Determination of Optimal Parameters for CD-SEM Measurement of Line Edge Roughness

Benjamin D. Bunday, Michael Bishop, and Donald McCormack
International SEMATECH, Austin, TX 78741
John S. Villarrubia, András E. Vladár, Ronald Dixson, Theodore Vorburger, and Ndubuisi G. Orji
National Institute of Standards and Technology,* Gaithersburg, MD 20899
John A. Allgair, Motorola assignee to International SEMATECH, Austin, TX 78741

ABSTRACT

The measurement of line-edge roughness (LER) has recently become a topic of concern in the litho-metrology community and the semiconductor industry as a whole. The Advanced Metrology Advisory Group (AMAG), a council composed of the chief metrologists from the International SEMATECH (ISMT) consortium's Member Companies and from the National Institute of Standards and Technology (NIST), has a project to investigate LER metrics and to direct the critical dimension scanning electron microscope (CD-SEM) supplier community towards a semiconductor industry-backed, standardized solution for implementation. The 2003 International Technology Roadmap for Semiconductors (ITRS) has included a new definition for roughness. The ITRS envisions root mean square measurements of edge and width roughness. There are other possible metrics, some of which are surveyed here. The ITRS envisions the root mean square measurements restricted to roughness wavelengths falling within a specified process-relevant range and with measurement repeatability better than a specified tolerance. This study addresses the measurement choices required to meet those specifications. An expression for the length of line that must be measured and the spacing of measurement positions along that length is derived. Noise in the image is shown to produce roughness measurement errors that have both random and nonrandom (i.e., bias) components. Measurements are reported on both UV resist and polycrystalline silicon in special test patterns with roughness typical for those materials. These measurements indicate that the sensitivity of a roughness measurement to noise depends importantly both on the choice of edge detection algorithm and the quality of the focus. Measurements are less sensitive to noise when a model-based or sigmoidal fit algorithm is used and when the images are in good focus. Using the measured roughness characteristics for UV resist lines and applying the ITRS requirements for the 90 nm technology node, the derived expression for sampling length and sampling interval implies that a length of line at least 8 times the node (i.e., 720 nm) must be measured at intervals of 7.5 nm or less.

Keywords: dimensional metrology, Line Edge Roughness (LER), Line Width Roughness (LWR), metrics, power spectral density (PSD), scanning electron microscopy (SEM).

1. INTRODUCTION

The sizes and shapes of MOS transistor gates define essential characteristics of the integrated circuit (IC). As gates shrink, even small departures from the intended shape and size may represent a significant fraction of the gate area. The phenomenon of variation of a linear feature's width (also called CD, for critical dimension) along its length is called linewidth roughness (LWR). The one edge equivalent, the phenomenon of the meandering of a single edge along its length, is called line edge roughness (LER).

Edge roughness needs to be measured and controlled because it has been observed to be detrimental to IC performance. [1] If a gate's width is not constant, the narrowest section of the channel will carry a disproportionate share of the current. Roughness is a statistical phenomenon; it can cause variations in channel length from transistor to transistor that

---

* Official contributions of the National Institute of Standards and Technology are not subject to copyright.

may in turn result in circuit timing issues. Off-state leakage currents and device drive currents are thought to be affected. [2] Since the polycrystalline silicon ("poly") gate acts as the mask for dopant implant, a rough edge will affect dopant distributions after diffusion. Roughness at wavelengths large compared to diffusion lengths affects the shape of the doped volume, while shorter wavelength roughness affects the dopant concentration gradient. Roughness metrology on resist is important because it is desirable to measure structures for screening purposes prior to etch. In turn, roughness in resist is thought to be caused by a multitude of variables, including aerial image fluctuations, resist material properties, acid diffusion, conditions during development, and possibly reticle roughness. [3] The extent to which roughness in resist transfers into silicon at etch probably depends upon the wavelength of the roughness, with short wavelengths likely transferring less than longer ones.

As some of the preceding examples indicate, there are good reasons to measure roughness wavelengths (or frequencies) as well as amplitudes. With the 2003 edition of the International Technology Roadmap for Semiconductors (ITRS) [4] metrology specifications are evolving in recognition of this. As the linewidth variation is the determining factor of roughness effect on product, the ITRS now specifies LWR over a window of spatial frequencies. Previous versions specified LER with no frequency window. The 2003 roadmap requires that the "LWR litho control," the amount of LWR tolerance in product for a given technology, be 8 % of the etched gate length (previous roadmap required 5 % of printed gate length). Precision (i.e., 3 standard deviation repeatability) is required to be no worse than 20 % of the manufacturing tolerance. Roadmap values for LWR control and required metrology precision can be seen in Fig. 1 below. The lithography roadmap definition for the LWR metric is 3 standard deviations (3σ) of total linewidth variation, evaluated along a distance that allows assessment of spatial wavelengths up to two times the technology node, while sampling the low-end spatial wavelengths down to a limit defined by x_j, the low-end-of-range of the drain extension found in the Thermal and Thin Film, Doping and Etching Technology Requirements Table in the ITRS. When LER is of concern, its metric is defined as LER=LWR/$\sqrt{2}$, assuming uncorrelated edges.

Technology Node	130nm	115nm	100nm	90nm	65nm	45nm	32nm	22nm	18nm
Year of Production	*2001*	*2002*	*2003*	*2004*	*2007*	*2010*	*2013*	*2016*	*2018*
DRAM 1/2 pitch [nm]	130	115	100	90	65	45	32	22	18
MPU printed gate length [nm]	90	75	65	53	35	25	18	13	10
MPU etched gate length [nm]	65	53	45	37	25	18	13	9	7
3σ LWR control, <8% of etched gate [nm]	5.2	4.2	3.6	3.0	2.0	1.44	1.04	0.72	0.56
Precision of LWR measurement [nm]	1.0	0.85	0.72	0.59	0.40	0.29	0.21	0.14	0.11
Precision of LER measurement [nm]	0.74	0.60	0.51	0.42	0.28	0.20	0.15	0.10	0.08
DRAM pitch [nm]	260	230	200	180	130	90	64	44	36
x_j (low end) [nm]	27	22	19	15	10	7	5	4	3
Length of Segment [nm]	260	230	200	180	130	90	64	44	36
Sampling Distance [nm]	13.5	11.0	9.5	7.5	5.0	3.5	2.5	2.0	1.5

Fig. 1. Selected sections of the 2003 International Roadmap for Semiconductors (ITRS) [4] for CD Metrology that are applicable to roughness. Note: the terms "Length of Segment" and Sampling Distance" here (tabulated from above ITRS quantities) are equivalent to "Evaluation Length" and "Sampling Interval" in traditional surface roughness standards [5].

Thus the range of spatial frequencies is $1/x_j$ to 1/pitch. The rationale for this range is as follows. The smallest spatial frequency (longest wavelength) is chosen to distinguish LWR from CD variation. If one considers a transistor gate that serially overlaps different active regions, the average gate length may be different over each active region. Changes in width from transistor to transistor are considered as CD variation, while changes in width within a single transistor are included in LWR. The upper end of the spatial frequency range (small wavelength) represents considerations of diffusion. Dopants will diffuse more under fast-varying roughness than under the rough parts of the gate.

In this study, we address some of the issues in the metrology of linewidth and line edge roughness. In Sec. 2 we discuss a number of metrics, paying particular attention to careful definition of the ones we use in this report, but also including a few metrics that we are not currently using, in order to illustrate some of the options. In Sec. 3 we address the question of how to choose the length of line to measure and the sampling distance in order for the measurement to meet the frequency range and precision requirements given in the ITRS. The roughness values so determined may still be subject

to measurement artifacts. In Sec. 4 we discuss the effect of noise on roughness determination. We have performed measurements on roughness test patterns. These measurements included studies of the repeatability of roughness measurements in the presence of varying amounts of noise and as a function of instrument focus. These measurements and the results are described in Sec. 5.

In this paper, we discuss roughness measurement in terms of SEM measurement, but the derivations and conclusions related to signal processing may also be applied to data obtained from other scanning or profiling techniques such as scanned probes. LER characterization techniques using atomic force microscopes have been reviewed by Orji et al.[9]

2. METRICS

It is conceptually useful to distinguish between a phenomenon and the various ways the phenomenon may be measured (various metrics). To take a familiar case by way of analogy, we observe that distributions of measured values have a center. There are a number of metrics, including the mean, the median, and the most likely value, that quantify the position of such a center. Similarly, we observe that fabricated structures exhibit random differences from their design. When the structures were intended to have straight edges and uniform width, we refer to this phenomenon as line edge or linewidth roughness (LER or LWR). For all of the metrics given below, we assume we have measured N edge positions, X_i, or widths, W_i, at measurement interval, Δ. We further assume that we subtract a line of best fit from the edge positions, or an average value from the widths, to produce edge residuals, x_i, or width residuals, w_i. (See an example in Fig. 4B.) That is,

$$x_i = X_i - (ai + b) \tag{1}$$

where a and b are determined by a linear least squares fit to the X_i vs. i curve. Similarly

$$w_i = W_i - \overline{W} \tag{2}$$

with $\overline{W}$ the average width defined in the usual way as $\sum W_i / N$.

