Line Edge Roughness

An introductory/intermediate course on
measuring, modeling and understanding LER

Gregg M. Gallatin

Center for Nanoscale Science and Technology
National Institute of Standards and Technology
Gaithersburg, MD 20899
gregg.gallatin@nist.gov
301-975-2140

Please read the following statement:

The full description of the procedures used in these course notes requires the identification of certain commercial products and their suppliers. The inclusion of such information should in no way be construed as indicating that such products or suppliers are endorsed by NIST or are recommended by NIST or that they are necessarily the best materials, instruments, software or suppliers for the purposes described.

Acknowledgements

Patrick Naulleau, Robert Brainard, Alex Liddle, Dimitra Niakoula, Elsayed Hassanein, Bruno La Fontaine, Tom Wallow, Adam Pawlowski, David van Steenwinckel, Robert Bristol, Pieter Kruit, Tor Sanstrom, Vivek Prabhu, Tony Novembre, Stuart Stanton, Victor Katsap, Yehiel Gotkis, Keith Standiford, Andrew McCullough…. And many others

References

It is impossible to list all the numerous references to LER data, LER modeling, LER metrology, LER effects, etc.
The references listed on the following pages are those papers which I myself have found to be useful.
I humbly apologize to the authors of any references not listed.

References

Data, General Properties, and Characterization of LER, Surface Roughness, and Resist Resolution:

He and Cerrina, J. Vac. Sci. Technol. 16 (1998)
William D. Hinsberg, et. al., Proc. SPIE 3999, 148 (2000)
Jonathan L. Cobb, et. al. Proc. SPIE 4688, 412 (2002)
J. A. Hoffnagle, et. al., Opt. Letts. 27, 1776 (2002)
J. A. Croon, et. al., IEDM (2002)
V. Constantoudis, et. al., J. Vac. Sci. Technol. B21, 1019 (2003)
William Hinsberg, et. al., Proc. SPIE 5039, 1 (2003)
Atsuko Yamaguchi and Osamu Komuro, Jpn. J. Appl. Phys. 42, 3763 (2003) Charlotte A. Cutler, Proc. SPIE 5037, 406 (2003)
Robert L. Brainard, et. al., Proc. SPIE 5374, 74 (2004)
Gerard M. Schmid, et. al., Proc. SPIE 5376, 333 (2004)
V. Constantoudis, et. al., J. Vac. Sci. Technol. B22, 1974 (2004)
Adam R. Pawloski, et. al., Proc. SPIE 5376, 414 (2004)
Gallatin, Proc. SPIE 5754 (2005)
Adam R. Pawloski, et. al., J. Microlith., Microfab., Microsys. 5(2), 023001 (2006)
Choi, et. al., Proc. SPIE 6519 (2007)
Kang, et. al., Proc. SPIE 6519 (2007)
Steenwinckel, et. al., Proc. SPIE 6519 (2007)
Gallatin, et. al., Proc. SPIE 6921 (2008)
Hassanein, et. al. Proc. SPIE 6921 (2008)
Steenwinckel, et. al., J. Micro/Nanolith., MEMS, MOEMS. 7(2) (2008)
Naulleau, et. al., J. Vac. Sci. Technol. 26 (2008)

… and many many others…

References

Metrology and LER Data Analysis:

Goldfarb, et. al., J. Vac. Sci. Technol. B22, 647 (2004)
Constantoudis, et. al., Proc. SPIE 5752 (2005)
Villarrubia and Bunday, Proc. SPIE 5752 (2005)
Katz, et. al., Proc. SPIE 6152 (2006)
Constantoudis, et. al., Proc. SPIE 6518 (2007)
Wang, et. al., Proc. SPIE 6518 (2007)
Naulleau and Cain, J. Vac. Sci. Technol. 25 (2007)
Wang, et. al., Proc. SPIE 6922 (2008)
Yamaguchi, et. al., Microelectronic Eng. 84 (2007)
Constantoudis and Gogolides, Proc. SPIE 6922 (2008)
Naulleau, et. al., J. Vac. Sci. Technol. 26 (2008)
Orji, et. al., Proc. SPIE 7042 (2008)
Bunday, et. al., Proc. SPIE 6922 (2008)

… and many others…

SEM and AFM Limitations, Modeling and Analysis:

Stedman, Journal of Microscopy 152 (1988)
Villarrubia, J. Research NIST 102 (1997)
Cazaux, Nanotechnology 15 (2004)
Villarrubia, et. al., Surf. Interface Anal. 37 (2005)
Villarrubia, et. al., Proc. SPIE 6518 (2007)
Chris Mack, Microlithography World, August 2008.

… and many others…

References

LER and LER-Related Resist Modeling:

P. C. Tsiartas, et. al., Macromolecules 30, 4656 (1997)
L. W. Flanagin, et. al., Macromolecules 32, 5337 (1999)
F. A. Houle, et. al., J. Vac. Sci. Technol. B18, 1874 (2000)
Gallatin, Proc. SPIE 4404, 123 (2001)
S. D. Burns, et. al., J. Vac. Sci. Technol. B20, 537 (2002)
Hiroshi Fukuda, J. Photopolymer Sci. and Tech. 15, 389 (2002)
F. A. Houle, et. al., J. Vac. Sci. Technol. B20, 924 (2002)
Hiroshi Fukuda, Jpn. J. Appl. Phys. 42, 3748 (2003)
Jonathan L. Cobb, et. al., Proc. SPIE 5037, 397 (2003)
Gallatin, et. al., J. Vac. Sci. Technol. B21, 3172 (2003)
William Hinsberg, et. al., Proc. SPIE 5039, 1 (2003)
D. Fuard, et. al., Proc. SPIE 5040, 1536 (2003)
Kozawa, et. al., J. Vac. Sci. Technol. 21 (2003)
Lei Yuan and Andrew R. Neureuther, Proc. SPIE 5376, 312 (2004)
William Hinsberg, et. al., Proc. SPIE 5376, 352 (2004)
Robert L. Brainard, et. al., Proc. SPIE 5374, 74 (2004)
Michael D. Shumway, et. al., Proc SPIE 5374, 454 (2004)
Kruit, et. al., J. Vac. Sci. Technol. 22 (2004)
Patrick P. Naulleau, Appl. Opt. 43, 788 (2004)
Okoroanyanwu and Lammers, Future Fab. Intl. 17, Chapter 5, Section 5 (2004)
Neureuther, et. al., J. Vac. Sci. Technol. B 24 (2006)
Rydberg, et. al., J. Microlitho., Microfab., Microsys. 5 (2006)
Kozawa, et. al., J. Vac. Sci. Technol. B. 25 (2007)

… **continued**…

References

Continued…LER and LER-Related Resist Modeling:

Gallatin, et. al., Proc. SPIE 6519 (2007)
Bristol, Proc. SPIE 6519 (2007)
Drygianakis, et. al., Proc. SPIE 6519 (2007)
Gotkis, IEPBN Meeting, Portland Oregon, 2008
Gotkis, IEUVI Resist TWG meeting, Feb 28, 2008
Saeki, et. al., Proc. SPIE 6923 (2008)
Kozawa, et. al., Applied Physics Express 1 (2008)
Chris Mack, LER Course, EUV Conference, Maui, (2008)

… and many many many others…

LER Effects on CD and Devices:

2003P. Oldiges, et. al., SISPAD (2000)
T. Linton, et. al., IEDM (2002)
J. A. Croon, et. al., ESSDERC (2003)
Asen Asenov, et. al., IEEE Trans. Electron Dev. 50, 1254 (2003)
Goldfarb, et. al., J. Vac. Sci. Technol. B22, 647 (2004)
Chandhok, Proc. SPIE 6519 (2007)
Tsikrikas, et. al., Proc. SPIE 6518 (2007)
Drygiannakis, et. al., Proc. SPIE 6518 (2007)
Ma, et. al., Proc. SPIE 6518 (2007)

… and many others…

OUTLINE

- **What is Line Edge Roughness (LER)?**

- **How is it measured and quantified?**
  - SEM vs AFM
  - σ values
  - PSD, Autocorrelation, HHCF,
  - Resist Blur, Correlation Length, Slopes, Roughness Exponent

- **How does it behave?**
  - Dependence on Dose, ILS, Blur, Frequency, Base Loading, PAG loading, …

- **The many causes of LER?**

- **Modeling LER…** *or at least the dominant contribution to LER*

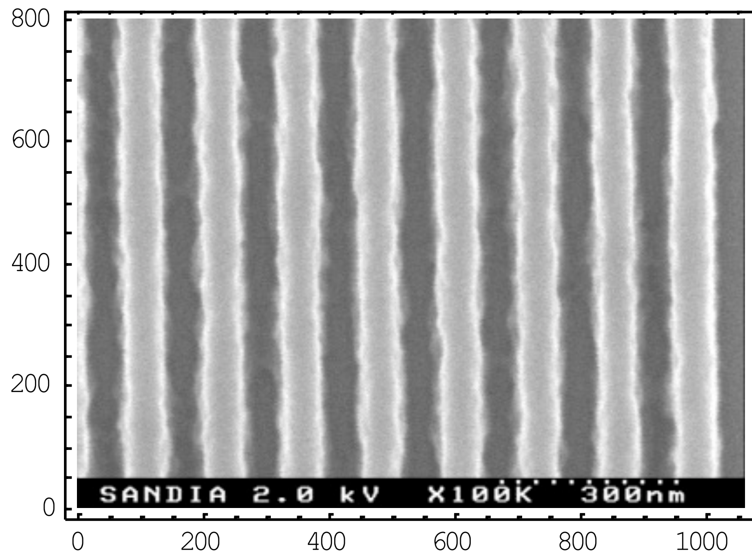
- **Consequences?**
  - The RLS tradeoff or Lithographic Uncertainty Principle
    - ➔ “Resolution, LER, Sensitivity…Pick any two”

- **How does it affect CD variation and device performance?**

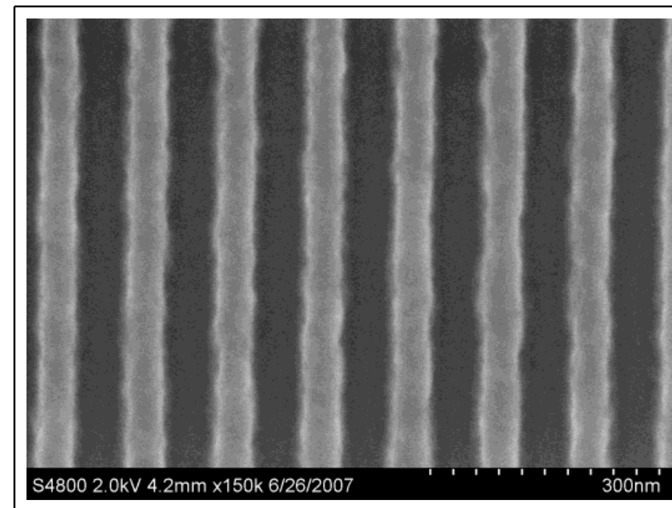
- **Can it be fixed?   Does it need to be fixed?**

LER Examples

Early EUV SEM
Sandia/Livermore

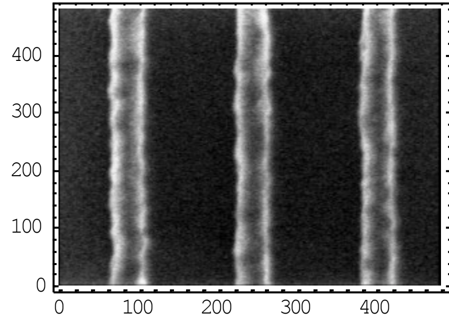


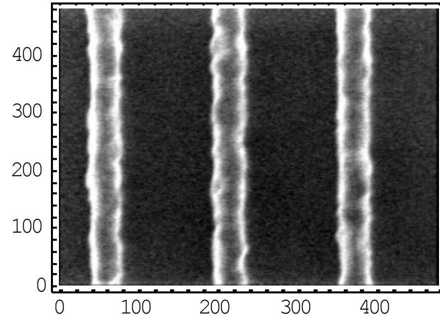
Recent EUV SEM
Imaged on MET at LBNL
Naulleau, et. al.
Resist: MET2D

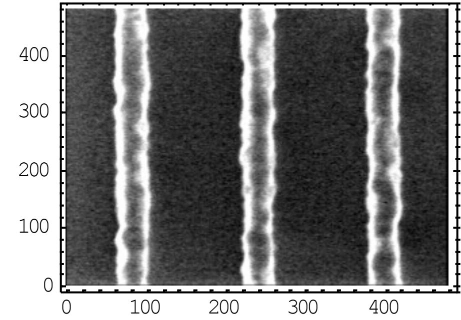


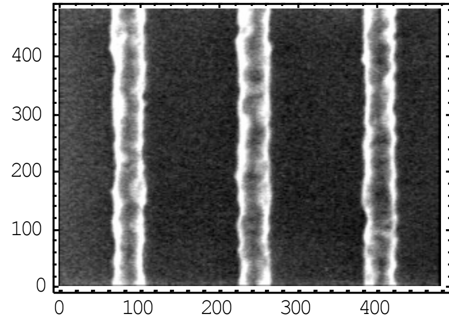
We will use this as an example case for studying LER analysis. For the most part the data from this SEM has been processed using the SuMMIT software (www.euvl.com/summit) program for doing LER analysis.

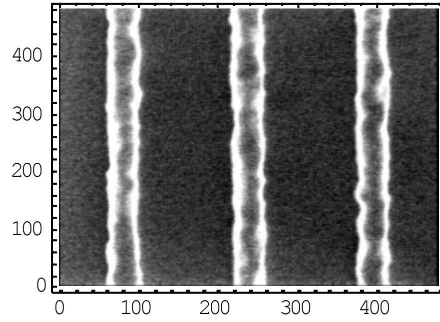
193 LER Examples

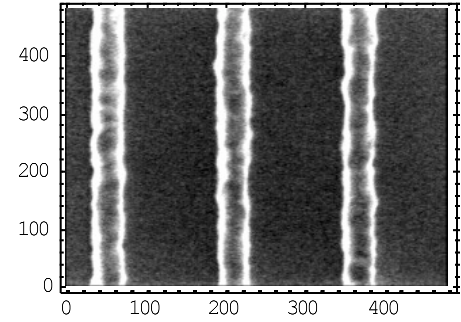


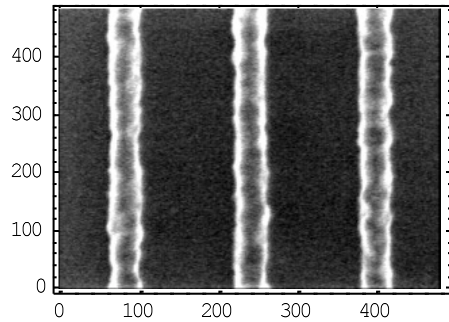


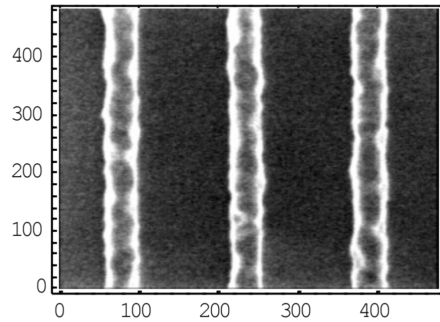


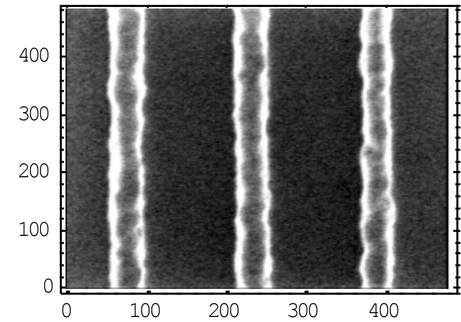


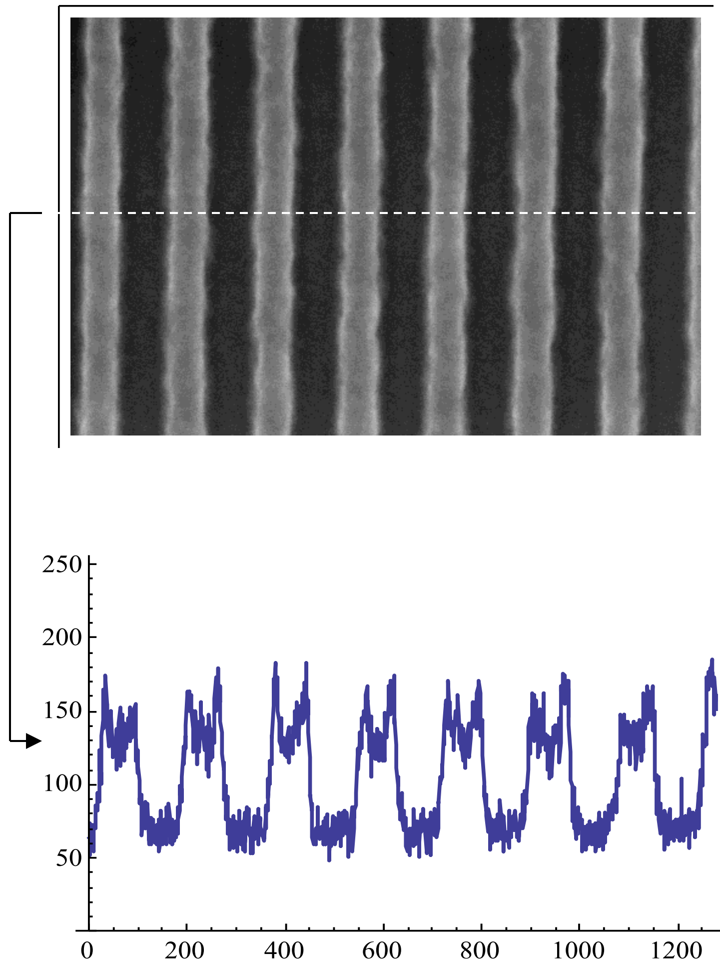












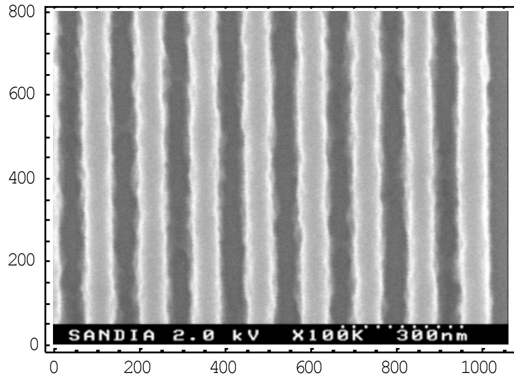
Edge Finding Algorithms

1. Threshold
2. Max slope
3. Threshold a polynomial fit to the image data in the region of the edge.
4. Max slope polynomial fit
5. Sigmoidal fit….
6. etc, etc, etc…

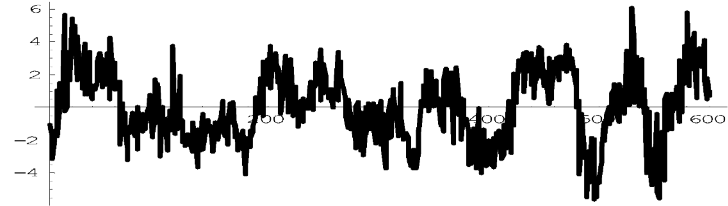
➔ Once the edge data is obtained subtract a best fit straightline from the edge data to **remove**
- Average edge position
- Edge “run-out”
  - Removes SEM image rotation
  - Want just the roughness, not the linear (“flat”) variation

Determining LER from a top-down CD SEM

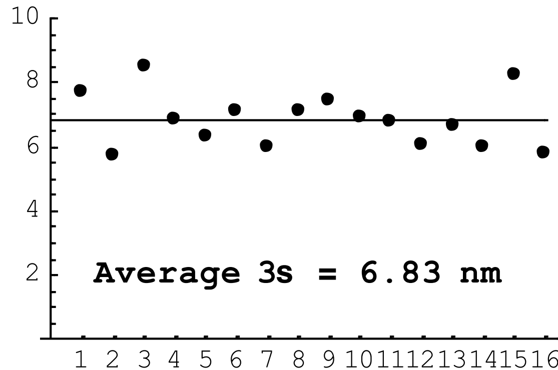
1. Apply edge finding algorithm and determine best fit straight line for each edge



2. Subtract straight line fit to obtain roughness residual

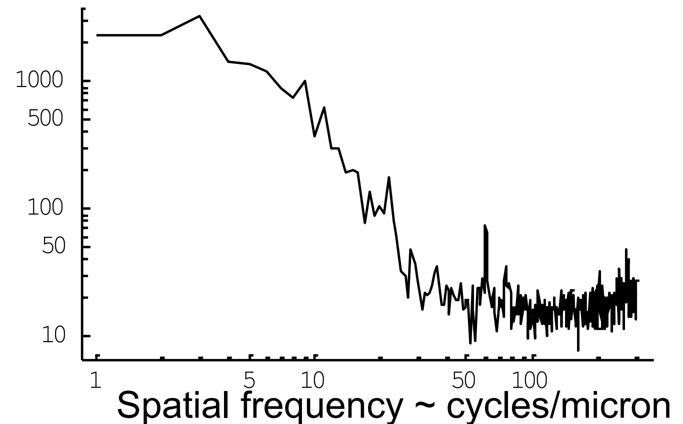


3. Compute things… such as 3σ per edge



NOTE: σ values computed by standard algorithms are generally biased and can be systematically larger or smaller than the "real" values. (see discussion of PSD).

…and/or FFT and square to obtain Frequency content



Spatial frequency ~ cycles/micron

From the 2007 ITRS Roadmap

MPU gate CD control requirements will stress many other aspects of lithography process control, including lenses, resist processing equipment, resist materials, and metrology. Process control, particularly for overlay and CD, is a major challenge. In addition, inherent roughness on small gate widths starts to add a non-negligible component to CD uniformity. It is unclear whether metrology, which is fundamental to process control, will be adequate to meet future requirements as needed for both development and volume manufacturing. Resist line edge roughness (LER) is becoming significant, as gate line width control becomes comparable to the size of a polymer unit. Next-generation lithography will

THE INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS: *2007*

feasible. Metrology will play a critical role in defining these lithography friendly design rules. The effects of line edge and line width roughness (LWR) are also becoming increasingly apparent in device performance; therefore, metrology tools need to be modified to accurately measure these variations as well. High frequency line width roughness affects dopant concentration profiles and affects interconnect wire resistance. Line width roughness at larger spatial frequency results in variations of transistor gate length over the active region of the device. This variation increases leakage of transistors and causes a variation of the speed of individual transistors, which in turn leads to IC timing issues. The line width and line edge roughness also provide a contribution to the CD uniformity error budget for small gate lengths and long LER/LWR correlation lengths. The CD uniformity component from LER/LWR is likely to drive the required LER/LWR numbers even more aggressively than in prior roadmaps (this will be addressed in the 2008 ITRS Update). Because of the particular challenges associated with imaging contact holes, the size of contact holes after etch will be

THE INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS: *2007*

From the 2007 ITRS Roadmap

A specific challenge for lithography modeling and simulation is to accurately predict the behavior of state-of-the-art photoresists over a wide range of imaging and process conditions. For these, better physical/chemical models must be developed to predict three-dimensional resist geometries after development and process windows, including effects such as line-edge roughness. Better calibration techniques are required both for model development and for customizing models implemented in commercial tools to appropriately describe the photoresists in question. Calibration obviously depends on the quality of input data, for example, CD measurements. Therefore, it is necessary to better understand and estimate measurement errors. Systematic errors should be dealt with by models of the measurement tools, such as CD-SEMs. With the growing importance of LWR and LER, lithography simulation needs to contribute to the assessment of their influence on device and interconnect performance (LER) and variability (LWR). Since the roughness of etched structures and not the resist pattern ultimately affects device performance, intimate coupling between resist and etching simulation is indispensable. Intimate links with etching simulation must also be established also to predict the geometry of non-ideal mask edges that frequently result from mask-making lithography steps.

*Low frequency line width roughness: (nm, 3 sigma) <8% of CD ******	1.2	1.1	1.0	0.8	0.8	0.7	0.6

***** *LWR_{Lf} is 3 σ σ deviation of spatial frequencies from 0.5 μm^{-1} to 1/(2*MPU ½ Pitch).*

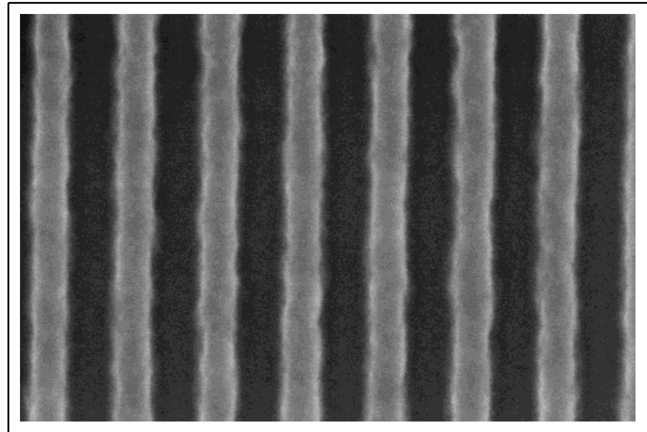
Note: Standard deviation is determined by biased estimate (corrected for SEM noise) of linewidth variation over a greater than or equal 2 μm measured at less than or equal 4 nm intervals.

