

Journal of
Micro/Nanolithography,
MEMS, and MOEMS

Nanolithography.SPIEDigitalLibrary.org

Multitaper and multisegment spectral estimation of line-edge roughness

Yao Luo
Serap A. Savari

Multitaper and multisegment spectral estimation of line-edge roughness

Yao Luo* and Serap A. Savari

Texas A&M University, Department of Electrical and Computer Engineering, College Station, Texas, United States

Abstract. Line-edge roughness (LER) has important impacts on the quality of semiconductor device performance, and power spectrum estimates are useful tools in characterizing it. These estimates are often obtained by taking measurements of many lines and averaging a classical power spectrum estimate from each one. While this approach reduces the uncertainty of the estimates, there are disadvantages to the collection of many measurements. We propose techniques with widespread application in other fields that simultaneously reduce data requirements and the uncertainty of LER power spectrum estimates over current approaches at the price of computational complexity. Multitaper spectral analysis uses an orthogonal collection of data windowing functions or tapers to obtain a set of approximately statistically independent spectrum estimates. The Welch overlapped segment averaging is an earlier approach to reduce the uncertainty of power spectrum estimates. There are known techniques to evaluate the uncertainty of power spectrum estimates. We simulate random rough lines using the Thorsos method. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMM.16.3.034001]

Keywords: power spectrum estimation; spectral leakage; line-edge roughness.

Paper 17045P received Apr. 7, 2017; accepted for publication Jul. 24, 2017; published online Aug. 22, 2017.

1 Introduction

Line-edge roughness (LER) is known to be a crucial factor in the yield of integrated circuit manufacturing (see, e.g., Ref. 1, pp. 82–92) and has been analyzed as part of the study of emerging patterning processes and techniques.^{2,3} For example, in extreme ultraviolet lithography, the LER resulting from photon shot noise and resist chemistry remains a concern for the use of this technology in high-volume manufacturing despite the ongoing development of resolution enhancement techniques.³ The critical-dimension scanning electron microscope (CD-SEM) is the standard metrology tool to obtain measurements for LER analysis (Ref. 4, pp. 109, 114), but it is also possible to use critical-dimension atomic force microscopes (CD-AFM) for this purpose (Ref. 4, p. 130). The standard deviation of sampled edge positions of a line does not completely describe its LER.⁵ The power spectrum is an important and more detailed LER characterization metric,^{6,7} which is related to transistor performance^{8,9} and is used in process monitoring.² However, it can be challenging to estimate the power spectrum of LER. First, the power spectrum estimates of LER are based on a finite number of sampled edge positions and they, therefore, suffer from a phenomenon called leakage where the spectral component at one frequency produces a spreading to other frequencies.¹⁰ References 11–13 have discussed the need to use data windowing functions or tapers¹⁴ to reduce the spectral leakage of LER spectrum estimates. Second, there is consensus in recent papers^{7,12,15–18} that it is necessary to reduce variance, i.e., the uncertainty of LER power spectrum estimates; in the newest of these papers,^{16–18} the authors additionally defined and considered different scenarios in the low-, middle-, and high-frequency regions.

The classical approach of LER power spectrum estimation is based on the periodogram,^{10,14} which is conceptually and computationally simple. However, the price for this simplicity is the relatively high uncertainty that can grow when the image noise increases. It is possible to reduce image noise by increasing the dose of CD-SEM or using CD-AFM, and one could decrease the power spectrum estimation uncertainty by averaging power spectrum estimates over many lines.^{15,17} However, these techniques may not be efficient because CD-SEM may cause sample damage as the dose level increases (see, e.g., Ref. 19 and Ref. 4, p. 42) and because CD-AFM is slow and has other disadvantages that affect measurement.^{4,19} We will consider two approaches from the more recent literature on spectral analysis, and we will evaluate their abilities to reduce the power spectrum estimation uncertainty given less measured data.

The remainder of the paper is organized as follows: in Sec. 2, we will discuss the Welch overlapped segment averaging (WOSA) spectrum estimate.²⁰ In Sec. 3, we will touch upon multitaper spectral analysis.^{10,21,22} In Sec. 4, we will write about our results on simulated random rough lines, and in the last section, we will offer concluding remarks.

2 Welch Overlapped Segment Averaging Spectrum Estimates

The power spectrum estimate of a random rough line is a random process, where the component at each frequency is a random variable. The bias of a power spectrum estimate refers to the difference between the mean values of these random variables and the true values of the power spectrum. Power spectrum estimates are most valuable when these random variables have low variance and low bias. The Bartlett method²³ is widely used to reduce variance in periodogram estimation by dividing the data samples into M disjoint

*Address all correspondence to: Yao Luo, E-mail: luoyao61@email.tamu.edu

segments of equal length and choosing the spectrum estimate to be the average of the periodograms associated with the M segments. In the LER literature, Ref. 15 suggests using a resampling method to reduce statistical noise, where assembled long lines based on random subsets of the available data are generated for power spectrum estimation. The modified periodogram¹⁰ has been recently introduced in the LER literature¹¹⁻¹³ to reduce the bias due to leakage. The WOSA spectrum estimate²⁰ is based on the modified periodogram and extends the Bartlett method to reduce both the bias and the variance in power spectrum estimation.

Suppose we sample a rough edge at N points with distance Δ between successive points. Let $w_k, k \in \{0, 1, \dots, N-1\}$ denote the values of the sampled edge positions and let \bar{w} be the mean of the edge positions. The sequence of real constants $h_k, k \in \{0, 1, \dots, N-1\}$, is called a data taper or window. The modified periodogram is defined as

$$\hat{S}_{\text{modified}}(f) = \Delta \left| \sum_{k=0}^{N-1} h_k (w_k - \bar{w}) e^{-i2\pi f \Delta k} \right|^2. \quad (1)$$

We require $\{h_k\}$ be a normalized taper with

$$\sum_{k=0}^{N-1} h_k^2 = 1 \quad (2)$$

for asymptotically unbiased power spectrum estimation. When $h_k = \frac{1}{\sqrt{N}}$ for all k , the window is called “rectangular” and Eq. (1) describes the periodogram. Let $S(f), |f| \leq 1/(2\Delta)$, denote the true power spectrum of interest, and let $H(f) = \Delta \left| \sum_{k=0}^{N-1} h_k e^{-i2\pi f \Delta k} \right|^2$. The mean of the modified periodogram estimator is¹⁰

$$E[\hat{S}_{\text{modified}}(f)] = \int_{-1/(2\Delta)}^{1/(2\Delta)} H(f - f') S(f') df'. \quad (3)$$

The convolution in Eq. (3) shows a potential bias called leakage, where the power at one frequency can spread to the others. This bias appears when a finite-sized sampling window is used for estimating the power spectrum of a process with a relatively high dynamic range and is reduced by a modified periodogram using a nonrectangular taper with a large central lobe, such as the Welch window.^{11,13} The k 'th element in the (unnormalized) Welch taper with size N is defined by

$$y(k, N) = 1 - \left(\frac{2k}{N} - 1 \right)^2, \quad k \in \{0, 1, \dots, N-1\}. \quad (4)$$

We normalize tapers when use them in power spectrum estimates.