Although major CD-SEM suppliers now offer various LER measurement solutions, definitions and sampling capabilities are not standardized. The measurement metrics typically offered are combinations of LWR, LER of a single edge, total LER (2 edges summed in quadrature), or range roughness, although leading edge tools are starting to offer more capability. The following list of metrics is not exhaustive, and not all are available on existing tools. We are particularly interested in the root mean square measures (Sec. 2.3) and the power spectral measures (Sec. 2.4) for the current report. The others are included in order to provide some of the flavor of what is possible.

2.1. Range

For a given sampled line segment the range is

$$\text{For LER:} \quad R_{\text{E}} = \max(x) - \min(x) \tag{3}$$

$$\text{For LWR:} \quad R_{\text{W}} = \max(w) - \min(w) \tag{4}$$

where the max and min operations are understood to cover all x_i and w_i within the specified segment. Here and throughout, the "E" and "W" subscripts refer to the edge and width measures.

2.2. Average roughness

$$\text{For LER:} \quad R_{\text{Ea}} = \frac{1}{N} \sum_{i=0}^{N-1} |x_i| \tag{5}$$

$$\text{For LWR: } R_{\text{Wa}} = \frac{1}{N}\sum_{i=0}^{N-1}\left|w_i\right| \tag{6}$$

These are arithmetic measures of average roughness, hence the "a" subscript. These definitions follow a published standard for surface roughness [5], though there is some question whether they might be better with N-2 and N-1 in their respective denominators.

2.3. Mean square roughness

$$\text{For LER: } R_{\text{E}q}^2 = \frac{1}{N-2}\sum_{i=0}^{N-1}x_i^{\,2} \tag{7}$$

$$\text{For LWR: } R_{\text{W}q}^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}w_i^{\,2} \tag{8}$$

These are quadratic measures of roughness, hence the "q" subscript. They are standard deviations, with the factors of *N*-1 and *N*-2 reflecting the number of degrees of freedom, as reduced by the 1 parameter (i.e., averaging) or 2 parameter (i.e., linear) fits in Eq. (2) and Eq. (1). The ITRS linewidth roughness definition appears to be essentially $3R_{\text{W}q}$.

Some instrument manufacturers supply a "total line edge roughness" metric.

$$R_T^2 = R_{\text{E}q\_\text{left}}^2 + R_{\text{E}q\_\text{right}}^2 \tag{9}$$

Here the "left" and "right" subscripts refer to the left and right edges of a line. For a line without taper the total roughness is related to $R_{\text{W}q}$ by $R_{\text{W}q}^2 = R_T^2 - 2cR_{\text{E}q\_\text{left}}R_{\text{E}q\_\text{right}}$ with c being the correlation coefficient between the two edges [6]. (See also Appendix I) If the left and right edges are uncorrelated, as is normally assumed, then the total roughness is equal to $R_{\text{W}q}$. However, we believe it is just as easy, and safer, to determine $R_{\text{W}q}$ from Eq. (8), which is valid even if there are nonzero correlations.

$R_{\text{W}q}$ is a function of $R_{\text{E}q\_\text{right}}$, $R_{\text{E}q\_\text{left}}$ and c. Examples of the effect of c upon $R_{\text{W}q}$ is given in Fig. 2 for the simple case of equal left and right edge roughness, reduced for illustrative purposes to a single sine wave each. The graph on the left shows non-correlated edges ($c=0$, sine waves 90° out of phase). $R_{\text{W}q}$ is the quadrature sum of each $R_{\text{E}q}$, so $R_{\text{W}q} = R_{\text{E}q}\sqrt{2}$. The middle graph shows perfectly correlated edges ($c=1$, sine waves in phase). $R_{\text{W}q} = 0$, since CD (distance between edges) is constant. The graph on the right shows perfectly anti-correlated edges ($c=-1$, sine waves 180° out of phase). $R_{\text{W}q} = 2R_{\text{E}q}$, as when two periodic functions add in superposition.
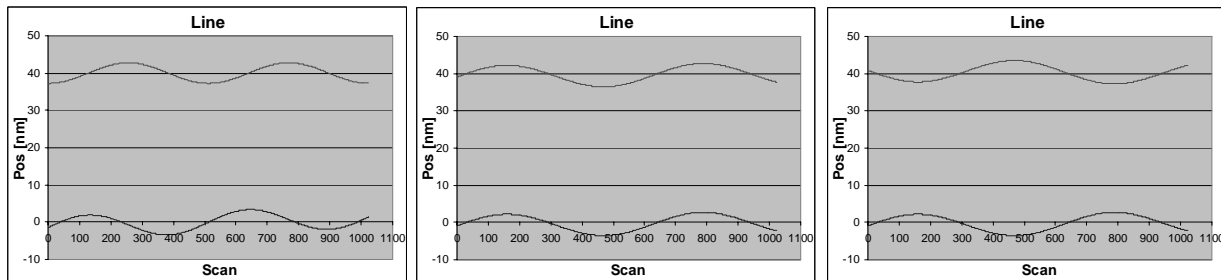

Fig. 2. Left: Non-correlated sine waves (90° out of phase). Center: Correlated sine waves (in phase). Right: Anticorrelated sine waves (180° out of phase).

Anticorrelated edges (at the right) are much more of an issue than correlated edges (middle), since anticorrelated edges result in more $R_{\text{W}q}$. Used with care, the correlation coefficient can be a good diagnostic. Correlation c might be

significantly non-zero due to non-random factors such as optical proximity effects, topography, granularity of resist or poly-Si, or due to a line being close to a random defect or particle. Before concluding, however, that a given nonzero value of c points to such an effect, consider that c determined from a finite sample is an estimate and is itself a random variable. A set of edges that, taken together, appear to have an average correlation coefficient close to 0 (e.g., the many edges forming the distribution in Fig. 3, with correlation centered near $c=0$) may nevertheless have individual members with rather large correlation coefficients (the wide tails). This implies that R_{Wq} is a better measure than R_{Eq} since correlation is included. Likewise, R_T should thus be avoided, as the correlation information is lost.

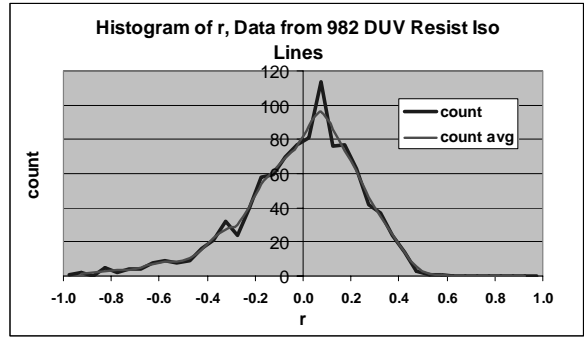**Histogram of r, Data from 982 DUV Resist Iso Lines**

Fig. 3. Histogram of correlation values observed in 982 different isolated resist lines, shown both smoothed and unsmoothed. Correlation coefficient is called "r" in the graph; it is the same as "c" mentioned above.

2.4. Amplitude Density Functions

All of the previous metrics have been attempts to characterize by a single number a distribution around a central value. It is also possible to characterize this distribution with a function, ADF(z). In this function, z is a value for the residual. "ADF" stands for amplitude density function. ADF(z)dz is the probability that a particular measured residual will lie between z and $z+$dz. (See Fig. 4 C, D, and E.) In practice ADF(z) is estimated by normalizing the histogram of the binned residuals:

$$\text{ADF}(z) = \frac{H(z, \Delta z)}{N \Delta z} \tag{10}$$

There are edge and width versions of the ADF function, depending upon which list of residuals, w_i or x_i, are used. Here $H(z, \Delta z)$ is the histogram, i.e., the number of residuals that lie between z and $z+\Delta z$. Δz is the bin size. The normalization insures that the sum of ADF(z)Δz over all the bins (or in the limit $\Delta z \to 0$, the integral of ADF(z)Δz over all z) is 1. For a good estimate the bin size must be judiciously chosen, small enough for good resolution but large enough for good statistics in each bin.