Semi-Spec

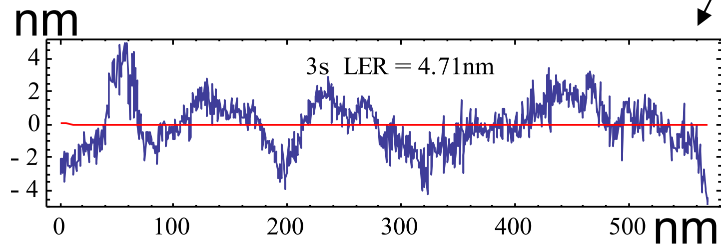
-2 micron long line ➔ Sets Minimum Spatial Frequency = ½ cycle/micron

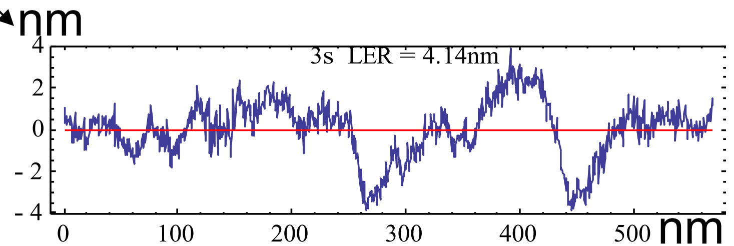
-10nm or less sample spacing ➔ Sets Maximum Spatial Frequency = 100 cycles/micron

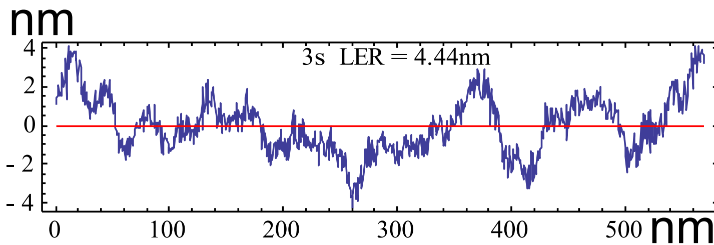
Example Edges

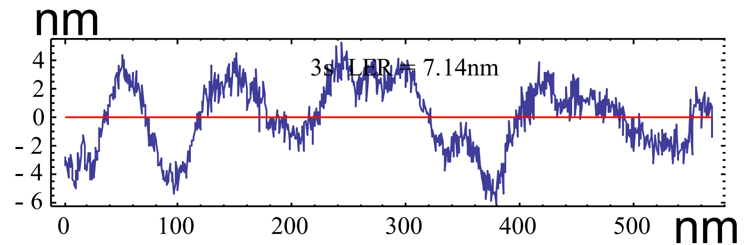


570nm
862 points









Computing σ

$$h(x) = \text{raw edge data}$$

$$H(x) = h(x) - \text{(best fit straight line)} \;\Rightarrow\; 0 = \frac{1}{L}\int_0^L dx\, H(x) = \frac{1}{L}\int_{-L/2}^{L/2} dx\, H(x)$$

$$\sigma^2 = \frac{1}{L}\int_0^L dx\, H(x)^2 = \frac{1}{L}\int_{-L/2}^{+L/2} dx\, H(x)^2$$

Data is discrete not continuous

$$\int_0^L dx = \int_{-L/2}^{+L/2} dx = \Delta x \sum_{n=1}^{N} \quad = \frac{L}{N}\sum_{n=1}^{N}$$

$$1\sigma = rms = \sqrt{\frac{1}{L}\int_0^L dx\, H(x)^2} = \sqrt{\frac{1}{L}\frac{L}{N}\sum_{n=1}^{N} H_n^{\,2}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} H_n^{\,2}}$$

$$= \sqrt{\frac{H_1^{\,2} + H_2^{\,2} + \cdots + H_N^{\,2}}{N}}$$

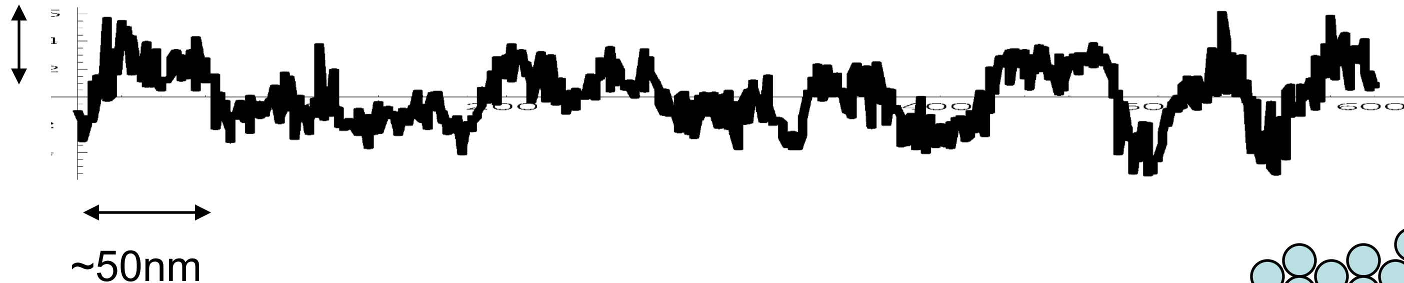
LER is commonly quoted as the 3σ value

NOTE: σ values computed this way are generally “biased”, i.e., they are systematically larger or smaller than the “real” values.

Low spatial frequency

➔ LER is long and slow compared to its amplitude

~6 nm Amplitude



Implication: LER is apparently not "atomistic" in origin

➔ Not driven by molecular weight.

Cutler, et. al., SPIE 03

Brainard, et. al., SPIE 04

Chemical and "atomistic" effects must contribute

but apparently they are not the dominant cause.

LER AutoCorrelation Function = ACF

$$\langle \ldots \rangle = \frac{1}{L}\int_{-L/2}^{+L/2} dx \ldots = \frac{1}{N}\sum_{n} \ldots$$

$$ACF(x, x') = \langle H(x) H(x') \rangle \qquad x' = x - s$$

$$= \frac{1}{L}\int_{-L/2}^{+L/2} dx\, H(x) H(x - s)$$

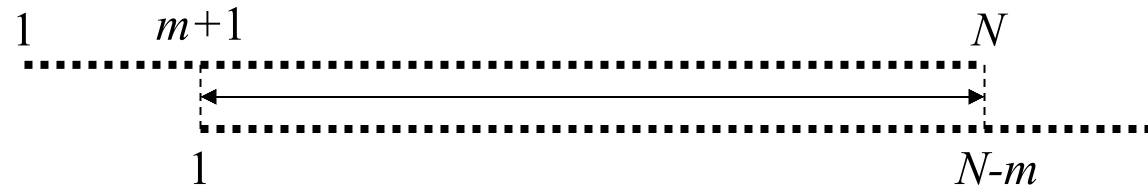
$$= \frac{1}{L}\sum_{n} \Delta x\, H(n\Delta x) H((n - m)\Delta x)$$

$$= \frac{1}{N}\sum_{n} H_n H_{n-m}$$

Shift, multiply, sum and normalize

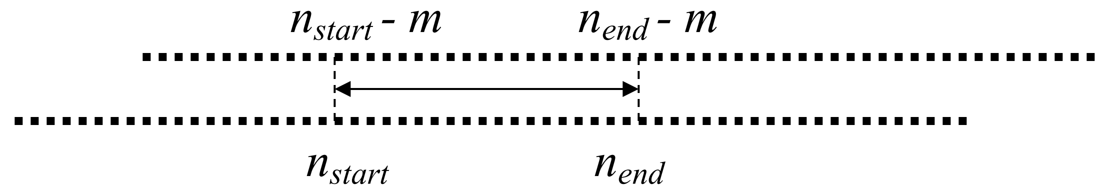
But the data has a finite length…
Need to adjust for this

Lop-off the ends



$$ACF_m = \frac{1}{N-m}\sum_{n=m+1}^{N-m} H_n H_{n-m} = \frac{H_{m+1}H_1 + H_{m+2}H_2 + \cdots + H_N H_{N-m}}{N-m}$$

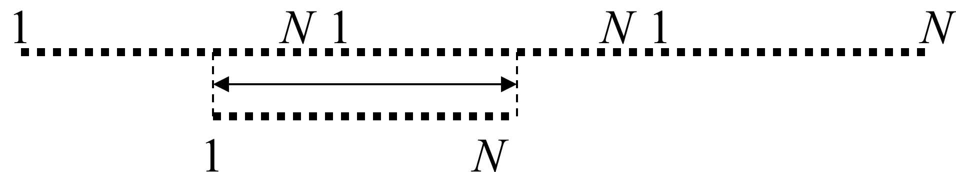
Take a fixed length out of the middle



$$ACF_m = \frac{1}{n_{end} - n_{start} + 1}\sum_{n=n_{start}}^{n_{end}} H_n H_{n-m} = \frac{H_{n_{start}} H_{n_{start}-m} + \cdots + H_{n_{end}} H_{n_{end}-m}}{n_{end} - n_{start} + 1}$$

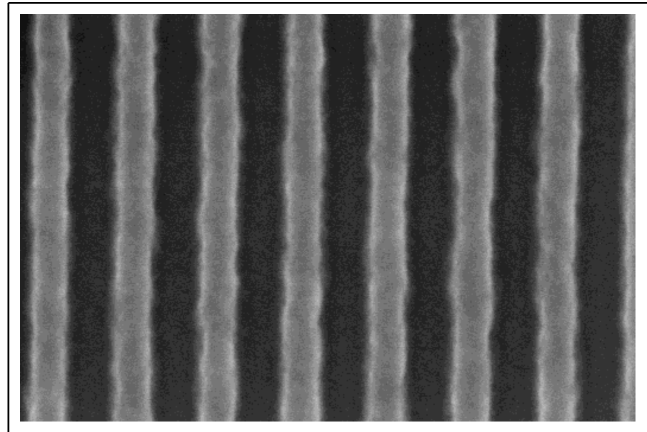
Assume Periodicity

Works well with FFT's

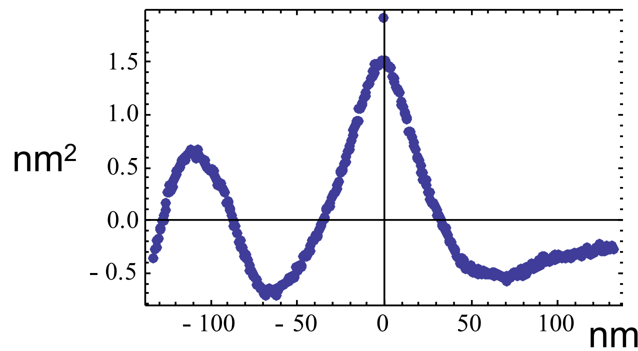


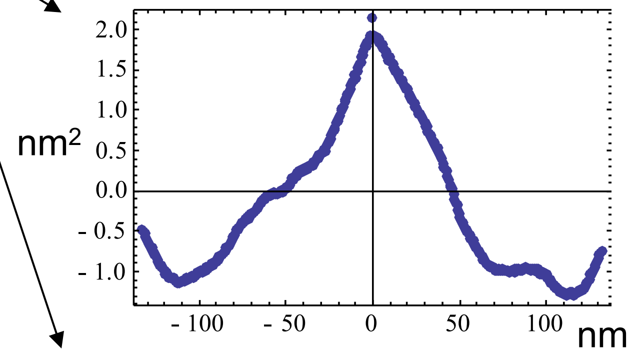
$$H_{n+N} = H_n \quad \Rightarrow \quad ACF_m = \frac{1}{N}\sum_{n=1}^{N} H_n H_{n-m}$$

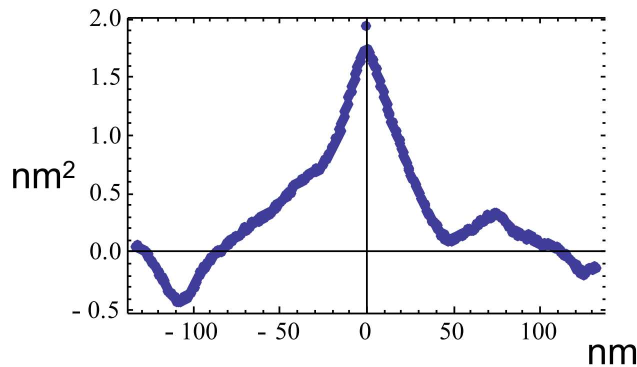
Use the middle 401 points to compute ACF

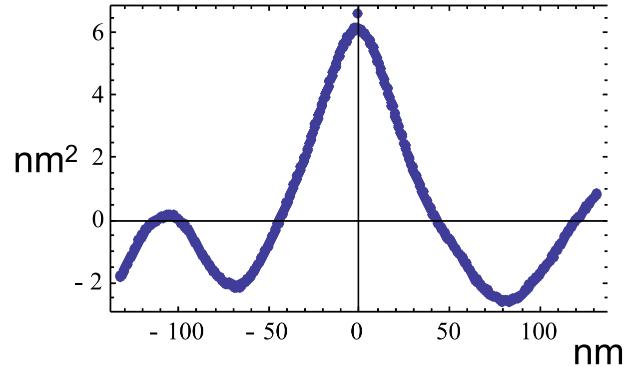


570nm
862 points

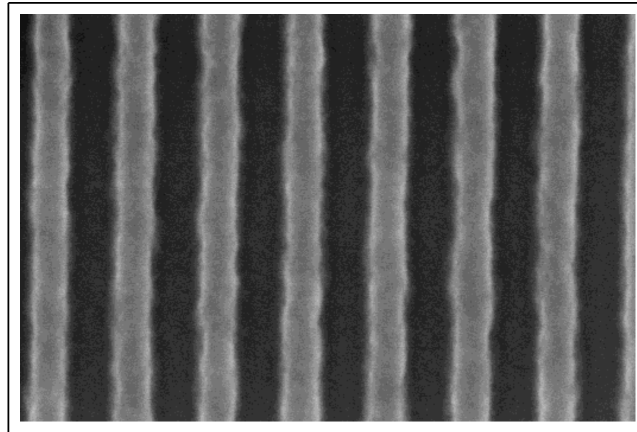




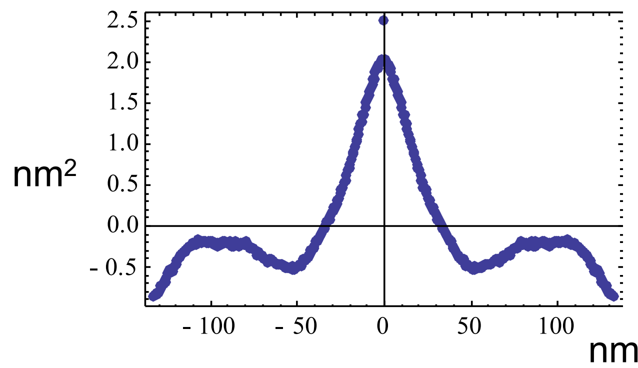


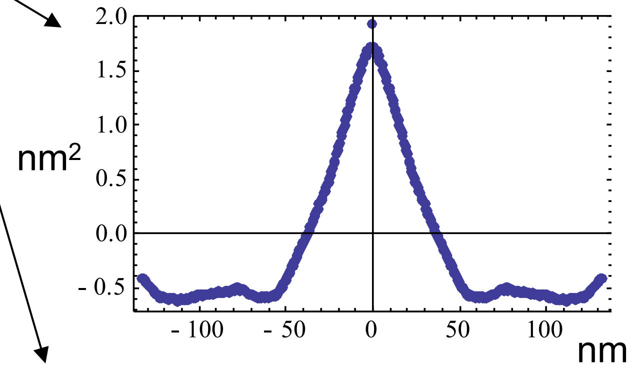


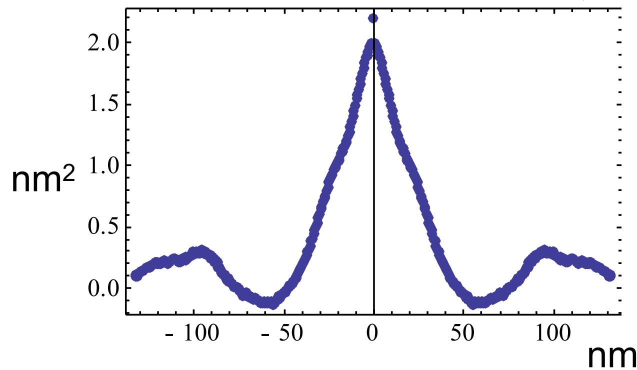
Assume periodicity to compute ACF

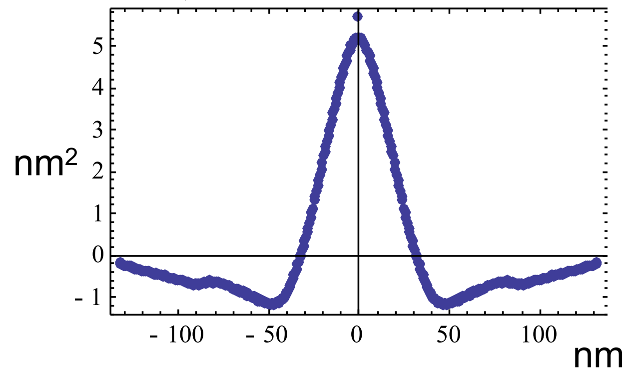


570nm
862 points

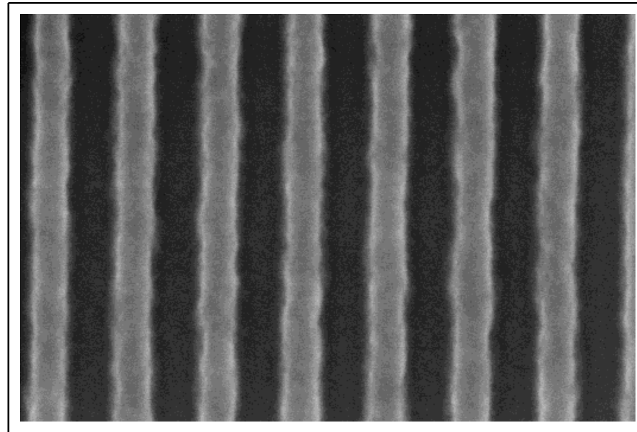








Assume periodicity
to compute ACF



570nm
862 points

nm²
2.5
2.0
1.5
1.0
0.5
0.0
- 0.5
- 100 - 50 0 50 100
nm

nm²
2.0
1.5
1.0
0.5
0.0
- 0.5
- 100 - 50 0 50 100
nm

nm²
2.0
1.5
1.0
0.5
0.0
- 100 - 50 0 50 100
nm

nm²
5
4
3
2
1
0
- 1
- 100 - 50 0 50 100
nm

ACF = AutoCorrelation Function

Fourier transform of the data $\tilde{H}(\beta) = \frac{1}{2\pi}\int dx\, H(x) e^{-i\beta x} \Rightarrow \tilde{H}(\beta)\,\text{units} = \text{length squared}$

$\text{Real} \Rightarrow \tilde{H}(\beta) = \tilde{H}(-\beta)^*$

$$\langle \ldots \rangle = \frac{1}{L}\int_{-L/2}^{+L/2} dx \ldots$$

$$ACF(x,x') = \langle H(x) H(x') \rangle \qquad \text{Let} \qquad x' = x - s$$

$$= \frac{1}{L}\int_{-L/2}^{+L/2} dx\, H(x) H(x-s) = \int d\beta\, d\beta'\, \tilde{H}(\beta)\tilde{H}(\beta') e^{-i\beta s} \underbrace{\frac{1}{L}\int_{-L/2}^{+L/2} dx\, e^{i(\beta+\beta')x}}_{\cong \frac{2\pi}{L}\delta(\beta+\beta')}$$

$$= \frac{2\pi}{L}\int d\beta \left|\tilde{H}(\beta)\right|^2 e^{\pm i\beta s}$$

$$= \frac{2\pi}{L}\int d\beta \left|\tilde{H}(\beta)\right|^2 e^{\pm i\beta (x-x')}$$

Acts like Dirac delta function

$$\Rightarrow ACF(x,x') = ACF(s) = ACF(x-x')$$

ACF = AutoCorrelation Function ➔ PSD = Power Spectral Density

PSD is the square of the FFT of the data with a "normalization" factor $2\pi / L$

NOTE: PSD units = length cubed

$$ACF(x-x') = \frac{2\pi}{L}\int d\beta \underbrace{\left|\tilde{H}(\beta)\right|^2}_{\text{Frequency content of the LER}} e^{\pm i\beta(x-x')} \quad \Rightarrow \quad \frac{2\pi}{L}\left|\tilde{H}(\beta)\right|^2 = PSD(\beta)$$

$$ACF(x,x) = ACF(x-x) = ACF(0) = \left\langle H(x)^2 \right\rangle = \sigma^2$$

$$\Rightarrow \quad ACF(x,x') = \sigma^2 f(x-x') \quad \text{with} \quad f(0) = 1$$

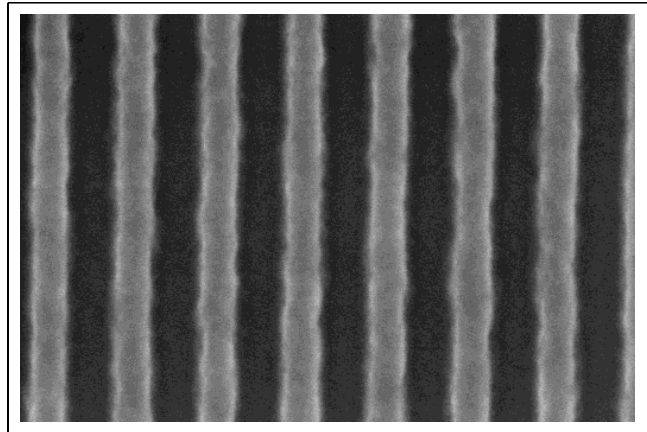
$$ACF(0) = \int d\beta \left(\frac{2\pi}{L}\left|\tilde{H}(\beta)\right|^2\right) = \int d\beta\, PSD(\beta) = \sigma^2$$

$$\frac{1}{\text{length}} \times \text{length}^3 = \text{length}^2$$

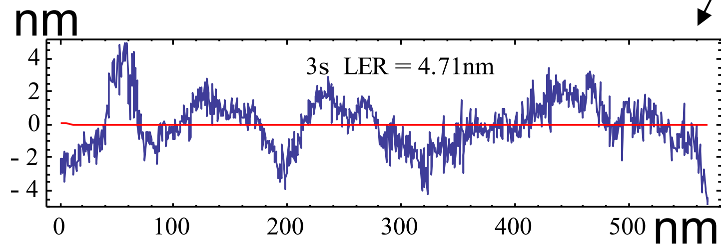
➔ Area under PSD = σ^2

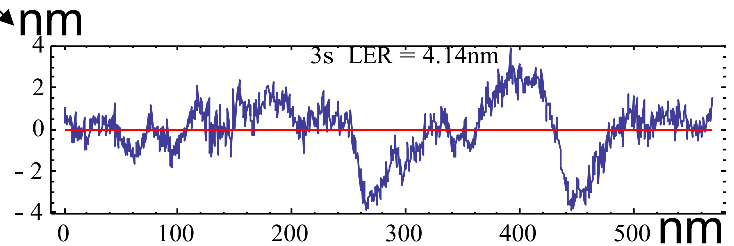
$$\sigma^2 = \int d\beta\, PSD(\beta) = \Delta\beta \sum_n PSD(n\Delta\beta) = \frac{2\pi}{L}\sum_n \frac{2\pi}{L}\left|\tilde{H}(n\Delta\beta)\right|^2 = \left(\frac{2\pi}{L}\right)^2 \sum_n \left|\tilde{H}(n\Delta\beta)\right|^2$$