The WOSA method divides the data samples into overlapping segments and uses the average of the modified periodograms associated with these segments. The nonrectangular windows reduce the bias due to spectrum leakage but place less weights on the end samples in each individual segment. The overlapping segments compensate for this effect and preserve the autocovariance information between adjacent segments,¹⁰ and therefore, reduce the variance in power spectrum estimation. The WOSA spectrum estimation has a

hardware implementation²⁴ and is thus potentially useful for in-line metrology.

To describe the WOSA spectrum estimate, we introduce some additional notation in Fig. 1. The T segments each have N_{seg} points and share a fraction r of their data points with the next segment, which leads to an “offset” of $D = N_{\text{seg}} \times (1 - r)$ points between neighboring segments. The number of overlapping segments T used in WOSA is

$$T = \frac{N - N_{\text{seg}}}{N_{\text{seg}} \times (1 - r)} + 1. \quad (5)$$

It is desirable to have more segments to reduce the variance in power spectrum estimation for a fixed N . Thus, N_{seg} should be chosen according to the lowest frequency resolution that can be tolerated. The parameter r is generally set to 50% to achieve nearly maximum variance reduction. There is a more recent variation of WOSA that permits the circular overlap of segments.²⁵ Using the circularly overlapped segments can increase T to T' , where T' is defined as

$$T' = \frac{N}{N_{\text{seg}} \times (1 - r)}. \quad (6)$$

For WOSA, the elements of the t 'th segment are $w_{k+D \times (t-1)}$ and $k \in \{0, 1, \dots, N_{\text{seg}} - 1\}$. Let

$$g_k^{(t)} = y(k, N_{\text{seg}}), \quad k \in \{0, 1, \dots, N_{\text{seg}} - 1\}, \quad t \in \{1, 2, \dots, T\},$$

$$h_k^{(t)} = \frac{g_k^{(t)}}{\sqrt{\sum_{k'=0}^{N_{\text{seg}}-1} |g_{k'}^{(t)}|^2}}. \quad (7)$$

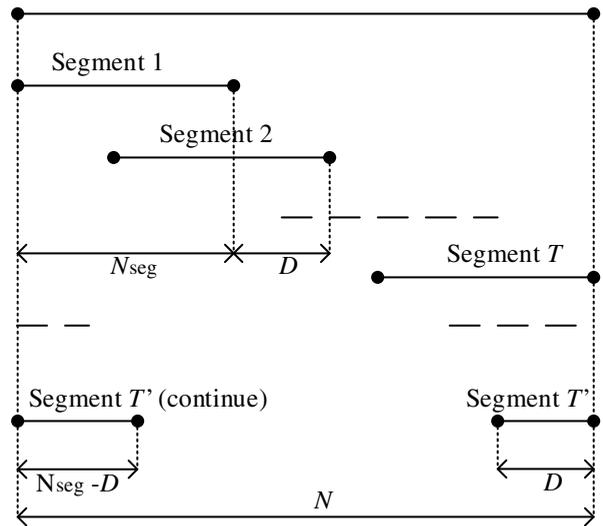


Fig. 1 T overlapped segments for WOSA and T' segments for WOSA via circular segments with offset D and segment size N_{seg} .

Then, the WOSA power spectrum estimate is

$$\hat{S}_{\text{WOSA}}(f) = \frac{1}{T} \sum_{t=1}^T \Delta \left| \sum_{k=0}^{N_{\text{seg}}-1} h_k^{(t)} [w_{k+D \times (t-1)} - \bar{w}] e^{-i2\pi f \Delta k} \right|^2. \quad (8)$$

For the WOSA variant with circular overlap, the elements of the t 'th segment are $w_{[k+D \times (t-1)] \bmod N}$. Thus, the first T segments are the same as those in WOSA. For $t > T$ suppose that the first v_t points come from the end of the line and the remaining $N_{\text{seg}} - v_t$ points come from the beginning of the line; here $v_t = N - \{[D \times (t-1)] \bmod N\}$. The tapers used for circular WOSA are

$$g_k^{(t)} = \begin{cases} y(k, N_{\text{seg}}), & 0 \leq k < N_{\text{seg}}, 1 \leq t \leq T \\ y(k, v_t), & 0 \leq k < v_t, t > T \\ y(k - v_t, N_{\text{seg}} - v_t), & v_t \leq k < N_{\text{seg}}, t > T \end{cases},$$

$$h_k^{(t)} = \frac{g_k^{(t)}}{\sqrt{\sum_{k'=0}^{N_{\text{seg}}-1} |g_{k'}^{(t)}|^2}}, \quad (9)$$

and the overall spectrum estimate is

$$\hat{S}_{\text{circular}}(f) = \frac{1}{T'} \sum_{t=1}^{T'} \Delta \left| \sum_{k=0}^{N_{\text{seg}}-1} h_k^{(t)} (w_{[k+D \times (t-1)] \bmod N} - \bar{w}) e^{-i2\pi f \Delta k} \right|^2. \quad (10)$$

3 Multitaper Spectrum Estimation

Multitaper methods²¹ are among the great advances in spectrum estimation. Multitaper methods have low uncertainty and are resistant to spectral leakage since they recover the information lost by single nonrectangular taper estimators through the use of a group of orthogonal tapers. As discussed in Ref. 10, multitaper methods work for various types of power spectra. They have widespread applications including neuroscience,²⁶ climate studies,²⁷ nuclear test-ban treaty verification,²⁸ and cognitive radio.²⁹ Multitaper spectral analysis uses an orthogonal collection of tapers on a finite sample of data to obtain a set of approximately statistically independent spectrum estimates. The overall spectrum estimate is either an average or a weighted average of the individual estimates. To provide more details about the earliest version of this technique, we introduce the following notation. Suppose we have the T orthogonal normalized tapers $h_k^{(t)}$, $k \in \{0, 1, \dots, N-1\}$, $t \in \{1, 2, \dots, T\}$. The basic estimator takes the arithmetic mean of the spectrum estimates associated with these tapers.

$$\hat{S}_{\text{multitaper}}(f) = \frac{1}{T} \sum_{t=1}^T \Delta \left| \sum_{k=0}^{N-1} h_k^{(t)} (w_k - \bar{w}) e^{-i2\pi f \Delta k} \right|^2. \quad (11)$$

To complete the specification of this estimate, we need to discuss the choice and number of orthogonal tapers.

Observe that the multitaper method is potentially useful for LER power spectrum estimation because it has some similarity to taking the average of modified periodogram spectrum estimates from a group of lines. The technique of

applying a group of orthogonal leakage-resistant tapers to a single line resembles the technique of applying a single-leakage resistant taper to a group of lines. Thus, the multitaper method may be able to reduce the uncertainty and leakage in LER spectrum estimation with a smaller group of lines.

