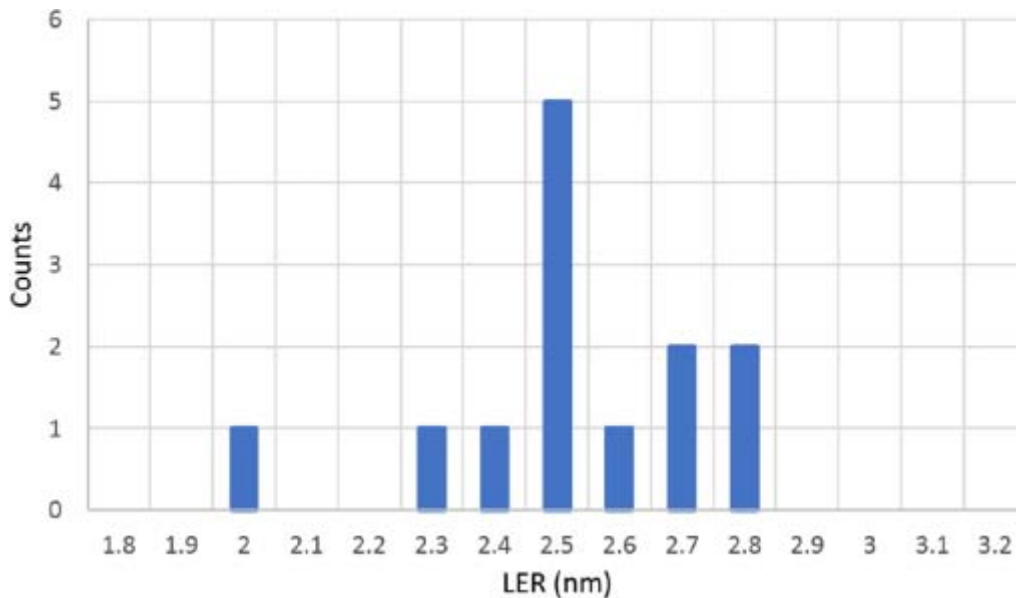
Are these two metrology requirements sufficient? We know that to measure roughness, we must define many additional parameters related to the image acquisition and to the measurement algorithms, such

as pixel size, FOV, frame averaging, digital filtering, and so on. The question is whether or not all these additional settings have an influence on the measurement outcome.

To answer this question, we performed a round robin exercise focusing on roughness metrology. We asked 13 different companies (ICM, equipment, and material suppliers) to measure the same exact wafer using their own internal best-known methods (BKM) and to feed back the results to us. The outcome of such an exercise is reported in Fig. 1. The results indicated a spread of the estimated roughness across the ICM industry of $\pm 30\% \pm 30\%$ (3σ).

Fig. 1

LER estimated by 13 different companies on the same wafer, each one using their own BKM. The results show a $3\sigma = \pm 30\% \pm 30\%$ uncertainty in the results.



The outcome of the roughness metrology round robin shows us that we have a fundamental problem. A $\pm 30\% \pm 30\%$ 3σ uncertainty on the measured LWR is clearly not acceptable if we need to compare our results to a specification or to other estimates coming from a different source. Obviously, this is not an issue if the measurement is only intended for process monitoring, but our focus here is not on process monitoring.

The fact that we found such a large variability clearly indicates that the ITRS constraints on roughness metrology need to be revised, as they do not to guarantee the highly needed standardization of roughness measurement procedures across the semiconductor industry. Hence, a more stringent set of metrology requirements needs to be defined to reduce the observed variability, ultimately caused by parameters known to have a dramatic influence on the measurement results. In the following sections, we investigate various acquisition and measurement parameters to define a specific procedure—iRP—capable to ensure a stable roughness metrology.

2.2.

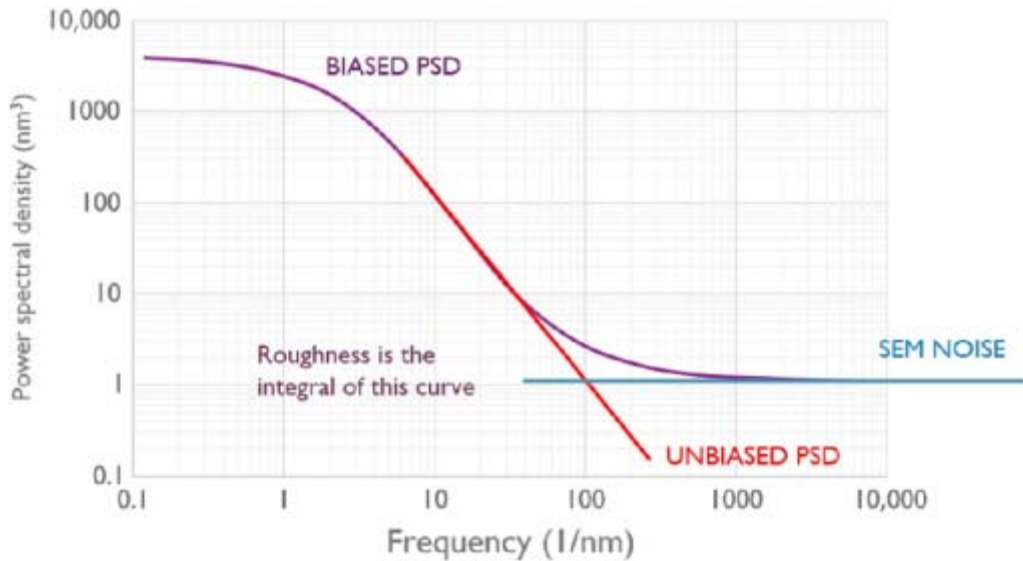
Unbiased Roughness Estimate

As briefly discussed in Sec. 2.1, a critical step for roughness metrology is to remove the influence of the SEM noise from the roughness measurement. The most common procedure used to this aim is the subtraction of the SEM noise in Fourier space.⁴