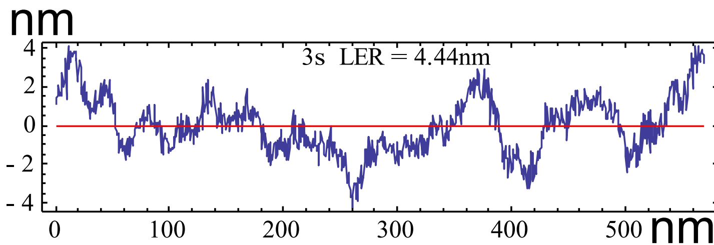
Example Edges

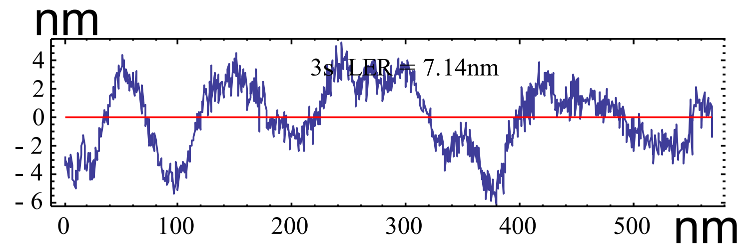


570nm
862 points



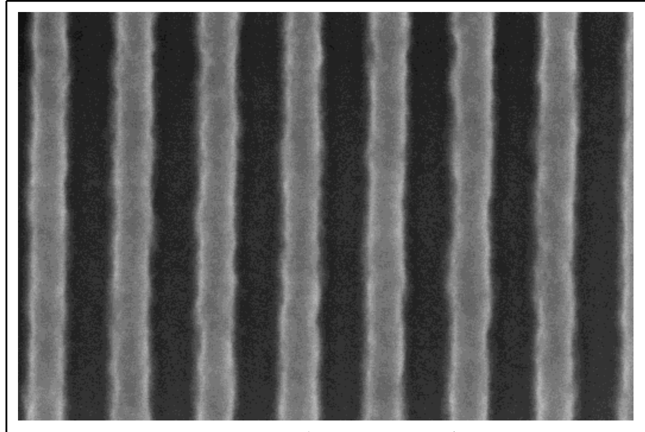






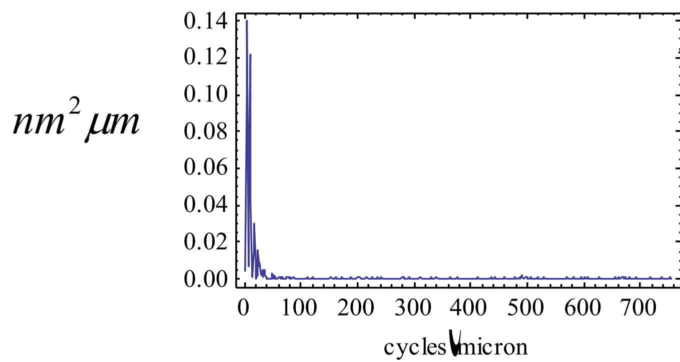
Compute PSD's using FFT

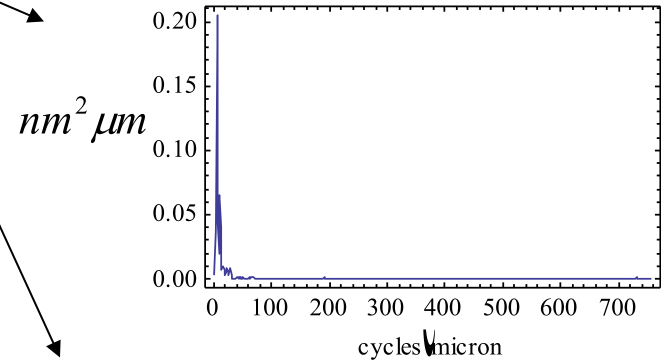
(Careful different FFT algorithms have different normalization factors)

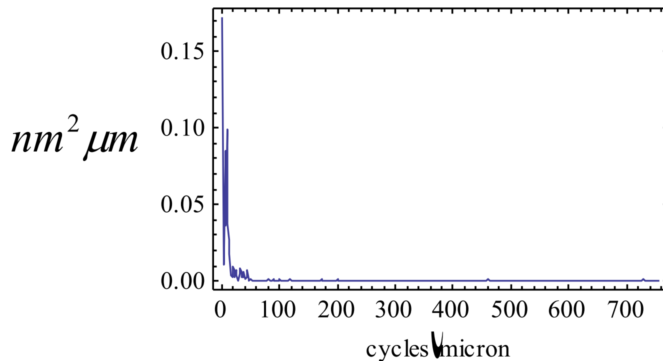


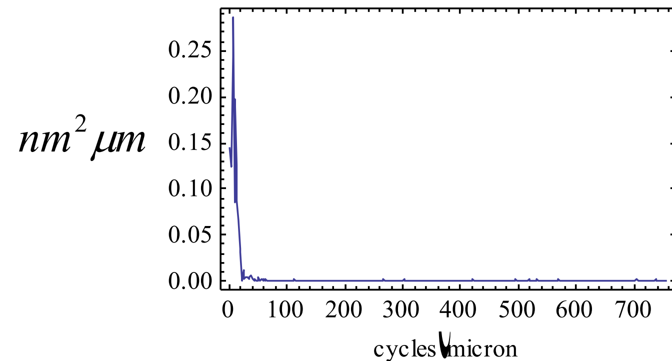
570nm

862 points

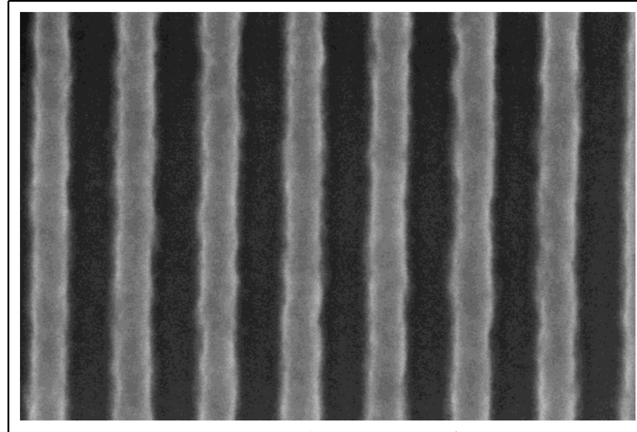




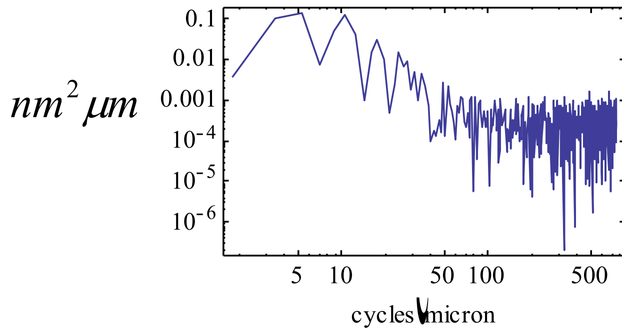


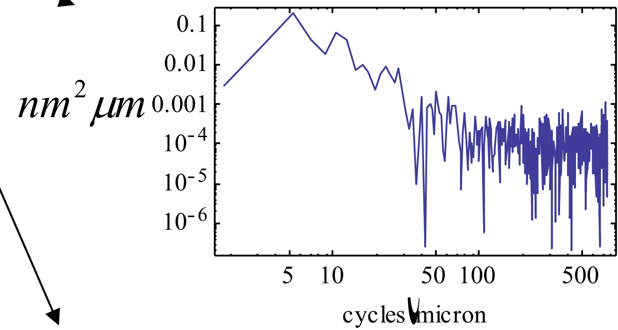


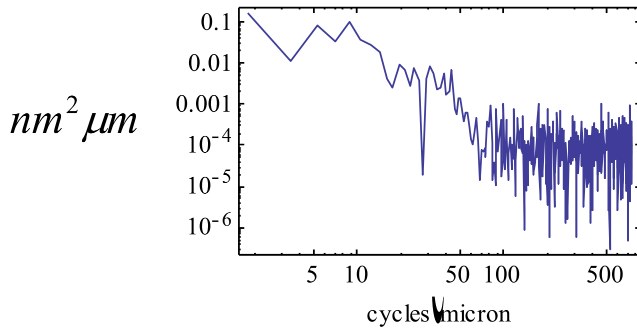
LogLog scale works much better for seeing PSD details

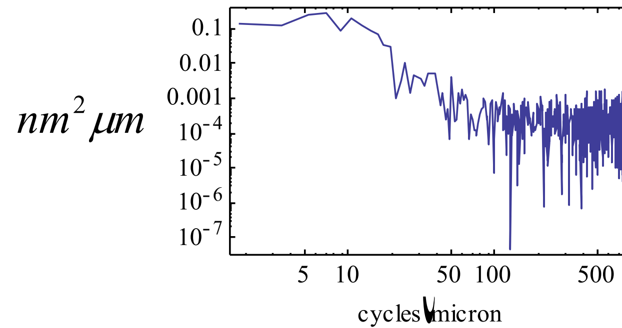


570nm
862 points

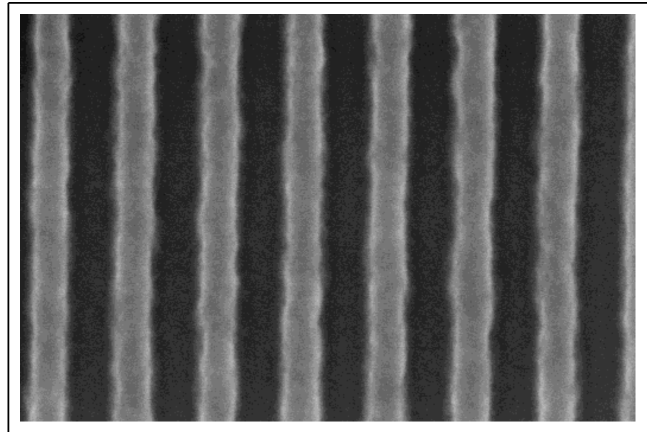




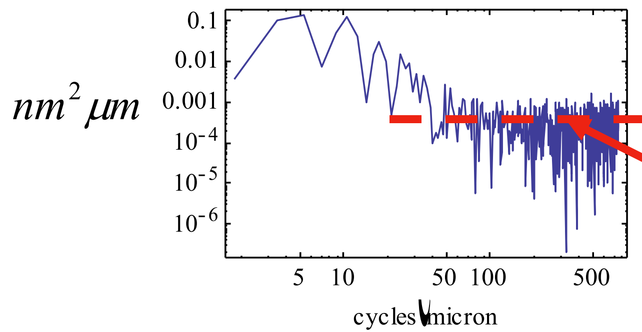




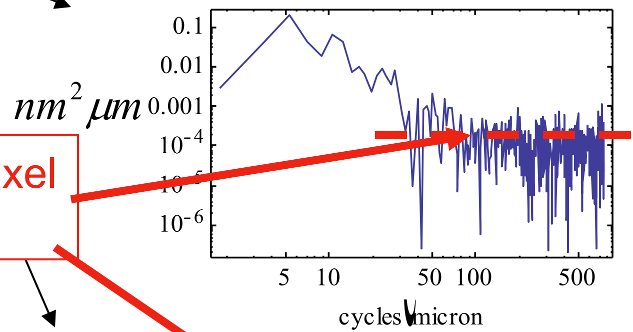
LogLog scale works much better for seeing PSD details

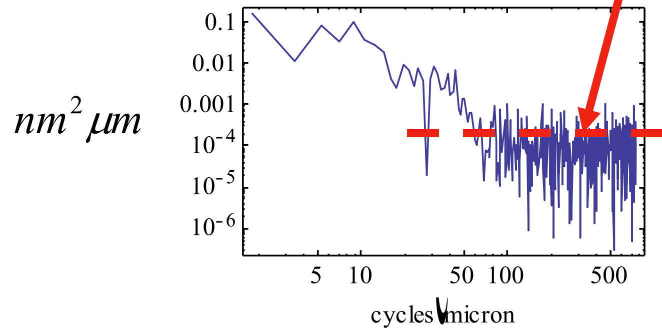


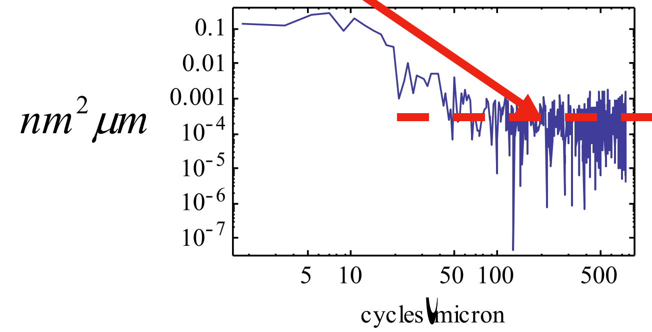
570nm
862 points



SEM Pixel-to Pixel White Noise

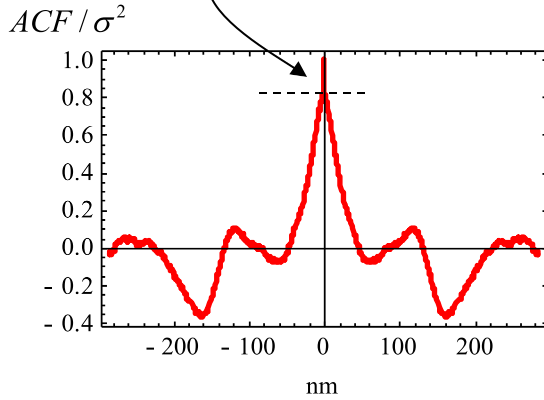


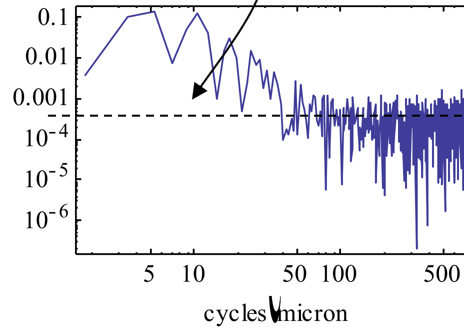


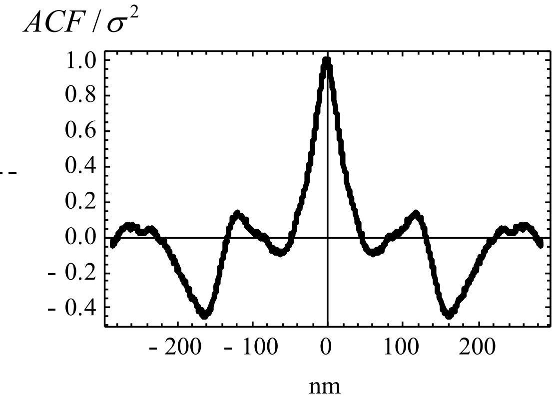


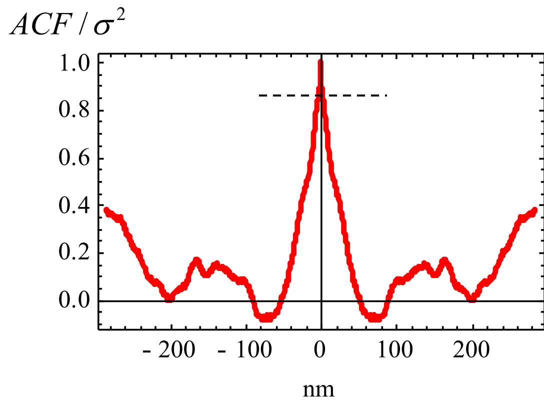
Spike in ACF comes from flat "white noise" background in the PSD

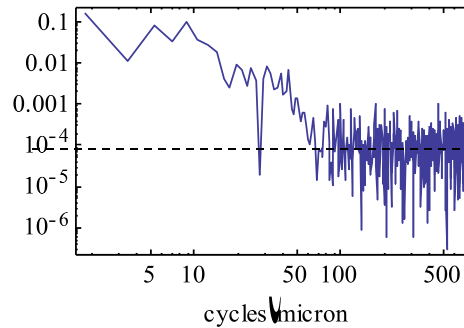
Subtract off the "white noise" background from the PSD to eliminate the spike

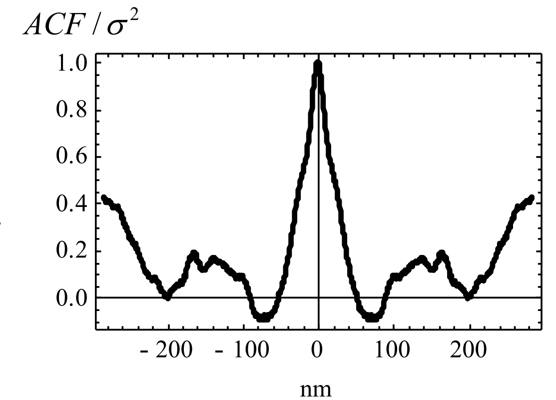




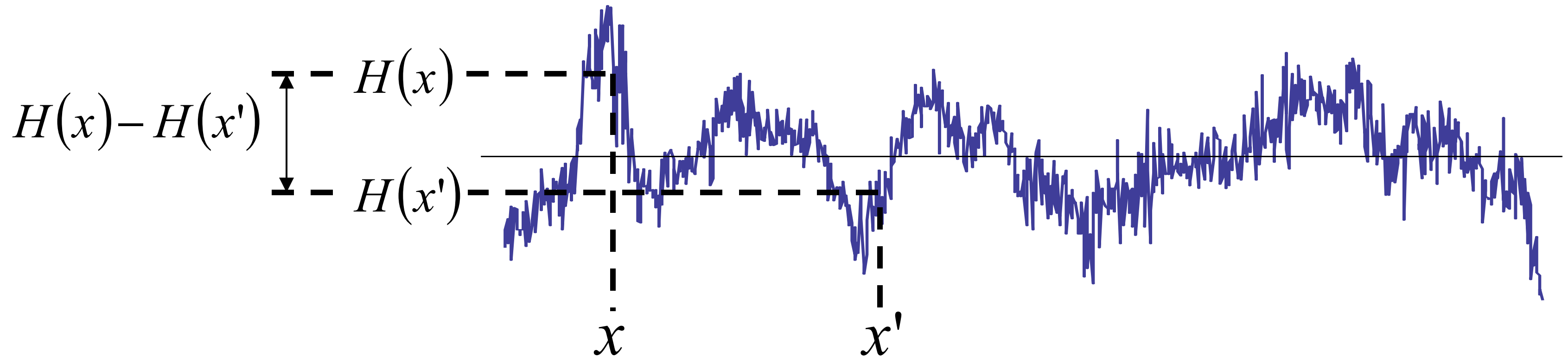








"HHCF" = Height Height Correlation Function



$$HHCF(x,x') = \left\langle \left(H(x)-H(x')\right)^2 \right\rangle$$

$$= \left\langle \left(H(x)^2 + H(x')^2 - 2H(x)H(x')\right) \right\rangle$$

$$= \left\langle H(x)^2 \right\rangle + \left\langle H(x')^2 \right\rangle - 2\left\langle H(x)H(x') \right\rangle$$

$$= \quad \sigma^2 \quad + \quad \sigma^2 \quad - \quad 2ACF(x-x')$$

$$= 2\sigma^2\left(1-f(x-x')\right)$$

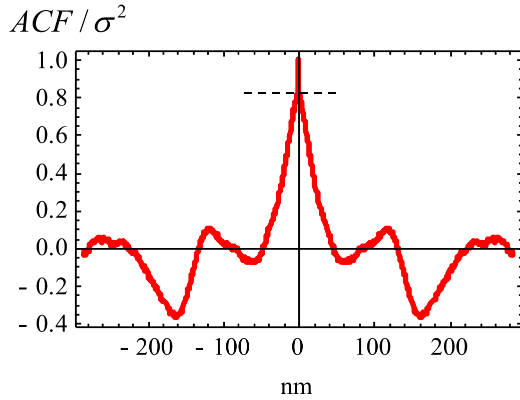
$$f(x) = \frac{ACF(x)}{\sigma^2}$$

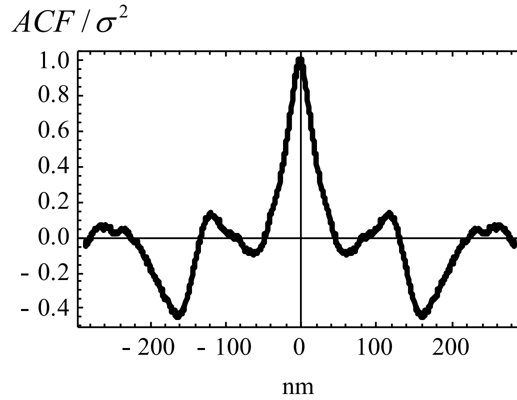
$$f(0) = 1$$

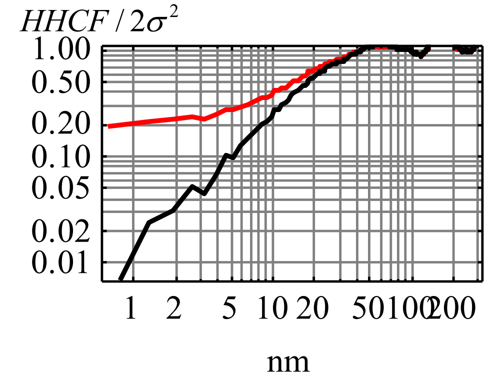
$$HHCF(x,x) = 2\sigma^2\left(1-f(0)\right) = 0$$

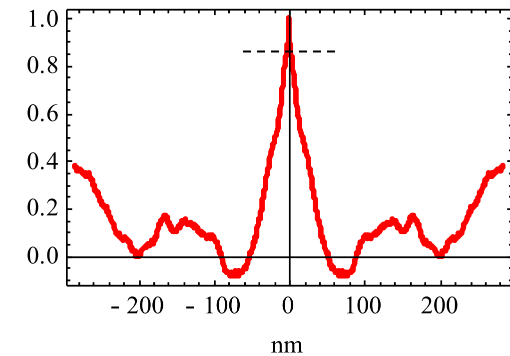
Raw ACF

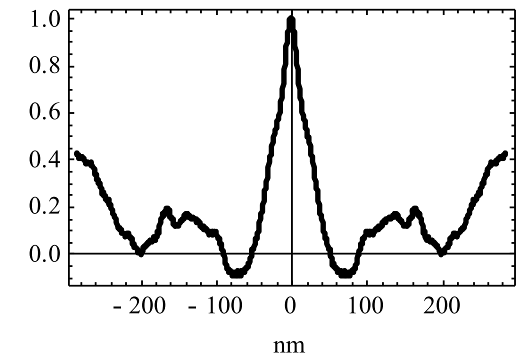
Background subtracted ACF

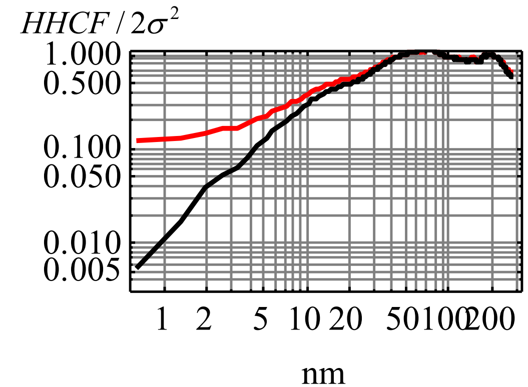




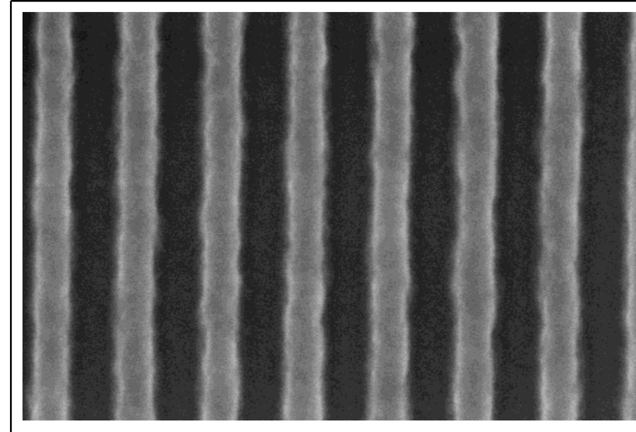


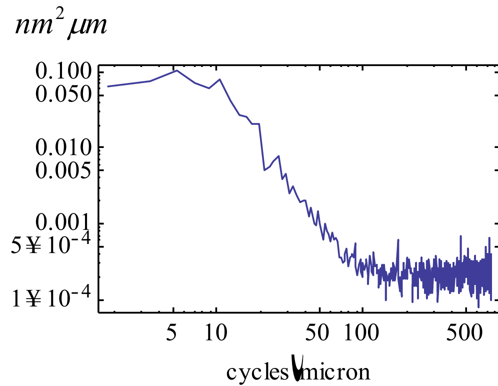


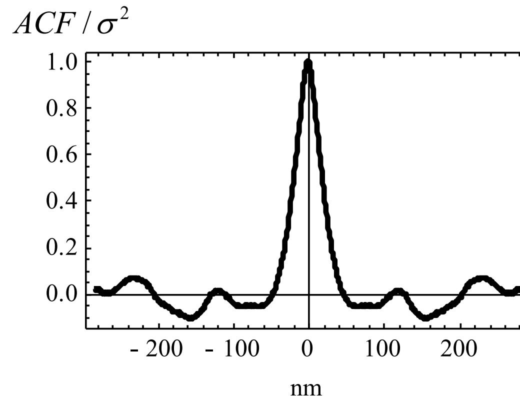


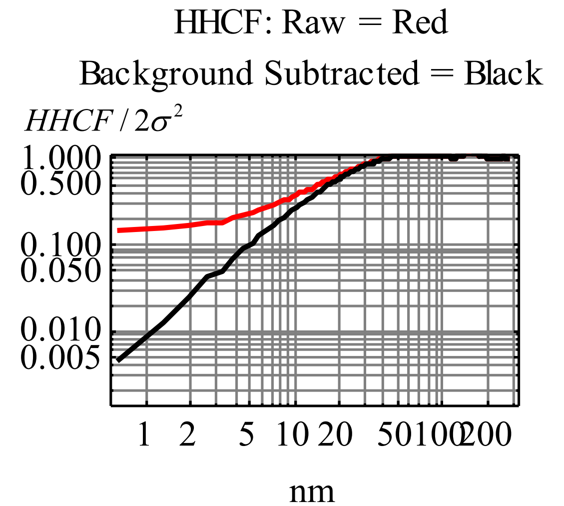


Beat down noise in PSD, ACF and HHCF by averaging the PSD's from all the edges









Standard shapes for the PSD and ACF

NOTE Units

Lorentzian PSD ⟷ **Exponential ACF**

$$\text{PSD}: \ \frac{2\pi}{L}\left|\widetilde{H}(\beta)\right|^2 = \sigma^2 \frac{\xi/\pi}{1+\xi^2\beta^2} \Rightarrow ACF(x) = \sigma^2 f(x) = \sigma^2 e^{-|x|/\xi} \quad \Rightarrow length^3 = \frac{length^2}{spatial\ frequency}$$