2.5. Power Spectral Density

The foregoing have all been amplitude measures of roughness. They contain no information about roughness wavelengths or characteristic sizes of roughness asperities in the direction parallel to the line. Because the effect of roughness on a device may depend upon its wavelength as well as its amplitude, it is important to have measures that include such information. One of these is the power spectral density, or PSD [7,8,9] which is related to the Fourier transform. We will have use for the PSD in Sec. 3. The coefficients of the discrete Fourier transform of a series, w_j, are given by [10]

$$C_k = \sum_{j=0}^{N-1} w_j e^{2\pi ijk/N} \qquad k = 0, ..., N-1 \tag{11}$$

These may be conveniently calculated using a fast Fourier transform (FFT) algorithm. Since the width residuals (or edge position residuals) are real valued, this transform has even symmetry around $k = 0$. The periodogram estimate of the power spectral density may then be defined in terms of the Fourier transform coefficients as

$$P(0) = P(f_0) = \frac{\Delta}{N^2}|C_0|^2$$

$$P(f_k) = P(f_k) = \frac{\Delta}{N^2}\left[|C_k|^2 + |C_{N-k}|^2\right] \qquad k = 1, 2, ..., N/2 - 1 \qquad (12)$$

$$P(f_c) = P(f_{N/2}) = \frac{\Delta}{N^2}|C_{N/2}|^2$$

where

$$f_k = \frac{k}{N\Delta} \qquad k = 0, 1, ..., N/2 . \qquad (13)$$

The middle expression in Eq. (12) sums contributions from both the positive and negative frequencies present in the Fourier transform, to produce a one-sided (positive frequencies only) PSD. The above definition follows Ref. 10 except for the factor of the sampling interval, Δ, which is required on dimensional grounds for real profiles. Like the ADF, the PSD is not a single number but a curve (Fig. 4G). The discrete form of Parseval's theorem is

$$\sum_{k=0}^{N-1} w_k^2 = \frac{1}{N\Delta}\sum_{k=0}^{N/2} P_k \qquad (14)$$

It is easy to see by comparison to the definition of R_{Wq} in Eq. (8) that the PSD is related to our rms LWR metric by

$$R_{Wq}^2 = \frac{1}{(N-1)N\Delta}\sum_{k=0}^{N/2} P_k \qquad (15)$$

The expression for R_{Eq}^2 is similar, except the factor of N-1 in the denominator becomes N-2 and the P_k must be defined in terms of the edge residuals instead of the width residuals. This means the area under the PSD curve is related to the rms measure of roughness. We can generalize this by summing only those P_k between specified limits corresponding to $f_{min} < f < f_{max}$ to determine the rms roughness contributed only by relevant frequencies.

2.6. Autocorrelation function and correlation length

The autocorrelation function can be calculated from the inverse Fourier transform of the PSD. Alternatively, it can be computed directly from the measured widths by $c(i) = \frac{1}{(N-1)R_{Wq}^2}\sum_k w_{i+k}w_k$. A similar definition applies for the edge correlation function except for the use of edge instead of width residuals and a factor of N-2 in the denominator instead of N-1. The amount, i, by which one copy of the curve is shifted with respect to itself before multiplying is referred to as the lag. It may be positive or negative. The autocorrelation has its maximum value of 1 at a lag of 0. For randomly rough (nonperiodic) edges produced by a stationary process the autocorrelation is expected to tend towards zero for increasing lag. (There are, however, practical issues in the estimation of correlation functions from finite length series. For instance, background subtraction can produce artifacts in the curve. [11,12]) The decrease in c occurs over a characteristic distance, called the correlation length. This length may be characteristic of a grain size or other physical phenomenon that sets a lateral distance scale for the roughness. As with line edge roughness, there are different metrics for the correlation length. It may be defined as the point at which the correlation decreases below a threshold such as 1/e, or it may be determined by fitting an exponential or other suitable function to the neighborhood around zero lag. (See Fig. 4H.)
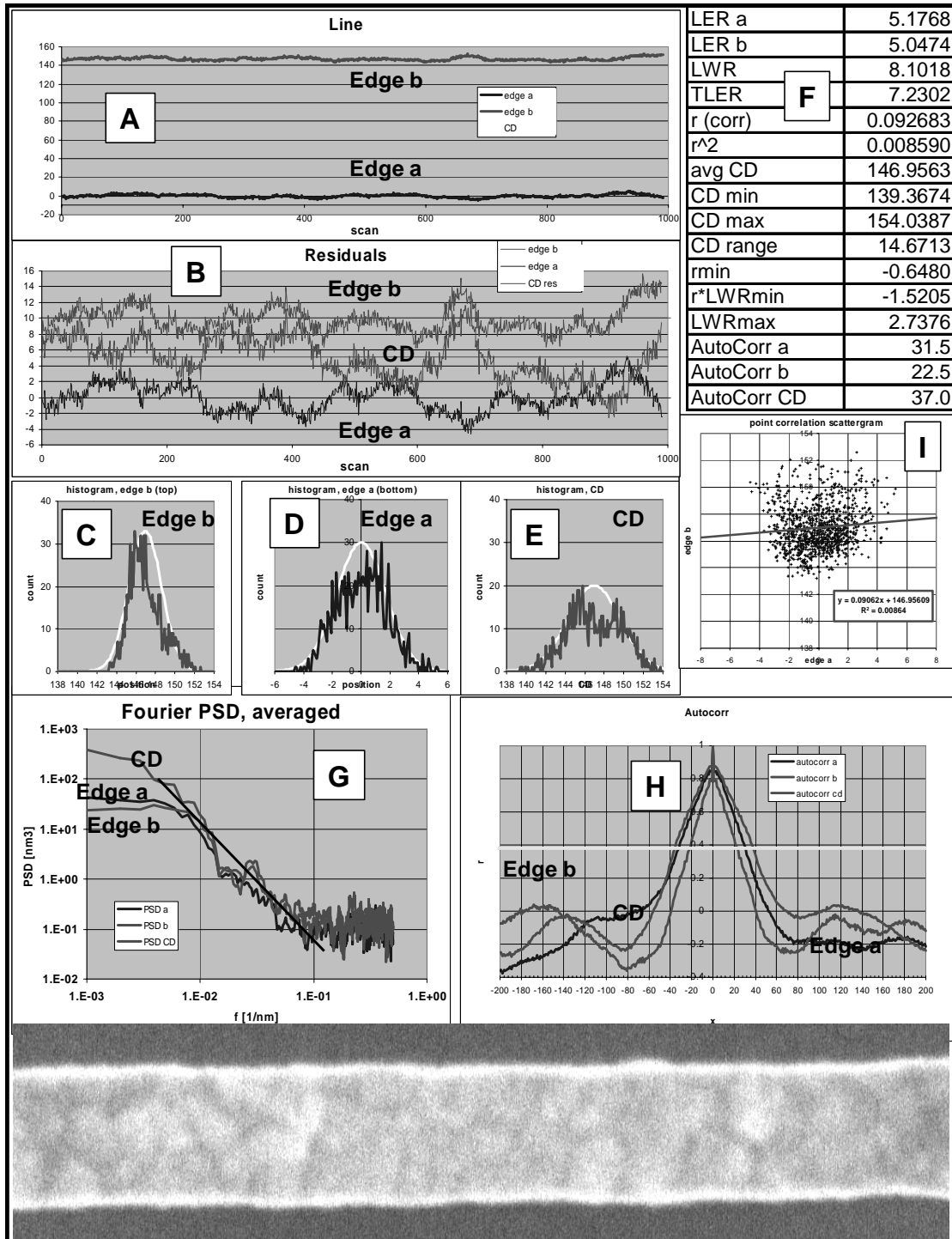
LER a	5.1768
LER b	5.0474
LWR	8.1018
TLER	7.2302
r (corr)	0.092683
r^2	0.008590
avg CD	146.9563
CD min	139.3674
CD max	154.0387
CD range	14.6713
rmin	-0.6480
r*LWRmin	-1.5205
LWRmax	2.7376
AutoCorr a	31.5
AutoCorr b	22.5
AutoCorr CD	37.0

Fig. 4. Example of roughness metrics applied to the line shown at the bottom. The case shown is one of our etched poly repeatability images of a dense line. A: Line edges, a on bottom, b on top. B: Blow-up of residuals, a on bottom, b on top, *w* in middle (labeled "CD"). C, D, & E: histograms of the edge locations a, b, and CD, respectively. All have a roughly Gaussian behavior (but this is not always so). F: Calculated metrics; note that r=0.093 for this example, so these edges are uncorrelated; this is not always so. G: Fourier Power Spectral Density. Note the $1/f^{2.3}$ slope on the log/log plot. H: Autocorrelation functions for edges a, b and CD. The 1/e (~0.36) crossing is a measure of the "correlation length". I: Scatter plot of correlation between edges. All dimensions are in nanometers.

3. HOW TO SAMPLE LINES FOR ROUGHNESS MEASUREMENT

The ITRS, as we saw in the introduction, requires us to estimate that part of the roughness that occurs between specified frequency limits, $f_{min} < f < f_{max}$, and it requires us to do so with a specified precision. Suppose we measure the widths (or edge positions) at N positions over a length, L, of the line. From these we subtract the average width, leaving us with width residuals, w_k, corresponding to positions $k\Delta$, $k = 0, 1, \ldots N$-1 with $\Delta = L/N$. The question we address in this section is, how should we choose Δ and N to meet the requirements?