This procedure is shown in Fig. 2. The biased power spectral density (PSD) obtained from a SEM image is clearly dominated by the SEM noise in the high-frequency regime. The SEM noise is a constant if it consists of white uncorrelated noise. If this is the case, then removing this constant from the bias PSD will then provide an unbiased PSD. Hence, the roughness calculated as the integral of the unbiased PSD will be immune to the SEM noise influence.

Fig. 2

The biased PSD shows a high frequency plateau caused by the CD SEM noise. Estimating the SEM noise in Fourier space (blue) allows to correct the PSD for the influence of the SEM noise (red). The integral under the unbiased PSD curve will provide the unbiased roughness estimate.



Although the need to correct for SEM noise is a fundamental requirement clearly stated by the ITRS, we must remark that none of the procedures used in the round robin exercise made such a correction. It is then essential that the roughness protocol explicitly requests the unbiasing of the SEM noise.

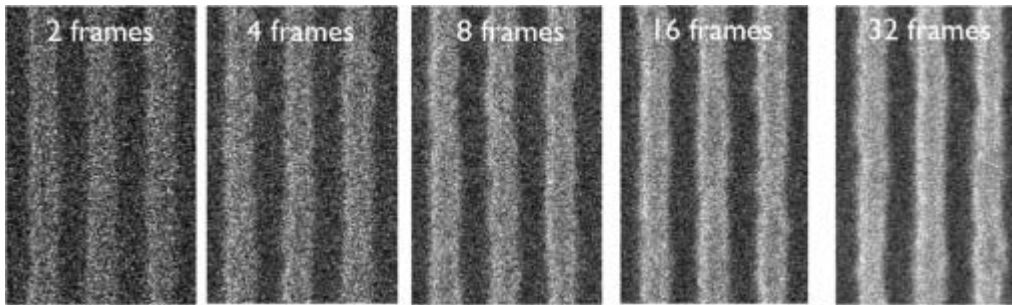
2.3.

Effect of Frame Averaging

To show how critical the unbiasing of the roughness estimate is and to understand if it is sufficient to guarantee a reliable roughness metrology, in this section we will compare biased and unbiased results as a function of averaging. In fact, the number of frames used to acquire an image has a direct impact on the level of SEM noise present in the data. The effect of frame averaging on image quality is clearly discernable in Fig. 3, even to an unexperienced eye.^{2,3}

Fig. 3

Impact of the averaging (number of frames) on image quality. As the number of frames decreases, the level of SEM noise in the images increases.

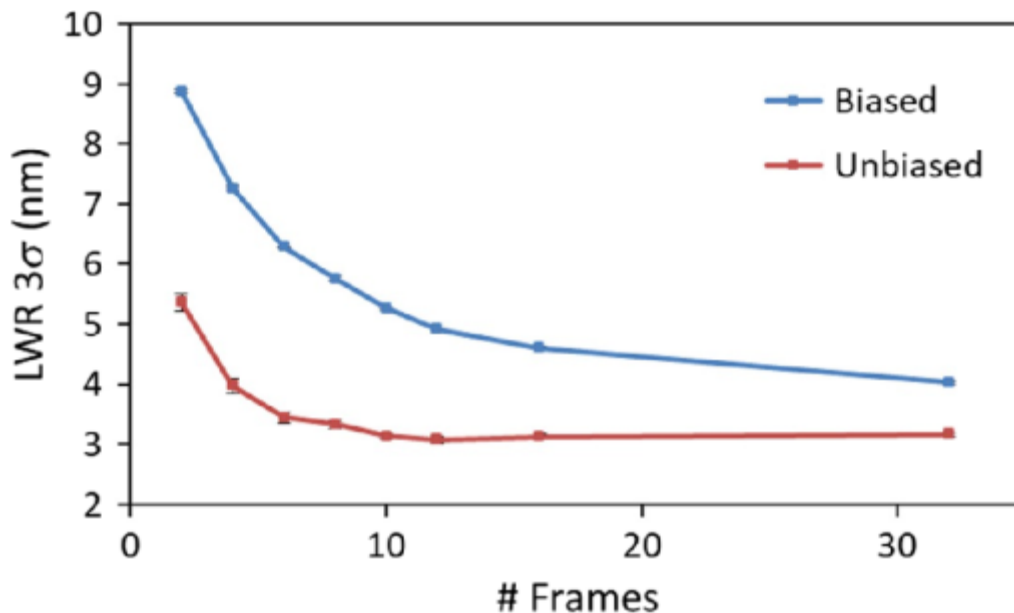


The main consequence of increased SEM noise is that, as the frame averaging is reduced, the edge detection algorithm will struggle to identify the edge,^{2,3} and, therefore, it will measure an artificially higher roughness, unrelated to the actual on-wafer roughness (which is obviously the same for all images), but caused by the specific signal-to-noise ratio level in each image.