There have been multiple classes of orthogonal tapers considered in the literature (see, e.g., Ref. 30). In his seminal paper, Thomson²¹ chose to use discrete prolate spheroidal sequences (DPSS), and we will begin by following his example. These orthogonal tapers continue to be popular because they are resistant to spectral leakage. They are also known as Slepian sequences because Slepian³¹ observed that they are the solution to the following famous time-frequency concentration problem:¹⁰ given a sampling frequency $1/\Delta$ and an "effective" bandwidth $W < 1/(2\Delta)$, find the sequence h_0, h_1, \dots, h_{N-1} with Fourier transform $H(f)$ defined over a continuous frequency domain $|f| \leq 1/(2\Delta)$ that maximizes λ given as

$$\lambda = \frac{\int_{-W}^W |H(f)|^2 df}{\int_{-1/(2\Delta)}^{1/(2\Delta)} |H(f)|^2 df}. \quad (12)$$

The first Slepian sequence corresponds to the largest λ . The second Slepian sequence maximizes λ among sequences orthogonal to the first Slepian sequence, and one can similarly construct an arbitrarily large set of Slepian sequences.^{21,31} There are many existing implementations for the construction of Slepian sequences. We use the R package multitaper 1.0–12 for Slepian sequence calculation that is based on a tridiagonal function method¹⁰ and LAPACK function calls.³²

The spectral leakage associated with using $\{h_k\}$ as a normalized taper for data sequence $\{w_k\}$ can be studied based on Eq. (12). As we discussed in Sec. 2, a taper with good leakage resistance should have a large central lobe within $|f| \leq W$.¹⁰ The first few Slepian sequences have good leakage resistance because they concentrate nearly all of the energy of $H(f)$ in the region $|f| \leq W$. How many orthogonal tapers can we select? The first few Slepian sequences obtained by optimizing Eq. (12) are known to have λ close to one, and they therefore offer good spectral leakage protection. The higher-order Slepian sequences may not help with spectral leakage. In practice, the number T of orthogonal tapers satisfies $T < 2NW$.²¹

Thomson²¹ also proposed a generalization of Eq. (11). The use of frequency-dependent weights $d_t(f)$, $t = 1, 2, \dots, T$ can reduce the bias from spectral leakage if the weights $\{d_t(f)\}$ are set to be close to 1 in the region where the spectrum is flat and are set to reduce the contribution from the higher-order tapers in the region where the spectrum has a large slope. The multitaper spectrum estimate with the weights $\{d_t(f)\}$ is given as

$$\hat{S}_{\text{multitaper}}(f) = \frac{1}{\sum_{t'=1}^T |d_{t'}(f)|^2} \sum_{t=1}^T |d_t(f)|^2 \Delta \left| \sum_{k=0}^{N-1} h_k^{(t)} (w_k - \bar{w}) e^{-i2\pi f \Delta k} \right|^2. \quad (13)$$

Riedel and Sidorenko²² proposed using the sinusoidal tapers for an alternate multitaper spectrum estimate. This

approach is interesting because the tapers have an analytic expression and because they approximate the solution to an optimization problem discussed by Papoulis³³ related to the bias error of a taper estimate as $N \rightarrow \infty$. The t 'th sinusoidal taper has the following expression:

$$h_k^{(t)} = \sqrt{\frac{2}{N+1}} \sin\left[\frac{\pi t(k+1)}{N+1}\right]. \quad (14)$$

One can again consider spectrum estimates based on an average or a weighted average of the individual single-tapered estimates.

The performance of Slepian sequences for this latter optimization problem on "local bias" depends on the choice of the effective bandwidth W . Increasing W generally increases the local bias but tends to reduce the variance from random errors since one can use more orthogonal tapers with λ value close to one.

4 Simulations

For our simulations we consider the periodogram, the modified periodogram using the Welch window, the original WOSA method with the Welch window and the variant with circular overlap²⁵ with the Welch window, Thomson's multitaper method using DPSS tapers with and without adaptive weights, and sinusoidal tapers without adaptive weights.

To study the performance of different spectrum estimators, we simulate random rough lines using the Thorsos method.^{34,35} We consider the K -correlation model or the Palasantzas power spectral density model³⁶

$$\text{PSD}(f) = \frac{\sqrt{\pi}\Gamma(\alpha + 0.5)}{\Gamma(\alpha)} \cdot \frac{2\sigma^2\xi}{[1 + (2\pi f\xi)^2]^{\alpha+0.5}}, \quad (15)$$

where σ represents LER, ξ is the correlation length, and α is the roughness (or Hurst) exponent. We follow Ref. 7 in choosing $\sigma = 1.5$ nm, $\xi = 25$ nm, and $\alpha = 0.75$, which are typical values observed in the experimental measurements. We generate 4096 or 20,480 edge positions for each line with the sampling distance $\Delta = 1$ nm and zero mean position \bar{w} . To incorporate some of the effects of leakage and SEM noise in our simulation, we use the middle 2048 edge positions for each line. These positions are then corrupted by additive white Gaussian noise.^{11,37-39} To evaluate the performance of LER power spectrum estimation, we consider confidence intervals, i.e., error bars, relative bias [see Eq. (18)], and spectrum concentration [see Eq. (12)]. We consider error bars to be the most important among these metrics for the application of LER metrology because it is related to random error. One generally can only analyze a relatively small number of lines where random errors are significant. For error bars and relative bias, we offer average results over all frequencies with $\sigma_{\text{noise}} = 0$ nm. For error bars, we also offer results under four other noise levels $\sigma_{\text{noise}} \in \{0.5, 1.0, 1.5, 2.0\}$ nm and for three frequency regions specified in a recent paper by Levi et al.,¹⁸ which defines the low-frequency region as all frequencies below $1/200$ nm⁻¹, the middle-frequency region as all frequencies between $1/200$ and $1/20$ nm⁻¹ and the high-frequency region as all frequencies above

$1/20$ nm⁻¹. For spectrum concentration, we offer results for the rectangular, Welch, DPSS, and sinusoidal tapers.

The spectrum estimate of a random rough line can be treated as a random process. For example, the periodogram or modified periodogram at each frequency f follows a χ^2 distribution for rough lines that follow a normal distribution, and this is a common approximation for the other spectrum estimates that we consider. We choose error bars based on an estimate of the 95% confidence interval at each frequency point.¹³ The lower bound $\text{CI}_{\text{lower}}(f)$ and upper bound $\text{CI}_{\text{upper}}(f)$ of the 95% confidence interval can be estimated by

$$\text{CI}_{\text{lower}}(f) = \frac{\nu\hat{S}(f)}{\chi_{1-0.025}^2(\nu)}, \quad (16)$$

$$\text{CI}_{\text{upper}}(f) = \frac{\nu\hat{S}(f)}{\chi_{0.025}^2(\nu)}, \quad (17)$$

where ν denotes the appropriate degrees of freedom (d.o.f.) for the χ^2 distribution. For power spectrum estimates obtained by averaging the periodograms or modified periodograms corresponding to M rough lines it is known that

