Issues in LER & LWR Metrology

J. S. Villarrubia

National Institute of Standards and Technology

I. Roughness Affects Device Performance

II. The usual (R_q) metric

III. Issues (and answers) in Measuring R_q

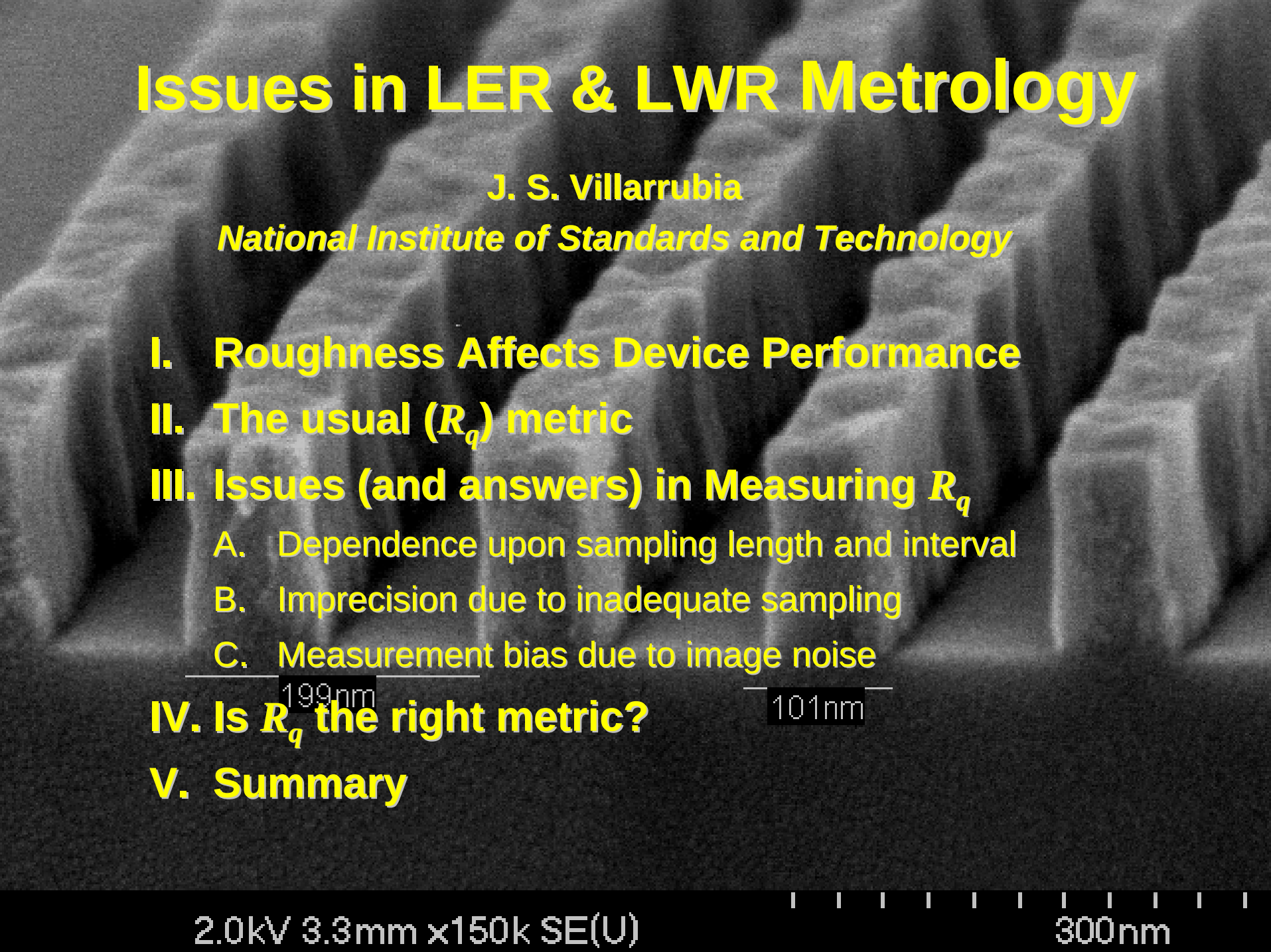
A. Dependence upon sampling length and interval

B. Imprecision due to inadequate sampling

C. Measurement bias due to image noise

IV. Is R_q the right metric?

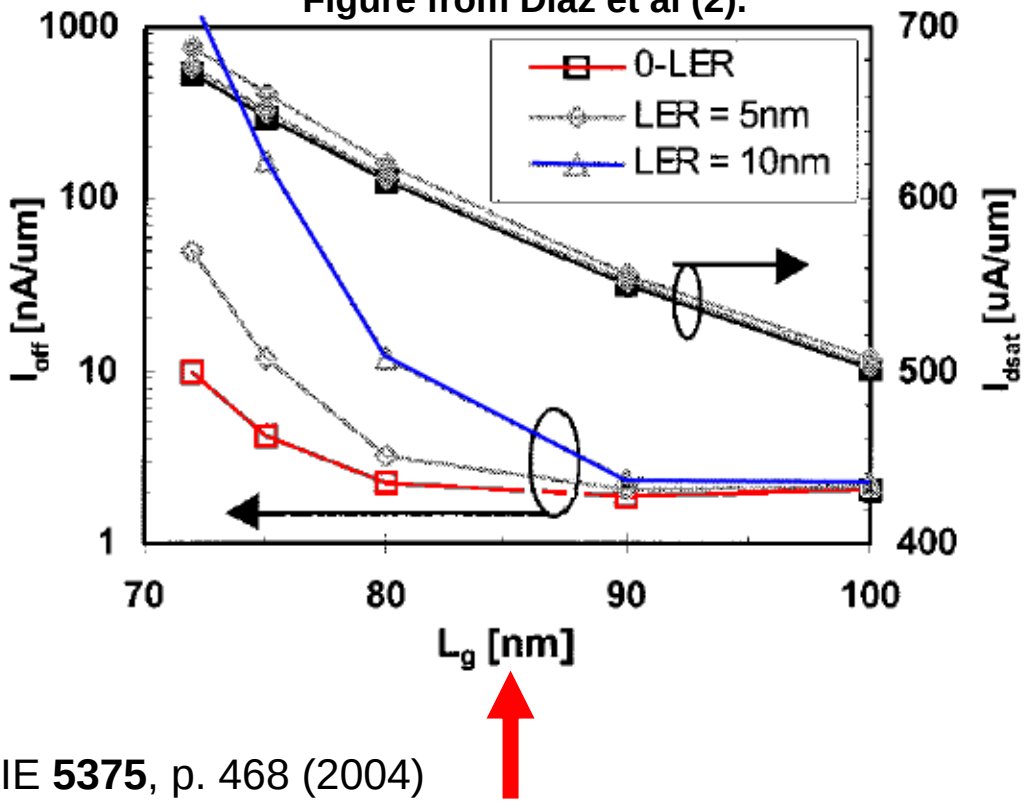
V. Summary



I. Roughness Affects Device Performance

- Edge or Width roughness has been associated with
  - Changes (and variability) in transistor threshold voltage
  - Significantly increased off state leakage
  - Long λ LER causes local variation in transistor to transistor performance