The impact of the averaging on the measured roughness is summarized in Fig. 4. As the number of frames is decreased, we observe that the biased LWR measurement increases drastically. By contrast, the unbiased LWR estimate, obtained by subtracting the SEM's flat white noise floor as described in Sec. 2.2, is lower in value and remains constant, at least down to about 10 frames in our case. This means that, for frame averaging larger than 10, the unbiasing procedure has less variability (is independent of the averaging) and it is expected to be more accurate than the biased measurement, being corrected for SEM noise.

Fig. 4

Impact of averaging (number of frames) on biased and unbiased LWR estimates. The biased LWR estimate increases dramatically when decreasing the frame number, while the unbiased LWR estimate stays constant down to a minimum number of frames (about 10 in this case). Figure from Ref. 4.

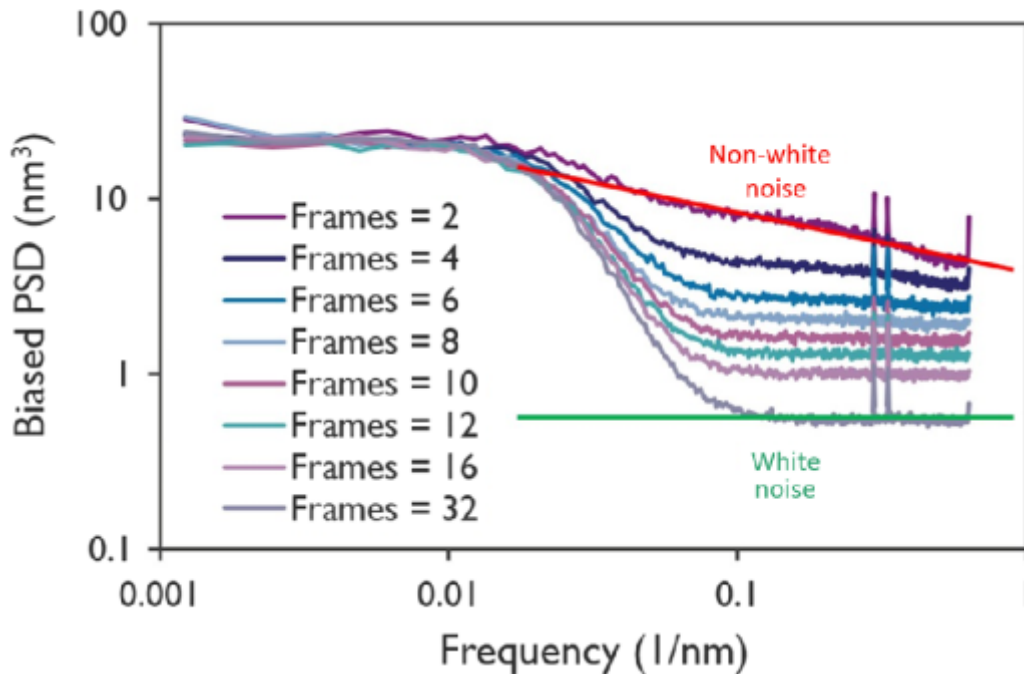


In the case of very limited averaging (frame number lower than 10), we observe that even the unbiased LWR estimate starts to pick up, indicating that the unbiasing procedure described in Sec. 2.2 starts to fail in the limit of exceedingly low integration settings.

To understand why the unbiasing procedure breaks down at a lower number of frames, the PSDs at various integration times are reported in Fig. 5. It is evident that, in the high frequency range, PSDs with a large number of frames have a flat noise floor, consistent with the white noise hypothesis required to justify the unbiasing procedure, as discussed in Sec. 2.2. By contrast, the high-frequency component appears to be sloped for small frame numbers, thus indicating that the white noise hypothesis is not satisfied anymore for very limited averaging and, subsequently, the unbiasing procedure described in Sec. 2.2 is not valid in this regime.

Fig. 5

PSDs for different averaging (number of frames). The high frequency noise contribution stays flat (white noise) down to a minimum number of frames. Below this limit, the high frequency portion of the PSD shows a sloped behavior, indicating a violation of the white noise assumption. Figure from Ref. 4.



In conclusion, the results discussed in this section provide two main findings. First and foremost, it is evident that unbiased roughness estimates are critical to reduce noise-induced variability and to provide more accurate and sensible results. Please consider that, without unbiasing, the measurement would be sensitive to any source of noise, not only integration time but also material dependencies or charging artifacts. A comparison of biased roughness between different resists, for example, is not appropriate even if the same measurement setting is used, as different resists will have different electron yield and, consequently, different levels of SEM noise. Considering that measurements reported across the semiconductor industry are very often biased, this provides a worrying perspective on the way critical decisions are sometimes made.

The second finding is that unbiasing by itself is not enough to guarantee a standardized roughness metrology. In the present work, for example, unbiasing is done using noise floor subtraction that, as explained in Sec. 2.2, is justified only if the white noise assumption is correct. However, the data in Fig. 5 show that this is not true in the low frame averaging regime. Hence, to have a sound and compatible roughness metrology, it is essential to define a complete protocol—the main goal of this paper—clearly indicating how to choose the critical parameters that could influence the results (one of them being the averaging, as demonstrated above).