We have already pointed out that we can generalize Eq. (15) by summing only those P_k between the specified frequency limits to determine the root mean square roughness contributed only by relevant frequencies. The first constraint upon our choice of N and L is determined by the requirement that the full frequency range of interest be contained within our PSD. The largest frequency in the PSD is the Nyquist frequency, $f_c = 1/(2\Delta)$. The requirement that this frequency be greater than or equal to f_{max} translates to $\Delta \leq 1/(2 f_{max})$. The smallest nonzero frequency component of the PSD is $1/L$. If this is to be less than or equal to f_{min} we must have $L \geq 1/f_{min}$. These conditions are necessary but perhaps not sufficient. Our line segment of length L is a finite sample, so the corresponding roughness estimate is subject to sampling error. If the corresponding uncertainty is not good enough, we may wish to impose yet stricter sampling requirements. There is an extensive literature concerning uncertainty in PSD estimation [13,14], and the details may depend upon how one chooses to process the measured data. For our present purpose, which is to elucidate rules of thumb for the choice of N and L, the differences between these do not concern us. For the sake of simplicity we therefore choose Bartlett's method of measuring a trace with length some integer multiple of $1/f_{min}$. This trace has mN points that will then be subdivided into m traces, each consisting of N points over the length $1/f_{min}$. The PSDs of these subdivided traces are summed. (An alternative is to smooth the PSD of the undivided trace. Different choices of windowing functions for the Fourier transform, or overlapping the subdivided segments are other variants.) Therefore we choose

$$\Delta = \frac{1}{2 f_{max}}$$
$$L = \frac{m}{f_{min}}$$

(16)

and increase m as required to improve the uncertainty.

The PSDs of the subdivided profile then each satisfy an equation like Eq. (14). Adding those m equations together yields

$$\sum_{i=1}^{m}\sum_{k=0}^{N-1} w_{ki}^{2} = \frac{1}{N\Delta}\sum_{i=1}^{m}\sum_{k=0}^{N/2} P_{ki}$$

(17)

where $i = 1,2,\ldots,m$ indexes the subdivisions. Because of the choices in Eq. (16), $P_{1,i}$ and $P_{N/2,i}$ correspond (for all i) to the frequencies f_{min} and f_{max} that form the bounds of the frequency window we want to measure. If the residuals are with respect to the average width (or best fit line for edges) for each segment individually, then P_{0i} will be identically equal to 0, so the sum over k includes exactly those roughness frequencies that we care about. The left side is just $m(N\text{-}1)r_{Wq}{}^{2}$ so

$$R_{Wq}{}^{2} = \frac{1}{m(N-1)N\Delta}\sum_{i=1}^{m}\sum_{k=0}^{N/2} P_{ki} = \frac{1}{(N-1)N\Delta}\sum_{k=0}^{N/2} \overline{P_k}$$

(18)

In the second form on the right, $1/m$ times the sum over i has been replaced by the average of the m PSDs, $\overline{P_k}$.

The uncertainty in R_q can be found straightforwardly if the uncertainties in the P's are known. The uncertainty in X as a result of an uncertain parameter, p, is $U_X = u_p |(dX/dp)|$. When there are many parameters with uncorrelated errors the uncertainties are added in quadrature. Applying this to Eq. (18) produces

$$u_{R_q^2}^2 = (2R_q u_{r_q})^2 = \left[\frac{1}{(N-1)N\Delta}\right]^2 \sum_{k=0}^{N/2} (u_{\overline{P_k}})^2 \tag{19}$$

The variance of each point in a single unaveraged periodogram is approximately the square of the spectrum. [13] That is, $u_{P_{ki}} \approx P_{ki}$. By averaging m spectra we reduce the uncertainty by a factor of $\sqrt{m}$. Replacing $u_{\overline{P_k}}$ with $\overline{P_k}/\sqrt{m}$ in Eq. (19), bringing the $(2R_q)^2$ to the right, and dividing by an extra factor of R_q^2 produces

$$\frac{u_{R_q}^2}{R_q^2} = \frac{1}{4mR_q^4}\left[\frac{1}{(N-1)N\Delta}\right]^2 \sum_{k=0}^{N/2} \overline{P_k}^2 \tag{20}$$

The reason for dividing by the extra factor of R_q^2 is that this now becomes an expression for the relative uncertainty, in which form the expression shortly becomes very simple. Substituting for R_q^2 from Eq.(15), most of the leading coefficients cancel to produce

$$\frac{u_{R_q}^2}{R_q^2} = \frac{\sum_{k=0}^{N/2} \overline{P_k}^2}{4m\left(\sum_{k=0}^{N/2} \overline{P_k}\right)^2} \tag{21}$$

The leading factor of $1/m$ means that if we pick m large enough, we can satisfy the measurement's precision requirement. Let us state that requirement in the form

$$\frac{u_{R_q}}{R_q} \leq \eta \tag{22}$$

For example, the ITRS requires that u_{R_q} be 20 % of the roughness tolerance for a given node. It seems reasonable therefore to set η to 0.2. Then when the roughness is at the tolerance, the statistical uncertainty also is, and the statistical uncertainty becomes better as the roughness gets smaller.

It is instructive to look at some special cases. Often the PSD approximately follows a power law, $\overline{P_k} = A/k^z$ (for $k>0$ and all i), over a frequency range of interest. If we substitute the power law into Eq. (21), the requirement for m becomes

$$m \geq \frac{\xi(N,z)}{4\eta^2} \quad \text{with} \quad \xi(N,z) = \frac{\sum_{k=1}^{N/2} \frac{1}{k^{2z}}}{\left(\sum_{k=1}^{N/2} \frac{1}{k^z}\right)^2} \tag{23}$$

Table 1: Properties of $\xi(z)$ and their implications

z	Limiting value of $\xi(z)$	Corresponding value of m for $\eta=0.2$
1.5	0.185	1.2
2.0	0.400	2.5
2.5	0.576	3.6
3.0	0.704	4.4

The ratio of sums, $\xi(N,z)$, for a given N and z is a dimensionless number less than or equal to 1. Its behavior for various values of N and z is illustrated in Fig. 5. It is desirable that the ratio be as small as possible, since this makes for a smaller value of m (i.e., the length of line we need to measure is fewer multiples of the basic length). Table 1 shows approximate limiting values for several values of z and the corresponding value of m when $\eta = 0.2$. We see that for $z>1$ the ratio never approaches 0. It approaches (quickly for z large, less so as z approaches 1) a nonzero limit as N increases. In the large N limit the sums become Riemann Zeta functions. (See Eq. 9.522 in Ref. 15.) For $z=1$ (a $1/f$ spectrum), the ratio approaches 0, but slowly (as $C/\ln(N)$, C a constant, see Ref. 15 Eq. 0.131.) For a flat, or white noise spectrum, corresponding to $z=0$), $\xi(0) = 2/N$.

Many of the resist edges measured for this study have z=2.5 or so for a substantial range of frequencies (see Fig. 4). Rounding up the value in Table 1, this means we need m=4. According to the ITRS table in Fig. 1, at the 90 nm node the highest and lowest frequencies of interest are 1/(15 nm) (i.e., the reciprocal of the low-end-of-range drain extension) and 1/(180 nm) (i.e., 1/DRAM pitch). Using these values in Eq. (23) results in the following measurement specification: We should measure a length of line at least L=720 nm at intervals of Δ=7.5 nm (i.e., 96 measurement positions) in order for uncertainty associated with sampling to be within the desired tolerance.
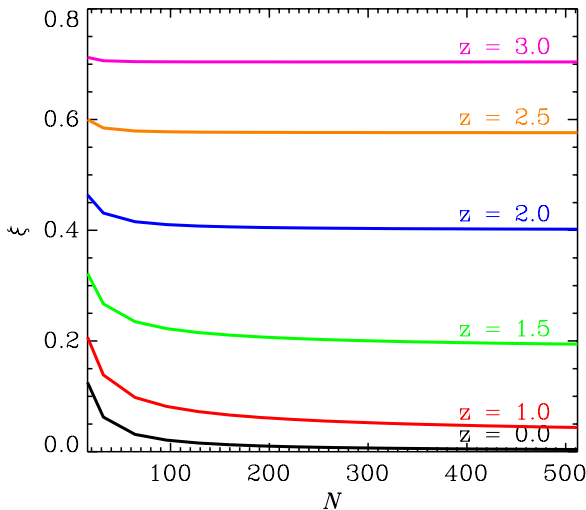

Fig. 5. Values of the ratio of sums in Eq. (23) for various z and N.

This length of line is 8 times the node length. The way to make sense of the requirement for long distance measurements is this: Closely spaced samples provide information about high frequency components of roughness. However, when the PSD is decreasing sufficiently quickly, the rms roughness is dominated by only a few low frequency components that have very high relative amplitudes. The extra knowledge provided by our closely spaced points does not improve our estimate of these. The only way to improve the estimate for these relevant components is to measure long distances to sample many wavelengths.