Table 1 Power spectrum estimates when $\sigma_{\text{noise}} = 0.5$ nm. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

No. of lines per estimate	Coverage rate /d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate /d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Periodogram (left) and modified periodogram with single Welch taper (right)				
26	0.9489/52	2.12/52	0.9499/52	2.12/52
Multisegment: classical WOSA with three segments per line				
14	0.9320/84	1.66/84	0.9743/56	2.09/56
Multisegment: circular WOSA with four segments per line				
14	0.9142/112	1.43/112	0.9697/70	1.84/70
Multisegment: classical WOSA with three segments per line				
26	0.9315/156	1.20/156	0.9742/104	1.49/104
Multisegment: circular WOSA with four segments per line				
26	0.9126/208	1.03/208	0.9693/130	1.32/130
Multitaper: six DPSS tapers per line and adaptive weights				
14	0.9493/168	1.12/168	0.9689/138	1.24/138
Multitaper: six sinusoidal tapers per line and nonadaptive weights				
14	0.9490/168	1.12/168	0.9687/138	1.24/138

$\nu \approx 2M$. For power spectrum estimates determined by averaging M classical WOSA method estimates or M circular WOSA method estimates with T segments per line or by averaging M multitaper estimates with T tapers per line, it is known that $\nu \leq 2MT$. It is of interest to determine the empirical fraction of instances where a true power spectrum value $S(f)$ falls into the estimated confidence interval $[CI_{\text{lower}}(f), CI_{\text{upper}}(f)]$; we call this fraction the coverage rate, and in theory, it should be 0.95. For the periodogram and modified periodogram, we choose $\nu = 2M$. For classical WOSA, circular WOSA, and the multitaper methods, we initially set $\nu = 2MT$ and subsequently lower this to match or surpass the average coverage rate of the periodogram or modified periodogram in the three different frequency regions specified earlier. By adjusting ν , we simultaneously change the widths of the estimated confidence intervals and the coverage rates.

In Table 1, we compile the average coverage rate and confidence interval width over all frequency points for eight

power spectrum estimation methods when $\sigma_{\text{noise}} = 0.5$ nm. Each simulated rough line consists of 2048 points and comes from a longer line with 4096 points with sampling distance $\Delta = 1$ nm. Each power spectrum estimate is calculated from the average of the individual spectrum estimates from 26 or 14 lines. We do 1000 simulations per technique. The appendix provides analogous results for the low-, middle-, and high-frequency regions specified earlier. The multitaper methods consistently offer the best error bar performance among the eight methods. The WOSA variant with circular overlap and the classical WOSA methods are the next most effective techniques. Finally, the periodogram and modified periodogram using a single taper provided the worst error bar performances among the eight techniques.

As we mentioned earlier, the bias and leakage of spectrum estimates are also of interest, and different techniques offer different trade-offs among performance metrics. Therefore, we will also mention the relative bias and the spectrum

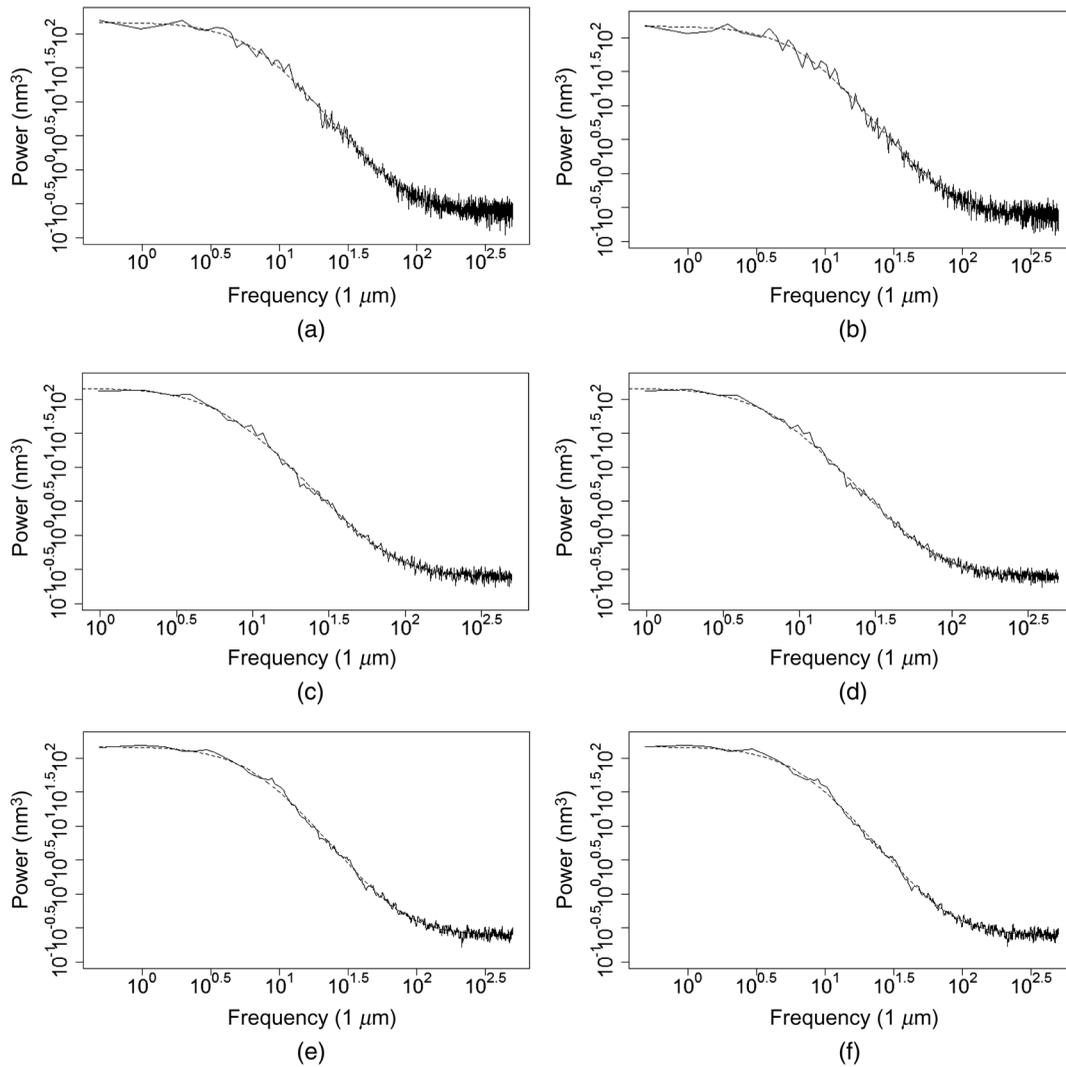


Fig. 2 Spectrum estimates: (a) periodogram and (b) modified periodogram with the Welch taper over 26 lines; (c, d) WOSA with three overlapped segments and with four circularly overlapped segments over 26 lines; (e) multitaper with DPSS tapers and adaptive weights, and (f) sinusoidal tapers without adaptive weights over 14 lines. Here $\sigma_{\text{noise}} = 0.5$ nm. The dashed and curved lines are the K -correlation model used in the Thorsos method to generate random roughness. The solid and curved lines are the power spectrum estimates.