2.4.

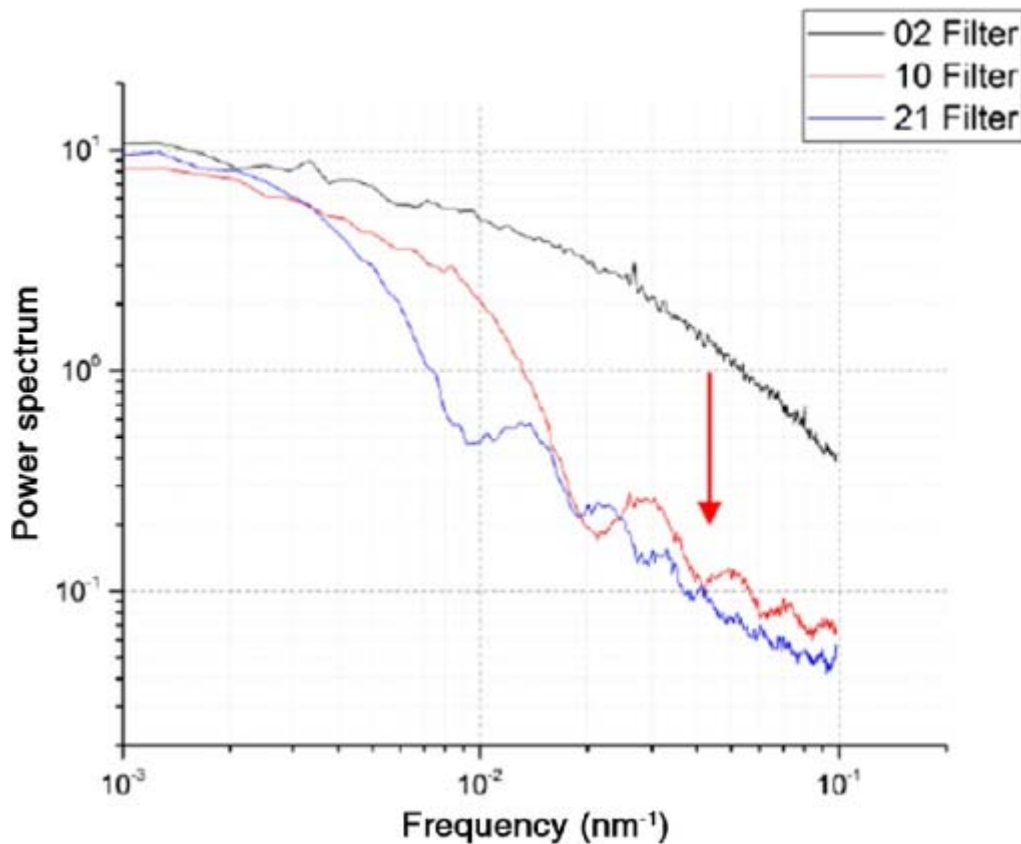
Effect of Filtering

It is a frequent practice in SEM metrology to reduce the impact of noise on the measurements results by applying digital filters. This common procedure is relatively harmless in critical dimension (CD) metrology, where filtering can improve precision at the cost of a slight loss in sensitivity, introducing a measurement bias that can anyway be corrected afterwards, using reference metrology. However, as we will see and as reported in the literature,⁵⁶⁻⁷ filtering has significant consequences on roughness metrology and, as in the case of image averaging, no quantitative guideline was provided by the ITRS in this respect.

The consequences of large digital filtering are shown in Fig. 6, where its impact on PSDs calculated from the same set of images is reported. By increasing the size of the filter, a dramatic change takes place in terms of PSD shape. Digital filtering is observed to literally erase the high- and mid-frequency information contained in the PSD, introducing artifacts (oscillations) that make any sort of PSD analysis fictitious. In addition, since roughness is the integral of the PSD, digital filtering will arbitrarily modify the measured roughness. In other words, when measuring roughness with the aid of digital filters (with or without unbiasing), we can basically obtain any results we want by changing the filter size, thus making our roughness metrology basically meaningless.

Fig. 6

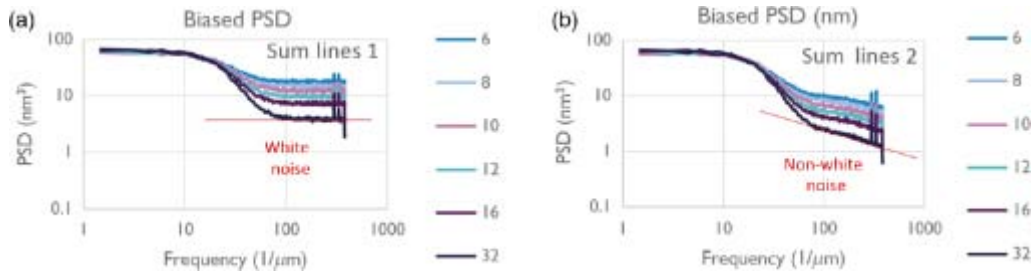
Impact of digital filtering in yy on PSD. The high frequency information is progressively removed and artifacts (oscillations) are introduced in the results, making meaningful roughness metrology impossible.