This section has outlined the means to satisfy the requirements for adequately sampling the roughness. However, the effect of noise on the roughness is an additional complication that still must be addressed. We discuss this in the next section.

4. EFFECT OF NOISE ON ROUGHNESS DETERMINATION

Determination of edge positions or linewidths requires use of an edge assignment algorithm. If there are random errors (noise) in the image, there will be a resulting random error in the edge assignments and linewidths, more or less error depending upon the sensitivity of the chosen algorithm to noise. Let us say the width determination error that results from a particular set of conditions to be ε_i and that these errors are distributed with a standard deviation σ_ε (Fig. 6). What is the effect on our roughness determination, for example using Eq. (8)? We will add the sum of squares of the measured widths instead of the actual widths and obtain a measured roughness given by

$$
\begin{aligned}
(N-1)R^2_{Wq\_meas} &= \sum_{i=0}^{N-1}\left[W_i + \varepsilon_i - \overline{(W+\varepsilon)}\right]^2 \\
&= \sum_{i=0}^{N-1}\left[w_i + \varepsilon_i - \overline{\varepsilon}\right]^2
\end{aligned}
\tag{24}
$$

The right side decomposes into 3 terms when expanded. The sum of the w_i^2 becomes $(N-1)R^2_{Wq}$ by definition. The sum of $\left(\varepsilon_i - \overline{\varepsilon}\right)^2$ becomes $(N-1)\sigma_\varepsilon^2$, also by definition. The expectation value of the cross term vanishes because the noise and the roughness are uncorrelated. Thus, the simple result:

$$
\left\langle R^2_{q\_meas}\right\rangle = R_q^2 + \sigma_\varepsilon^2
\tag{25}
$$

where < > denotes the expectation value. We have left out the "W" in the subscript because the same result can be shown to apply for the case of edge roughness, Eq. (5), provided σ_ε is understood to be the standard deviation of the distribution of edge errors instead of width errors.
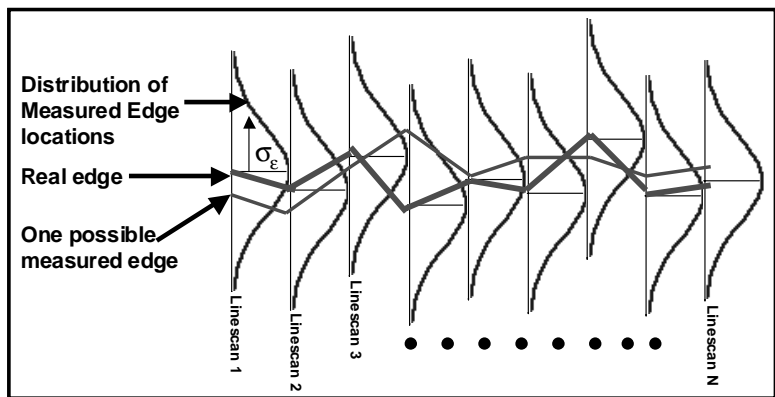

Fig. 6. Schematic for problem of considering how noise effects the apparent measured position of a real line edge. The thick meandering line represents the real edge. The slimmer meandering line represents one measured edge out of infinite possibilities. The Gaussian curves overlaid on each linescan represent the distribution of measured edge locations.

This result means random errors in the edge roughness bias the measurement—they add root sum of squares fashion with the true roughness to bias the measured result. If the σ_ε are known (for example, by estimation from repeated measurements of the same sample, or by estimation of the noise floor in the PSD), σ_ε^2 can be subtracted from the

measured square of the roughness to obtain a corrected estimate of the actual squared roughness. Such correction is reasonable when the edge assignment repeatability is smaller than or comparable to the roughness. However, when it becomes much larger than the roughness, discerning the roughness above the noise "background" is likely to become increasingly difficult.

Noise also affects the repeatability of the roughness measurement. Since $R_{\mathrm{W}q}$ is essentially a standard deviation, its repeatability is the standard deviation of a standard deviation. The algebra is therefore a bit more involved than usual [16], but assuming σ_ε is small compared to the roughness, the result is

$$\sigma_{R_{\mathrm{W}q\_meas}} = \frac{\sigma_\varepsilon}{\sqrt{N-1}} \tag{26}$$

We have not derived an expression for the case when σ_ε is comparable to or greater than R_q. However, the following relationship matches simulation results for the more general case, and it agrees with Eq. (26) in the appropriate limit:

$$\sigma_{R_{\mathrm{W}q\_meas}} \cong \frac{\sigma_\varepsilon}{\sqrt{2(N-1)}} \sqrt{1 + \frac{R_{\mathrm{W}q}^2}{\sigma_\varepsilon^2 + R_{\mathrm{W}q}^2}} \tag{27}$$

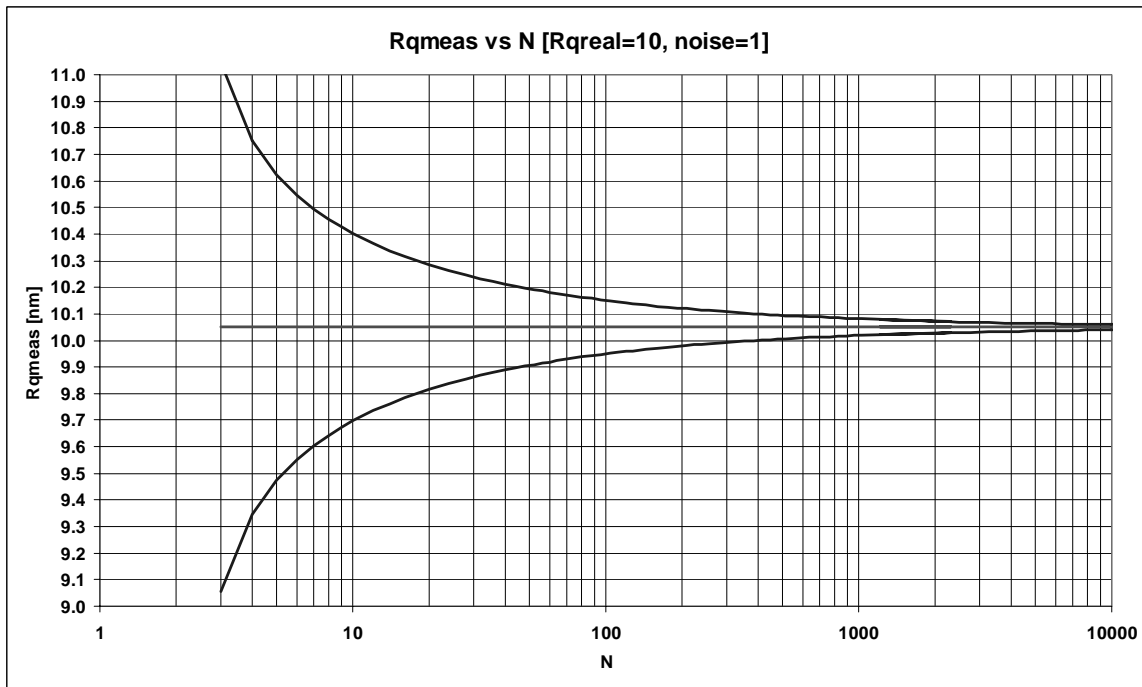The behavior of Eq. (25) through Eq. (27) is illustrated graphically in Fig. 7.

Fig. 7. Example of bias and uncertainty, i.e., Eq. (25) through Eq. (27), as functions of the number of linescans for a LWR measurement with noise equal to 10 % of $R_{\mathrm{W}q}$. The central line is the expected value for infinitely many repeated measurements. Note that it is biased with respect to the true value (10 in this example). The spread between the outer lines represents the expected scatter (±1 standard deviation) for a single LWR measurement.

Some CD-SEMs allow for "binning", i.e. measuring an edge's position at the site of N "bins" with results from n linescans averaged together. This has the effect of reducing σ_ε as $\sqrt{n}$, with the penalty of less spatial frequency sensitivity.

With the repeatability of the measurement of R_{Wq} known in the limit of $\sigma_\varepsilon << R_{Wq}$, the ramifications to meeting the ITRS LWR measurement precision specifications are calculated below.