Gaussian PSD ⟷ **Gaussian ACF**

$$\frac{2\pi}{L}\left|\widetilde{H}(\beta)\right|^2 = \sigma^2 \frac{\xi}{2\sqrt{\pi}} e^{-\xi^2\beta^2/4} \Rightarrow ACF(x) = \sigma^2 f(x) = \sigma^2 e^{-(x/\xi)^2} \quad \Rightarrow length^3 = \frac{length^2}{spatial\ frequency}$$

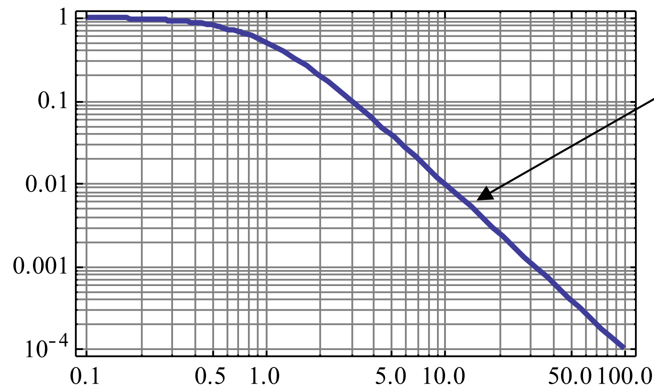
$$\text{Power Spectrum}: \ \left|\widetilde{H}(\beta)\right|^2 = \frac{L}{2\pi}\left(\sigma^2 \frac{\xi/\pi}{1+\xi^2\beta^2}\right) \quad \Rightarrow length^4 = \frac{length^2}{(spatial\ frequency)^2}$$

$$\left|\widetilde{H}(\beta)\right|^2 = \frac{L}{2\pi}\left(\sigma^2 \frac{\xi}{2\sqrt{\pi}} e^{-\xi^2\beta^2/4}\right) \quad \Rightarrow length^4 = \frac{length^2}{(spatial\ frequency)^2}$$

$$f(x=\xi) = \frac{1}{e} \cong 0.37 \qquad \xi = \text{"correlation length"} \qquad \Rightarrow ACF(x=\xi) \approx 0.37\sigma^2$$

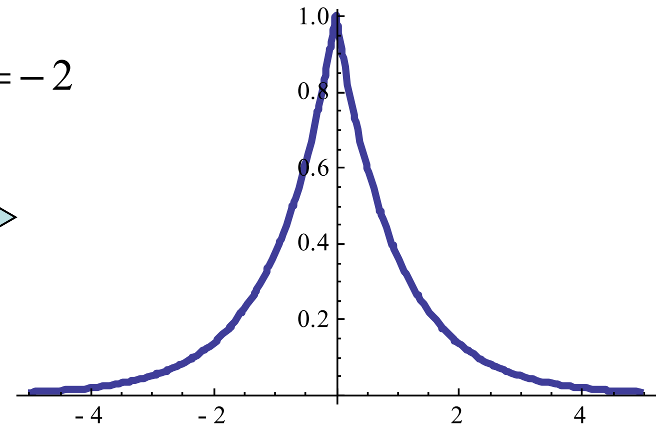
$$HHCF(x) = 2\sigma^2(1-f(x)) \Rightarrow HHCF(x=\xi) = 2\sigma^2\left(1-\frac{1}{e}\right) \approx 0.63(2\sigma^2)$$

Lorentzian PSD

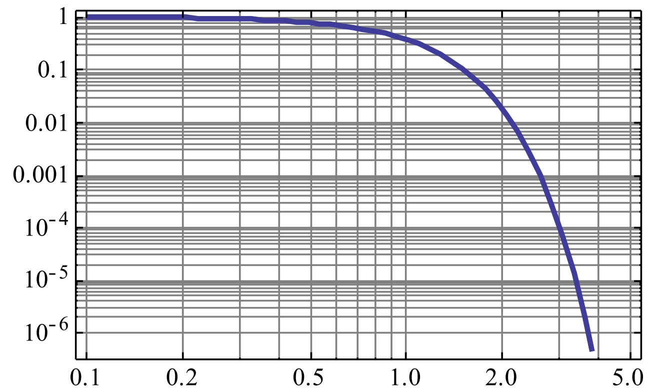


loglog slope = − 2

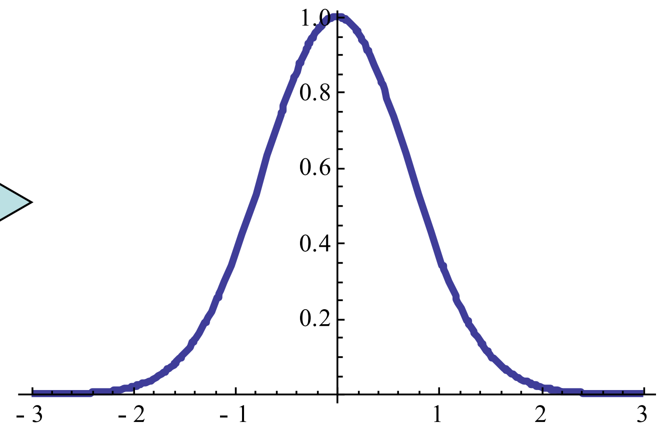
Exponential ACF



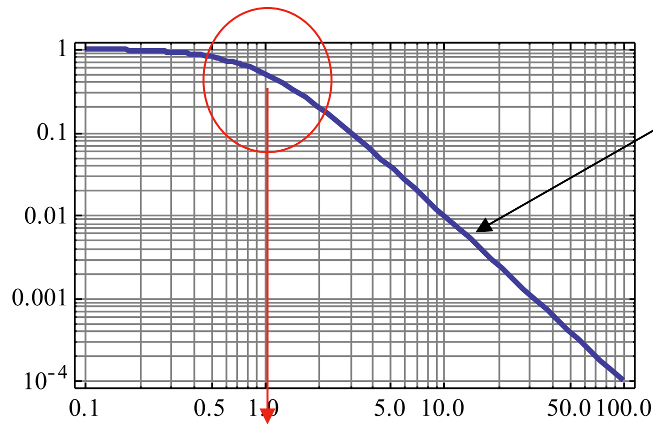
Gaussian PSD



Gaussian ACF



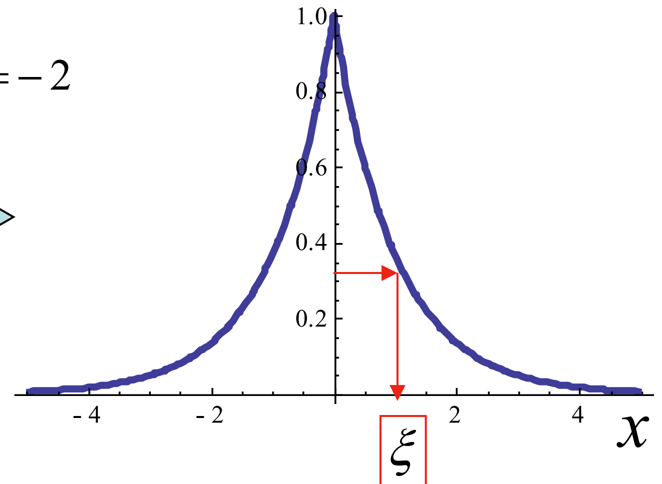
Lorentzian PSD



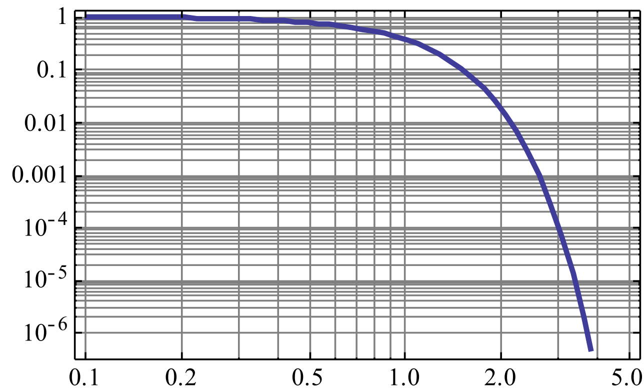
loglog slope = − 2

$1/\text{correlation length} = 1/\xi$

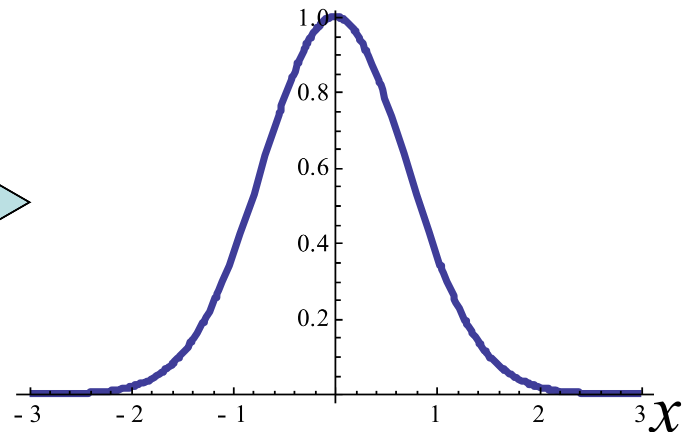
Exponential ACF



Gaussian PSD



Gaussian ACF



PSD loglog slope γ ←→ HHCF scaling at small $x - x'$

$$HHCF(x) = 2\left(\sigma^2 - \sigma^2 f(x)\right) = 2\left(ACF(0) - ACF(x)\right) = 2\frac{2\pi}{L}\int d\beta \left|\tilde{H}(\beta)\right|^2 \left(1 - e^{i\beta x}\right)$$

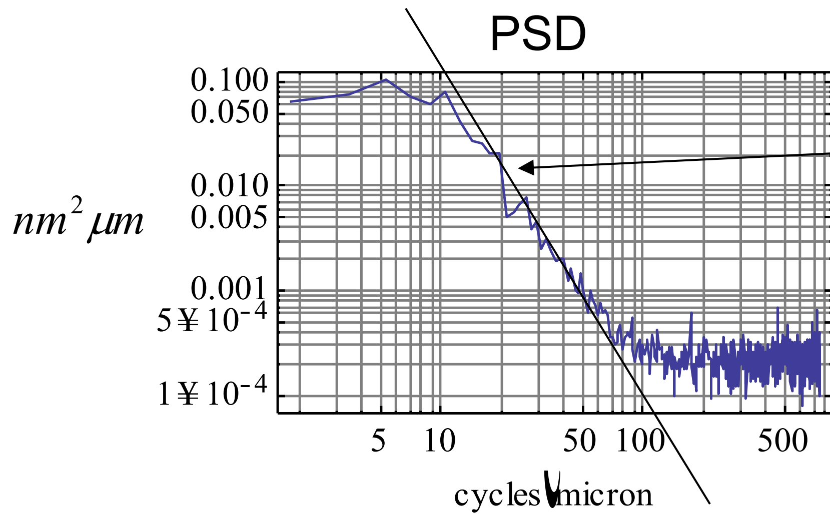
$$\left|\tilde{H}(\beta)\right|^2 \propto \frac{1}{1+(\xi\beta)^\gamma} \qquad \Rightarrow \text{For } \beta\xi >> 1 \quad \frac{\partial \ln(PSD(\beta))}{\partial \ln(\beta)} = -\gamma = \text{"}PSD \text{ loglog slope"}$$

$$HHCF(x) \propto \int d\beta \frac{1-e^{i\beta x}}{1+(\xi\beta)^\gamma} = \frac{1}{x}\int dq \frac{1-e^{iq}}{1+\left(\frac{\xi}{x}q\right)^\gamma} \text{ after letting } q = \beta x$$

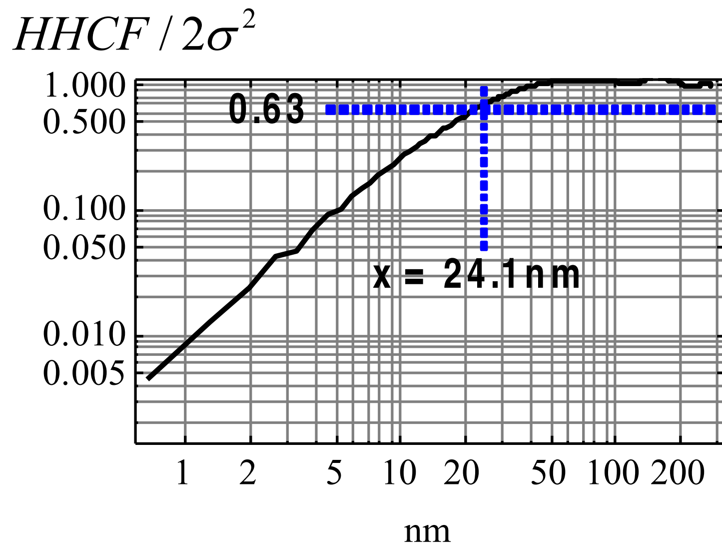
$$\text{For } x << \xi \quad \Rightarrow \quad 1+\left(\frac{\xi}{x}q\right)^\gamma \approx \left(\frac{\xi}{x}q\right)^\gamma \quad \Rightarrow \quad HHCF(x << \xi) \propto x^{\gamma-1}$$

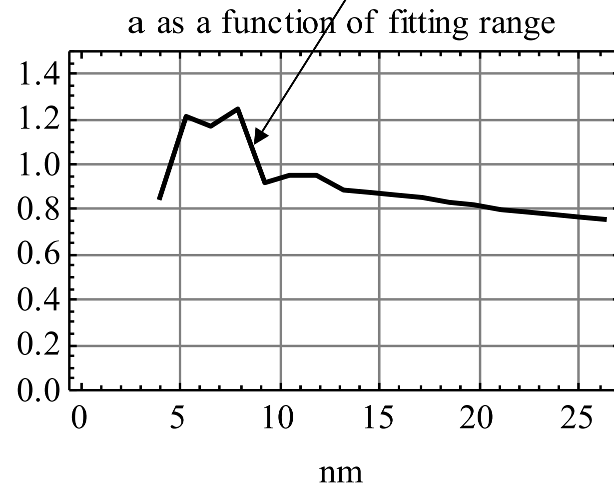
$$\text{Critical exponent } \alpha \text{ defined by } HHCF(x << \xi) \propto x^{2\alpha} \qquad \text{Thus } \gamma = 2\alpha + 1$$

$$PSD \text{ loglog slope at high frequency} = 2 \times \text{critical exponent} + 1$$

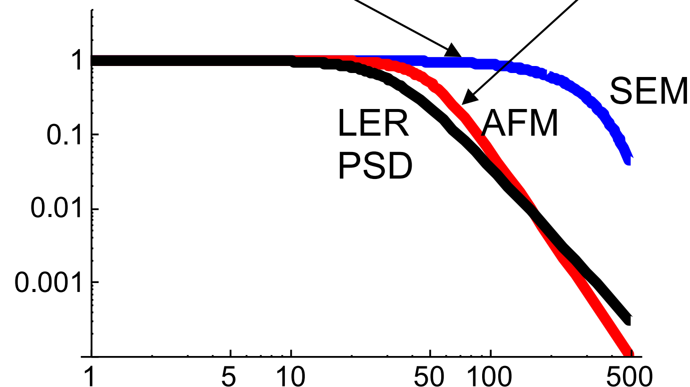


Loglog slope = –3 $\Rightarrow \quad \gamma = 3 \quad \Rightarrow \quad \alpha = 1$





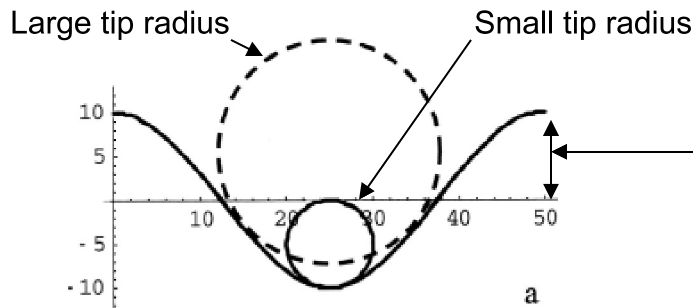
SEM	**AFM**
Probe ~ Gaussian e-beam Radius ~ 2 to 5nm	Probe ~ “Mechanical” tip Radius ~ 10’s of nm
➔ Fundamental frequency cutoff ~ 1/2nm to1/5nm ~500/μm to 200/μm	➔ Fundamental frequency cutoff ~ 1/10’s of nm << 100/μm
Cutoff frequency is much higher than the observed shoulder in LER PSD’s ➔ SEM resolution is not significantly filtering the LER	Cutoff frequency is close to the observed shoulder in LER PSD’s ➔ AFM tip is significantly filtering the LER



AFM tip has a finite radius

➔ Tip filters out high frequency content of the roughness

➔ Makes a smoothly varying surface look "grainy"



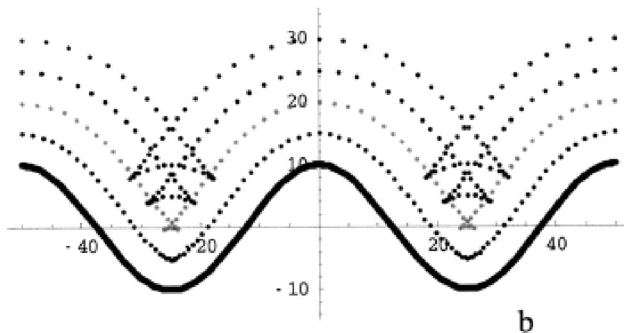


FIG. 2. (a) Model sinusoidal surface used to illustrate AFM probe filtering effects. For the wave amplitude and period shown, the tip probe represented by the smaller circle (full line) allows full access to the surface dip. The tip probe described by the larger circle (dotted line) cannot describe depressed areas properly. (b) Effect of increasing tip radii on the resulting imaged surface. Dotted traces closer to the surface correspond to smaller tip sizes.

Amplitude at which tip filtering becomes significant

$$Amplitude = \frac{1}{\beta^2 R_{tip}}$$

$$\left| Amplitude \right|^2 = \frac{1}{\beta^4 R_{tip}^{\ 2}}$$

➔ High frequency content of the PSD is filtered out.

➔ Filtered PSD loglog-slope ~ –4

AFM measurement of LER

Goldfarb, et. al., JVSTB 22 (2004)

Patterned wafer is cleaved and turned on edge

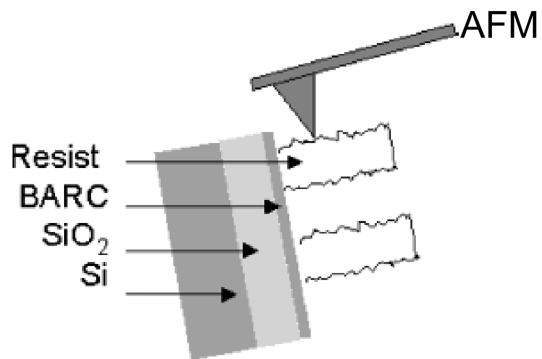
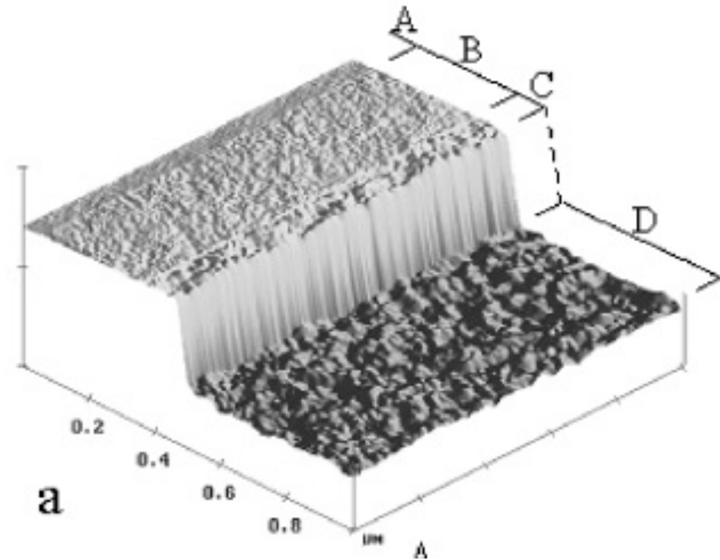


FIG. 1. AFM technique used to image sidewalls of multilayer systems, showing sample orientation.



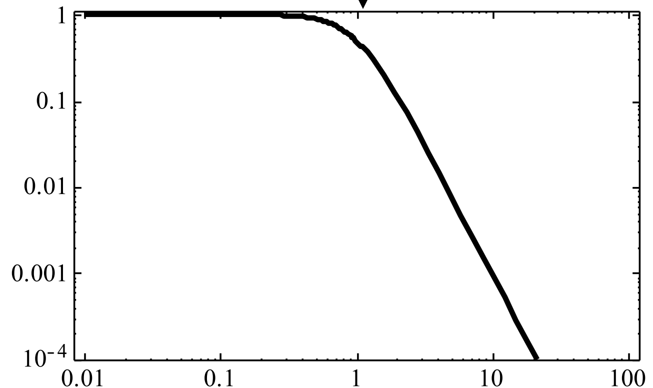
A note on the addition of a flat (constant) noise background to the "ideal" PSD

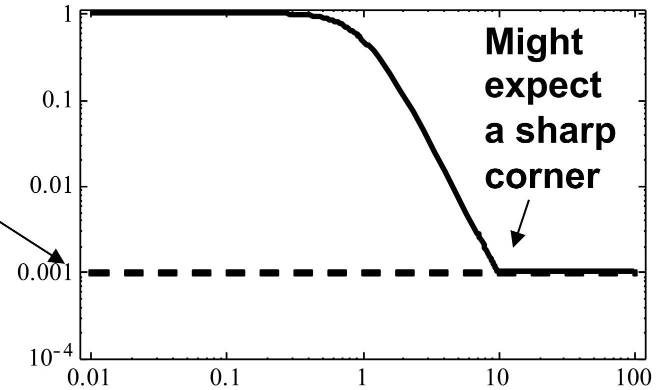
SEM images contain "white" noise ➔ Noise in each pixel is uncorrelated to noise in other pixels
➔ White Noise ➔ Flat PSD

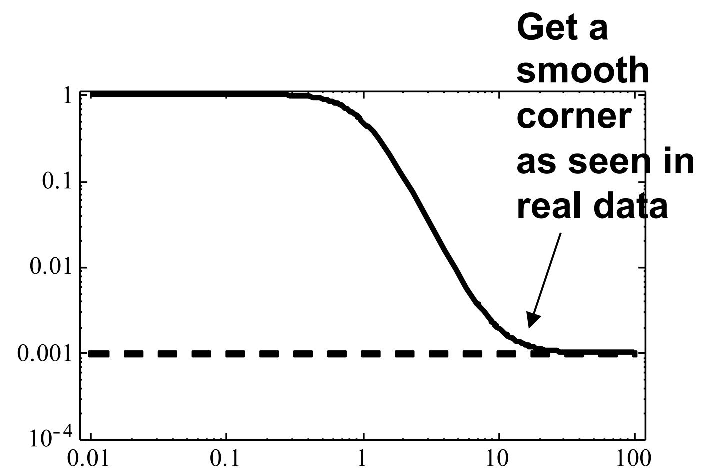
$$PSD(\beta) = \frac{1}{1+(\xi\beta)^{\gamma}} \Rightarrow \frac{1}{1+(\xi\beta)^{\gamma}} + B$$