In some cases, however, a small amount of digital filtering is critically needed to make standard edge detection algorithms work. Even mild filtering has an impact on the results, as shown in Fig. 7. The impact is less dramatic when compared to the situation discussed in Fig. 6, but we observe that mild digital filtering does modify the high-frequency behavior of the PSD, making the white noise hypothesis invalid and the unbiasing procedure in principle inapplicable.

Fig. 7

(a) Unfiltered PSD results versus (b) filtered PSD results with filter 2 pixels in yy for different frame averaging (from 6 to 32 frames).



To summarize, filtering has a very negative impact on roughness metrology, and it should not be used. This is indeed possible, due to the introduction of the recently introduced filterless algorithms.⁸

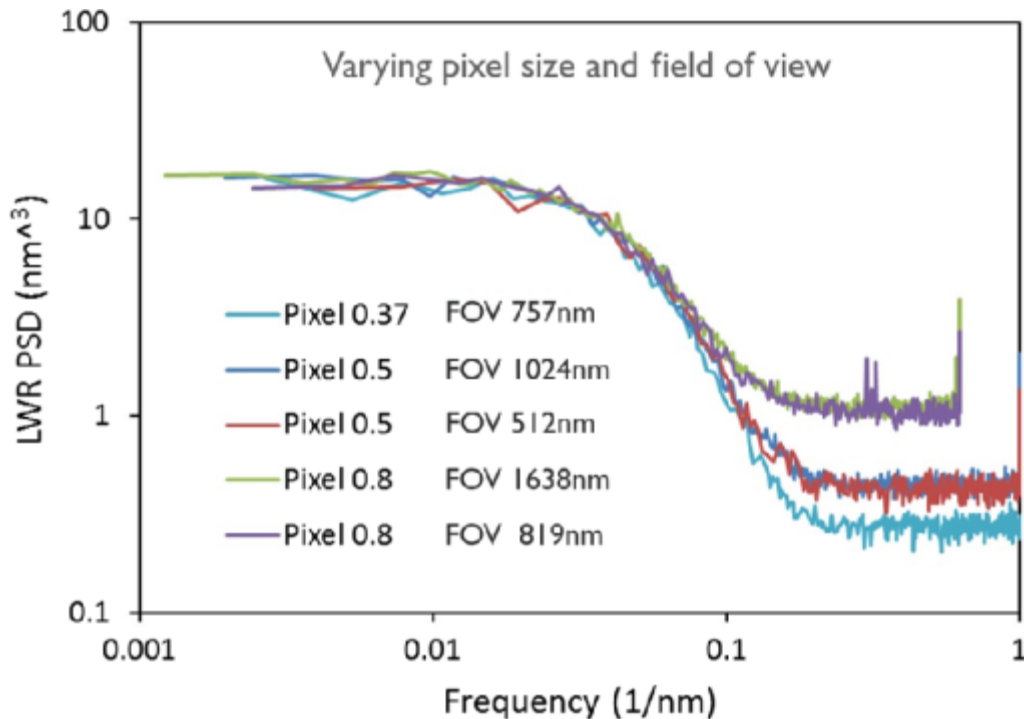
Effect of Pixel Size and FOV

Another critical parameter that will have a strong impact on roughness metrology is the pixel size. Also in this case, even though the influence of pixel size on measured roughness is quite intuitive, the ITRS limited its recommendations to a yy pixel size smaller than 10 nm.

In Fig. 8, we report the impact of pixel size and FOV on measured PSDs. As the pixel size changes, the high frequency plateau is observed to change too, and the impact of the SEM noise becomes less influential when the pixel size is smaller. Ultimately, this means that biased roughness results, if measured with different pixel size, will not be consistent, as each pixel size is characterized by a different amount of SEM noise.

Fig. 8

PSD results for different pixel sizes and FOV. The high frequency flat floor depends on pixel size, while FOV does not appear to modify the shape of the PSD curves. Figure from Ref. 4.