Table 2 Overall results from 1000 simulations with one spectrum estimate coming from 26 lines for the periodogram and the modified periodogram with a Welch window. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each 95% confidence interval is computed in terms of a χ^2 distribution with 52 d.o.f. as discussed in Ref. 13.

σ_{noise}	SNR	Coverage rate of periodogram	Average width of confidence interval (nm ³)	Coverage rate of Welch window	Average width of confidence interval (nm ³)
0		0.9156	1.92	0.9498	1.91
0.5	9.542	0.9489	2.12	0.9499	2.12
1.0	3.522	0.9489	2.74	0.9496	2.74
1.5	0	0.9496	3.78	0.9499	3.78
2.0	-2.499	0.9496	5.22	0.9495	5.22

Table 3 Overall results from 1000 simulations with one spectrum estimate coming from 14 lines. The results for the classical WOSA method assume Welch windows, $r = 50\%$, and three segments per line. For the variant with circular overlap, assume Welch windows, $r = 50\%$ and four segments per line. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

σ_{noise} (nm)	SNR (dB)	Coverage rate d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Classical WOSA					
0		0.9739/56	1.89/56	0.9310/84	1.51/84
0.5	9.542	0.9743/56	2.09/56	0.9320/84	1.66/84
1	3.522	0.9739/56	2.68/56	0.9313/84	2.14/84
1.5	0	0.9743/56	3.67/56	0.9315/84	2.93/84
2	-2.499	0.9744/56	5.06/56	0.9317/84	4.04/84
Circular WOSA					
0		0.9692/70	1.67/70	0.9133/112	1.29/112
0.5	9.542	0.9697/70	1.84/70	0.9142/112	1.43/112
1	3.522	0.9694/70	2.37/70	0.9135/112	1.84/112
1.5	0	0.9697/70	3.24/70	0.9126/112	2.51/112
2	-2.499	0.9698/70	4.46/70	0.9132/112	3.46/112

concentration associated with spectrum estimation. We define $\epsilon_{\text{bias}}(f)$ as follows:

$$\epsilon_{\text{bias}}(f) = \frac{\hat{S}(f) - S(f)}{S(f)}. \quad (18)$$

This expression is related to the ϵ_{alias} and $\epsilon_{\text{leakage}}$ parameters defined in Ref. 11. The spectrum concentration is evaluated using Eq. (12) in terms of for a given bandwidth W . To evaluate ϵ_{bias} , we obtain each spectrum estimate from the average of 10,000 individual spectrum estimates with 2048 points per line and sampling distance $\Delta = 1$ nm. The 2048 points of each line are taken from a longer line with 4096 points generated using the Thorsos method applied to the Palasantzas model. We obtain λ for each taper using simulated sequences of 20,480 points with the middle $N = 2048$ points assigned with taper values and the remaining points assigned a value of zero.

In Fig. 2, we illustrate six power spectrum estimates assuming $\sigma_{\text{noise}} = 0.5$ nm. We will next offer more details about the simulations for the various power spectrum estimates.

For the simulations using the periodogram and the modified periodogram with a Welch window every spectrum

Table 4 Overall results from 1000 simulations with one spectrum estimate coming from 26 lines. The results for the classical WOSA method assume Welch windows, $r = 50\%$, and three segments per line. For the variant with circular overlap, assume Welch windows, $r = 50\%$, and four segments per line. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

σ_{noise} (nm)	SNR (dB)	Coverage rate d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Classical WOSA					
0		0.9742/104	1.35/104	0.9310/156	1.09/156
0.5	9.542	0.9742/104	1.49/104	0.9315/156	1.20/156
1	3.522	0.9737/104	1.90/104	0.9305/156	1.54/156
1.5	0	0.9739/104	2.62/104	0.9311/156	2.11/156
2	-2.499	0.9744/104	3.60/104	0.9317/156	2.90/156
Circular WOSA					
0		0.9690/130	1.19/130	0.9135/208	0.93/208
0.5	9.542	0.9693/130	1.32/130	0.9126/208	1.03/208
1	3.522	0.9693/130	1.69/130	0.9123/208	1.32/208
1.5	0	0.9694/130	2.32/130	0.9131/208	1.82/208
2	-2.499	0.9696/130	3.20/130	0.9128/208	2.50/208

Table 5 Average simulation values of ϵ_{bias} from 10,000 randomly generated lines for the periodogram, the modified periodogram with the Welch taper, the classical WOSA method with three segments per line, and the variant with circular overlap and four segments per line when $\sigma_{\text{noise}} = 0$ nm. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷

	Periodogram	Single Welch taper	Classical WOSA	Circular WOSA
ϵ_{bias}	0.0836716	-0.0000699	0.0000342	0.0001851

estimate comes from the average of 26 lines; we consider the 95% confidence intervals based on a chi-squared distribution with 52 d.o.f. We do 1000 simulations per technique. We report the average widths of confidence intervals and coverage rates over all frequency points in Table 2. The estimated confidence intervals based on the modified periodogram with the Welch window are, on average, nearly equal in width as those based on the periodogram, but the periodogram tends to have worse coverage rates because of its higher potential of spectral leakage.

The simulation values of ϵ_{bias} for the periodogram and the modified periodogram with the Welch taper are $\epsilon_{\text{bias}} = 0.0836716$ and $\epsilon_{\text{bias}} = -0.0000699$, respectively. The simulation values of λ for the rectangular taper in a periodogram and the Welch single taper in a modified periodogram assuming $N = 2048$ and $W = 4/2048$ nm⁻¹ are $\lambda = 0.974749$ and $\lambda = 0.999627$, respectively.

We next report results on the WOSA method and the variant with circular overlap. We choose the Welch taper and $r = 50\%$. Every segment has $N_{\text{seg}} = 1024$ points. We average over 14 lines (see Table 3) or 26 lines (see Table 4) to obtain one spectrum estimate. As discussed in Eqs. (5) and

(6), for a line of $N = 2048$ points, we use three segments per line for the classical WOSA method and four segments per line for the variant with circular overlap. The simulations with the WOSA methods consistently attain better performance than those for the single-taper spectrum estimates. As expected, we attain better performance by averaging over 26 lines than by averaging over 14 lines. The WOSA method with circular overlap is better than the classical WOSA method in terms of confidence interval widths and coverage rates. However, we will see that the classical WOSA method has a smaller bias than the circular variation for the simulated data we consider.

We report the simulation values of ϵ_{bias} for the classical WOSA methods and the variant with circular overlap in Table 5. The results for the periodogram and the single Welch taper are also listed here for comparison. While the WOSA methods have better variance than the periodogram and the modified periodogram with a single Welch taper for a line with the same number of samples and distance between samples, the frequency resolution of WOSA methods is worse because of the decreased length of individual segments.