"Ideal" PSD

SEM White Noise PSD Is added to Data PSD

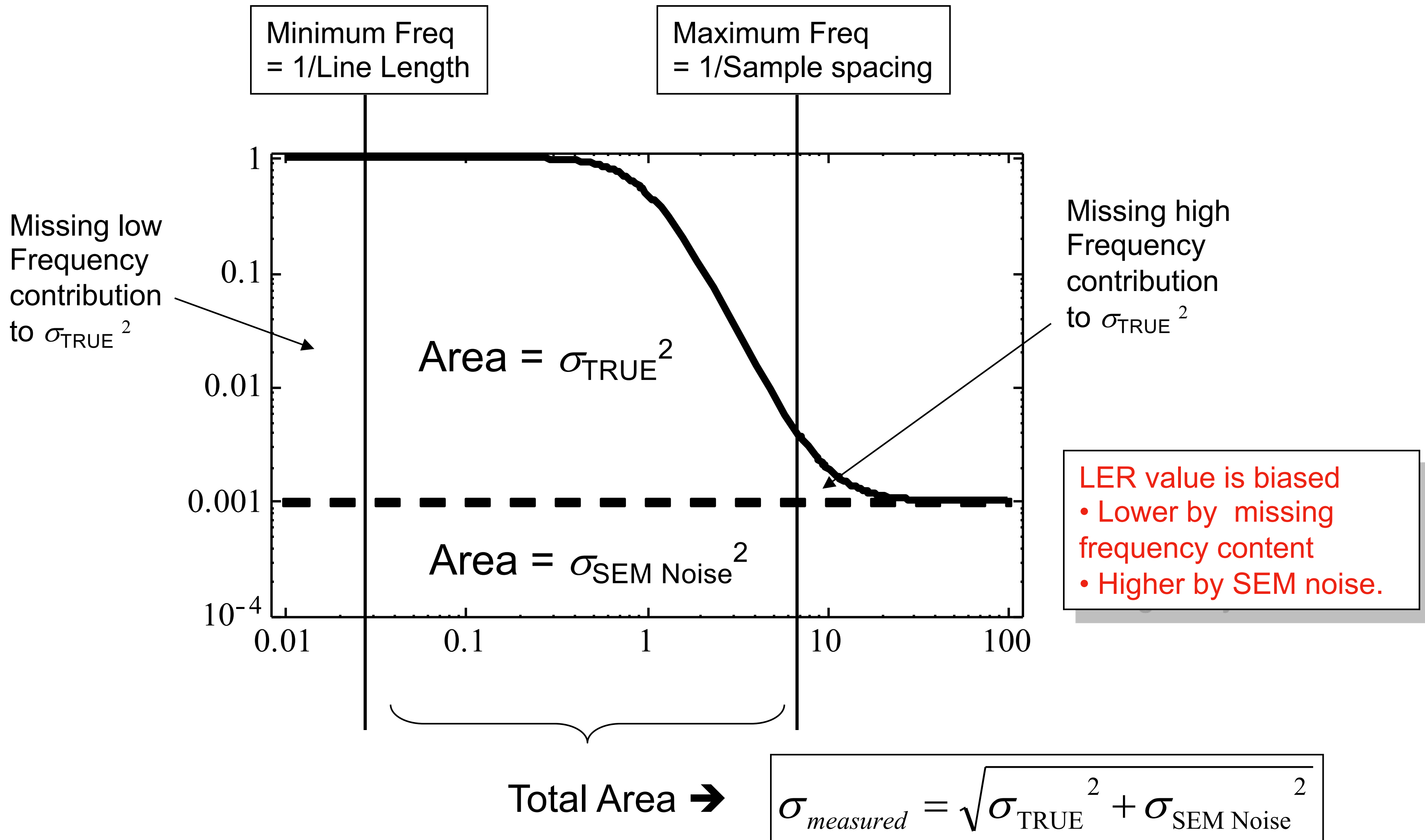






Bias in LER measurements: $\sigma = \sqrt{\text{Area under PSD}}$

See Villarubia and Bunday refs



SVL = Sigma Versus Length

Full length σ^2

$$\sigma^2 = \frac{1}{L}\int_{-L/2}^{+L/2} dx\, H(x)^2$$

Compute σ^2 over variable length l

$$SVL(l)^2 = \frac{1}{l}\int_{-l/2}^{+l/2} dx\left(H(x) - \frac{1}{l}\int_{-l/2}^{+l/2} dx\, H(x)\right)^2$$

$$= \frac{1}{l}\int_{-l/2}^{+l/2} dx\, H(x)^2 - \frac{1}{l^2}\left(\int_{-l/2}^{+l/2} dx\, H(x)\right)^2$$

Take average to determine mean behavior of SVL

$$SVL(l)^2 = \frac{1}{l}\int_{-l/2}^{+l/2} dx\left\langle H(x)^2\right\rangle - \frac{1}{l^2}\int_{-l/2}^{+l/2} dxdx'\left\langle H(x)H(x')\right\rangle$$

$$= \sigma^2 - \frac{1}{l^2}\int_{-l/2}^{+l/2} dxdx'\,\sigma^2 f(x-x')$$

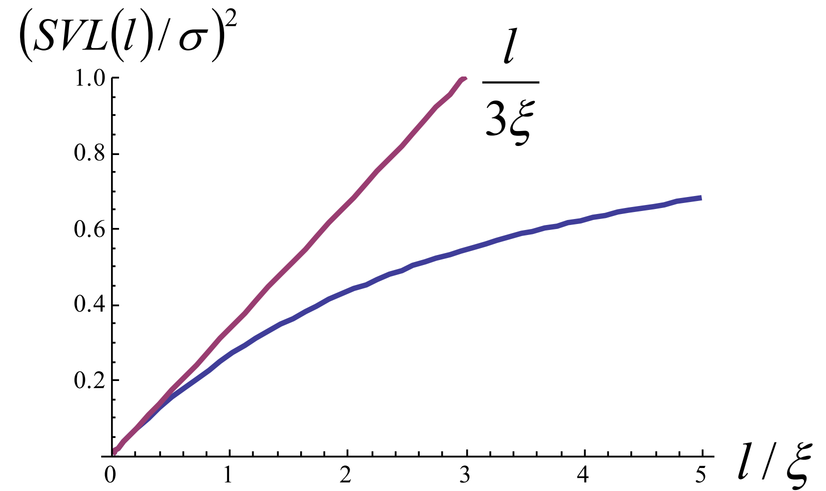
$$= \sigma^2\left(1 - \frac{1}{l^2}\int_{-l/2}^{+l/2} dxdx'\, f(x-x')\right)$$

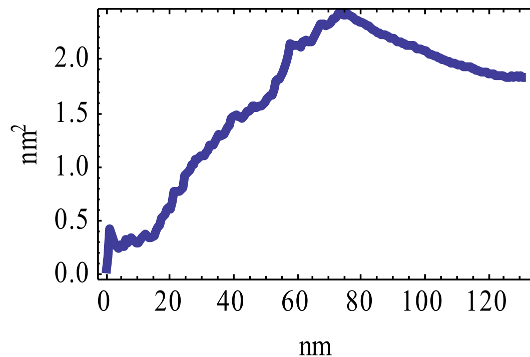
Statistical Average behavior of SVL

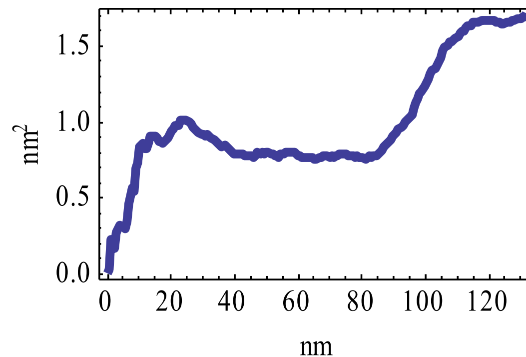
$$SVL(l)^2 = \sigma^2\left(1 - \frac{1}{l^2}\int_{-l/2}^{+l/2} dx dx' f(x-x')\right)$$

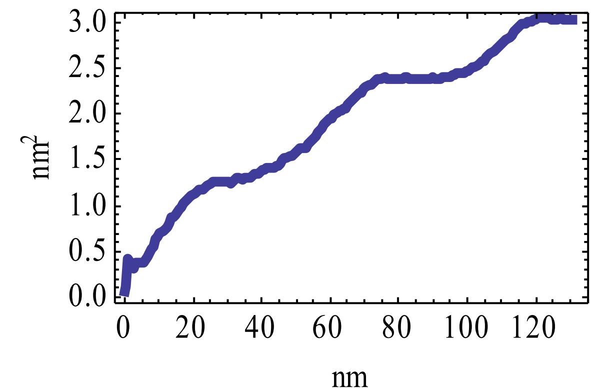
For $f(x-x') = e^{-|x-x'|/\xi}$ get

$$SVL(l)^2 = \sigma^2\left(1 - 2\frac{\xi}{l}\left(e^{-l/\xi} + \left(\frac{\xi}{l} - 1\right)\left(e^{-l/\xi} - 1\right)\right)\right)$$







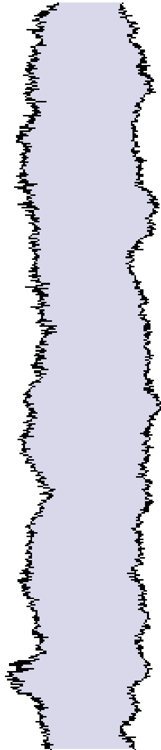


Line Width "Roughness" (LWR) = Linewidth variation along a single line

Uncorrelated Edges

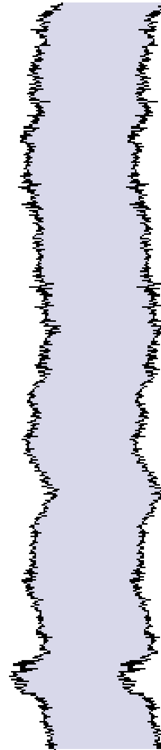
$$LWR = \sqrt{\sigma_{left}^{2} + \sigma_{right}^{2}}$$