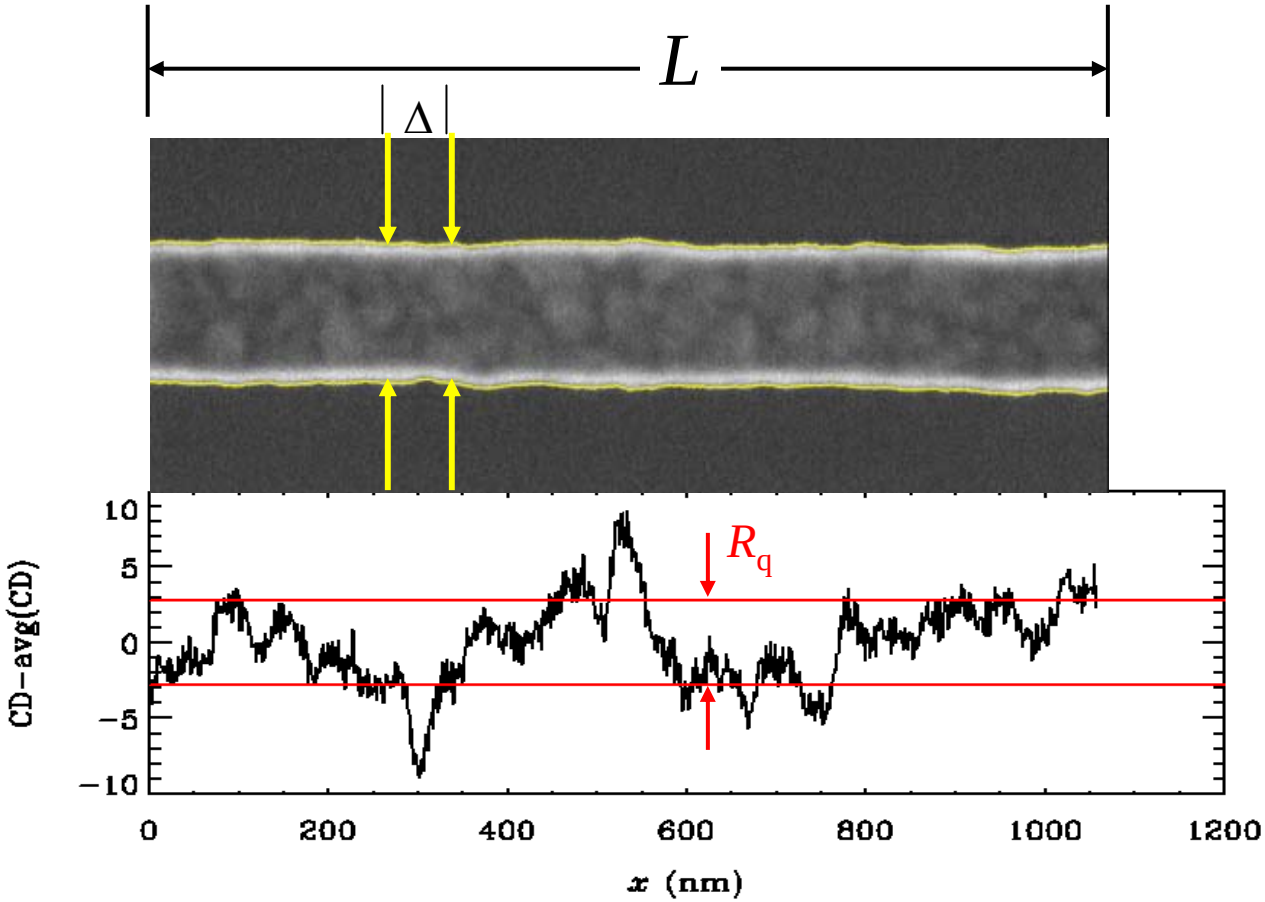
Figure from Diaz et al (2).



See for example,
(1) Yamaguchi et al., Proc. SPIE **5375**, p. 468 (2004)
(2) Diaz et al., IEEE Electron Dev. Lett. **22**(6) p 287 (2001)
(3) Xiong & Bokor, SPIE **4689** p. 733 (2002)

II. The R_q metric

R_q^2 is the Roughness *Variance*



Measure CD at *N* discrete intervals, Δ, and compute: $R_q^2 = \frac{1}{N-1}\sum_{i=0}^{N-1}\left(w_i - \overline{w}\right)^2$

R_q is a standard deviation roughness metric

- It (or $3R_q$) is used almost universally in the semiconductor industry.
- Specifications in the ITRS are stated in terms of $3R_q$.

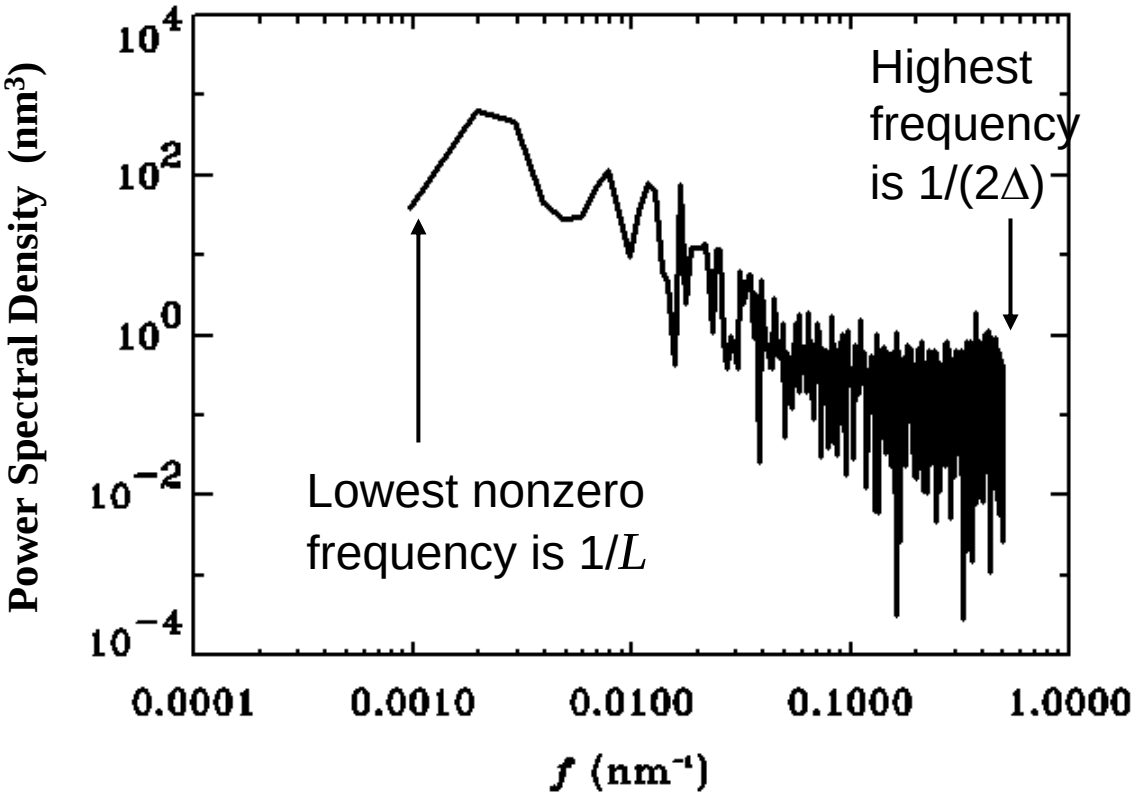
Table 117a Lithography Wafer Metrology Technology Requirements—Near-term