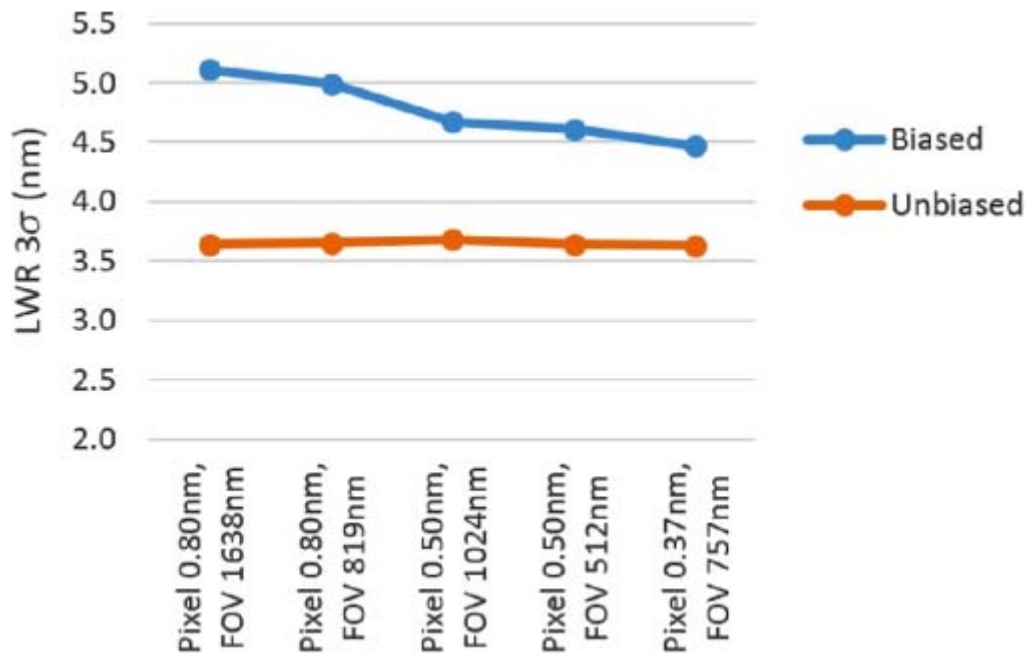


By contrast, the impact of FOV on roughness is observed to be negligible. In [Fig. 8](#), the PSD curves with different FOV do overlap, if the pixel size remains the same. This means that the SEM noise level is not influenced by the FOV, and the image size can be adjusted according to the needs, provided that the pixel size remains constant and the number of pixels is changed instead.

The influence of pixel size on the high frequency SEM noise, shown in [Fig. 8](#), indicates that unbiasing is needed when comparing data acquired with different pixel resolution. In [Fig. 9](#) biased and unbiased LWR results, obtained by the PSD curves in [Fig. 8](#), are compared. In the biased case, the difference caused by the pixel size is quite evident, while the FOV does not seem to have a major impact. The unbiased measurements, by contrast, show a very low variability (range=0.05nmrange=0.05 nm) as compared to the biased ones (range=0.64nmrange=0.64 nm).

Fig. 9

Biased and unbiased LWR results obtained from the PSDs in [Fig. 8](#). The impact of pixel size on variability is evident in the biased results, while FOV (changes in the number of pixels) has only a marginal influence on the results. The unbiased results show a very limited variability, demonstrating that than un-biasing procedure removes any influence caused by the pixel size.SD results for different pixel sizes and FOV. The high frequency flat floor depends on pixel size, while FOV does not appear to modify the shape of the PSD curves. Figure from Ref. [4](#).



Although in principle, unbiasing the PSD will remove the pixel size influence, in practice, if the pixel size is too large, the unbiasing procedure will become less effective, as the high frequency SEM noise plateau and the low frequency floor will get closer and estimations of the SEM noise will be prone to errors.⁹ In conclusion, pixel size has a strong influence on roughness metrology, and large pixel sizes should be avoided, as they introduce a critical amount of SEM noise. By contrast FOV changes (obtained by altering the number of pixels forming the image) are not going to influence roughness metrology. Unbiasing the measurement is critical but, as in the case of averaging, it is not enough to solve all the issues, as unbiasing will not work for large pixel sizes. It is then critical to define a specific pixel size in the measurement protocol, to avoid uncertainty.

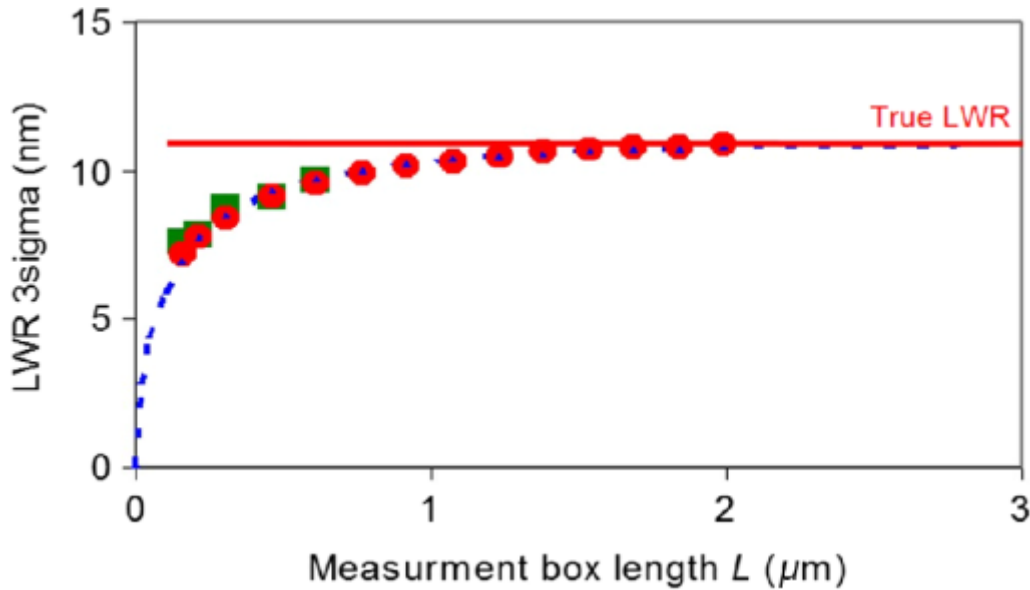
2.6.

Effect of Measurement Box Length

As discussed in Sec. 2.1, besides removing the SEM noise, the ITRS mandated to use a measurement box length of $2\mu\text{m}^2$ μm or longer. The reason for this is that—among the other factors discussed above—the roughness measurement is also function of the box length (the longer the line segment being sampled, the larger the number of low frequency components getting added, the larger the resultant measured roughness). This is intuitively understandable as, when the box length tends to zero, the roughness measurement will tend to zero as well, while for an infinite box length, the measurement will be asymptotically accurate. The variation of the roughness measurement as a function of the box length is shown in Fig. 10. So, in order to ensure a stable LWR metrology, as well as to avoid reporting artificially small numbers, a minimum box length of $2\mu\text{m}^2$ μm was mandated by the ITRS.