We conclude this section with some remarks on the multitaper methods. The choices of NW and the number of tapers used are not unique. In examining error bar performance, we use six NW = 4 DPSS tapers and six sinusoidal tapers (see Fig. 3). We produce one spectrum estimate by averaging individual estimates from 14 lines; we initially consider 168 d.o.f and subsequently lower this by 20 or 30 to match or surpass the coverage rates of the first two techniques. The results are summarized in Table 6 and demonstrate the superiority of the multitaper method over the periodogram, the modified periodogram with a Welch window, and the WOSA methods.

The effective bandwidth and the number of tapers can be adjusted to obtain different trade-offs between bias and

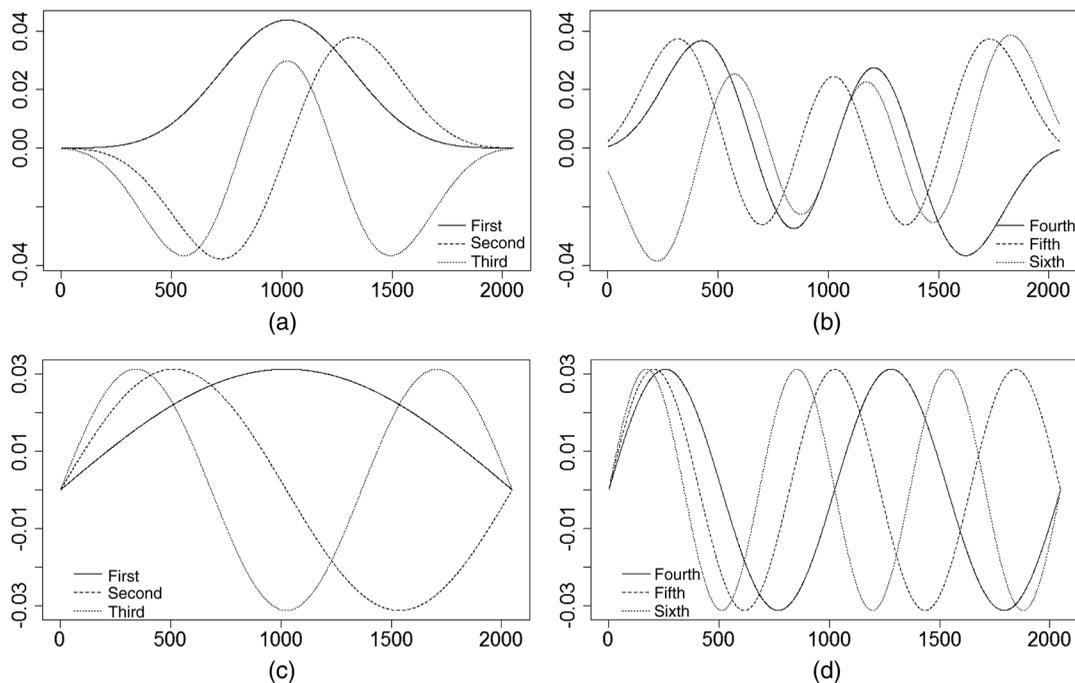


Fig. 3 (a) The first to third and (b) the fourth to sixth lowest order Slepian sequences of length 2048 with NW = 4. (c) The first to third and (d) the fourth to sixth lowest order sinusoidal tapers of length 2048.

Table 6 Overall results from 1000 simulations with one spectrum estimate coming from 14 lines for the multitaper method with Slepian tapers and sinusoidal tapers. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

σ_{noise} (nm)	SNR (dB)	Coverage rate d.o.f. = 168	Average width of confidence interval (nm ³)/d.o.f. = 168	Coverage rate/d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Slepian (DPSS) tapers					
0		0.9359	1.01	0.9516/148	1.08/148
0.5	9.542	0.9493	1.12	0.9689/138	1.24/138
1.0	3.522	0.9486	1.45	0.9684/138	1.61/138
1.5	0	0.9492	2.00	0.9687/138	2.22/138
2.0	-2.499	0.9487	2.77	0.9686/138	3.07/138
Sinusoidal tapers					
0		0.9490	1.01	0.9625/148	1.08/148
0.5	9.542	0.9490	1.12	0.9687/138	1.24/138
1	3.522	0.9497	1.45	0.9691/138	1.61/138
1.5	0	0.9495	2.00	0.9689/138	2.21/138
2	-2.499	0.9488	2.76	0.9683/138	3.06/138

spectrum concentration. Table 7 lists the average simulation value of ϵ_{bias} over all frequencies for the sinusoidal tapers and the DPSS tapers. There is a tendency for ϵ_{bias} to increase with the number of tapers, but we have seen that the multitaper methods reduce the variance of the estimates when one has access to relatively few lines. The sinusoidal tapers have

Table 7 Average simulation values of ϵ_{bias} from 10,000 randomly generated lines for the DPSS tapers and the sinusoidal tapers when $\sigma_{\text{noise}} = 0$ nm. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ The corresponding results for the periodogram and the modified periodogram with a single Welch taper are $\epsilon_{\text{bias}} = 0.0836716$ and $\epsilon_{\text{bias}} = -0.0000699$, respectively.

T	Sinusoidal	DPSS (NW = 2)	DPSS (NW = 4)	DPSS (NW = 4, adaptive)	DPSS (NW = 8)
1	-0.0000686	0.0000004	-0.0000724	-0.0000724	-0.0001705
2	-0.0000087	0.0011432	0.0002295	0.0002295	0.0004815
3	0.0000186	0.0111442	0.0002141	0.0002140	0.0006259
4	0.0002364	0.0535332	0.0003480	0.0003240	0.0008320
5	0.0003881	0.1404533	0.0006003	-0.0011971	0.0010770
6	0.0006640	0.2512091	0.0026047	-0.0035863	0.0013110
7	0.0008845	0.3641771	0.0142636	-0.0038643	0.0014573

smaller ϵ_{bias} than the DPSS tapers since the sinusoidal tapers approximate the minimum bias tapers.

Table 8 lists the spectrum concentration results in terms of λ for the first seven NW = 4 DPSS tapers and the first seven sinusoidal tapers with $N = 2048$, $W = 4/2048 \text{ nm}^{-1}$, and sampling distance $\Delta = 1$ nm. As expected, the DPSS tapers offer better results because they are obtained from the original spectrum concentration problem.