Year of Production	*2003*	*2004*	*2005*	*2006*	*2007*	*2008*	*2009*
Technology Node		*hp90*			*hp65*		
DRAM ½ Pitch (nm)	*100*	*90*	*80*	*70*	*65*	*57*	*50*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)	*120*	*107*	*95*	*85*	*76*	*67*	*60*
MPU/ASIC Un-contacted Poly ½ Pitch (nm)	*107*	*90*	*80*	*70*	*65*	*57*	*50*
MPU Printed Gate Length (nm)	*65*	*53*	*45*	*40*	*35*	*32*	*28*
MPU Physical Gate Length (nm)	*45*	*37*	*32*	*28*	*25*	*22*	*20*
Line width roughness (nm, 3 σ) <8% of CD ***	◆ 3.6	3.0	2.6	2.2	2.0	1.8	1.6
Wafer CD metrology tool precision (nm) * (P/T=.2) for LWR***	0.72	0.592	0.512	0.4	0.4	0.352	0.32
Maximum CD measurement bias (%) [B]	10	10	10	10	10	10	10

III. Issues in measuring R_q

Issue #1: Dependence upon sampling length and interval

Power Spectral Density

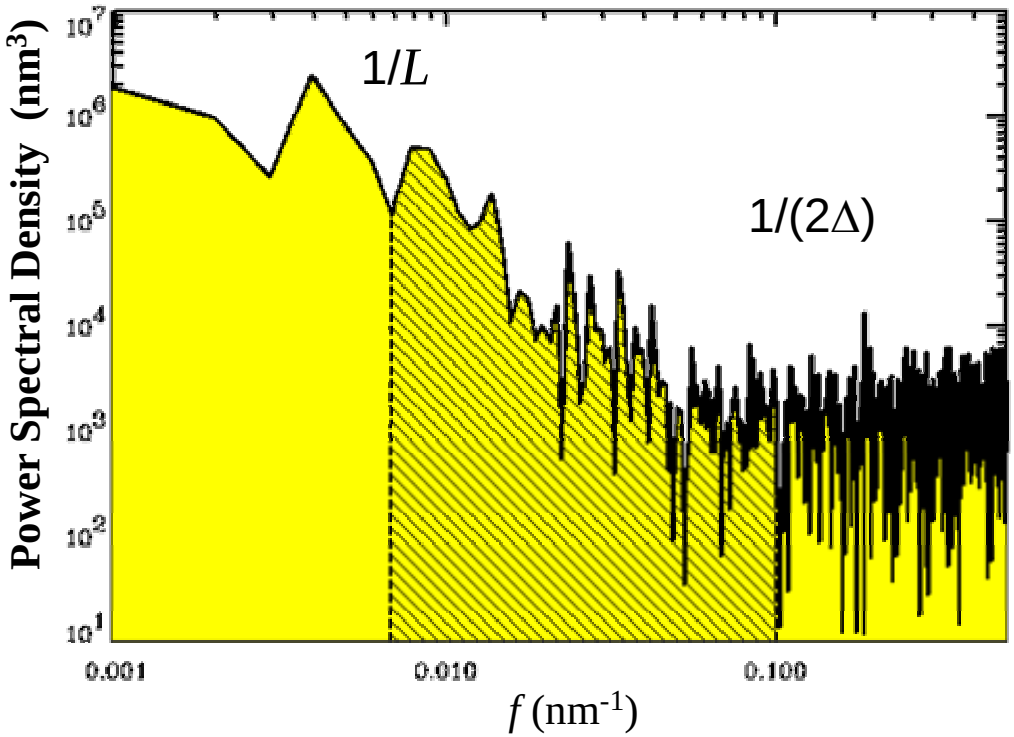


Apart from normalization, PSD is the Fourier transform squared.