Technology Node	130nm	115nm	100nm	90nm	65nm	45nm	32nm	22nm	18nm
Year of Production	*2001*	*2002*	*2003*	*2004*	*2007*	*2010*	*2013*	*2016*	*2018*
3s LWR control, <8% of etched gate [nm]	5.2	4.2	3.6	3.0	2.0	1.44	1.04	0.72	0.56
1s LWR control, <8% of etched gate [nm]	1.7	1.4	1.2	1.0	0.7	0.48	0.35	0.24	0.19
3s Precision of LER measurement [nm]	0.74	0.60	0.51	0.42	0.28	0.20	0.15	0.10	0.08
1s Precision of LER measurement [nm]	0.25	0.20	0.17	0.14	0.09	0.07	0.05	0.03	0.03

Linescan Reproducibility, σ_ε [nm]	Minimum # of Linescans to Meet ITRS Precision Spec								
0.25	4	4	5	6	10	16	29	57	92
0.50	7	9	11	15	31	57	107	220	361
0.75	12	17	22	31	66	125	237	491	810
1.00	19	28	37	54	115	220	419	871	1437
1.50	40	59	81	118	256	491	939	1956	3231
2.00	69	103	141	208	452	871	1667	3475	5742

Fig. 8. Minimum number of linescans necessary to meet ITRS LWR measurement precision specifications, for different values of σ_ε.

As we have just seen, the effect of noise on roughness metrics is different than on many other more familiar measurements. For example, measuring CDs means measuring average positions of edges. In a CD-SEM, this involves performing a number of linescans to collect an image. Linescans may be binned to improve signal to noise, or the CDs of many linescans may be averaged. Positive excursions from the average value cancel negative excursions to first order, thereby diluting the effect of outliers. Roughness, on the other hand, is a dispersion. Both positive and negative excursions from the mean position of the edge *add* to the total roughness. Since noise always adds to the apparent roughness, the average of *N* repeated measurements does not tend to the correct value as *N* goes to infinity. This places a premium on reducing random errors in edge positions or linewidths before the determination of roughness begins. Some of the strategies for noise reduction employed with other measurements have a steeper cost when measuring roughness. For example, binning of linescans reduces the effective spatial resolution along the edge, thereby impairing measurement of high spatial frequency components of roughness.

In the following section we experimentally investigate some of the variables that appear to impact the magnitude of random errors in edge assignment. These include signal to noise, choice of edge algorithm, and quality of focus.

5. EXPERIMENTAL

5.1. Instruments/Samples/Measurements in general

The measured features consisted of either resist on poly-Si, or etched 1500 Å Poly-Si on 20 Å gate oxide. The resist used was 2700 Å Sumitomo PAR-810 193 nm resist [17]. Structures were exposed using ISMT's AMAG-4L reticle. The targets imaged were isolated and dense lines in scatterometry patterns of 150 nm line width with 450 nm pitch for the dense case, and 1500 nm pitch for the isolated case. Also, special test structures with induced roughness at designed wavelengths of 50 nm, 200 nm, and 250 nm [18] were imaged.

All data in this work were collected in the form of tiff images of features for offline analysis. The CD-SEM used was a recent model CD-SEM in ISMT's ATDF (ISMT's fab). As the resolution of the CD-SEM is on the order of 2-2.5 nm, the measurement technique was designed to sample 1 nm pixels to assure proper Nyquist sampling of all observable roughness features. Thus, each image scanned a 1000 nm × 1000 nm region, with 960 pixels in the x-direction and 1024 pixels in the y-direction. All filtering was turned off to acquire the raw signal in the imaging. The averaging factor was set to 1 linescan per bin.

Slightly different conditions were used for etched poly than for resist. For etched poly imaging, the accelerating voltage was 800 V with 10 pA beam current and 3.2 μs beam dwell time per pixel. For resist imaging, the accelerating voltage was 500 V with 10 pA beam current and 2.3 μs beam dwell time per pixel. One interesting point here is that the beam

condition for measuring LER on resist was significantly different than the typical condition used to measure CDs, where 193 nm resist shrinkage must be minimized; a pixel dwell time of ~0.8 μs is much better for reducing shrinkage of this resist. However, as mentioned in previous work [18], the LER measurement is less prone to shrinkage / reproducibility issues by virtue of the fact that all features along the same sidewall shrink in the same direction (versus the CD measurement case, where two edges shrink in opposite directions).

From the acquired images, edges were located using offline software. Four edge detection algorithms were used:

- Maximum derivative (MAXD), based on a threshold algorithm.

- Regression-to-baseline (REGR), based on fitting the waveform peak edge with a line and intersecting with a baseline value.

- Sigmoidal fit (SIGM), based on fitting a sigmoidal mathematical function to the peak of the waveform.

- Model-Base Library fit (MBL), based on physics simulations.

Note that the first three were designed to mimic those used in typical current-model CD-SEMs. All four edge algorithms have been described in a recent related work [19]. The edge algorithms were applied to each image along each linescan (no binning). The results files included edge locations of all edges in an image, PSDs of each edge and linewidth, and autocorrelation functions of each edge and linewidth.

Images were acquired of many targets for a larger study, but the ones used in this work are poly and resist images of the following experiments:

- Repeatability—five images were taken (static mode) of the same segment of line, at optimal image quality, at the standard conditions described above.

- Image integration time variation—images were taken (static mode) of the same segment of line, at optimal image quality, with varying beam dwell time per pixel.

- Image focus variation—images were taken (static mode) of the same segment of line, at varying image quality due to intentional defocusing, at the standard conditions described above.

- Long line length—six 1000 nm × 1000 nm images were taken in "tiled" fashion with 100 nm overlap; matching the images together allows for a ~5500 nm long image of lines to be analyzed, with 1 nm pixel size in both x- and y-directions.

5.2. Repeatability and Image Integration Time Variation

Using the principles described in the noise discussion above, we wish to evaluate the performance and optimal operating conditions of our edge detection algorithms. From the repeatability images, edge locations between consecutive runs were correlated to one another to compensate for possible image shift between images. Linescans were thus matched up, and a variance of the residuals of each linescan's edge location from the line of best fit for each edge was calculated. The square root of the average of these linescan variances yields the σ_ε value, defined earlier as the 1σ repeatability of edge location on a single linescan. With 1024 linescans, the error of the measurement of this value is quite small (a few linescans were "lost" due to image shifting between observations—the calculation was done over the linescans which each image had in common). This was executed for each of the four edge detection algorithms for the standard repeatability condition. Also, from Eq. (25), if the actual R_q can be assumed to be unchanged between two images, the relative values of σ_ε between the two images can be calculated with the two values of R_{q_meas}. This was done through the series of images with varying image integration time and focus. Since the absolute value of σ_ε was found for the repeatability condition, the relative value for σ_ε for each of the other images can be thus referenced back to the value at the repeatability condition, and thus made absolute, also. Results for computed values of σ_ε are shown in Fig. 9. Also, relative noise was measured on the different images through image integration time, using the *Measure* software package by Spectel Research [17]. The noise decreases as ~$t^{-1/2}$, as should be expected.

To verify that the assumption of ~constant R_q is valid, an extra image at the nominal condition was taken at end of the integration time runs, and the measurement reproduced the one from the beginning of the runs.
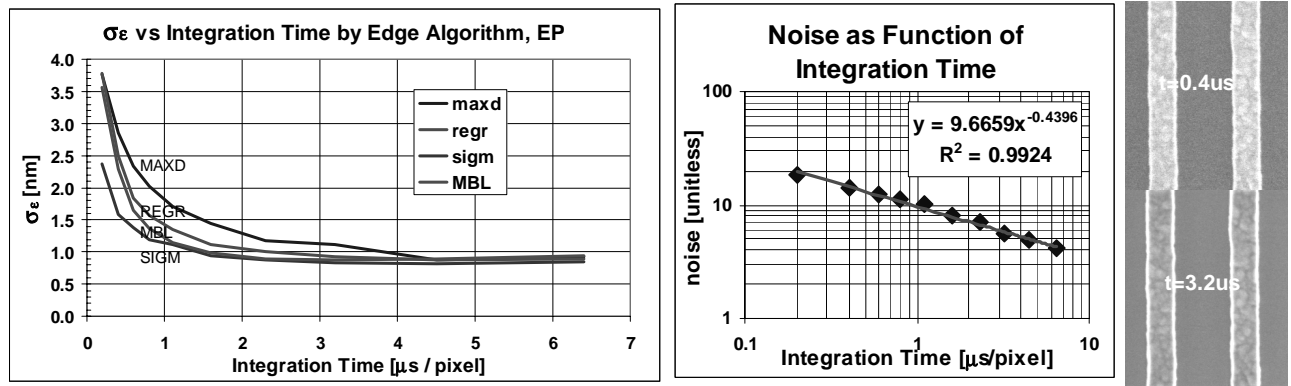
Fig. 9. Left: σ_ε as function of image integration time for different edge algorithms on etched poly. Right: more time means less noise relative to signal. Note that 3.2 μs is the nominal value we use for roughness measurement on etched poly (EP). Results for higher values of pixel time could be computed, but problems with the data collection prevented this. Note that the curves begin a slight upward trend at higher pixel times, where charging or vibrations can cause an "edge blurring" effect, which raises σ_ε. At the nominal 3.2 μs pixel time, the order of the performance of the edge algorithms, in increasing σ_ε, is MBL (best, at $\sigma_\varepsilon=\sim0.8$ nm) closely followed by sigmoidal, with regression and max derivative performing with higher σ_ε.