5 Conclusions

Most of the metrology literature that mentions power spectrum estimation refers only to the periodogram, which is over a century old. We propose using the more modern multitaper and multisegment spectral estimation techniques. For LER metrology over a relatively small group of lines, we assess the effectiveness of a spectrum estimate by the widths of confidence intervals and by the experimental coverage rates. For a broader discussion of the trade-offs among different techniques, we also considered the relative bias and spectrum concentration of the various estimates. We investigated the performance of the periodogram, the modified periodogram with the Welch taper, the Welch's overlapped segment averaging method and the variant with circular overlap with the Welch taper, and the multitaper methods using DPSS tapers with adaptive weights and sinusoidal tapers without adaptive weights for the spectrum estimation of random rough lines at five different noise levels. The multitaper methods offer the smallest average error bars among these techniques at a given coverage rate while the Welch's overlapped segment averaging method is not quite as effective but may be a better candidate for in-line metrology because of an existing hardware implementation. These results applied not only when we consider an average performance over all frequencies but also in the low-, the middle-, and the high-frequency ranges. The average widths of confidence intervals invariably increased with increasing noise, so improvements in power spectrum estimation may be even more important at high noise levels. In recent work, we used the multitaper method in an attempt to reduce metrology errors in LER estimation from simulated low-dose SEM images.⁴⁰ For a future direction of research, we observe that the WOSA variant with circular overlap is potentially interesting for estimating the

Table 8 The simulation values of λ for the first seven $NW = 4$ DPSS tapers and the first seven sinusoidal tapers with $N = 2048$, $W = 4/2048 \text{ nm}^{-1}$, and sampling distance $\Delta = 1 \text{ nm}$. The corresponding values of λ for the rectangular window in a periodogram and the Welch single taper in a modified periodogram under same setting are $\lambda = 0.974749$ and $\lambda = 0.999627$, respectively.

t	1	2	3	4	5	6	7
DPSS	1.000000	1.000000	0.999999	0.999973	0.999520	0.993893	0.947127
Sinusoidal	0.999747	0.998883	0.997332	0.994229	0.989002	0.976673	0.937289

power spectrum of contact edge roughness since the edge points of a contact hole have approximately a circular shape, i.e., the first and last sampled edge positions are close in space.⁴¹

Appendix A: Simulations for the Low-Frequency Region

In Table 9, we compile the average coverage rate and confidence interval width over points in the low-frequency region for eight power spectrum estimation methods when $\sigma_{\text{noise}} = 0.5 \text{ nm}$. The multitaper methods offer the best error

Table 9 Power spectrum estimates when $\sigma_{\text{noise}} = 0.5 \text{ nm}$. The low-frequency region is defined as all frequencies below $1/200 \text{ nm}^{-1}$. The parameters of the Palasantzas model are $\text{LER} = 1.5 \text{ nm}$, correlation length = 25 nm , and roughness exponent = 0.75 .⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

No. of lines per estimate	Coverage rate/d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate/d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Periodogram (left) and modified periodogram with single Welch taper (right)				
26	0.9370/52	102.06/52	0.9361/52	102.97/52
Multisegment: classical WOSA with three segments per line				
14	0.9070/84	80.39/84	0.9550/56	100.69/56
Multisegment: circular WOSA with four segments per line				
14	0.8888/112	68.65/112	0.9534/70	88.51/70
Multisegment: classical WOSA with three segments per line				
26	0.9066/156	57.77/156	0.9580/104	71.61/104
Multisegment: circular WOSA with four segments per line				
26	0.8840/208	49.72/208	0.9474/130	63.42/130
Multitaper: six DPSS tapers per line and adaptive weights				
14	0.9158/168	53.68/168	0.9434/138	59.52/138
Multitaper: six sinusoidal tapers per line and nonadaptive weights				
14	0.9187/168	53.57/168	0.9444/138	59.40/138

bar performance among the eight methods in the low-frequency region.

Appendix B: Simulations for the Middle-Frequency Region

In Table 10, we compile the average coverage rate and confidence interval width over points in the middle-frequency region for eight power spectrum estimation methods when $\sigma_{\text{noise}} = 0.5 \text{ nm}$. The multitaper methods offer the best error bar performance among the eight methods in the middle-frequency region.

Table 10 Power spectrum estimates when $\sigma_{\text{noise}} = 0.5 \text{ nm}$. The middle-frequency region is defined as all frequencies between $1/200$ and $1/20 \text{ nm}^{-1}$. The parameters of the Palasantzas model are $\text{LER} = 1.5 \text{ nm}$, correlation length = 25 nm , and roughness exponent = 0.75 .⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

No. of lines per estimate	Coverage rate/d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate/d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Periodogram (left) and modified periodogram with single Welch taper (right)				
26	0.9471/52	9.96/52	0.9496/52	9.89/52
Multisegment: classical WOSA with three segments per line				
14	0.9320/84	7.85/84	0.9737/56	9.83/56
Multisegment: circular WOSA with four segments per line				
14	0.9139/112	6.75/112	0.9699/70	8.70/70
Multisegment: classical WOSA with three segments per line				
26	0.9323/156	5.67/156	0.9743/104	7.03/104
Multisegment: circular WOSA with four segments per line				
26	0.9133/208	4.88/208	0.9695/130	6.25/130
Multitaper: six DPSS tapers per line and adaptive weights				
14	0.9459/168	5.30/168	0.9666/138	5.87/138
Multitaper: six sinusoidal tapers per line and nonadaptive weights				
14	0.9463/168	5.28/168	0.9665/138	5.86/138

Table 11 Power spectrum estimates when $\sigma_{\text{noise}} = 0.5$ nm. The high-frequency region is defined as all frequencies above $1/20$ nm⁻¹. The parameters of the Palasantzas model are LER = 1.5 nm, correlation length = 25 nm, and roughness exponent = 0.75.⁷ Each confidence interval is computed in terms of a χ^2 distribution as discussed in Ref. 13. The d.o.f. are reported within the table.

No. of Lines per estimate	Coverage rate /d.o.f.	Average width of confidence interval (nm ³)/d.o.f.	Coverage rate /d.o.f.	Average width of confidence interval (nm ³)/d.o.f.
Periodogram (left) and modified periodogram with single Welch taper (right)				
26	0.9492/52	0.26/52	0.9501/52	0.26/52
Multisegment: classical WOSA with three segments per line				
14	0.9323/84	0.20/84	0.9746/56	0.25/56
Multisegment: circular WOSA with four segments per line				
14	0.9145/112	0.17/112	0.9699/70	0.22/70
Multisegment: classical WOSA with three segments per line				
26	0.9317/156	0.14/156	0.9743/104	0.18/104
Multisegment: circular WOSA with four segments per line				
26	0.9129/208	0.12/208	0.9695/130	0.16/130
Multitaper: six DPSS tapers per line and adaptive weights				
14	0.9501/168	0.14/168	0.9694/138	0.15/138
Multitaper: six sinusoidal tapers per line and nonadaptive weights				
14	0.9496/168	0.14/168	0.9692/138	0.15/138

Appendix C: Simulations for the High-Frequency Region

In Table 11, we compile the average coverage rate and confidence interval width over points in the high-frequency region for eight power spectrum estimation methods when $\sigma_{\text{noise}} = 0.5$ nm. The multitaper methods offer the best error bar performance among the eight methods in the high-frequency region.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF), Award No. ECCS-1201994. The second author is grateful to Y. Borodovsky and C. A. Mack for helpful conversations and to T. Groves for various correspondences.