$$\approx \sqrt{2}\, LER$$



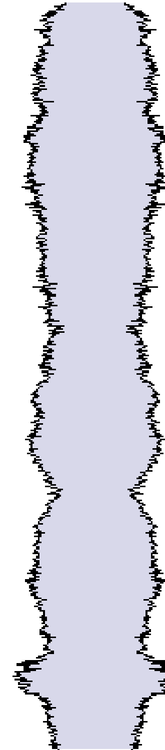
"Correlated" Edges

$$LWR = 0$$

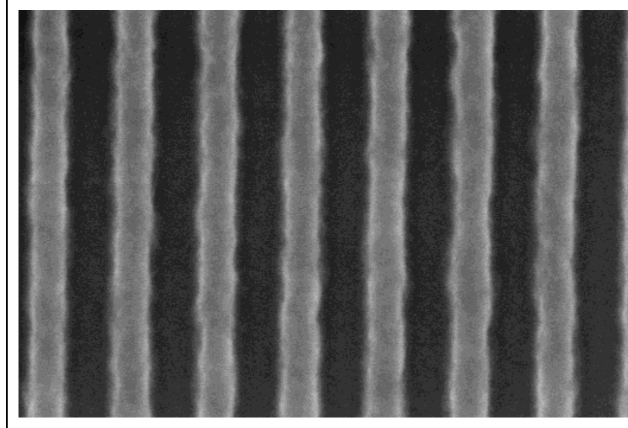


"Anticorrelated" Edges

$$LWR = 2\, LER$$

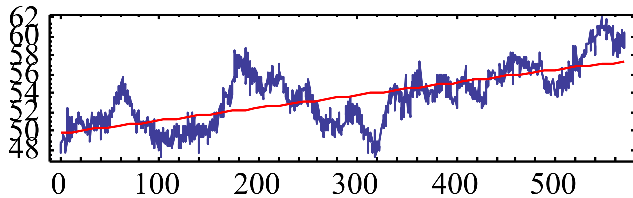


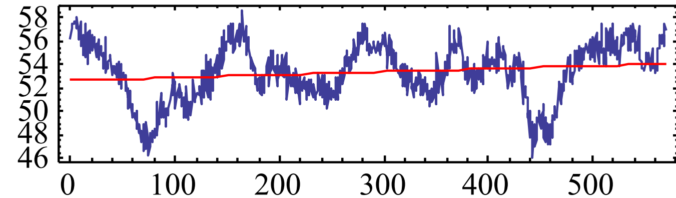
Line Width Roughness

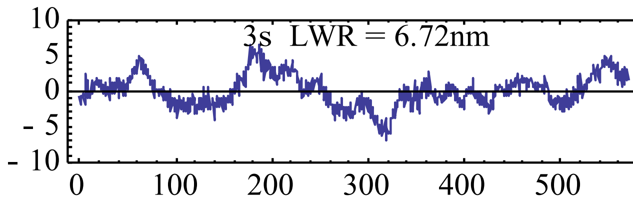


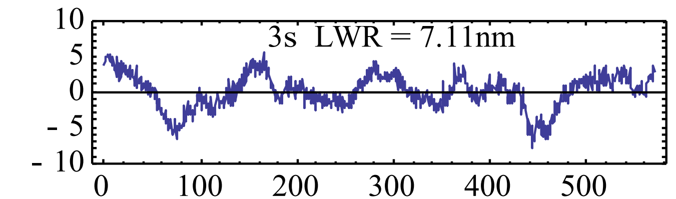
Linear CD variation = 7.66nm

Linear CD variation = 1.44nm

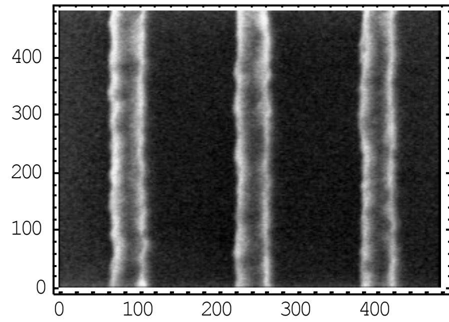


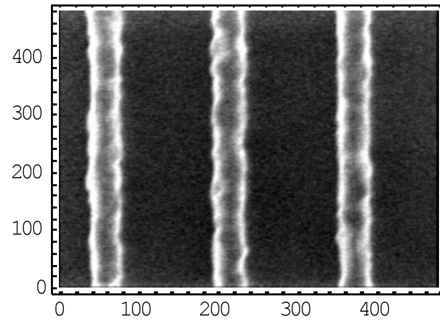


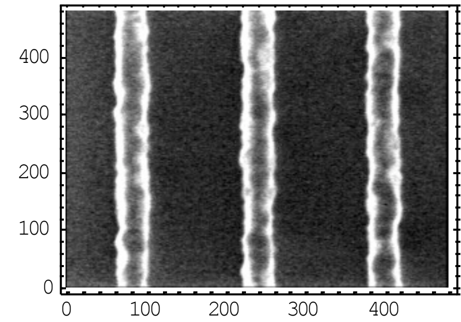


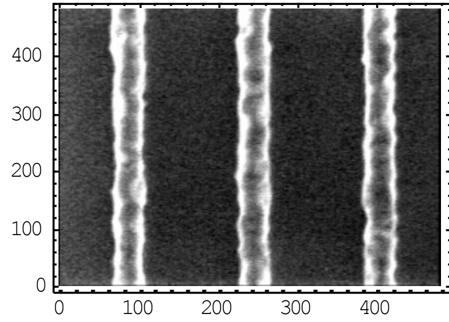


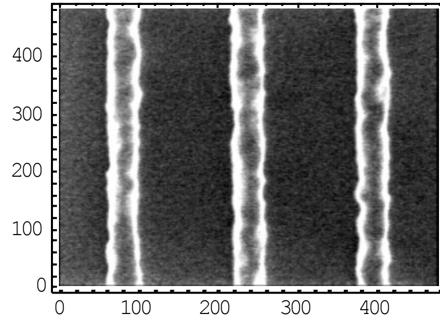
193 Example

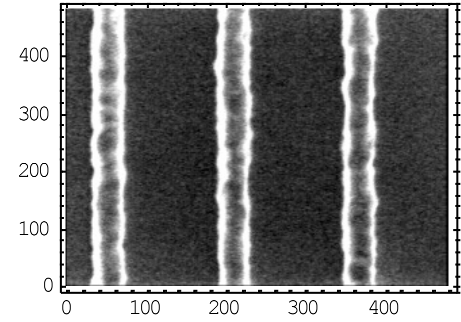


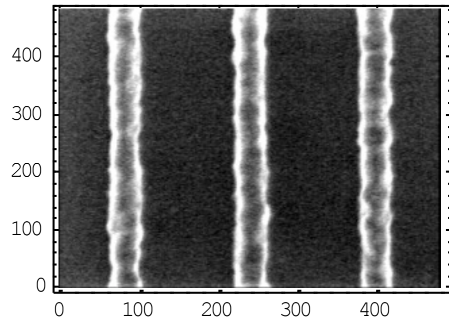


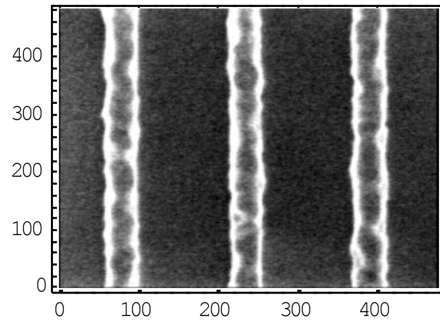


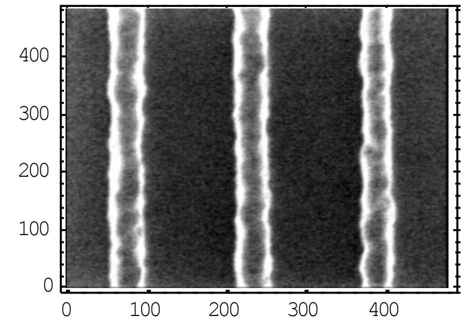




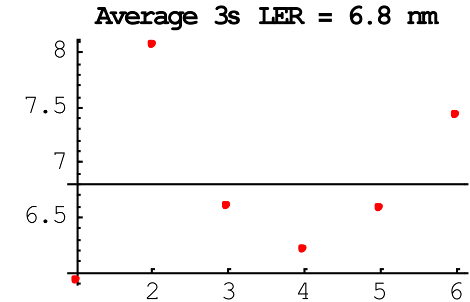


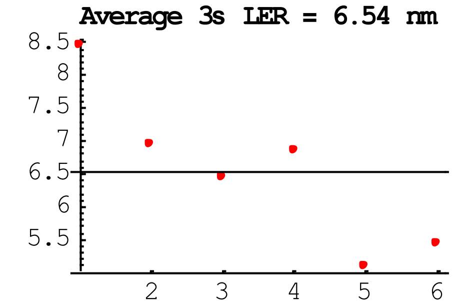


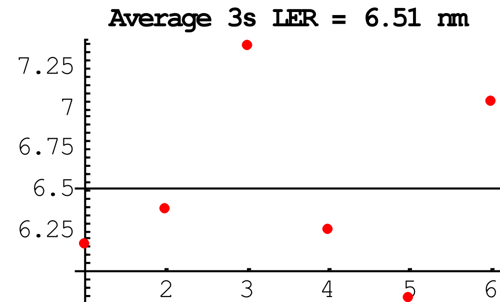


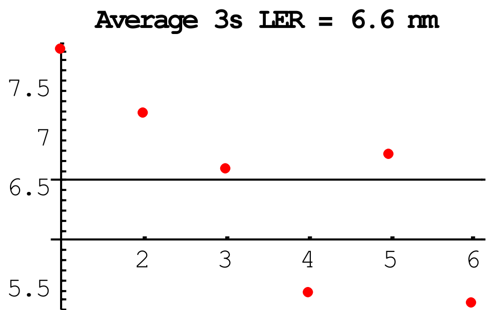


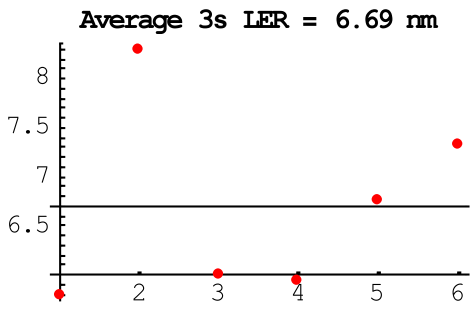
Edge by Edge 3σ

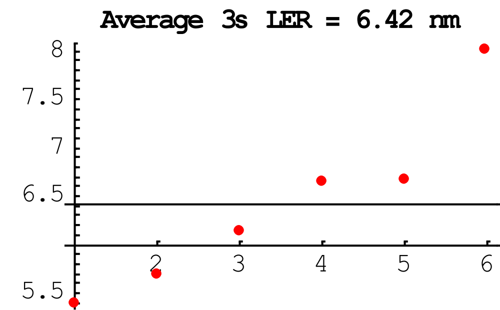


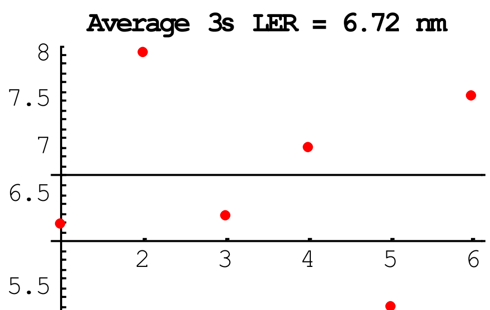


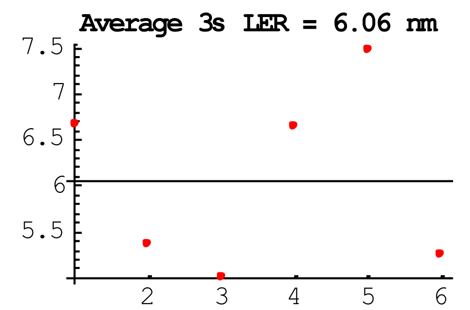


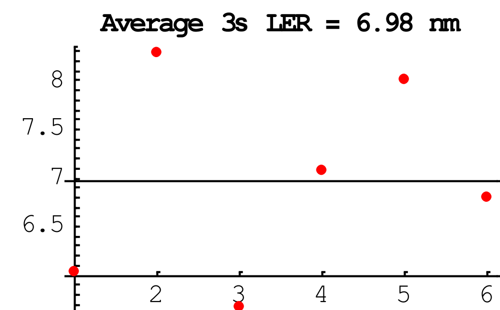




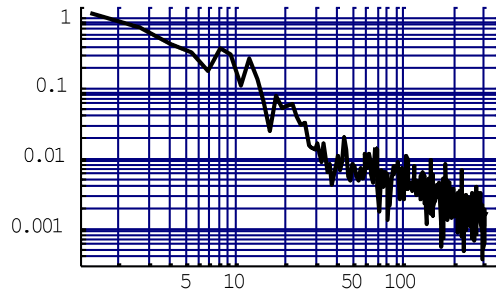


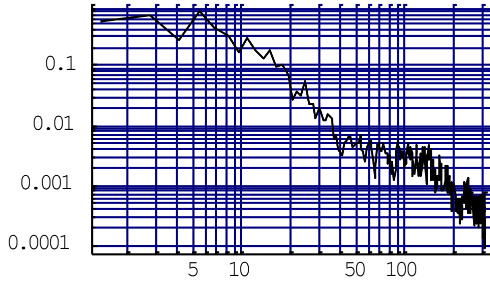


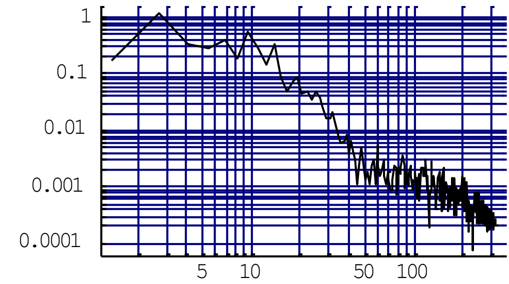


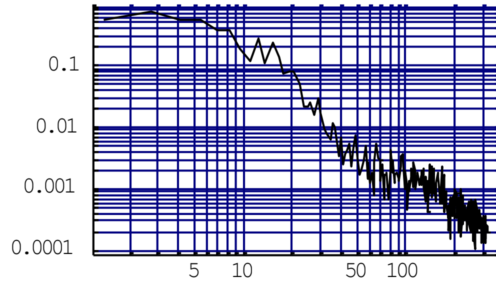


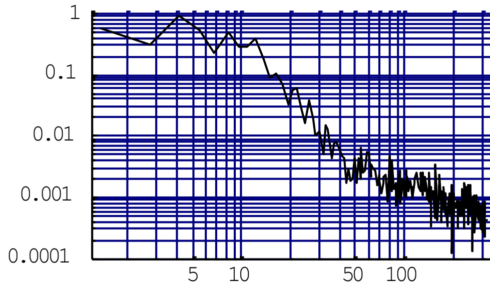
Edge by Edge PSD’s

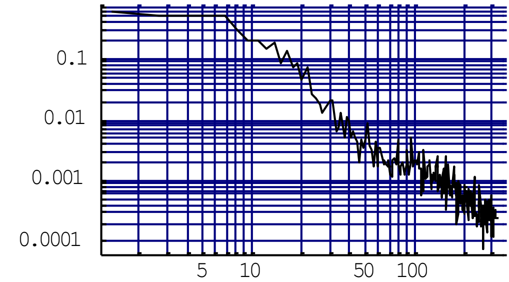


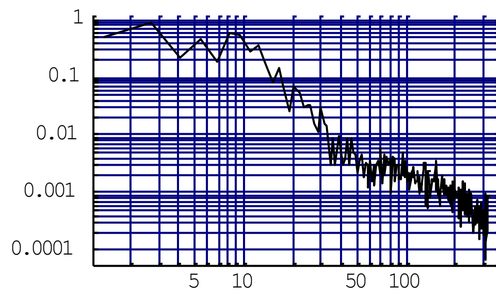


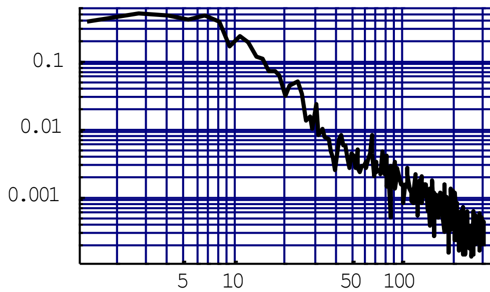


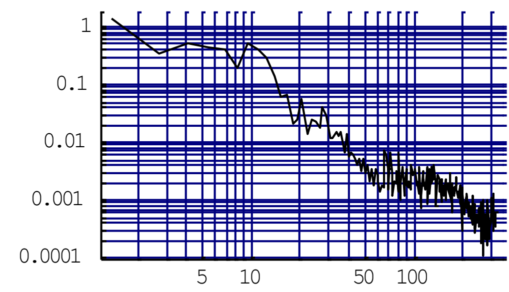












193 data from Pawloski, et. al., JM3 2006

Pawloski et al.: Line edge roughness and intrinsic bias…

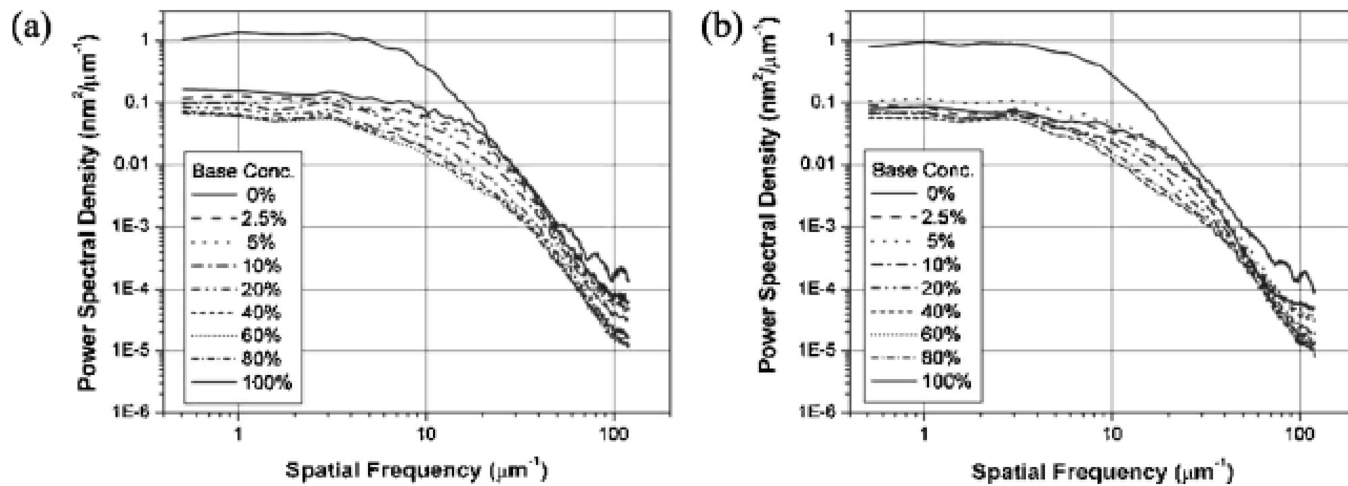
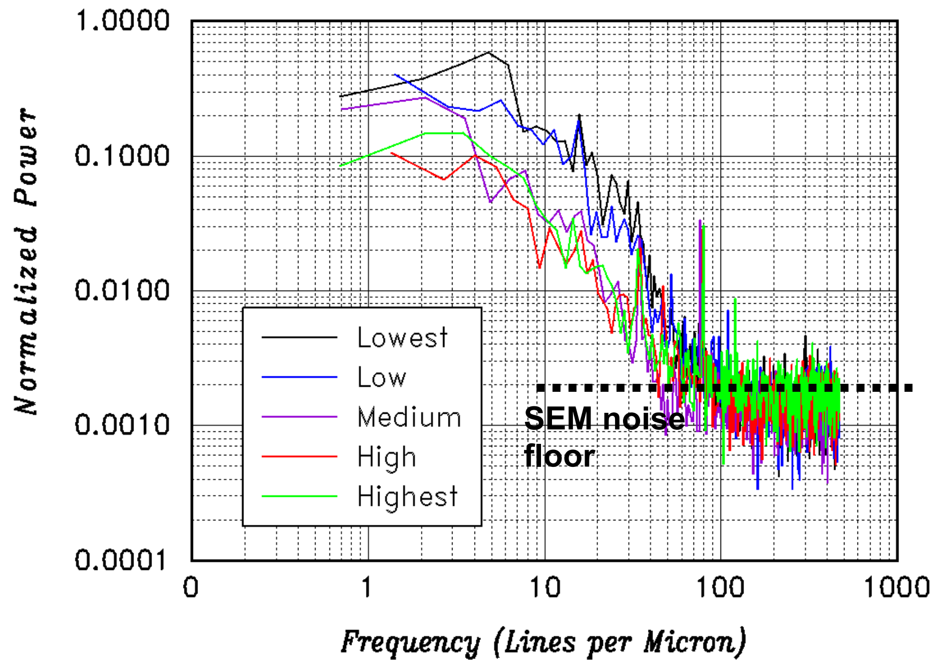
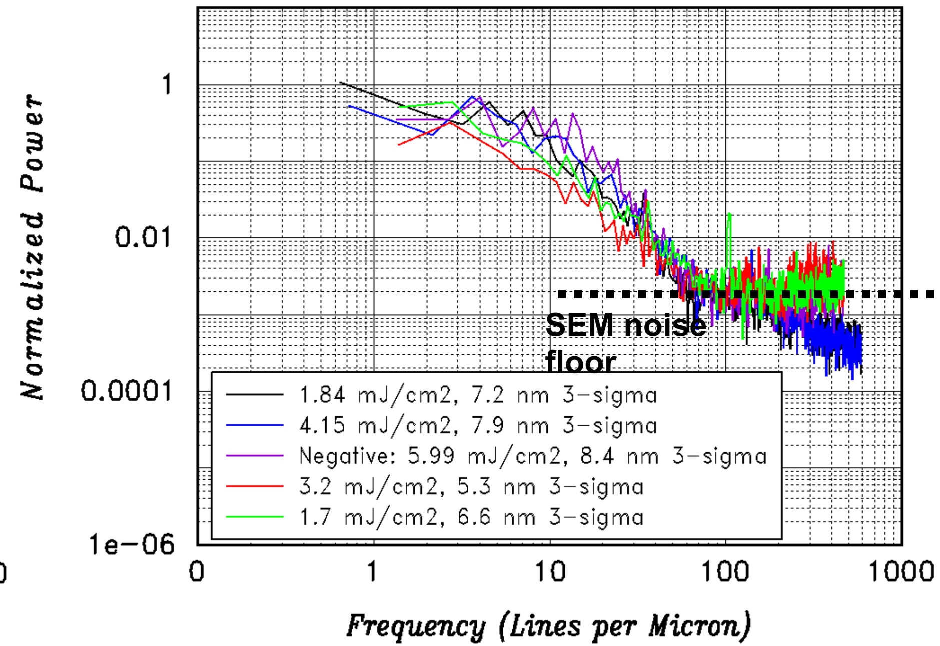


Fig. 14 Power spectral density as a function of spatial frequency for resists of various base concentration in the (a) copolymer and (b) terpolymer resist systems.

EUV Data from Cobb,et. al. SPIE 02.





General Behavior of LER

…."known" LER Scaling Laws….

- LER ~ 1 / Dose$^{1/2}$

  … Usually attributed to… and/or… called "shot noise"

- LER ~ 1 / Image edge slope ~ 1 / Image Contrast ~ 1 / ILS

  … Holds in the regime of low image contrast.

- LER power spectra (PSD) ~ Generic Modified Lorentzian

  Shape ~ 1 / (1 + $(\xi\beta)^{\gamma}$) with $\gamma \approx 3$

  … Dominated by low frequency content

  ➔ LER autocorrelation length is "large" ~ 20-30nm

  … Determines how LER impacts CD variation

LER ~ 1 / Dose$^{1/2}$

Lots of references…..

For example: from Brainard, et. al., SPIE 2004

sections are consistent with the conclusion that the LER *vs.* E_{size} behavior at both DUV and EUV are defined by the Poisson statistics of shot noise.[8]

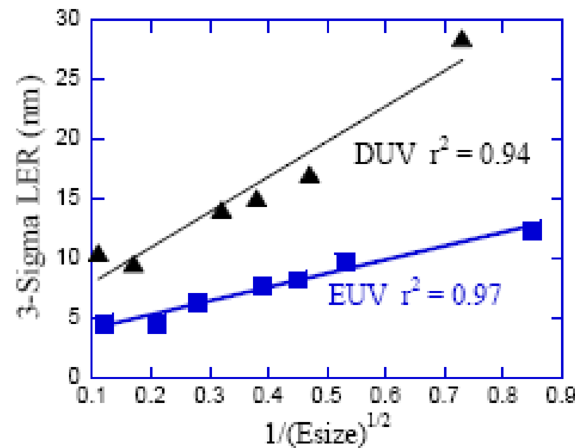
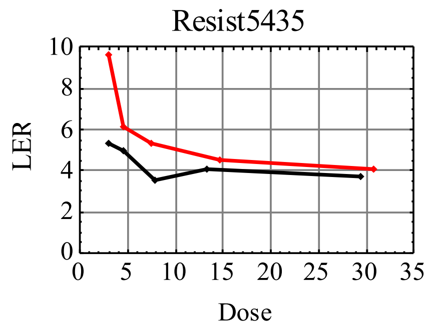
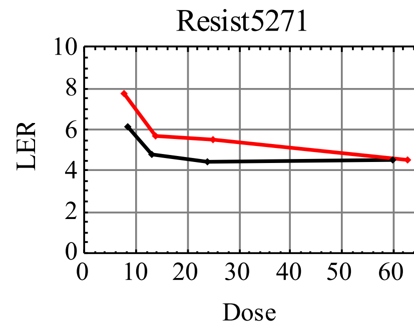
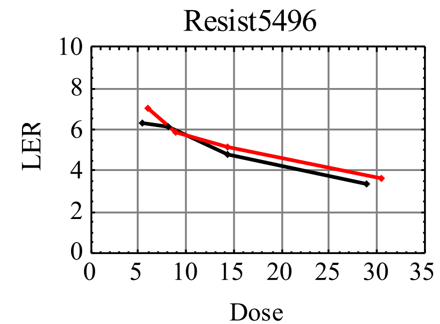


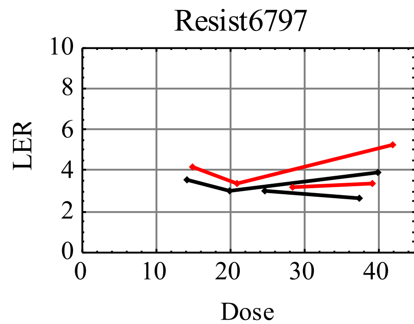
Figure 4. LER *vs.* $(E_{size})^{-1/2}$ for DUV and EUV exposure of the seven EUV-2D type resists with seven levels of added base.

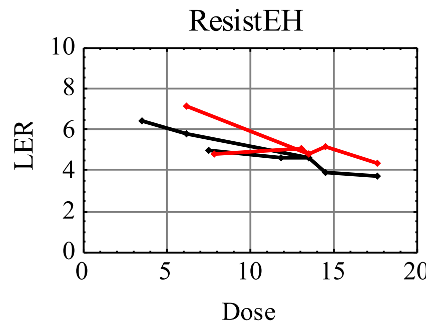
LER versus Dose: Recent EUV data (“RLS” project supported by Sematech)











NOTE: LER “saturates” at high dose

(high dose ~ high baseloading)

Resists:

“5435” ~ EUV2D (High Ea Phenolic)
“5271” ~ MET2D (High Ea Phenolic)
“5496” ~ Low Ea Phenolic
“6797” ~ Aliphatic 193nm resist
“EH” ~ High Ea Phenolic

LER ~ 1/Image Log Slope =1 / ILS

$$ILS = \frac{\partial I(x)/\partial x}{I(x)} = \frac{\partial}{\partial x}\ln(I(x))$$

Lots of references…..

For example: from Pawloski, JM3 2006

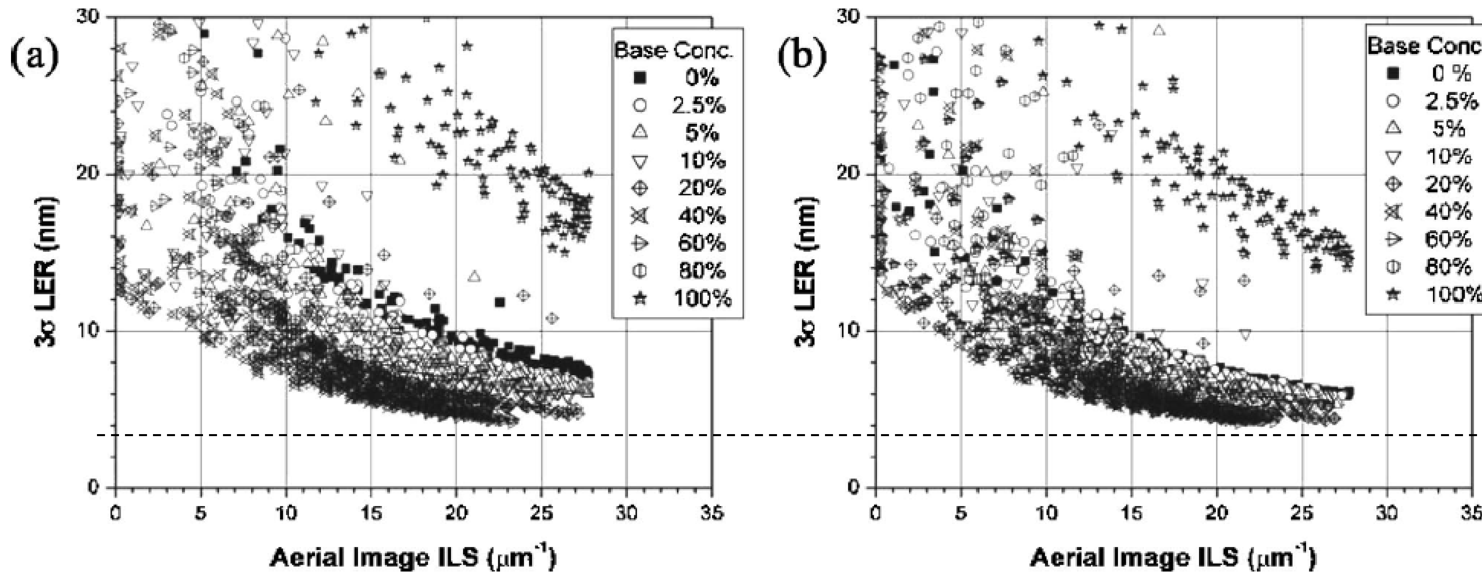


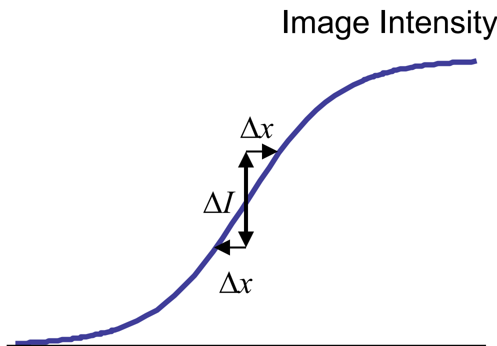
Fig. 7 Experimentally measured 3σ LER as a function of the aerial image ILS for resists with various base concentrations in the methacrylate (a) copolymer and (b) terpolymer systems.

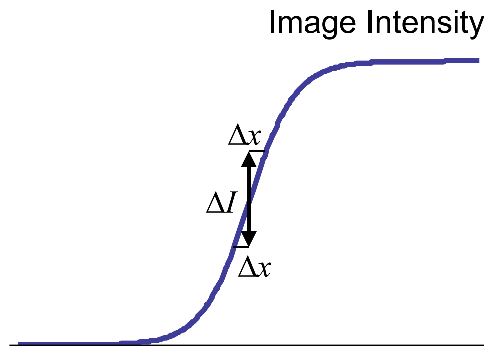
NOTE Saturation at high ILS

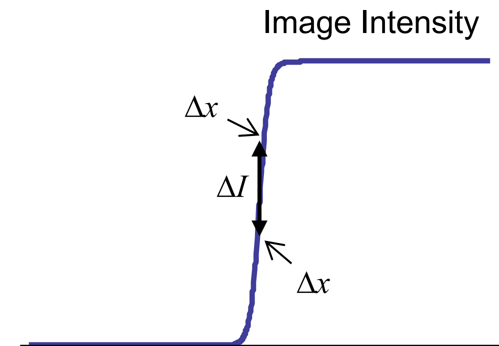
“Known” Scaling Law

$$LER \propto \frac{1}{\left(\frac{\partial I(x)/\partial x}{I(x)}\right)} \sim \frac{1}{\left(\frac{\partial \ln(I)}{\partial x}\right)} \sim \frac{1}{ILS} \propto \frac{1}{\text{Image Slope}}$$

“Why the dependence on 1/ILS ??”







Given ΔI...

Low ILS ➔ Low slope
➔ Large Δx
➔ High LER

Medium ILS ➔ Medium slope
➔ Medium Δx
➔ Medium LER

High ILS ➔ High slope
➔ Low Δx
➔ Low LER

Implication: LER →0 as ILS → ∞...Data ➔ LER saturates at high ILS

Pawloski et al.: Line edge roughness and intrinsic bias...

LER Frequency Content

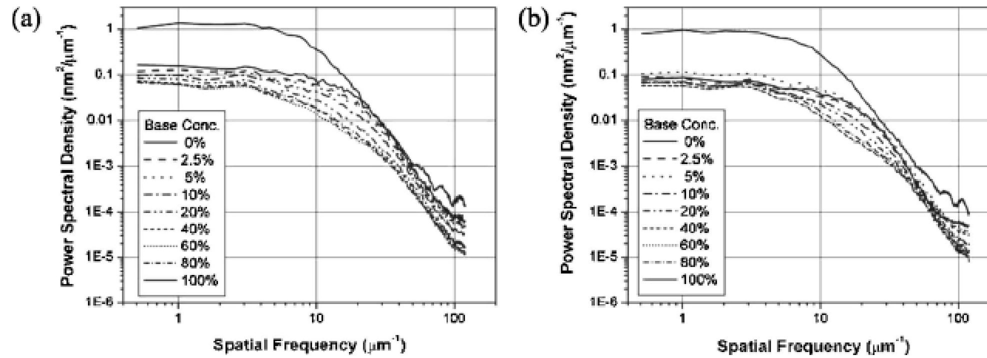
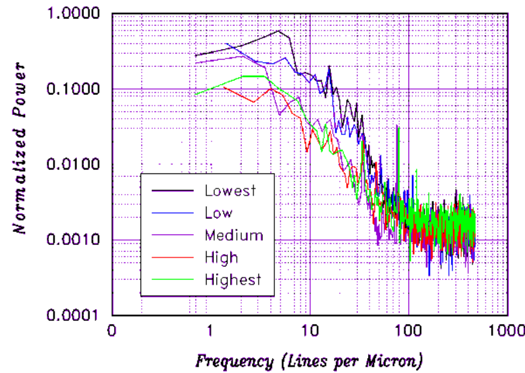
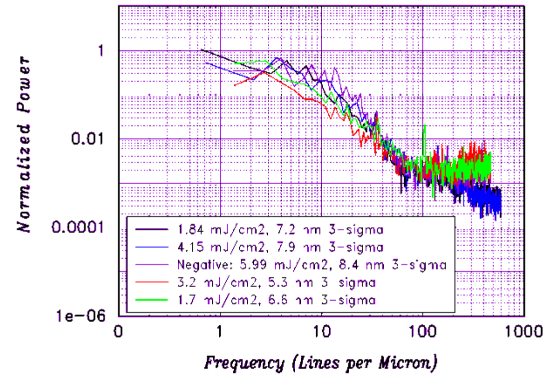


Fig. 14 Power spectral density as a function of spatial frequency for resists of various base concentration in the (a) copolymer and (b) terpolymer resist systems.

Pawloski, et. al. JM3, 2006





Cobb, et. al. SPIE, 2002

The key question

"Why is LER dominated by low spatial frequencies?"

Alex Liddle

What causes LER?

Many contributors

- Quantum Mechanics
- Shot Noise
- Chemical Statistics
- Reaction Diffusion Statistics
- Speckle
- Mask Roughness
- Mesoscopic Structure (non-uniformity in resist properties)
- …

Modeling LER?

	Advantages	Disadvantages
Analytical Models	Nominally straightforward to understand, use and interpret. Work on the macroscopic level. Approximations used are explicit, i.e., you know explicitly what is and is not included in the model.	Difficult to incorporate microscopic details of the chemistry, etc. …
Numerical Models	"Easily" incorporate details of the chemistry and physics on the molecular/atomic scale. Allow you to work up from the microscopic scale. Provide a direct connection of the fundamental chemistry and physics to the statistics of resist exposure and development	Detailed modeling of 3D resist over large volumes is prohibitive. Difficult/Impossible to get to the macroscopic scale of LER, i.e., the LER coherence length ~ many 10's of nm or more.

Lots of models have been generated by lots of different people

Here we study a “3 step” analytical model that explains

- The various scaling laws
- The “saturation” of LER with Dose and ILS
- The low frequency content of the LER
- The shape of the PSD.

This model

- Is fundamentally analytic and linear or quasi-linear
- Covers the Quantum, Shot Noise, Chemical Statistics and Reaction-Diffusion contributors to LER

LER Model we will study here:

Generated by tracking the process steps of a chemically amplified resist.

1. Exposure: Releases acid from "photo acid generator" (PAG) compound in the resist

   LER Model: Probability of acid generation ~ Image intensity

   ➔ Sensitivity ("Dose-to-Size")

2. Post Exposure Bake (PEB)….. "bake wafer at 90°C for 120seconds"

   LER Model: Acid diffuses and "deprotects" resist polymer

   ➔ Resist Resolution (Intrinsic Resist "Blur")

3. Development: Converts deprotection density into final resist profile

   LER Model: Statistical Fluctuations in deprotection produce roughness

   ➔ LER (3σ roughness)

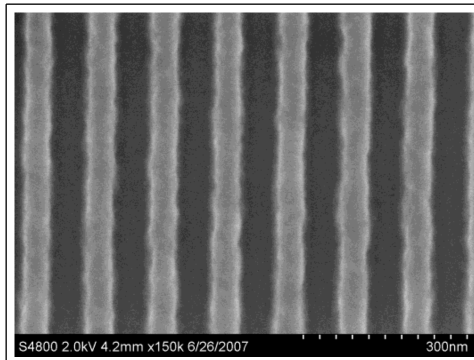
Model details: Gallatin SPIE 2005
Model is very similar to Fukuda JPST 02, Fukuda JJAP 03, Brainard SPIE 04 and others.
Same approach used for CD evaluation: Fuard SPIE 03, Shumway SPIE 04, Naulleau AO 04, etc...

Exposure according to Mr. Schrodinger

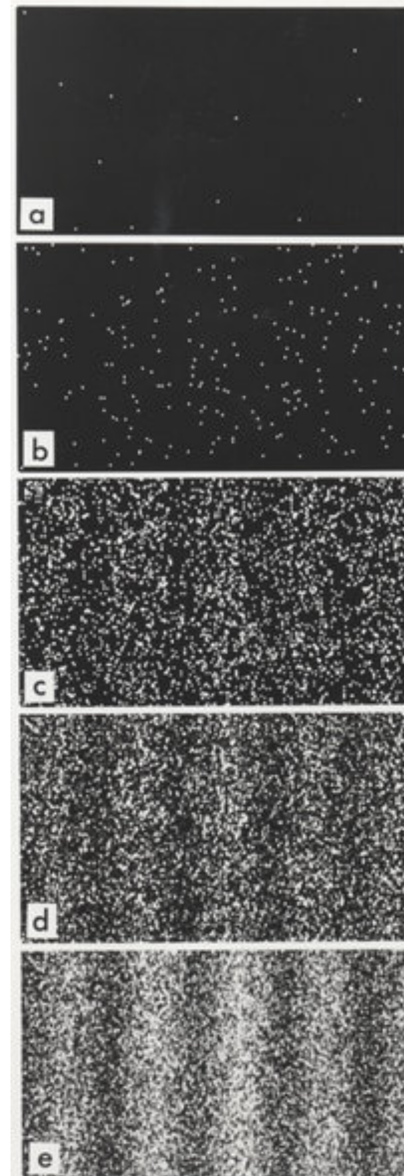
You control the image intensity
but not where interactions take place
in the photoresist.