Relationship between PSD and R_q

Parseval's theorem relates roughness to area under the PSD curve

$$\left(N-1\right)R_q^2 = \text{Area}$$

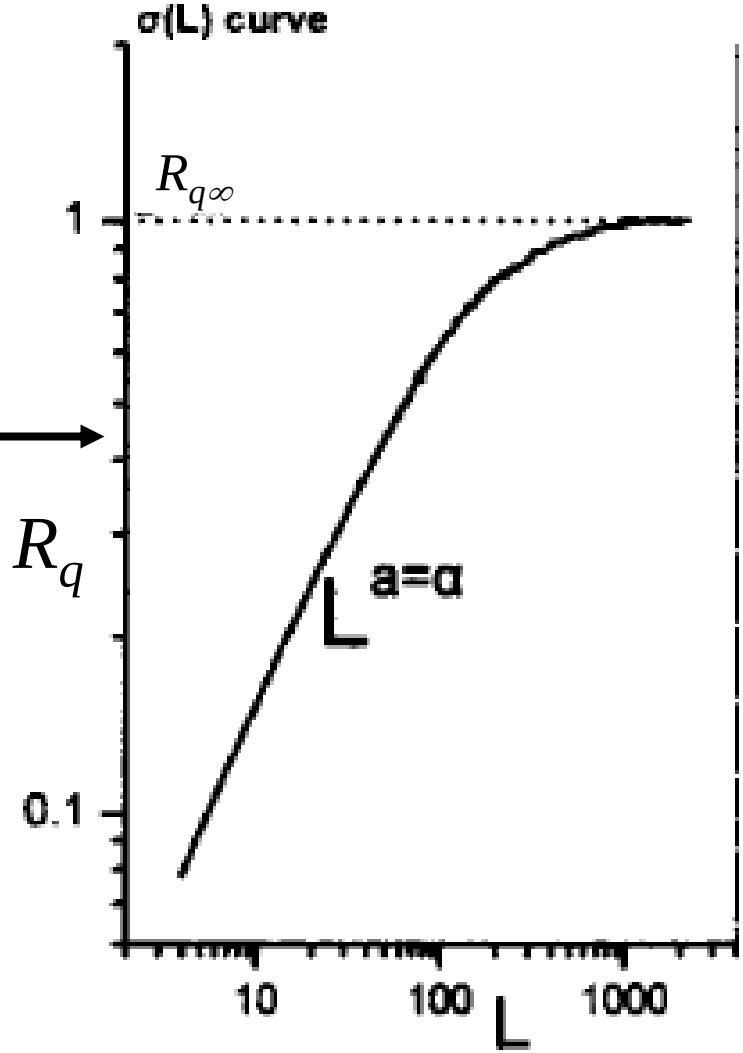


- Any real measurement will have $\Delta > 0$ and $L < \infty$
- Only roughness frequencies between $1/L$ and $1/(2\Delta)$ will have been sampled.
- Obviously, the area (i.e., roughness) depends upon these limits.

Sampling Dependencies

- R_q generally depends upon L as seen in this figure reproduced from Constantoudis et al.

Figure (with vertical axis relabeled by me) from Constantoudis et al., JVST B **22**, 1974-1981 (2004).



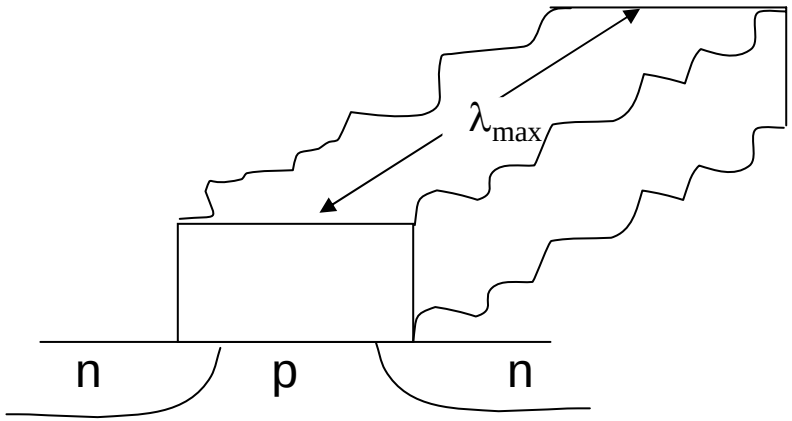
- R_q also generally depends upon Δ.

Common fallacies: (1) Comparing part to whole of segment. (2) Comparing meas. from instruments with different settings.

Know Which Wavelengths to Measure

L and Δ should be chosen based upon which roughness wavelengths affect device performance.

Choosing λ_{max}



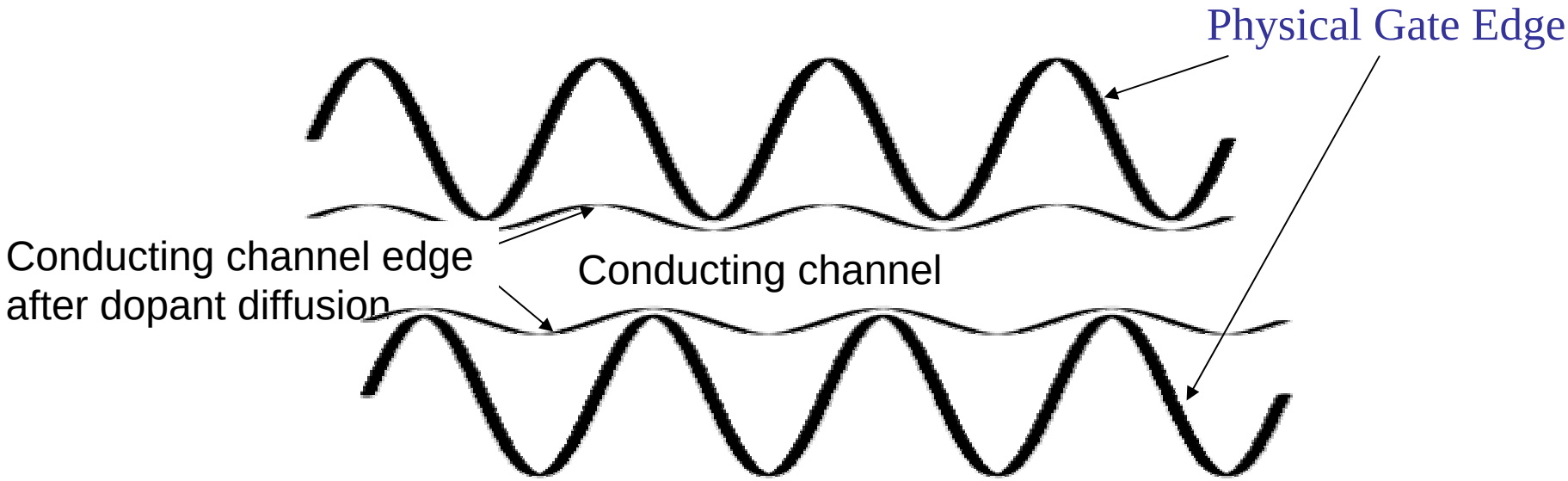
ITRS: Longest Roughness λ_{max} = DRAM pitch (180 nm for 90 nm node). I.e., it's defined as a multiple of the transistor's *short* dimension.

Why not a multiple of the transistor's *long* dimension?

Know Which Wavelengths to Measure

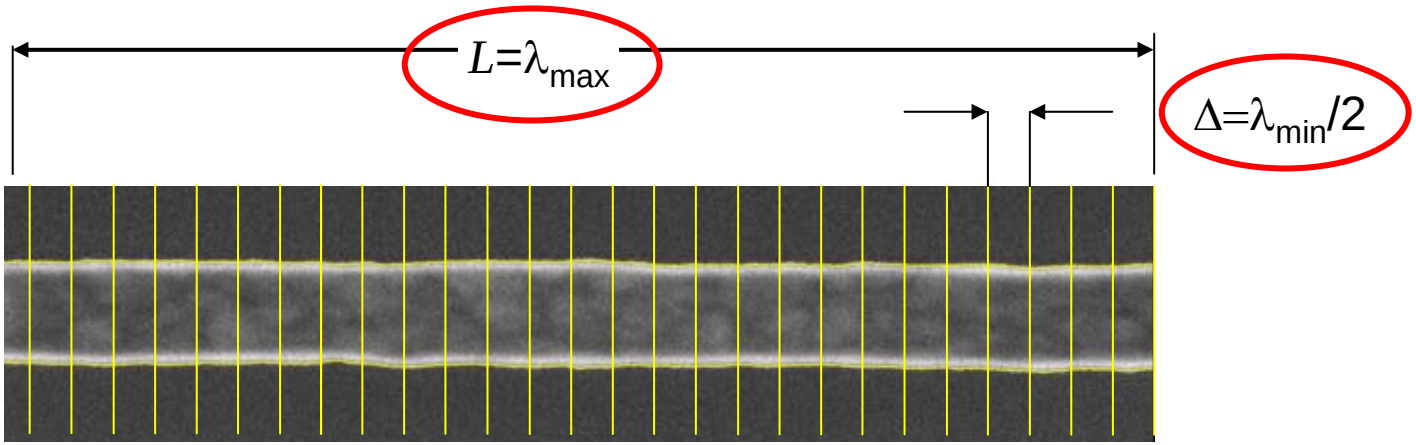
Choosing λ_{min}

The ITRS currently does not specify a shortest wavelength. This is a problem because it means the quantity we are to measure is not defined.