Fig. 10

Impact of the measurement box length on the roughness accuracy. The dashed line is the analytical prediction (Eq. 1), while the markers are the experimental results.



However, although the box length requirement was widely adopted throughout the industry, one unintended consequence of its success was that, to satisfy such a requirement, a rectangular scan acquisition mode was implemented. This approach used asymmetric pixel size in XX and YY, with a typical pixel size in YY of about 5 nm. However, because of the negative impact of large pixel sizes (see Sec. 2.5), this caused the proliferation of image acquisition settings characterized by an elevated level of SEM noise. In addition, as unbiasing is not exactly popular, this led to overestimates of the measured roughness throughout the industry. In other words, the solution originally suggested by ITRS, aimed to get a stable roughness metrology as well as to avoid reporting artificially small numbers, backfired in the long term, ultimately resulting in artificially large roughness numbers.

In order to address this issue, let us start by discussing whether the requirement is still valid. In a previously published work,¹⁰ we calculated the functional form of the roughness with respect to measurement box length

$$(1)$$

$$\sigma = \sigma_{inf} \cdot \left[1 - 4\xi \left\{ L + [-1 + \exp(L/\xi)] \xi \right\} / 2L^2 \right]$$

$$\sigma = \sigma_{inf} \cdot \left[1 - 4\xi \left\{ L + [-1 + \exp(L/\xi)] \xi \right\} / 2L^2 \right]$$

We remark that the roughness measurement, besides being a function of the box length L , is also function of another parameter, the correlation length ξ ,¹⁰ which has been neglected in the ITRS discussion. In fact, the asymptotic convergence to σ_{inf} goes as L/ξ . The correlation length ξ is related to resist resolution, a parameter that has obviously improved since the time ITRS defined its requirement, and if we look at published data, we observe that ξ has dropped from 35¹⁰ down to 7 nm¹¹ in the last 10 years. This means that, if we want to keep the same accuracy level, we can drop the measurement box length requirement from 2 μ m down to 400 nm.

In conclusion, the box length should not be a fixed requirement, but needs to be connected to the correlation length ξ . The box length should be about 50 times the correlation length to guarantee an accuracy on the order of a few percent. In addition, rectangular scans, when needed, should be implemented by changing the FOV in terms number of pixels while keeping constant the pixel size, to avoid the inconsistency caused by the pixel size changes, as discussed in Sec. 2.5.

Accuracy

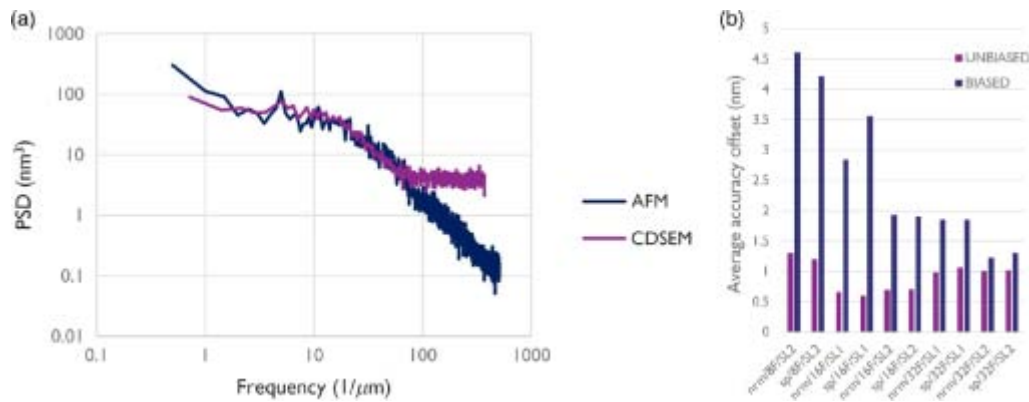
We used AFM to compare the roughness metrology accuracy in the biased and unbiased cases.^{12,13} More than 200 AFM images were acquired to meet the statistical requirements. Five different focus conditions were used to compare the sensitivity of the roughness measurement by AFM and SEM. To avoid scanning artifacts, the AFM images were obtained by scanning parallel to the features. All AFM results were deconvolved in both XX and YY. Nondeconvoluted data clearly showed artifacts related to the tip smoothing effect in both images and PSD spectra. These artifacts were removed by the deconvolution process. Probe shape calibration is done on standard sample (NCDMO-70X from VLSI). After etching, 20 nm isolated lines were measured. A carbon nanotip (10 to 15 nm diameter) was used for the measurements. The AFM and SEM acquisition settings used the same pixel size and number of pixels to avoid discrepancies.

SEM measurements with a variety of settings were used to test the impact on variability. Measurements with different averaging (8, 16, and 32 frames), different scan strategies (normal and special), and different digital filtering (sum line 0 and 1) were acquired.