Consequences:

- Line Edge Roughness (LER)



- Contact Hole Size Variation
- … etc…

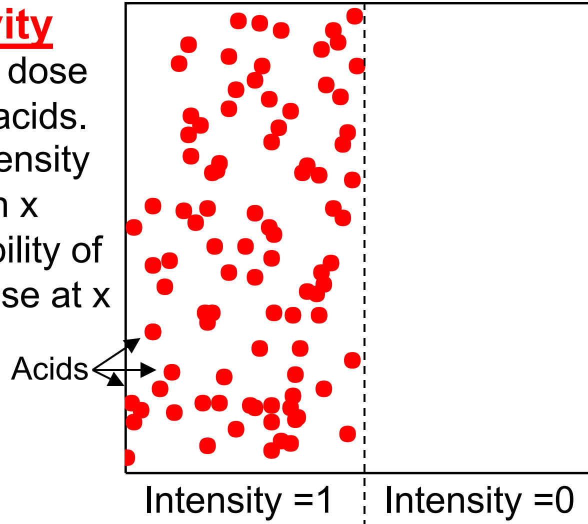


Time

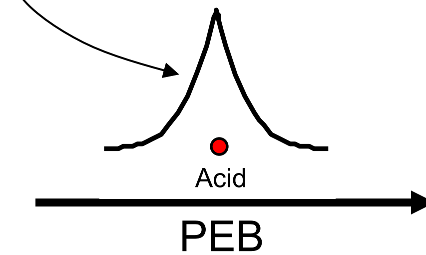
LER Model → RLS Model

Gallatin SPIE 2005

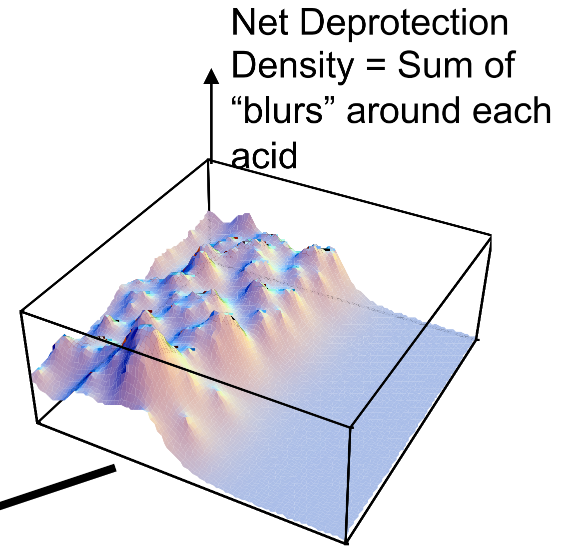
Sensitivity
Exposure dose releases acids. Image intensity at position x
→ Probability of acid release at x



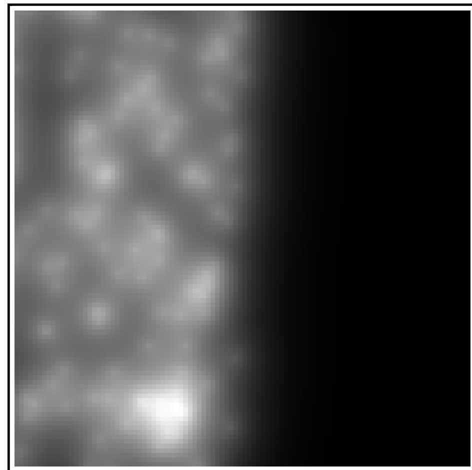
Resolution
Diffusion/deprotection "blur" develops around each acid during PEB



Net Deprotection Density = Sum of "blurs" around each acid



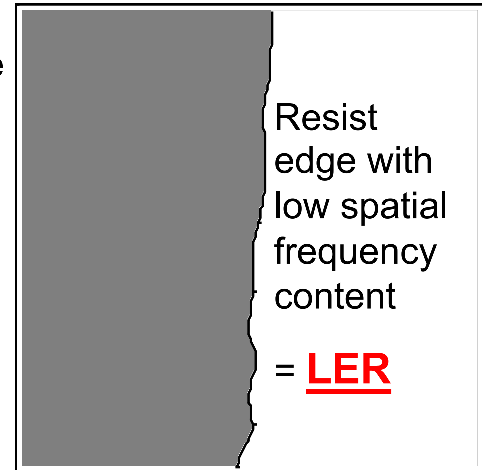
Grayscale plot net deprotection density



Resist edge position occurs at a fixed value of deprotection
~Critical Ionization Model

Development

Burns,et.al., JVST B 2002



Exposure

$\rho_{PAG}(\vec{r},t) = \text{PAG density at position } \vec{r} \text{ at time } t$

$$\rho_{PAG}(\vec{r},t) = \rho_{PAG}(\vec{r},0) - \rho_{Acid}(\vec{r},t)$$

$$\frac{\partial \rho_{Acid}(\vec{r},t)}{\partial t} = \alpha\, Q\, v\, I(\vec{r})\rho_{PAG}(\vec{r},t) = \alpha\, Q\, v\, I(\vec{r})\big(\rho_{PAG}(\vec{r},0) - \rho_{Acid}(\vec{r},t)\big)$$

$\alpha = \text{resist absorptivity (} \exp[-\alpha T] = \text{transmitted intensity, } T = \text{resist thickness)}$

$Q = \text{"Quantum Efficiency"} = \text{\# acids generated/absorbed photon.}$

$v = \text{photon - PAG interaction volume}$(Only PAGs within volume v surrounding the position of absorption can be affected)

$I(\vec{r},t) = \text{Image intensity (\# photons/(area} \times \text{time))}$

$$\text{Solution}: \rho_{Acid}(\vec{r},t) = \rho_{PAG}(\vec{r},0)\big(1 - \exp[-\alpha\, Q\, v\, I(\vec{r})t]\big) = \rho_{PAG}(\vec{r},0)\big(1 - \exp[-\alpha\, Q\, v\, E(\vec{r})]\big)$$

$E(\vec{r}) = I(\vec{r})t = \text{Dose (\# photons/area)}$

$\alpha\, Q\, v = \text{Dill C (area/\# photons)} \quad \Rightarrow \quad 1/\text{C} = \text{Saturation Dose} = E_{sat}(\text{\# photons/area})$

Exposure Statistics

Previous slide ➔ Exposure is deterministic. **IT IS NOT**

Quantum Mechanics ➔ Probability of photon absorption ~ Image Intensity

$\rho_{Acid}(\vec{r},t)$ should be interpreted as a probability

Probability of a PAG to release an acid at position $\vec{r} = \left(1 - e^{-CE(\vec{r})}\right)$

Probability of a PAG not to release an acid at position $\vec{r} = e^{-CE(\vec{r})}$

Let $a = 1 \Leftrightarrow$ acid released

$a = 0 \Leftrightarrow$ acid NOT released

Probabitity distribution for an acid to be generated at $\vec{r}$

$$P_{acid}(\vec{r},a) = \delta_{a,1}\left(1 - e^{-CE(\vec{r})}\right) + \delta_{a,0}e^{-CE(\vec{r})}$$

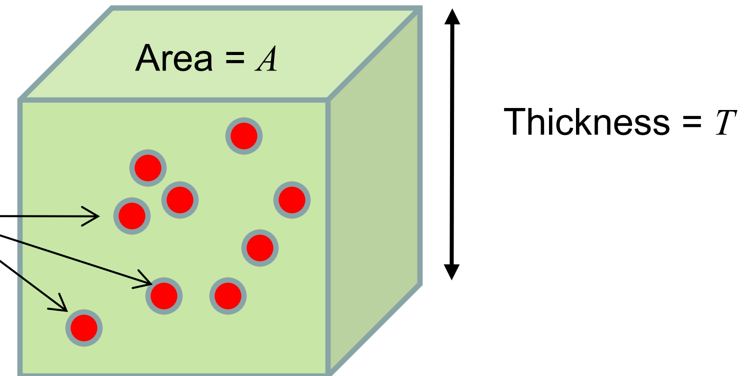
Net probability distribution

$$P_{acid}^{net}(\vec{r}_1,a_1,\vec{r}_2,a_2,\cdots,\vec{r}_N,a_N) = P_{acid}(\vec{r}_1,a_1)P_{acid}(\vec{r}_2,a_2)\cdots P_{acid}(\vec{r}_N,a_N)$$

Combine

- Distribution of Released acids
- PAG distribution

N PAGs uniformly distributed throughout volume $V = AT$



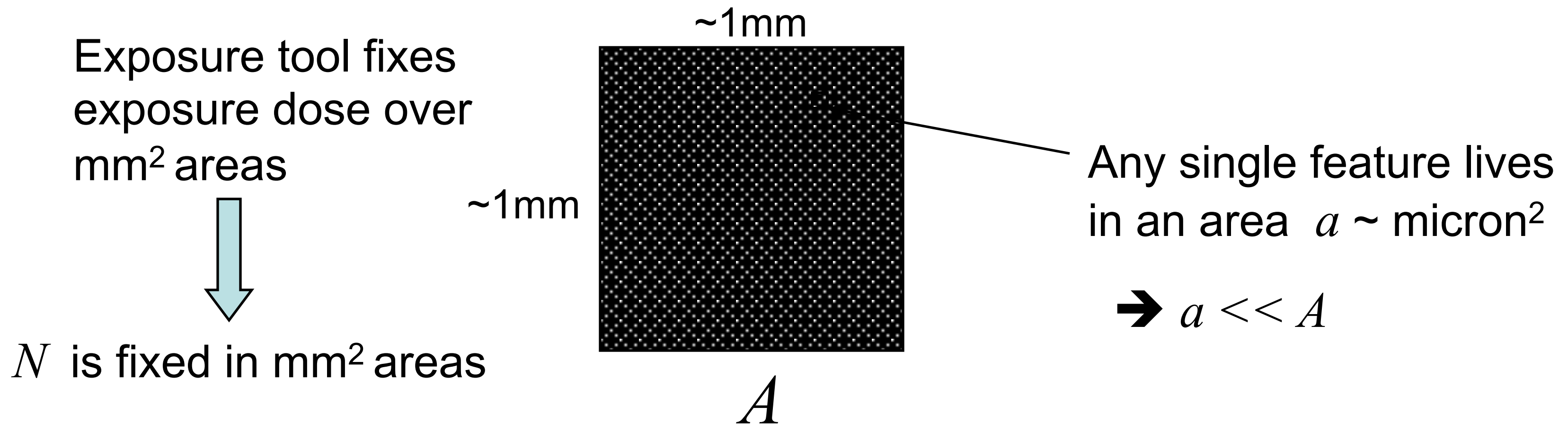
$$\Rightarrow \text{Probability distribution for one PAG } = \frac{1}{V}$$

$$\Rightarrow \quad \text{Joint probability distribution for all } N \text{ PAG's} = P_{PAG} = V^{-N} = (AT)^{-N}$$

$$N = \text{the total number of PAG molecules loaded into the resist volum } V = AT$$

$$\Rightarrow \quad \rho_{PAG}(\vec{r}, 0) = \frac{N}{V} = \frac{N}{AT}$$

Poisson Statistics



Quantum Mechanics ➔ Image in resist builds up one absorption at a time

Probability for releasing a fixed number of acids $P(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) = P(\vec{r}_1)P(\vec{r}_2)\cdots P(\vec{r}_N)$

Probability of getting n acids in area $a << A$ $P_n(a << A) \rightarrow Poisson$

Fixed number of acids the macroscale ➔ Poisson statistics on the microscale.

"shot noise"

PEB Reaction Diffusion "blur"

- Acids deprotect polymer during PEB
- Assume acids act independently (deprotection is not saturated)
- Solve chemical kinetics rate equations for the deprotection generated by a single acid located at $\vec{r} = 0$

Rate at which polymer is deprotected

$$\frac{\partial \rho_P(\vec{r},t)}{\partial t} = -k\,\rho_{Acid}(\vec{r},t)\rho_P(\vec{r},t) + \text{higher order terms}$$

Rate at which acid diffuses around.

$$\frac{\partial \rho_{Acid}(\vec{r},t)}{\partial t} = D\vec{\nabla}^2 \rho_{Acid}(\vec{r},t)$$

NOTE: The position dependence, i.e., image information, is encoded in the acid distribution

Density of Protected Polymer = $\rho_P(\vec{r},t)$

Deprotected Density + Protected Density = Initial Polymer Density

Density of Deprotected Polymer = $\rho_D(\vec{r},t) = \rho_P(\vec{r},0) - \rho_P(\vec{r},t) \equiv \rho_{P0} - \rho_P(\vec{r},t)$

$$k = \text{deprotection rate constant (Units = "volume"/time)}$$

PEB Reaction-Diffusion Equations ➔ $$\rho_D(\vec{x},t) = 1 - \exp\left[-k\int_0^t dt\; e^{D\,t\,\vec{\nabla}^2}\,\delta(\vec{x})\right]$$

Deprotection Density

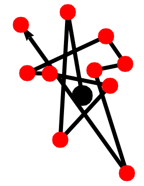
Deprotection Rate Units $= \dfrac{nm^d}{\sec}$

$d = \#\text{Dimensions}$

Diffusion Range = Resist "Blur"

$$R = \sqrt{D\,t}$$

Acid Diffusion



3D: $$\rho_D(\vec{r},t) = 1 - \exp\left[-\frac{kt}{4\pi R^2 r}\left(1 - erf\left(\frac{r}{2R}\right)\right)\right]$$

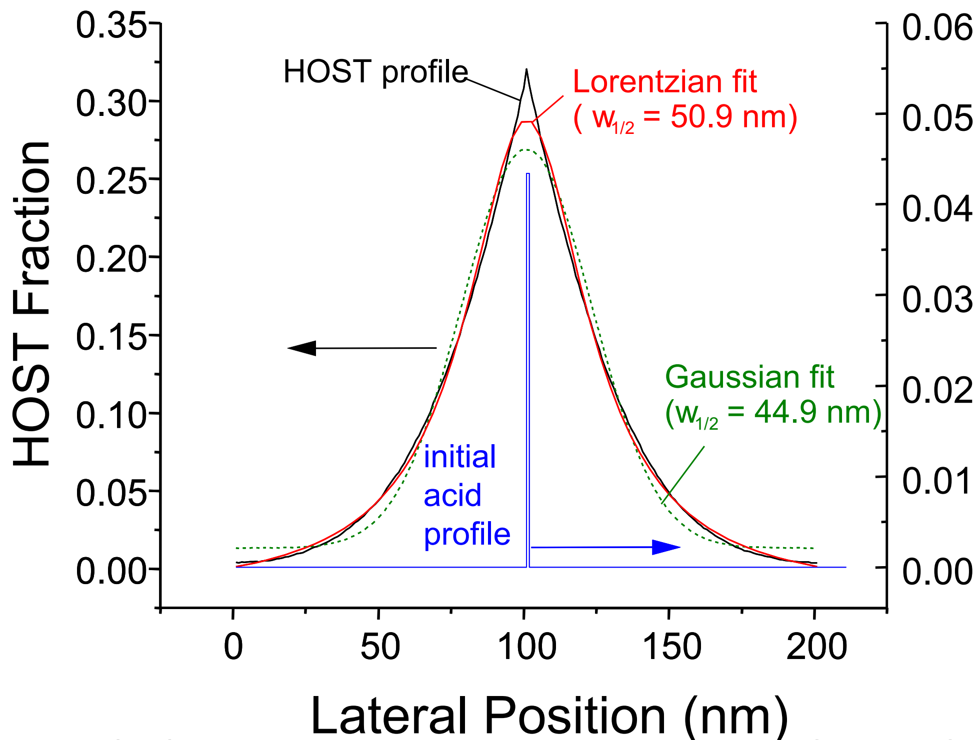
2D: $$\rho_D(\vec{r},t) = 1 - \exp\left[-\frac{kt}{4\pi R^2}\Gamma\left(0,\left(\frac{r}{2R}\right)^2\right)\right]$$

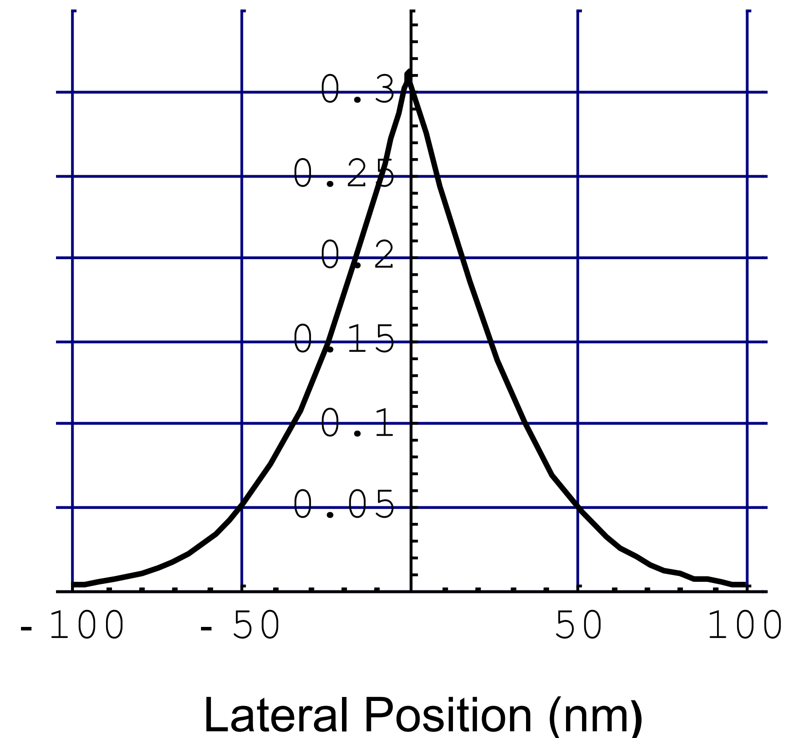
$\Gamma(a,b) = \text{Incomplete Gamma Function}$

1D: $$\rho_D(\vec{r},t) = 1 - \exp\left[-kt\left(\frac{1}{\sqrt{\pi}R}e^{-r^2/2R^2} - \frac{r}{2R^2}\left(1 - erf\left(\frac{r}{2R}\right)\right)\right)\right]$$

Evaluate 1D result

- Matches full numerical simulation Hinsberg, et. al, SPIE 03
- And experimental shape Hoffnagle, Opt. Letts. 02



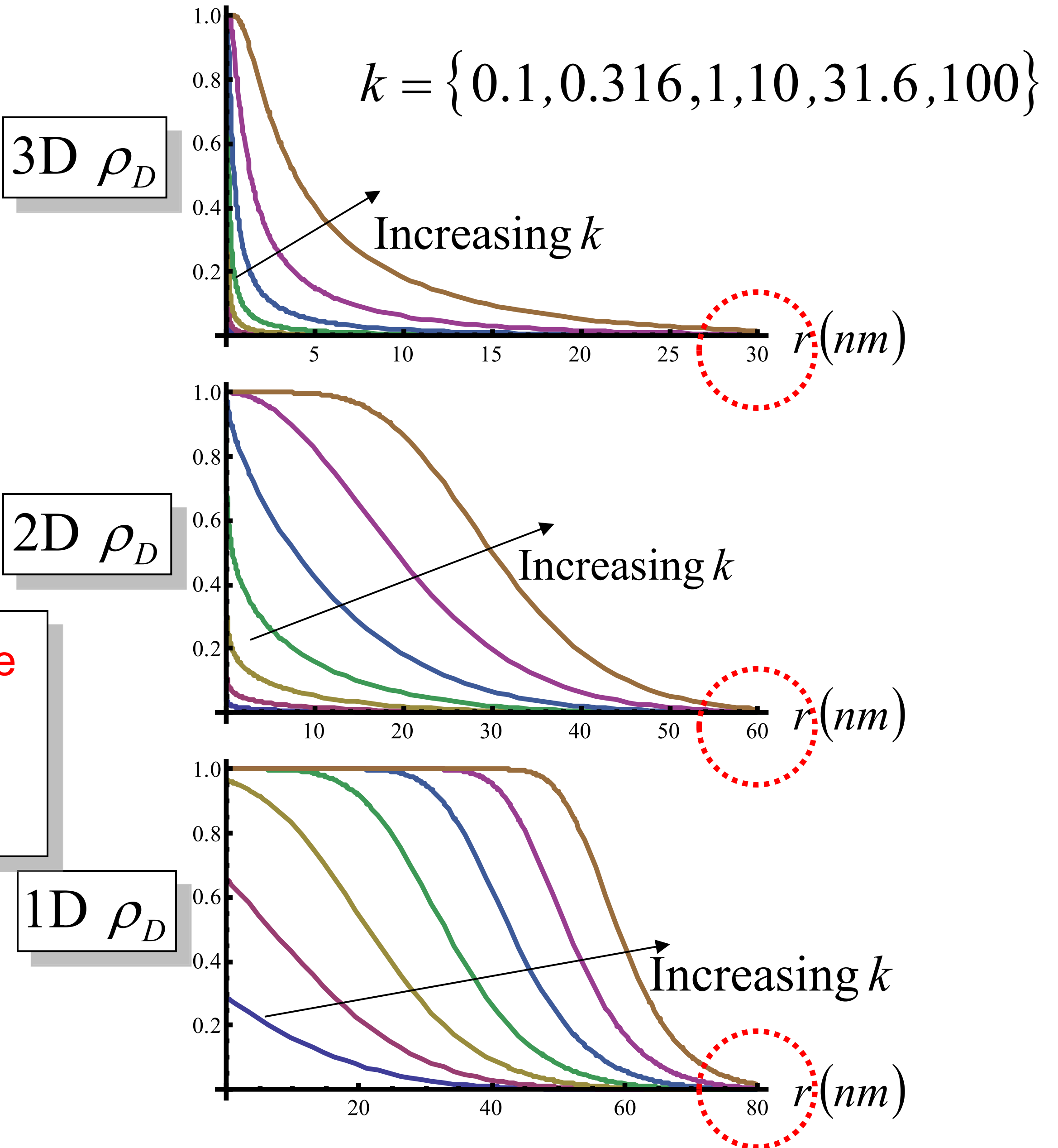


Graphs of Deprotection Density

For all 3 Graphs

$$R = 15nm$$

The value of k and the space dimensionality both have a very strong effect on the effective resist blur.



Kang, et. al., SPIE 6519 (2007)

Development

- Highly nonlinear process ➔ Difficult to model in detail
  - Algorithms
    - String algorithm (2D)
    - Surface algorithm (3D)
    - Ray algorithm (same idea works in any D)
    - Cell algorithm
    - ….

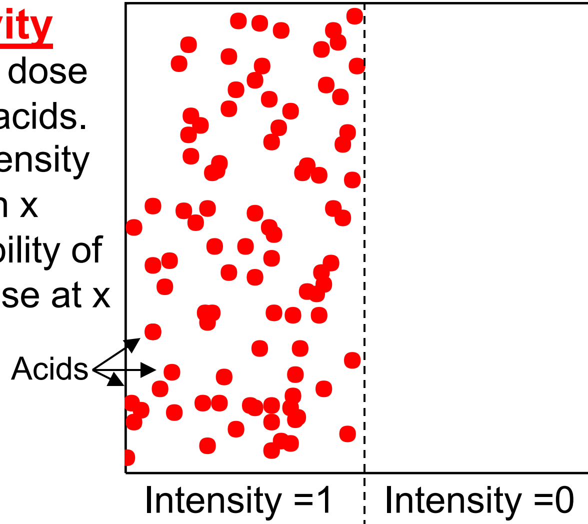
Here we use a naïve simplified version of the “critical ionization criteria”

Surface on which the net deprotection = constant ➔ Resist surface

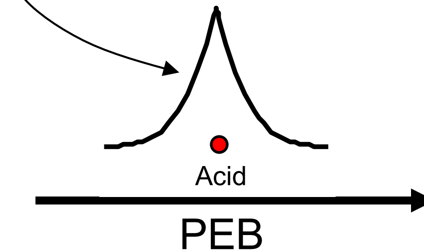
LER Model → RLS Model

Gallatin SPIE 2005

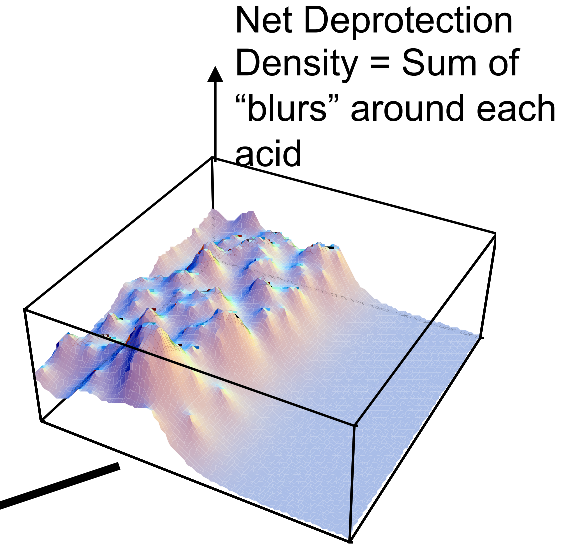
Sensitivity
Exposure dose releases acids. Image intensity at position x → Probability of acid release at x



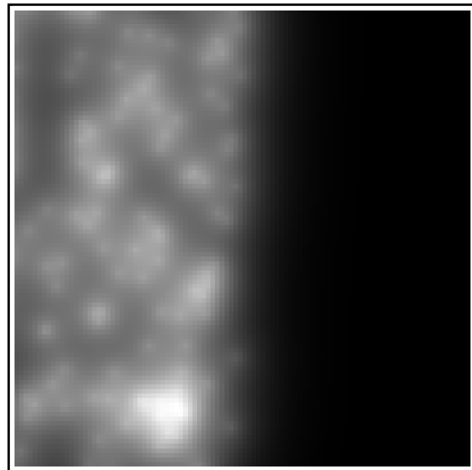
Resolution
Diffusion/deprotection "blur" develops around each acid during PEB



Net Deprotection Density = Sum of "blurs" around each acid



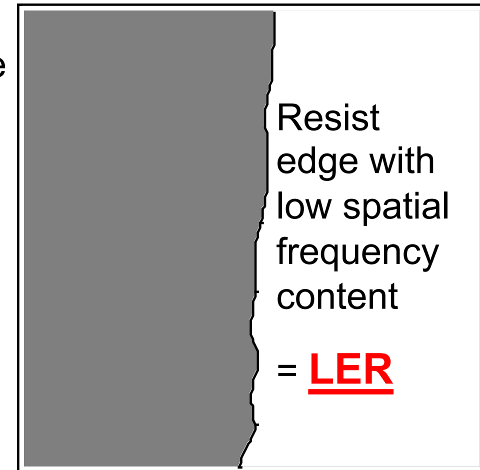
Grayscale plot net deprotection density



Resist edge position occurs at a fixed value of deprotection
~Critical Ionization Model

Development

Burns,et.al., JVST B 2002



$$\text{Net Deprotection} = \sum_n a_n \rho_D\left(\vec{r} - \vec{r}_n\right)$$

$\vec{r}_n = \text{PAG positions}$

$$a_n = \begin{cases} 1 \text{ if acid } n \text{ is released} \\ 0 \text{ if acid } n \text{ is not released} \end{cases}$$

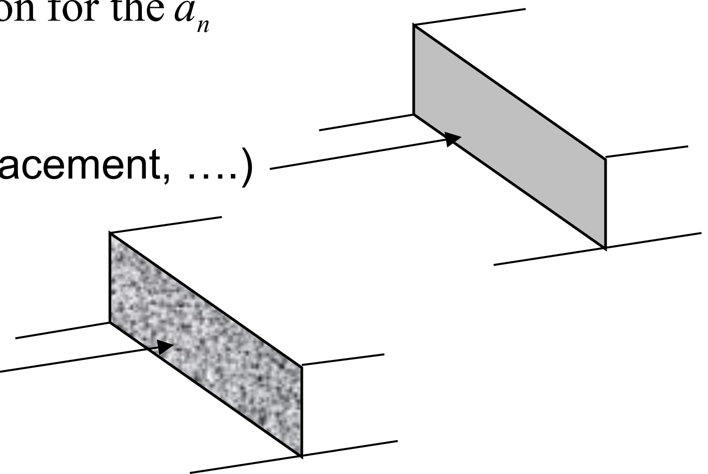
Resist surface defined implicitly by $\sum_n a_n \rho_D\left(\vec{r}_{surface} - \vec{r}_n\right) = \text{constant}$

Surface depends on the particular values of a_n and $\vec{r}_n$

Statistics given by $\begin{cases} P_{PAG} = \text{Probability distribution for the } \vec{r}_n \\ P_{acid}^{net} = \text{Probablitiy distribution for the } a_n \end{cases}$

Average ➔ Smooth resist surface (CD, edge placement, ….)