Notice that from the image on the left in Fig. 9 that the sigmoidal and MBL edge algorithms demonstrate better performance, as these algorithms make more use of the information in the entire waveform. Thus, they are better suited to LER measurement (and, arguably, CD measurement as well, edge location bias issues aside, since better noise performance is desirable for the case of CD measurement also). Notice that σ_ε does not go to 0 in the limit of large pixel time. At values higher than an optimum value (between 3 μs and 4 μs in this case) it begins to trend slowly upwards again, perhaps due to vibration, charging, or contamination. Notice that in our experience with these samples, $3\sigma_\varepsilon$ is already close to 3 nm, whih is the roughness tolerance for the 90 nm technology node. This means the roughness "background" produced by noise is already comparable to the roughness we need to measure.

5.3. Defocus Variation

This analysis was repeated for the set of images of varying defocus in Fig. 10. We see that noise in the intensity values (graph on the right) remains relatively constant with defocus, while σ_ε increases dramatically. This effect was seen in simulations a year ago [19], where it was attributed to smearing of the edge. (Edge detection is improved when edges are sharp because the contrast between the intensity gradient signal and false gradients due to noise is greatest then.)
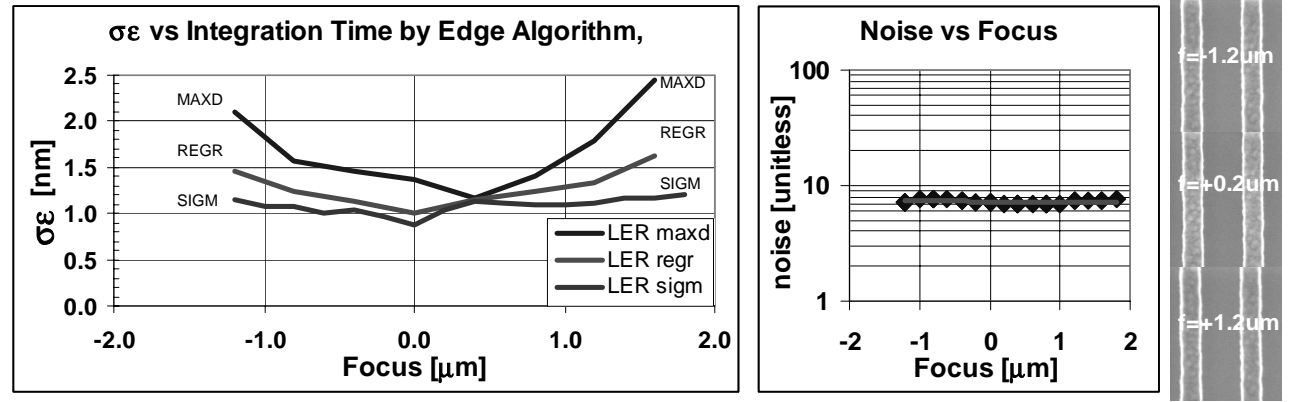

Fig. 10. σ_ε as function of image defocus for three edge algorithms on etched poly. At the nominal 0 μm focus, the order of the performance of the edge algorithms, in increasing σ_ε, is sigmoidal (best, at $\sigma_\varepsilon=\sim0.8$ nm), with regression close behind and max derivative performing with higher σ_ε. Also note that noise does not change measurably with focus. The noise graph is scaled the same as the noise vs. integration time results to demonstrate this.

Since defocus effectively increases the electron beam spot size, it also degrades the resolution. Thus, defocus has two effects on roughness measurement. On the one hand, it increases the sensitivity of the measurement to image noise, tending to increase the apparent (i.e., measured) roughness. On the other hand, loss of resolution causes high frequency roughness to be missed, tending to decrease the apparent roughness. Neither of these is good. Though their opposite signs mean that fortuitous cancellation of errors can occur, measurements in better focus measure more of the actual roughness and less of the noise.

5.4. Illustration of Sampling Statistics on Long Tiled Lines

As an example of calculating uncertainty for a PSD, we have analyzed one of our long tiled lines. The edges were located for the entire line, and these were segmented into smaller pieces. Let's assume that the frequency window of interest is 2 nm through 250 nm. The 4 μm long line is segmented into 16 of these 250 nm pieces. We then calculate R_{Wq} and the PSD for each of the 16 segments. The standard deviation of these sixteen R_{Wq}'s is 0.74 nm. The standard deviation of the mean is therefore $0.74 \, \text{nm} / \sqrt{16} = 0.185 \, \text{nm}$.

From the PSD of each segment, we then calculate u_{Rq} using Eq.(21). The value is 0.187 nm. Thus we have good agreement between the predicted and actual observed standard deviation using multiple segments of the same line.

Note that the rms average R_{Wq} of the segments is 2.91 nm, while the R_{Wq} of the 4 μm long (unsubdivided) line is 3.27 nm. The reason for the difference is that the measurement on the long line includes some long wavelength roughness that is excluded from the measurement of the shorter segments. If the length and number of the segments is chosen as dictated by Eq. (16), these neglected longer wavelengths will be precisely those wavelengths that we intend to neglect because they are outside of the range of roughness that is relevant to the manufacturing process.

From the PSD results on the right in Fig. 11, it can also be seen that the average PSD of the multiple segments overlays the PSD of the longer line, as expected, only the frequency window of the function for the segments is a subset of the frequency window for the longer line. Also, the curve is much more smooth since it is averaged.
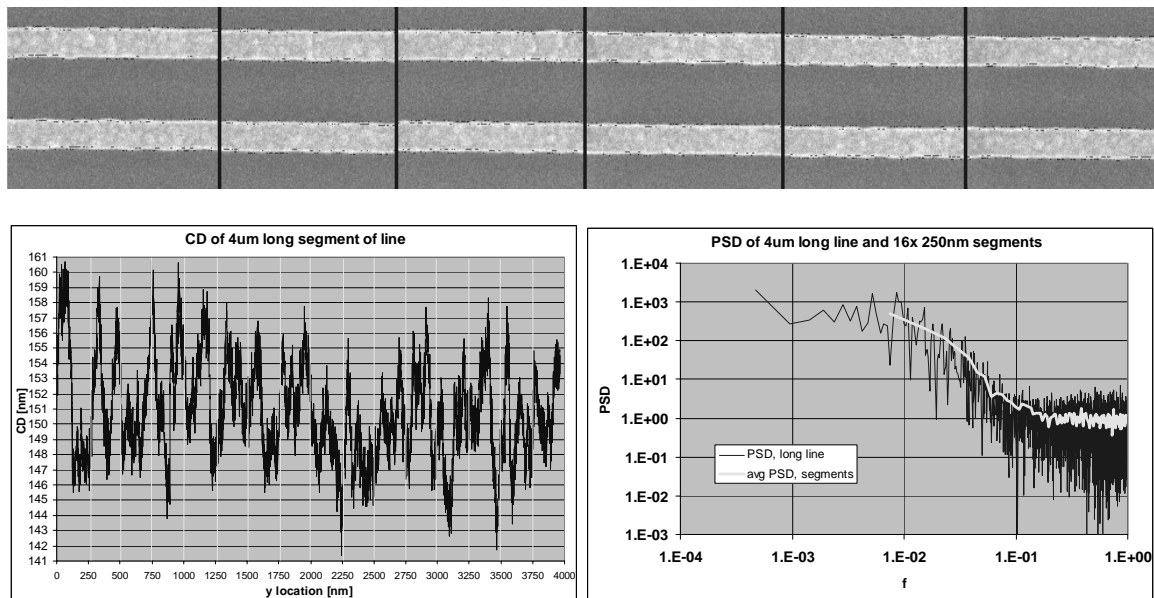

Fig. 11. Top: Tiled image of etched poly dense lines. Lower Left: CD of bottom line of tiled image; the bright lines show where this function was segmented into 16 smaller images. Bottom Right: PSD of entire 4 μm line (dark) and average PSD of 16 250 nm segments (yellow).

6. SUMMARY AND CONCLUSIONS

There are a number of ways that LER and LWR can be quantified. We discussed some of the available metrics in Sec 2. The various metrics capture different aspects of roughness, some, for example, emphasizing aspects of roughness amplitude while others measure characteristic roughness wavelengths or correlation lengths. It is not yet always known which of these measures captures those aspects of roughness that are most relevant to device performance. Indeed, different devices or different processes may require different metrics. It therefore seems advantageous for measuring tools to offer the manufacturing process designer flexibility in the form of a suite of metrics. CD-SEMs [20] used for monitoring production could then be programmed to use whichever metric or set of metrics had proven most useful during process development. Of course, with more than one metric it will be important to have consistent and standardized definitions for each one in order to avoid confusion.

We have paid particular attention in this report to the root mean square measures of edge and linewidth roughness, because these are given prominence in the 2003 ITRS. The repeatability of determining edge positions and adequate sampling are significant determinants of the quality of a rms roughness measurement.