The SEM and AFM PSD are compared in Fig. 11(a). The average accuracy offset between SEM and AFM for all the various SEM conditions, for both biased and unbiased roughness metrology, are reported in Fig. 11(b). As expected, the biased results are strongly dependent on the SEM measurement setting, with an accuracy offset varying from 1 to 5 nm. The unbiased results, by contrast, show a limited variability, from 0.6 to 1.4 nm.

Fig. 11

(a) The SEM and AFM PSD are compared. (b) Accuracy offset of 3σ LER between AFM and SEM for various SEM conditions, in the biased and unbiased cases. The xx-axis describes different scan strategies (nm, sp), 8, 16, or 32 frames of integration (8F, 16F, 32F), and the number of sum lines of 1 and 2 (SL1, SL2).



In conclusion, these results indicate that, as expected, the unbiasing procedure can greatly reduce the variability caused by the different SEM settings. Nevertheless, a residual variability of 1 nm range still exists in the unbiased case, which is far from small if we must measure roughness in the order of 1 nm or less, thus demonstrating once again the need to create a well-defined protocol.

2.8.

imec Roughness Protocol 2018

Finally, using all the information discussed in the previous chapters, we decided to define the iRP (2018), shown in Table 1. The goal is to achieve a much overdue alignment of the metrology for such a critical parameter across the industry.

Table 1

The imec roughness protocol. If any of the options is used, it should be clearly reported with the data.
imec roughness protocol 2018 (iRP 2018)

Parameter	Setting	Options
Pixel size XX	0.8 nm	
Pixel size yy	0.8 nm	
Beam current	8pA	
Landing energy	500 eV (ADI) / 800 eV (AEI)	
Pixel # XX	2048	(a)
Pixel # yy	2048	(b)
Frame averaging	16 (ADI) / 32 (AEI)	
Filtering XX	None	(c)
Filtering yy	None	
Threshold	60%	
Minimum # of images	50	
Images files	TIF uncompressed	
PSD unbiasing	always	

(a): Pixel # in x can change.

(b): Pixel # in y can change. Box length should be $50 \times \xi$.

(c): Filtering in x should not be used. If this option is not available, filter in x should be 5 pixels.

Following the discussion Sec. 2.5, we fixed the pixel size to 0.8 nm in both XX and YY. The beam parameters are defined with different landing energy for after lithography inspection (ADI) and after etch inspection (AEI). The standard setting in terms of number of pixel is $2048 \times 2048 \times 2048$, covering an FOV of about $1.6 \mu\text{m} \times 1.6 \mu\text{m}$, thus covering a broad range of cases. The number of frames is defined differently for ADI and AEI, trying to avoid the low-averaging regime and limiting the dose for the ADI case. No filter is allowed in either XX or YY. Threshold is fixed to 60%. Any analysis should be performed on at least 50 uncompressed TIF images, as PSD unbiasing can only be performed if the PSD is statistically sound. Finally, and most importantly, unbiasing of the roughness is strictly mandatory. The third column of the iRP indicates which parameters are allowed to vary, as detailed in the footnote. The number of pixels in XX and YY can change, as the FOV has no impact on the roughness metrology, as shown in Sec. 2.5. The FOV in YY should be set based on the correlation length. Although in principle filtering should never be used, we need to remark that not all the analysis tools have access to filterless algorithms. We decided that it is better, at least for now, to limit the use of filtering, if it must be done, to a specific setting (5 pixel in xx). This recommendation could be revised in the future, depending on the progress made from the metrology suppliers in this area. Finally, we stress that, if any of the options allowed in the iRP are used, it is critical that this information is provided with the data to avoid confusion.

3.

Conclusions

In this paper, we investigated the robustness of roughness metrology in ICM nowadays. Our round robin exercise indicated the existence of an unacceptably large variability across the industry ($3\sigma = \pm 30\%$). We studied then how much such a variability is influenced by the choice of setting parameters for which we currently have no stringent guidelines (averaging, pixel size, digital filtering,

FOV, and more). We found that most of them have a tremendous impact on the results. Specifically, averaging and pixel size have an enormous impact on SEM noise, and digital filtering can make roughness metrology basically meaningless. FOV changes, by contrast, are relatively harmless, but only if they are implemented by changing the number of pixels, not the pixel size. We re-evaluated as well the ITRS requirement on $2\mu\text{m}^2$ μm measurement box length, finding it obsolete in most cases, and that, when needed, it should explicitly account for the correlation length.

We found that the impact of these measurement parameters can be mitigated if unbiased metrology is used. Unbiased metrology is in fact a critical, unavoidable tool to ensure a sound roughness metrology. However, our results indicated that to be able to have an effectively standardized metrology, unbiased roughness metrology is necessary but not sufficient, as such methodology is justified only in certain ranges (not valid for large pixel size, low averaging regimes, or in case of digital filtering).

In conclusion, our results clearly indicate the need to have a stringent protocol, rigorously defining the measurement and analysis settings needed to achieve a standardized roughness metrology. Hence, we outlined here what we called the iRP, intended to replace the limited set of recommendations by the ITRS with a consistent, more comprehensive set of guidelines.

In the future, our goal is to drive adoption of the roughness protocol across the industry, as well as to drive the development of the analytical tools needed for unbiased metrology, and to overcome current limitations such as those related to the white noise assumption or digital filtering.

Acknowledgments

We would like to thank all the people at imec, Hitachi, Fractilia, Demokritos, and SuMMIT who contributed to this work with fruitful discussion and suggestions.