A reasonable choice (for roughness in poly gates) might be a dopant diffusion length (~10 nm), since roughness at wavelengths shorter than this will be smoothed by diffusion.

Choose *L* and Δ Accordingly



L = length of line segment to measure
Δ= measurement interval

L should equal longest wavelength you care about
Δ should be ½ the shortest wavelength you care about

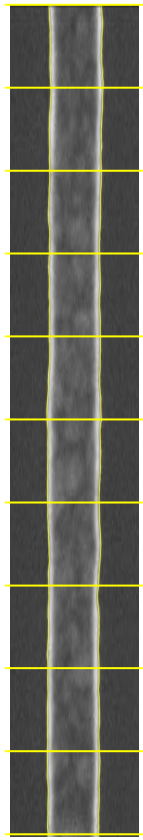
Issue #2 Imprecision due to inadequate sampling

Sampling Error

- Roughness is random
  - Every line segment is different from every other
  - Roughness varies from sample to sample
- How do we sample to determine roughness with small enough sampling error?
  - Let us say small enough means

$$\sigma_{R_q} / R_q \leq \eta$$

η = 0.2 for example



R_q = 2.0 nm

1.3

2.3

1.8

2.0

3.5

1.3

2.3

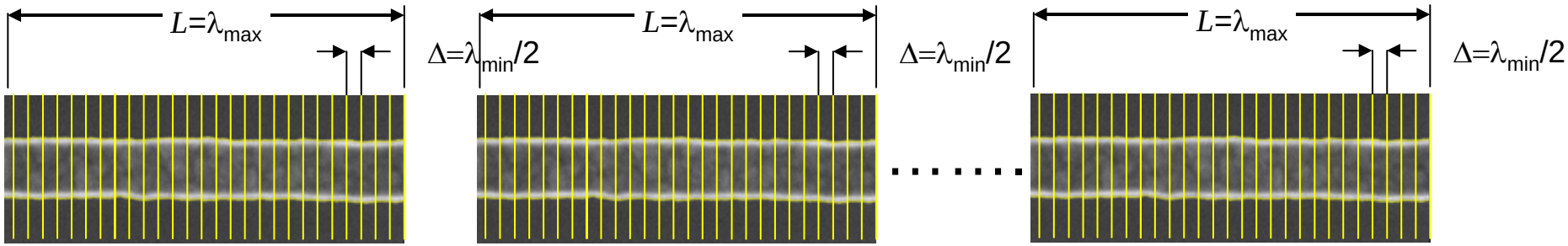
1.0

1.5

σ_{Rq} = 0.7

σ_{Rq} / R_q = 35%

Approach: Measure More Segments



Segment 1 Segment 2 Segment *m*

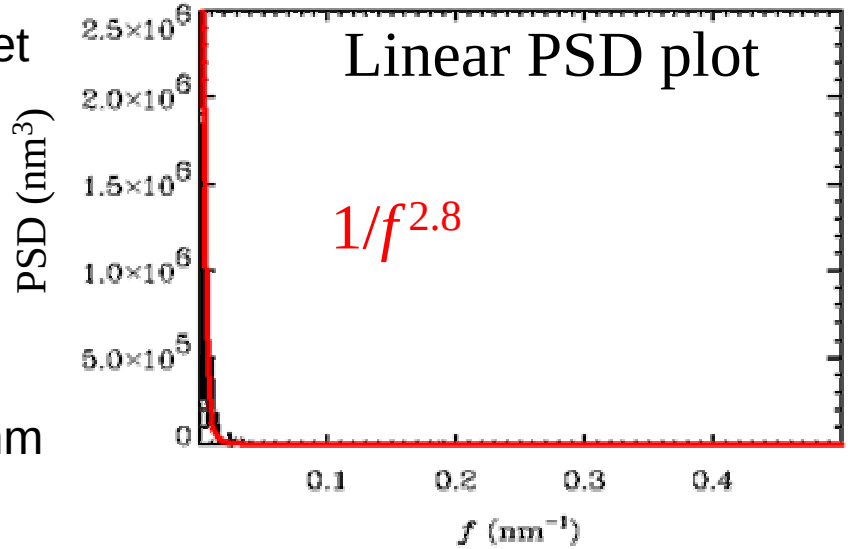
1. Measure *m* different segments. *m* depends on PSD.
2. Average the R_q^2 to obtain spec'd precision

$$m \geq \frac{\left(\sum_{k=0}^{N/2} \overline{P_k}\right)^2}{4\eta^2 \left(\sum_{k=0}^{N/2} \overline{P_k}\right)^2}$$

For the derivation, see Bunday et al., Proc. SPIE **5375**, p. 515 (2004)

Some Numbers for a Typical Measurement

- Using PSDs we measured, we get the following requirements for 90 nm node:
  - Δ = 7.5 nm (This is $\lambda_{min}/2$.)
  - L_{int} = 180 nm (This is λ_{max}.)
  - m = 4
- Total measured $L = mL_{int}$ = 720 nm (8 × node).