Edge position repeatability, what we have called here σ_ε, may be ascertained by repeated measurements of the same part of a line. Once the measurements are corrected for any drift that may have occurred between them, the average rms differences at each point on the line provide a measure of σ_ε. This metric should be considered a basic "building block" of CD-SEM (or any scanning probe-based) metrology. The magnitude of σ_ε is important because it adds in quadrature with the true rms roughness to produce the measured roughness. This means the measured roughness is always larger than the true value by an amount determined by σ_ε, unless σ_ε is known or measured so a correction can be applied. This is unlike the case with most familiar measurements, in which noise is equally likely to cause negative errors as positive ones. In our sample measurements, this roughness "background" was already comparable to the roughness tolerance specified in the ITRS for the 90 nm technology node. Currently the only way to reduce this background is to average edge positions determined from repeated linescans at each measurement location.

Factors that must be optimized to improve edge detection repeatability include the edge detection algorithm, image resolution, and noise. Noise is a function of integration time, electron dose (charge per unit area, which is current × time / irradiated area), and sample charging. Resolution is a function of the interaction of the beam with the sample, the beam control settings, such as focus and stigmation, and other tool-related variables such as stage vibration. Selection of a tool with better resolution and control over filtering are thus very important. In illustrative measurements of resist lines performed for this study, the best value of σ_ε was not much better than 1 nm for single unbinned linescans. This repeatability level dictates that roughness measurements be based upon a minimum of approximately 50 linescans in order to meet the precision requirements of the roadmap for the current node. The number of linescans needed will increase with each succeeding node, unless the edge detection repeatability improves proportionally.

ITRS specifications imply the need to measure root mean square roughness between specified frequency limits with a specified repeatability. Roughness is a random phenomenon. For a particular sampled length of line to be adequately representative of other, unmeasured lengths, the sampled length must be long enough to be representative. We have derived expressions for the length of line that should be measured and the sampling distance along that line in order to meet the ITRS requirements (or any other desired spatial frequency window and repeatability). The solution indicates that for edges with power spectra that slope downward more steeply than $1/f$ the roughness is determined mainly by low spatial frequency components. Reducing the uncertainty of these requires long lengths of line to be sampled. With lines with roughness like those in our test samples, lengths equal to 8 times the technology node are required. The number of needed measurement positions within this length of line is close to 100. It may be desirable for future versions of the ITRS to take note of this requirement, as well as the need for better edge detection repeatability.

7. APPENDIX I: THE RELATIONSHIP BETWEEN R_{Wq} AND R_{Eq}

Consider the two edges of a long segment of untapered line. At any position along the line, the edges have positions x_1 and x_2, respectively. Over the line segment being considered, the edges have lines of least square fit of f_1 and f_2

respectively. By an untapered line we mean the left and right edges are parallel to each other on average, so that for the long segment under consideration the fits have approximately equal slope. We consider the large N limit, so we may ignore differences between N, N-1, and N-2. In this approximation the LER of each edge is:

$$R_{Eq1} = \sqrt{\frac{\sum(x_1 - f_1)^2}{N}} \quad R_{Eq2} = \sqrt{\frac{\sum(x_2 - f_2)^2}{N}} \tag{28}$$

All sums are over all sampled positions of each edge, i.e. at each scan line. Also note the statistical definition of the linear correlation coefficient (c):

$$c = \frac{\sum(x_1 - f_1)(x_2 - f_2)}{\sqrt{\sum(x_1 - f_1)^2 \sum(x_2 - f_2)^2}} \tag{29}$$

The CD at any position along the line is $x_1 - x_2$ and R_{Wq} is the standard deviation of the CD. Again, using the definition of the standard deviation, the R_{Wq} can be expressed as:

$$R_{w_q}^2 = \frac{\sum\left[(x_1 - x_2) - (\overline{x_1 - x_2})\right]^2}{N} \approx \frac{\sum\left[(x_1 - f_1) - (x_2 - f_2)\right]^2}{N}$$
$$= \frac{\sum(x_1 - f_1)^2}{N} + \frac{\sum(x_2 - f_2)^2}{N} - 2\frac{\sum(x_1 - f_1)(x_2 - f_2)}{N} \tag{30}$$

In the first line, $\overline{x_1 - x_2} = \overline{x_1} - \overline{x_2}$ is the average CD. In the second form we use the fact that for parallel edges the average CD is also the difference between the linear fits to the separate edges. In the final form, the first two terms are, by identity, R_{Wq1}^2 and R_{Eq2}^2. The third term can be strategically split:

$$R_{w_q}^2 = R_{Eq1}{}^2 + R_{Eq2}{}^2 - 2\frac{\sum(x_1 - f_1)(x_2 - f_2)}{\sqrt{\sum(x_1 - f_1)^2 \sum(x_2 - f_2)^2}} \frac{\sqrt{\sum(x_1 - f_1)^2 \sum(x_2 - f_2)^2}}{N} \tag{31}$$

Resubstituting the identities of c, R_{Eq1} and R_{Eq2}, yields the desired result:

$$R_{Wq} = \sqrt{R_{Eq1}{}^2 + R_{Eq2}{}^2 - 2cR_{Eq1}R_{Eq2}} \tag{32}$$

ACKNOWLEDGEMENTS

REFERENCES

1. Xiong, S., Bokor, J. "Study of Line Edge Roughness in 50 nm Built MOSFET Devices." Proceedings of SPIE 2002, 4689: p733-741.

2. Diaz, C., Tao, H., Ku, Y., Yen, A., Yound, K. "An Experimentally Validated Analytical Model for Gate Line-Edge Roughness (LER) Effects on Technology Scaling." IEEE Electron Device Letters, v.22, No 6: June, 2001, p287-289.

3. Ercken, M., Storms, G., Delvaux, C., Vandenbroeck, N., Leunissen, P., Pollentier, I. "Line Edge Roughness and its Increasing Importance." Proceedings of ARCH Interface 2002.

4. International Technology Roadmap for Semiconductors (ITRS), 2003 Edition, http://member.itrs.net.

5. ASME B46.1-2002, "Surface Texture (Surface Roughness, Waviness, and Lay)" (American Society of Mechanical Engineers, 2003)

6. P. R. Bevington and D. K. Robinson, Data Reduction and Error Analysis for the Physical Sciences, 2nd Edition, (McGraw-Hill, Inc., New York, 1992) pg. 42.

7. ASTM F1811-97 *Standard Practice for Estimating the Power Spectral Density Function and Related Finish Parameters from Surface Profile Data* (American Society for Testing and Materials, West Conshohocken, PA, 1997).

8. E. Marx, I.J. Malik, Y.E. Strausser, T. Bristow, N. Poduje, and J.C. Stover, "Power Spectral Densities: A Multiple Technique Study of Different Si Wafer Surfaces", J. Vac. Sci. Technol., **B20,** 31 (2002).

9. N.G. Orji, M.I. Sanchez, J. Raja, and T.V. Vorburger, "AFM Characterization of Semiconductor Line Edge Roughness", in *Applied Scanning Probe Methods,* B. Bhushan, H. Fuchs, and S. Hosaka, eds. (Springer Verlag, Berlin, 2004) chap. 9.

10. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, (Cambridge, Cambridge University Press) 1988.

11. E. L. O'Neill and A. Walther, "A problem in the determination of correlation functions," J. Opt. Soc. Am., **67** p 1125-1126 (1977).

12. E. R. Freniere, E. L. O'Neill, and A. Walther, "Problem in the determination of correlation functions. II," J. Opt. Soc. Am., **69**, p634-635 (1979).

13. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, (Englewood Cliffs, N.J.: Prentice Hall) 1975, p. 547.

14. G. M. Jenkins and D. G. Watts, *Spectral Analysis and its Implications*, (San Francisco, CA, Holden-Day), 1968.

15. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products,* (New York, NY, Academic Press) 1980.

16. The algebra is omitted for space reasons, but is available upon request from the authors.

17. Certain commercial products are identified in this report in order to describe the experimental and analytical procedures adequately. Such identification does not imply recommendation or endorsement by NIST or International SEMATECH, nor does it imply that the items identified are necessarily the best available for the purpose.

18. Bunday, B., Bishop, M., Villarrubia, J., and Vladar, A. "CD-SEM Measurement of Line Edge Roughness Test Patterns for 193 nm Lithography". Proceedings of SPIE **5038**, 2003, pp 674-688.

19. Villarrubia, J. S., Vladár, A. E., and Postek, M. T., "A Simulation Study of Repeatability and Bias in the CD-SEM," Proceedings of SPIE **5038**, 2003, pp 138-149.

20. B. D. Bunday, M. Bishop, and J. Allgair. "Results of Benchmarking of Advanced CD-SEMs at the 90 nm CMOS Technology Node". Proceedings of SPIE **5375**, 2004, to be published.