References

- H. J. Levinson, *Principles of Lithography*, Vol. 198, SPIE Press, Bellingham, Washington (2010).
- L. Sun et al., "Line edge roughness frequency analysis for SAQP process," *Proc. SPIE* **9780**, 97801S (2016).
- R.-H. Kim et al., "Application of EUV resolution enhancement techniques (RET) to optimize and extend single exposure bi-directional patterning for 7 nm and beyond logic designs," *Proc. SPIE* **9776**, 97761R (2016).
- B. Su, E. Solecky, and A. Vaid, *Introduction to Metrology Applications in IC Manufacturing*, Vol. T101, SPIE Press, Bellingham, Washington (2015).
- V. Constantoudis et al., "Quantification of line-edge roughness of photoresists. II. Scaling and fractal analysis and the best roughness descriptors," *J. Vac. Sci. Technol.* **21**(3), 1019–1026 (2003).
- A. Hiraiwa and A. Nishida, "Discrete power spectrum of line width roughness," *J. Appl. Phys.* **106**(7), 074905 (2009).
- T. Verduin, P. Kruit, and C. W. Hagen, "Determination of line edge roughness in low-dose top-down scanning electron microscopy images," *J. Micro/Nanolithogr. MEMS MOEMS* **13**(3), 033009 (2014).
- E. Baravelli et al., "Impact of line-edge roughness on FinFET matching performance," *IEEE Trans. Electron Devices* **54**(9), 2466–2474 (2007).
- A. Asenov, S. Kayaz, and A. R. Brown, "Intrinsic parameter fluctuations in decanometer MOSFETs introduced by gate line edge roughness," *IEEE Trans. Electron Devices* **50**(5), 1254–1260 (2003).
- D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*, Cambridge University Press, Cambridge (1993).
- C. A. Mack, "Systematic errors in the measurement of power spectral density," *J. Micro/Nanolithogr. MEMS MOEMS* **12**(3), 033016 (2013).
- B. D. Bunday and C. A. Mack, "Influence of metrology error in measurement of line edge roughness power spectral density," *Proc. SPIE* **9050**, 90500G (2014).
- C. A. Mack, "More systematic errors in the measurement of power spectral density," *J. Micro/Nanolithogr. MEMS MOEMS* **14**(3), 033502 (2015).
- G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*, Holden-Day, San Francisco (1968).
- A. Hiraiwa and A. Nishida, "Spectral analysis of line edge and line-width roughness with long-range correlation," *J. Appl. Phys.* **108**(3), 034908 (2010).
- L. Sun et al., "Application of frequency domain line edge roughness characterization methodology in lithography," *Proc. SPIE* **9424**, 942404 (2015).
- E. Dupuy et al., "Spectral analysis of sidewall roughness during resist-core self-aligned double patterning integration," *J. Vac. Sci. Technol. B* **34**(5), 051807 (2016).
- S. Levi et al., "Edge roughness characterization of advanced patterning processes using power spectral density analysis (PSD)," *Proc. SPIE* **9782**, 97820I (2016).
- L. Azarnouche et al., "Unbiased line width roughness measurements with critical dimension scanning electron microscopy and critical dimension atomic force microscopy," *J. Appl. Phys.* **111**(8), 084318 (2012).
- P. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.* **15**(2), 70–73 (1967).
- D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE* **70**(9), 1055–1096 (1982).
- K. S. Riedel and A. Sidorenko, "Minimum bias multiple taper spectral estimation," *IEEE Trans. Signal Process.* **43**(1), 188–195 (1995).
- M. S. Bartlett, "Smoothing periodograms from time series with continuous spectra," *Nature* **161**(4096), 686–687 (1948).
- K. K. Parhi and M. Ayinala, "Low-complexity Welch power spectral density computation," *IEEE Trans. Circuits Syst. Regul. Pap.* **61**(1), 172–182 (2014).
- K. Barbé, R. Pintelon, and J. Schoukens, "Welch method revisited: non-parametric power spectrum estimation via circular overlap," *IEEE Trans. Signal Process.* **58**(2), 553–565 (2010).
- P. Mitra and H. Bokil, *Observed Brain Dynamics*, Oxford University Press, New York (2007).
- M. E. Mann and J. M. Lees, "Robust estimation of background noise and signal detection in climatic time series," *Clim. Change* **33**(3), 409–445 (1996).
- S. J. Gibbons, F. Ringdal, and T. Kværna, "Detection and characterization of seismic phases using continuous spectral estimation on incoherent and partially coherent arrays," *Geophys. J. Int.* **172**(1), 405–421 (2008).
- S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.* **23**(2), 201–220 (2005).
- I. K. Fodor and P. B. Stark, "Multitaper spectrum estimation for time series with gaps," *IEEE Trans. Signal Process.* **48**(12), 3472–3483 (2000).
- D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—V: the discrete case," *Bell Labs Tech. J.* **57**(5), 1371–1430 (1978).
- E. Anderson et al., *LAPACK Users' Guide*, SIAM, Philadelphia (1999).
- A. Papoulis, "Minimum-bias windows for high-resolution spectral estimates," *IEEE Trans. Inf. Theory* **19**(1), 9–12 (1973).
- E. I. Thorsos, "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.* **83**(1), 78–92 (1988).

35. C. A. Mack, "Generating random rough edges, surfaces, and volumes," *Appl. Opt.* **52**(7), 1472–1480 (2013).
36. G. Palasantzas, "Roughness spectrum and surface width of self-affine fractal surfaces via the K-correlation model," *Phys. Rev. B* **48**(19), 14472–14478 (1993).
37. B. D. Bunday et al., "Determination of optimal parameters for CD-SEM measurement of line-edge roughness," *Proc. SPIE* **5375**, 515 (2004).
38. J. S. Villarrubia and B. D. Bunday, "Unbiased estimation of linewidth roughness," *Proc. SPIE* **5752**, 480 (2005).
39. A. Hiraiwa and A. Nishida, "Image-noise effect on discrete power spectrum of line-edge and line-width roughness," *Jpn. J. Appl. Phys.* **50**, 016602 (2011).
40. Y. Luo and S. A. Savari, "Reduction of metrology error for line-edge roughness measurement from low-dose SEM images," in *61st Int. Conf. on Electron, Ion, and Photon Beam Technology and Nanofabrication (EIPBN)*, <http://eipbn.omnibooksonline.com> (2017).
41. V.-K. Murugesan-Kuppuswamy, V. Constantoudis, and E. Gogolides, "Contact edge roughness: characterization and modeling," *Microelectron. Eng.* **88**(8), 2492–2495 (2011).

Yao Luo is currently pursuing her PhD in electrical engineering at Texas A&M University. She received her BS degree in electronic science and technology from Southeast University, China, in 2013. Since then, she attended Texas A&M University and has been working under the supervision of Dr. Serap Savari. Her current research interests include statistical signal processing and data compression with applications in VLSI fabrication.

Serap A. Savari is on the faculty of Texas A&M University. She has served on the program committee of the annual Data Compression Conference since 2000. From 2002 to 2005, she was an associate editor for the *IEEE Transactions on Information Theory*. She has served on the program committees for several conferences and workshops in information theory, and she joined the program committee of the 33rd European Mask and Lithography Conference (EMLC 2017) in January.