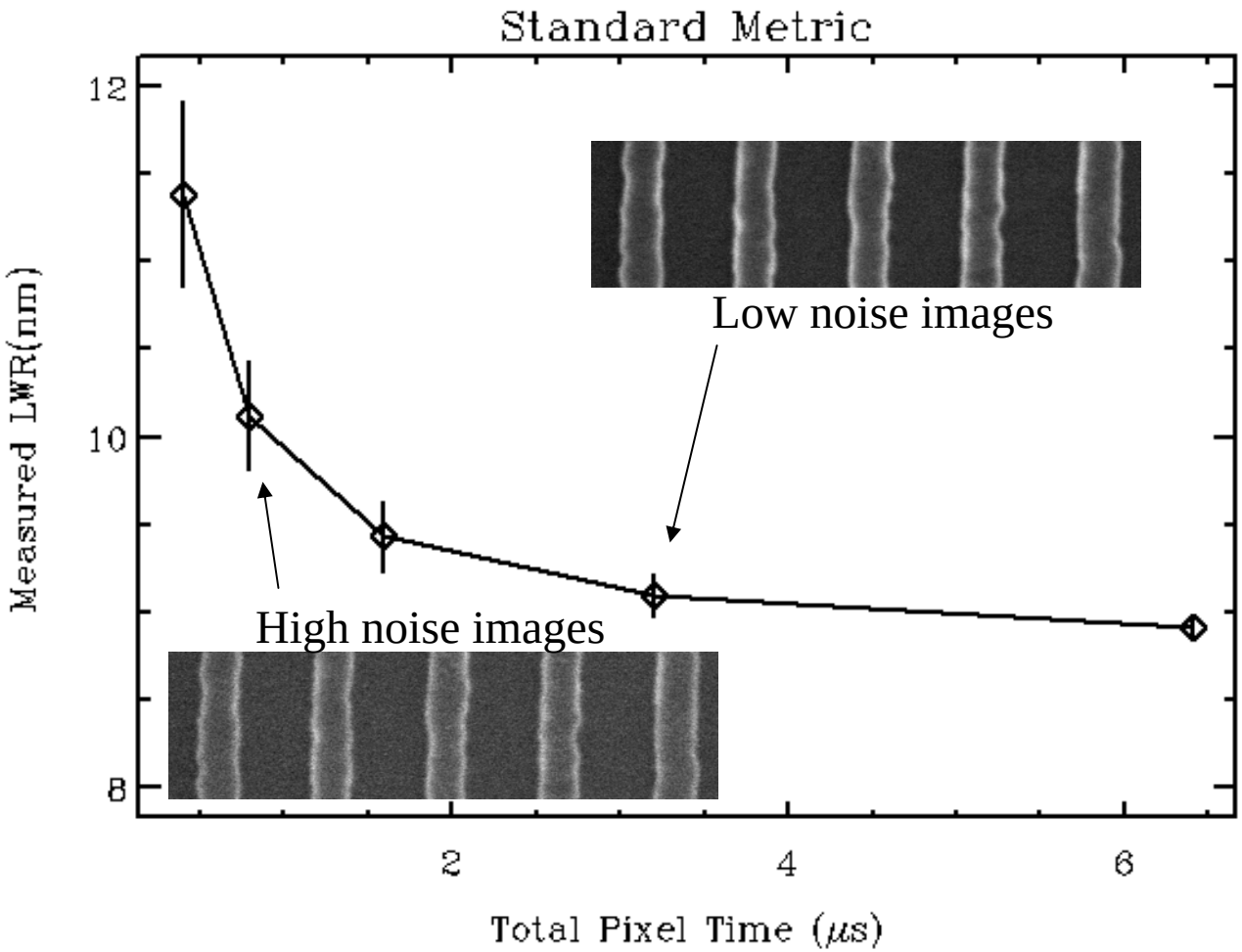


- Length of line to measure depends upon PSD
- Semiconductor samples we've measured have roughness proportional to $1/f^2$ to $1/f^3$.
- For roughness in this range, you must measure 6 × node to 10 × node to keep sampling error < 20%

Issue #3 Measurement bias due to image noise

Why do we think LWR measurements are subject to bias?

Images of the **same** spot have more or less LWR (using a simple standard deviation metric) depending upon whether they have more or less noise.



The Bias is a Large Effect

$$\left\langle R_q^{\,2} \right\rangle = R_{\mathrm{t}}^{\,2} + \sigma_{\varepsilon}^{2}$$

σ_ε is the CD "static repeatability" under LWR measurement conditions.

It is likely to be worse in a LWR measurement than in a CD measurement, because you must limit averaging along the line (to avoid averaging away short-wavelength roughness that you want to measure).

We get σ_ε of about 0.8 nm. ($3\sigma_\varepsilon$ = 2.4 nm)
ITRS spec. says LWR must be < 2.6 nm (3σ), i.e., about the same

What can we do about it?
A proposed measurement procedure

1. Decide the total electron dose (dwell time × beam current × # frames) you want to use.
2. Take *two* images, each at *half* of this dose.
3. Determine CDs and LWR variances from the two images.
4. From the CD differences between the two images you can estimate σ_ε^2 as:

$$\sigma_\varepsilon^2 = \frac{1}{2N}\sum_{i=1}^{N}\left(W_{i1} - W_{i0}\right)^2$$

5. Our new metric: The best estimate of roughness variance is the average variance from our two images, minus σ_ε^2

$$R_q^2 = \frac{R_{01}^2 + R_{02}^2}{2} - \sigma_\varepsilon^2$$

For details see Villarrubia & Bunday, Proc. SPIE **5752** (2005) submitted.

Bias in R_0 and R_q

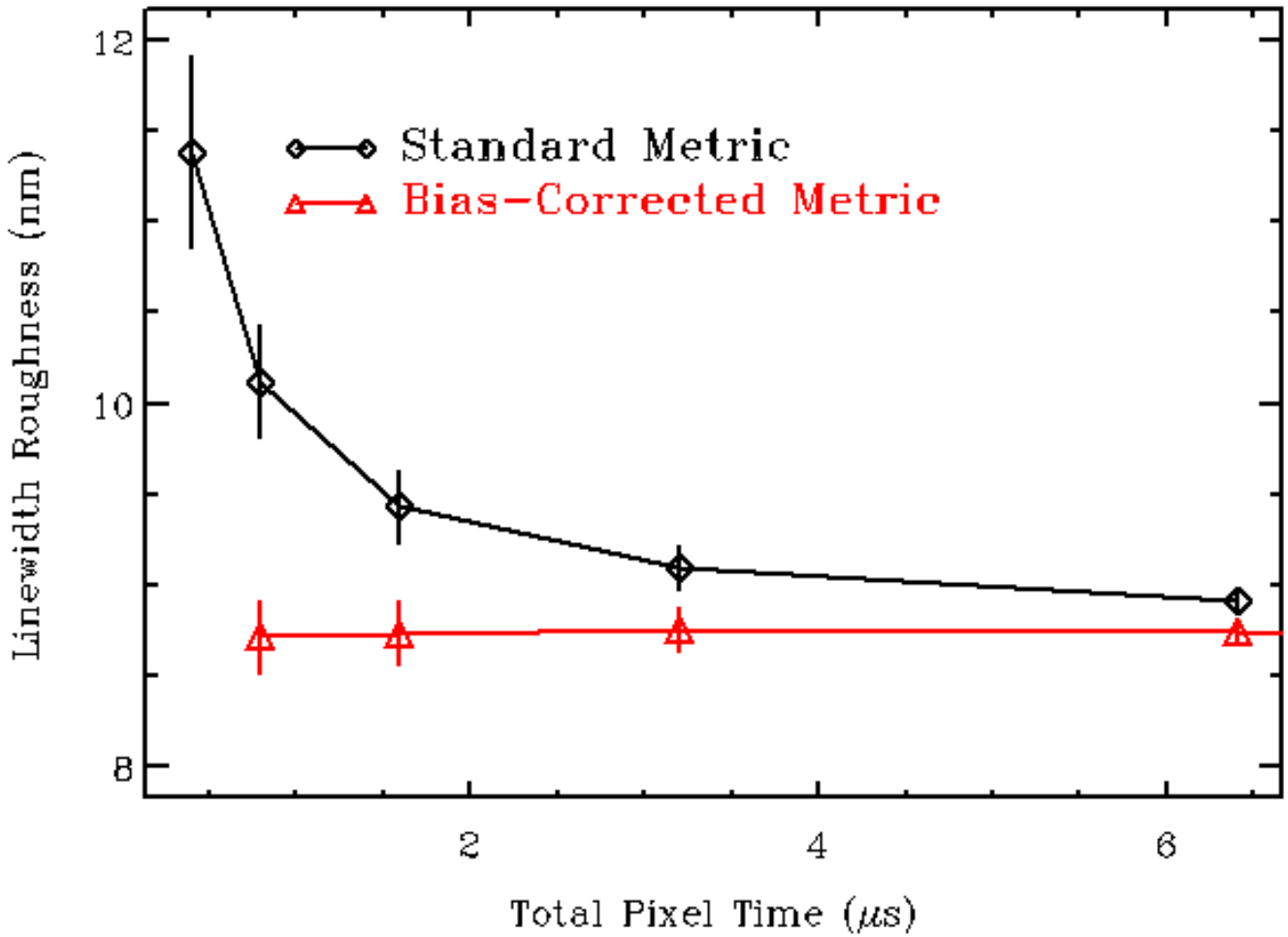


Figure from Villarrubia & Bunday, Proc. SPIE **5752** (2005) submitted.

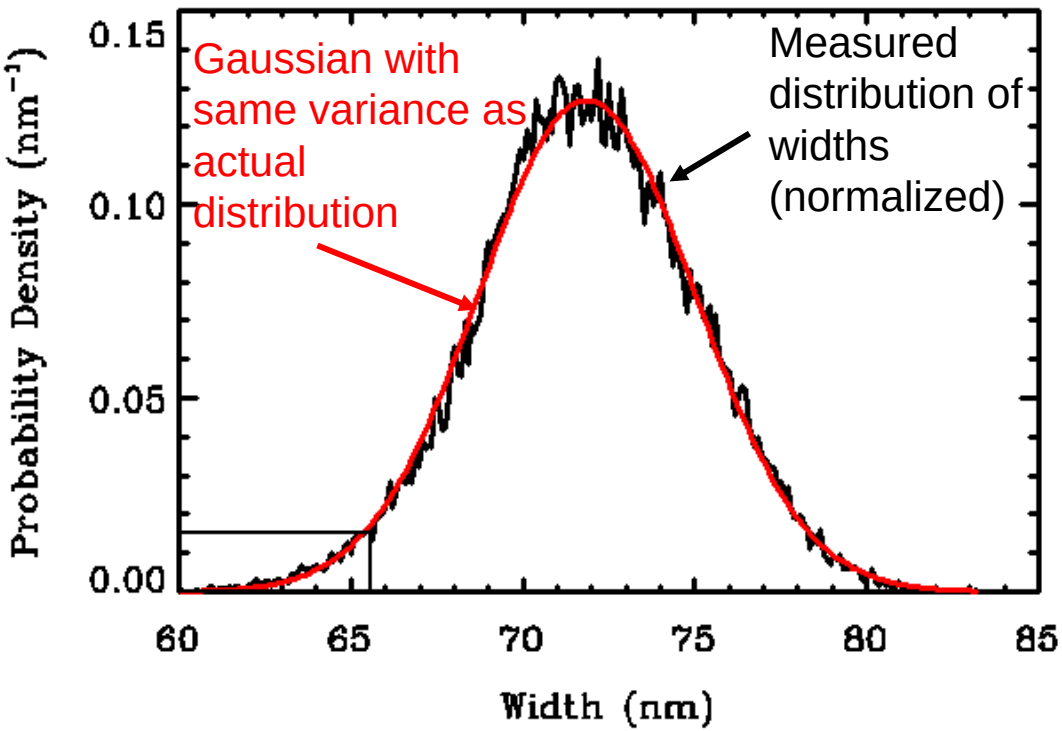
IV. Is R_q the right metric?

Observations and Questions about R_q

- We've seen that R_q sums all λs between $\lambda_{min} = 1/L$ and $\lambda_{max} = 1/(2\Delta)$
- This means λs are sorted into two kinds.
  - Those outside the range (0 weight)
  - Those inside the range (equal weights)
- Do all roughness λs in fact have either no impact on device performance or equal impacts? This seems unlikely.
- Therefore: the best PSD-based metric may require a more complicated weighting function. (The weights would no doubt be process-specific: Contact hole edge roughness affects performance differently than poly gate edge roughness.)

Amplitude Density Function

- The best metric may not be PSD- (or standard deviation-) based at all. Consider the amplitude density function (ADF):



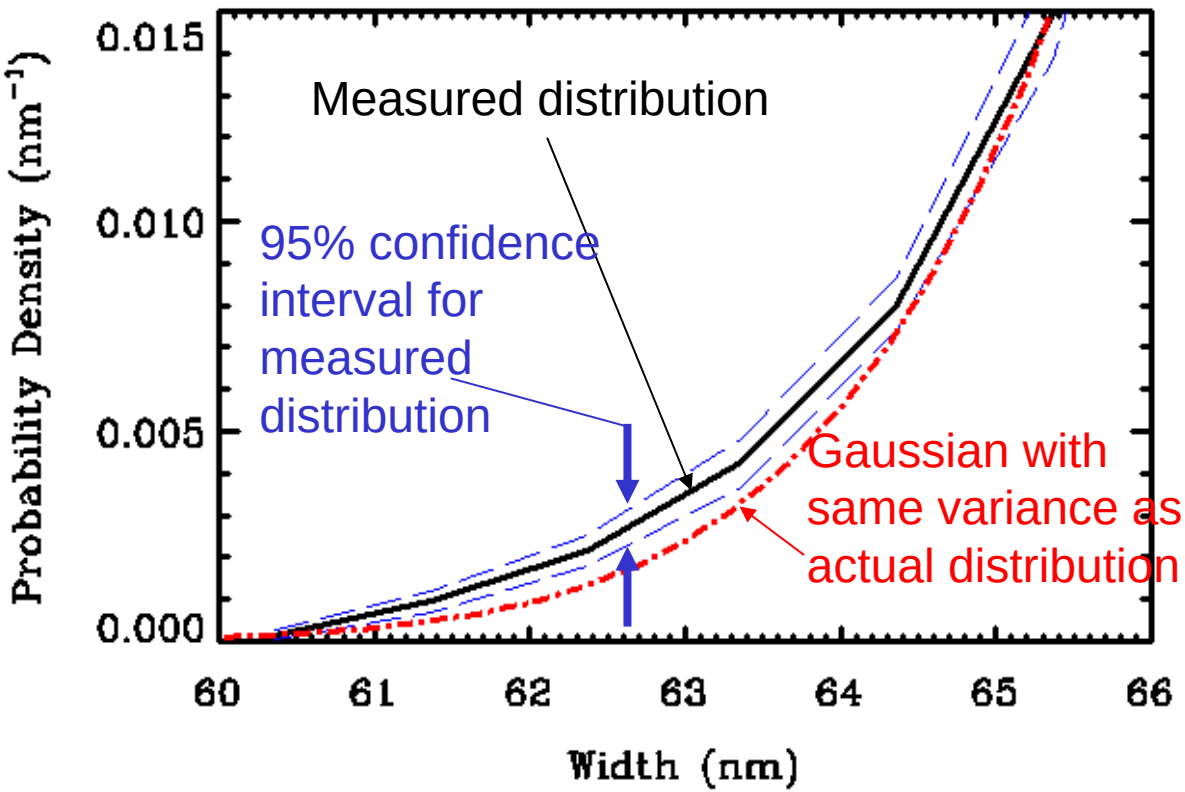
The ADF is just the width histogram normalized to unit area.

It is the width probability distribution.

It is usually close to Gaussian.

Amplitude Density Function

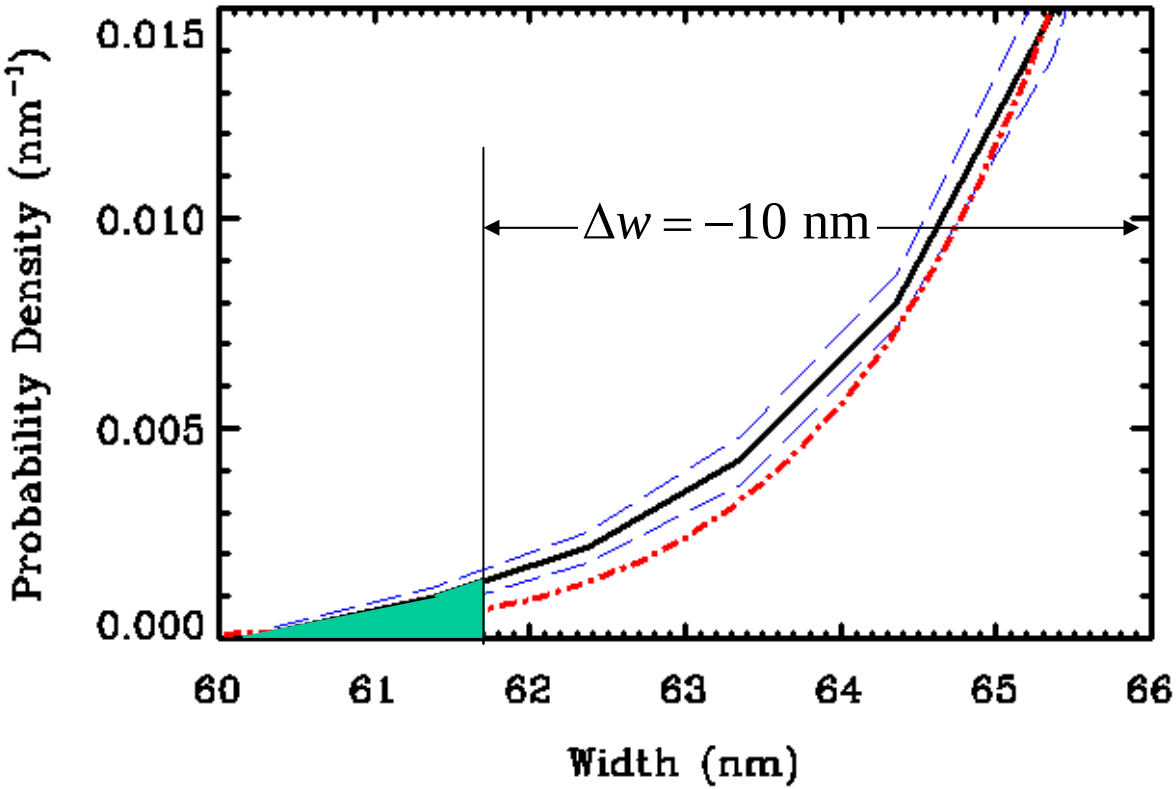
- But real distributions often differ from ideal (Gaussian) distributions in the tails—as does this one.



Suppose it's the tail of the distribution that is where all the action is?

E.g., Suppose a few unusually narrow gates are the yield killers. –or suppose leakage current is determined almost entirely by the channel's narrowest length.

Possible metrics based on ADF



A probability metric:

$$\Delta w = w - \overline{w}$$

$$P_{10} = P\left(\Delta w < -10 \text{ nm}\right) = 0.1\%$$

Or, **a percentile metric:**

$$\Delta w\left(P = 0.1\%\right) = -10 \text{ nm}$$

The Gaussian fit underestimates the probability by about 50%.
Clearly, if this metric is what we care about, we should estimate it directly from the tail of the measured ADF, not indirectly from R_q.

Summary I

- We have addressed several measurement issues for the usual "3σ" roughness metric
  - Problem #1: LWR changes when you change L or Δ.
  - Solution: Choose L and Δ "on purpose"
  - Problem #2: Sampling error. LWR is random. The patch you measure may differ from the average.
  - Solution: Measure many different patches. How much is enough?—We know how to answer that.
  - Problem #3: Measurement is significantly biased. "Noise roughness" adds to actual LWR.
  - Solution: At least 2 repeats for each patch allows "Noise roughness" to be measured independently and a correction applied. –We know how to do this too.

Summary II

- Our favorite standard deviation measure of roughness may not be the best
  - Other metrics exist.
  - Metrics should be process driven.
  - Too little is currently known about how LWR affects the process to be certain of the proper metric.
  - A plausible case can be made for some alternative metrics—e.g., some based upon amplitude density function. I showed you two:
    - Probability for a stated width deviation
    - The width deviation that corresponds to a stated probability
  - We should remain open-minded about metrics until more is known about the relationship between LWR and device performance.

Thank You…

To the organizers for inviting me to speak to you today.

…To NIST's Office of Microelectronics Programs and the Manufacturing Engineering Laboratory's Nanomanufacturing Program for funding my time.

To those who have collaborated with me on work referenced in this talk, especially:

--Ben Bunday (SEMATECH)

--András Vladár (NIST)

--Michael Postek (NIST)

To you, for your attention