RMS Fluctuations ➔ Edge roughness = LER



…do the math…

- Average or Mean resist surface $\vec{r}_{surface}$ is given implicitly by

$$\text{constant} = \tau = \bar{\rho}_{PAG} \int_V d^3 r \, \rho_D\left(\vec{r}_{surface} - \vec{r}\right)\left(1 - \exp\left[-\alpha Q v E(\vec{r})\right] - \frac{\bar{\rho}_{Base}}{\bar{\rho}_{PAG}}\right)$$

Total resist volume

- Resist edge roughness autocorrelation function

$$ACF\left(\vec{r}_{edge}, \vec{r}_{edge}'\right) = \left(\frac{1}{\bar{\rho}_{PAG}}\right)\left(\frac{1}{\alpha Q v}\right)^2 \frac{\int_V d^3 r \, \rho_D\left(\vec{r}_{edge} - \vec{r}\right)\rho_D\left(\vec{r}_{edge}' - \vec{r}\right)\left(1 - \exp\left[-\alpha Q v E(\vec{r})\right] - \frac{\bar{\rho}_{Base}}{\bar{\rho}_{PAG}}\right)}{\left(\int_V d^3 r \, \rho_D\left(\vec{r}_{edge} - \vec{r}\right)\left(\partial_\perp E(\vec{r})\right)\exp\left[-\alpha Q v E(\vec{r})\right]\right)^2}$$

Derivative in direction perpendicular to the edge

Model ➔ Make Approximations ➔ LER Scaling Laws

The scaling laws are approximations to integrals in the full model prediction.
Model predicts "bimodal" scaling depending on the relative sizes of the optics and resist blurs.

Resist resolution is better than Image resolution

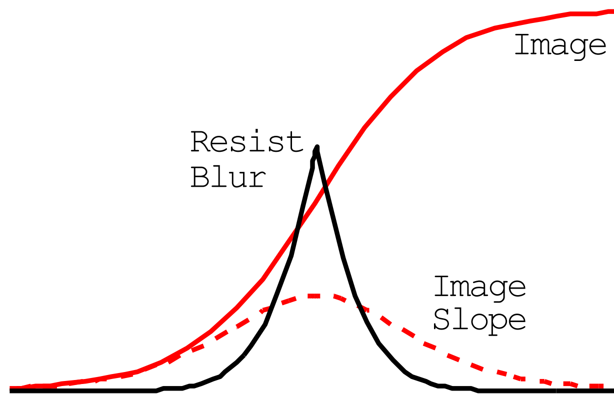
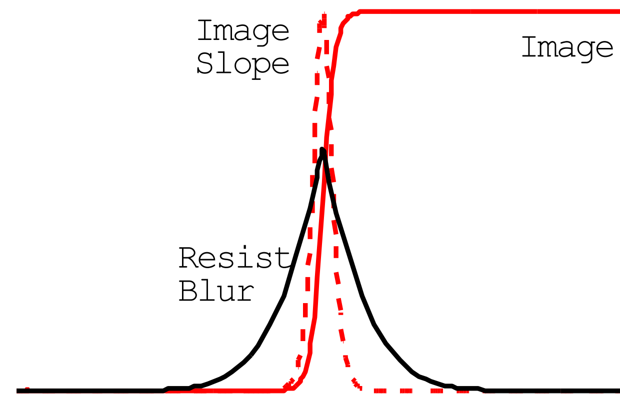


Image resolution is better than Resist resolution



$$1\sigma\ \text{LER} \approx \left(\frac{I_{edge}}{\partial I_{edge}}\right)\sqrt{\frac{1}{\bar{\rho}_{\text{PAG}}\alpha\, Q\, v\, E_{size} R^3}}$$

$$\underbrace{\left(\frac{I_{edge}}{\partial I_{edge}}\right)}_{\frac{1}{ILS}}$$

$$1\sigma\ \text{LER} \approx \sqrt{\frac{1}{\bar{\rho}_{\text{PAG}} R\, \alpha\, Q\, v\, E_{size}}}$$

R = Deprotection "blur" radius (nm)
E_{size} = Dose-to-Size (#photons/nm^2)
$I(x)$ = Image intensity
α = resist absorptivity (1/nm)
T = resist thickness (nm)
Q = quantum efficiency
v = photon-PAG interaction volume (nm^3)
$\bar{\rho}_{\text{PAG}}$ = PAG loading density (#PAG/nm^3)

Scaling Laws ➔ Resolution, LER and Sensitivity are tightly coupled

Model mixes and hence couples the process steps to LER

1. Exposure: Image intensity ~ Probability distribution for acid release.
   Acid release positions are statistical ➔ **Sensitivity**

2. PEB: Acids diffuse and deprotect the resist
   PEB diffusion range ➔ **Resolution (resist "blur")**

3. Development: Spatial distribution of deprotection determines final resist profile
   Line edge statistics ➔ **LER**

Model Scaling Law

$$\sigma_{LER} \approx \left(\frac{I}{\partial I}\right)_{edge} \sqrt{\frac{T}{\bar{\rho}_{PAG}\,\alpha\, Q\, v\, E_{size} R^3}}$$

LER (σ_{LER}), Dose (E_{size}), "Blur " (R^3)

I = Image Intensity

E_{size} = Sizing Dose (#photons/nm^2)

T = Resist Thickness (nm)

α = Absorptivity (nm^{-1})

Q = Quantum Efficiency

v = photon - PAG interaction volume (nm^3)

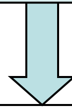
R = PEB Diffusion Range (nm)

$\bar{\rho}_{PAG}$ = PAG loading (#PAG/nm^3)

Rearrange the scaling law

➔ "The Resolution, LER, Sensitivity (RLS) Tradeoff"

$$Blur^3 \times LER^2 \times Dose \sim \text{Constant}$$



For a standard chemically amplified resist and process cannot have blur, LER and dose all small at the same time.

"You can't always get what you want"… Mick Jagger

This type of behavior has been found by many researchers:

Lammers, et al., SPIE 2007, Bristol, et al., SPIE 2007, Brainard, et al., SPIE 2004, Gallatin SPIE 2005, ...

LER Model → Analytical form of the PSD

Formula below is valid for small k. For large k, compute PSD perturbatively for an analytic solution or do the integral numerically.

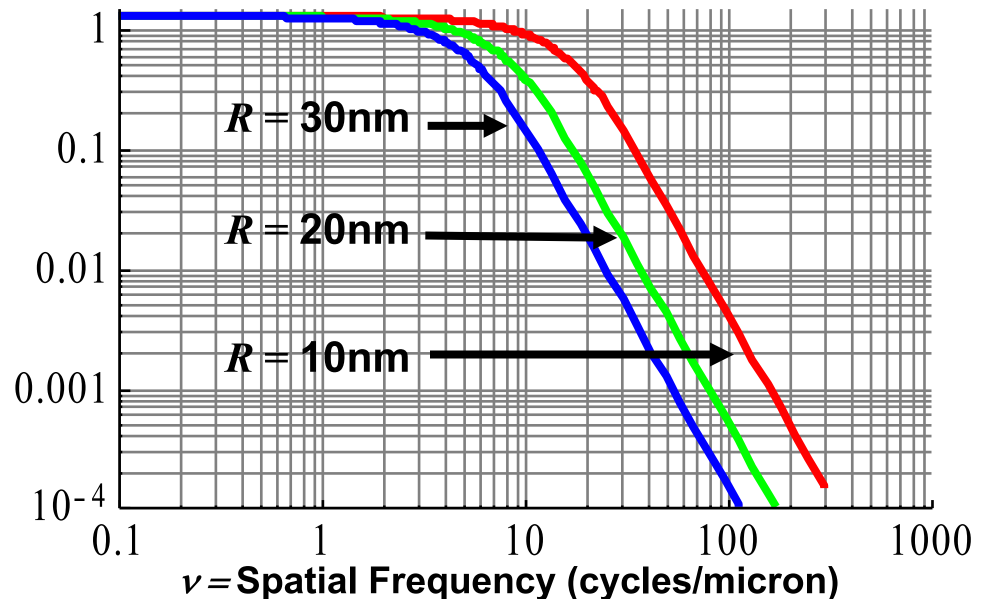
$$PSD(\beta) = norm \times \frac{1}{(R\beta)^3}\begin{bmatrix} 2(R\beta)e^{-2(R\beta)^2}\left(\sqrt{2\pi} - 2\sqrt{\pi}e^{(R\beta)^2}\right) \\ + 2\pi\left(1 - 2(R\beta)^2\right)erf(R\beta) \\ + \pi\left(4(R\beta)^2 - 1\right)erf\left(\sqrt{2}R\beta\right) \end{bmatrix}$$

Normalization factor ~ $\sigma_{LER}{}^2$

PSD "shape": Depends only on $R\beta = R2\pi\nu$

Can determine resist parameters from roughness data

- rms roughness → σ_{LER} → $norm$
- Intrinsic resist "blur" R is determined by fitting the analytic PSD "shape" to |FFT(data)|^2 / $norm$



Model comparison to recent EUV LER data

- Combined effort of LBL, CNSE, and Rohm and Haas with SEMATECH funding
- Exposures were done on the 0.3 NA MET at Berkeley
- Four different resists with 4 or 5 different base loadings each
  Resist Names: "5435" (EUV2D) "5271" (MET2D) "5496" "EH" (Open Source Resist)
- Features imaged: 50 nm and 60 nm 1-to-1 lines/spaces
- CD and LER data through dose and focus at each base loading
- LER and PSDs computed from average of left and right edge data at each focus, dose, and base loading condition
- LER computed from both filtered and unfiltered PSD data
  - "Best Dose" is as indicated in the graphs and tables
  - "Best Focus" is set to 0, by definition
- Resist blur values, R, are fit using the analytical PSD formula:
  Resist "blur" R is the only fitting parameter

Example Result: Resist "5435", Base Loading G, 50 nm 1-1 lines/spaces

LER(unfiltered)

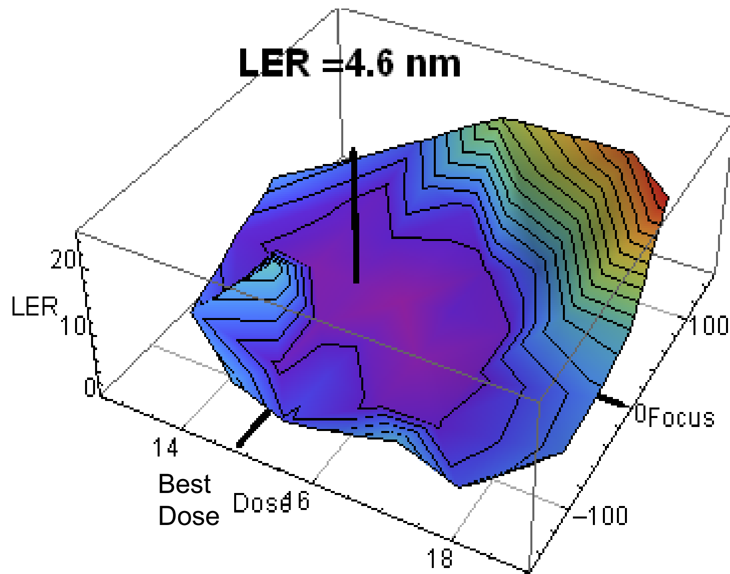
Best Focus: 0; Best Dose: 14.82

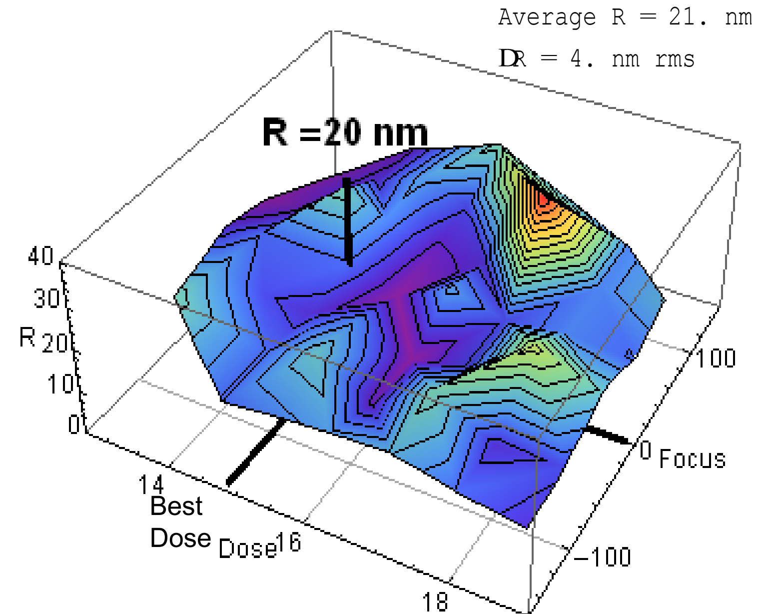
Dose\Focus	-150	-100	-50	0	50	100	150
12.8				5.5	8.8		
13.44		8.4	6.	5.2	5.5		
14.11			11.7	4.4	5.	5.1	7.4
14.82	9.9	5.	4.6	4.6	5.3	8.2	
15.56	6.8	6.3	5.	4.	5.1	5.7	15.8
16.34	8.4	4.9	4.3	5.2	5.1	14.2	
17.15	10.7	5.2	4.7	4.9	8.2		
18.01	7.3	6.5	6.8	10.7	16.5	23.	
18.91	14.7	8.6	8.8	11.5	25.		

R ("blur")

Best Focus: 0; Best Dose: 14.82

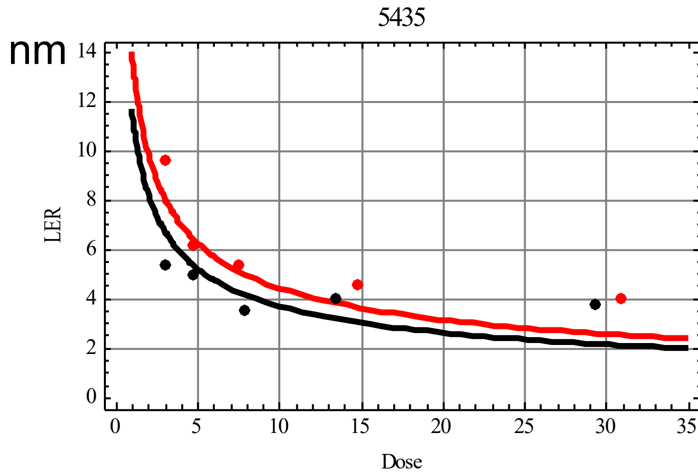
Dose\Focus	-150	-100	-50	0	50	100	150
12.8				18	16		
13.44		25	22	20	17		
14.11			19	20	25	18	17
14.82	20	20	18	20	21	22	
15.56	21	24	21	16	17	19	25
16.34	22	17	16	22	18	30	
17.16	23	25	19	19	38		
18.01	20	18	29	21	21	21	
18.91	18	19	24	20	23		

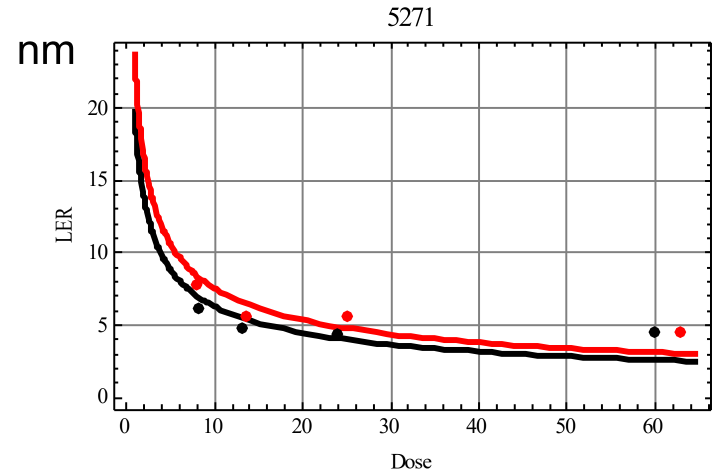


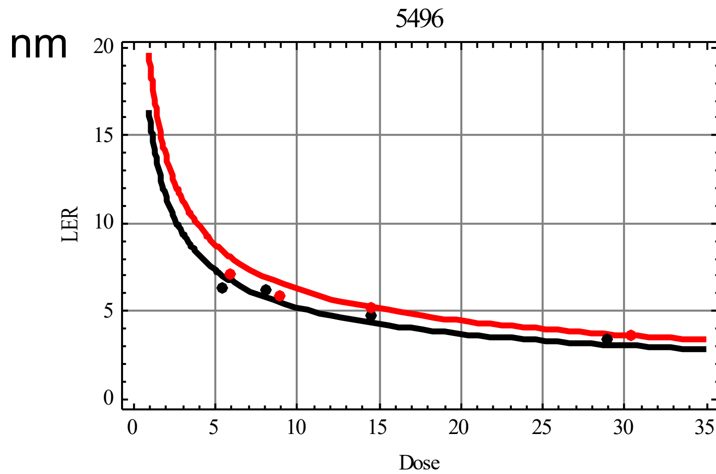


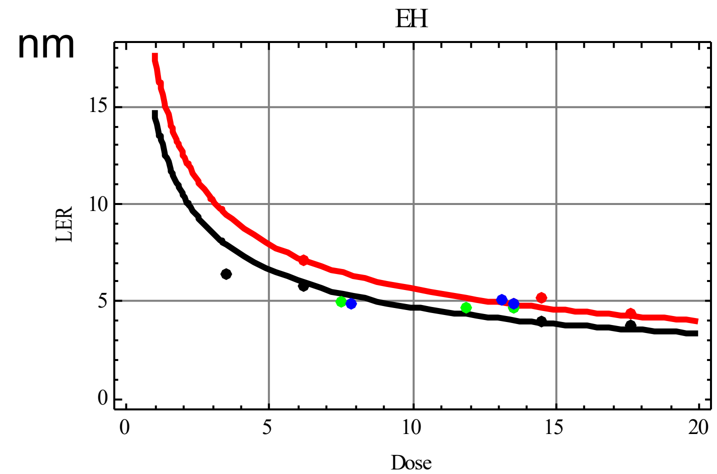
LER data compared to the scaling law

Red = 50 nm dense L/S
Black = 60 nm dense L/S
Dots = data, Curve = model using measured values of ILS, PAG density, C, Dose, Blur,….

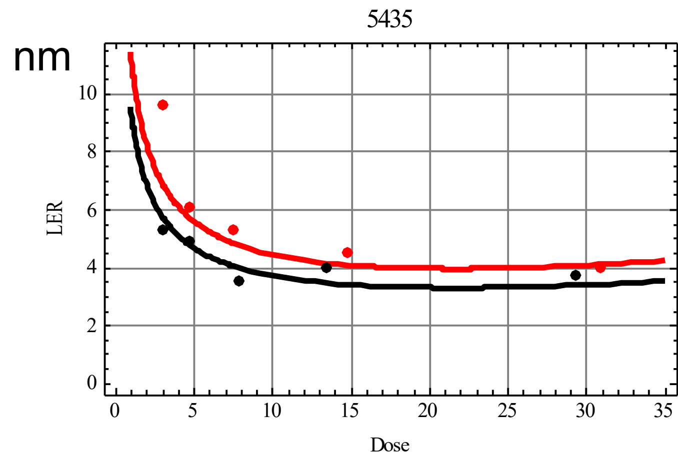
nm


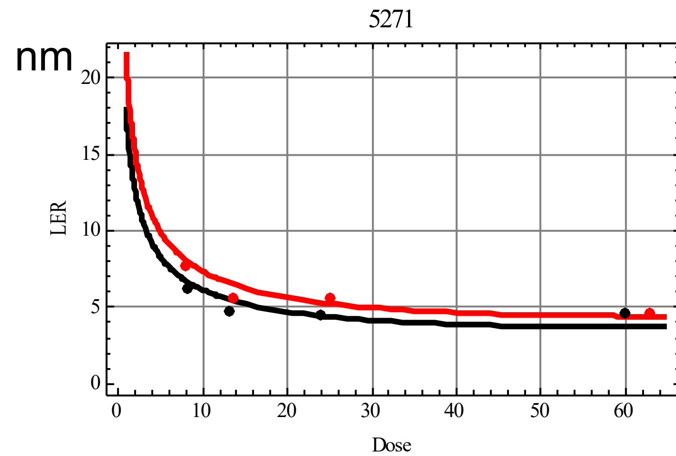
nm


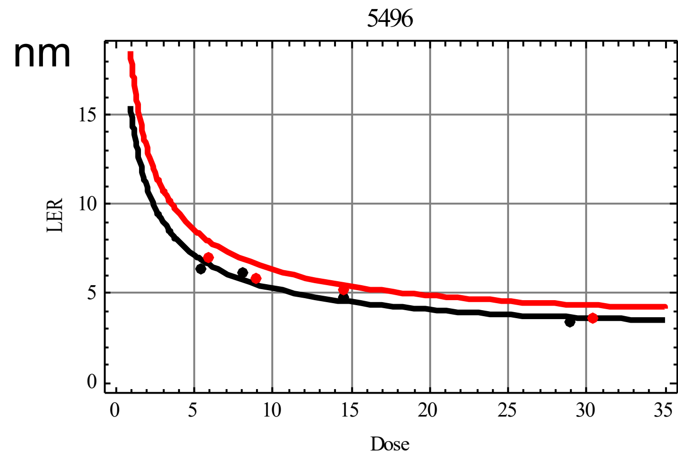
nm


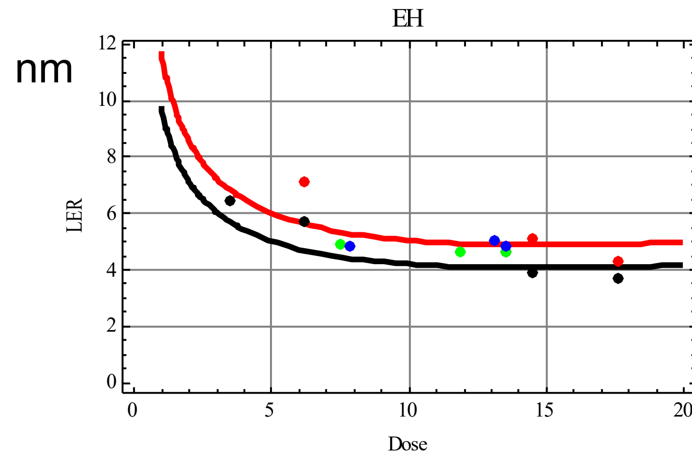
nm


LER data compared to the scaling law with saturation effects included

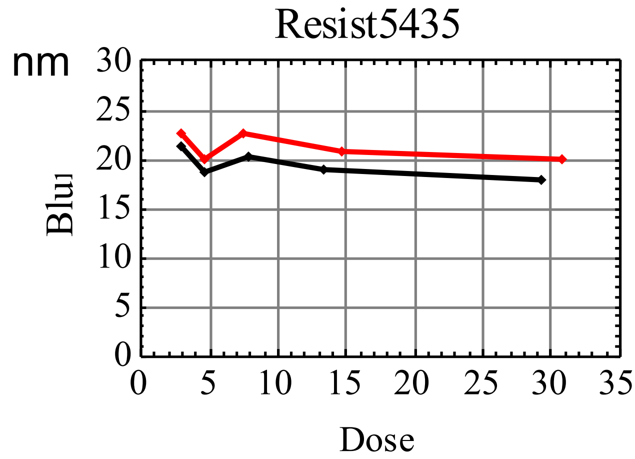


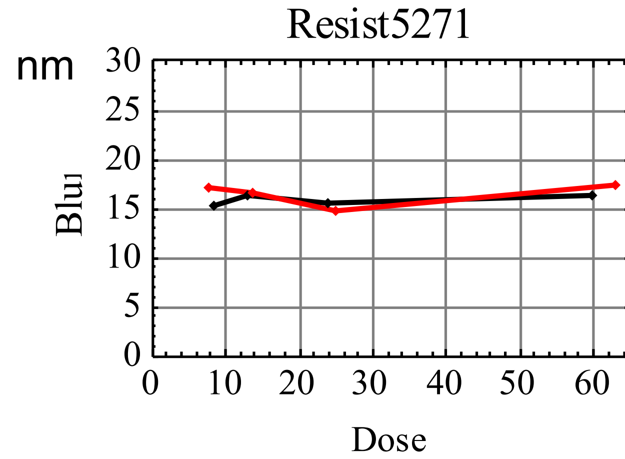


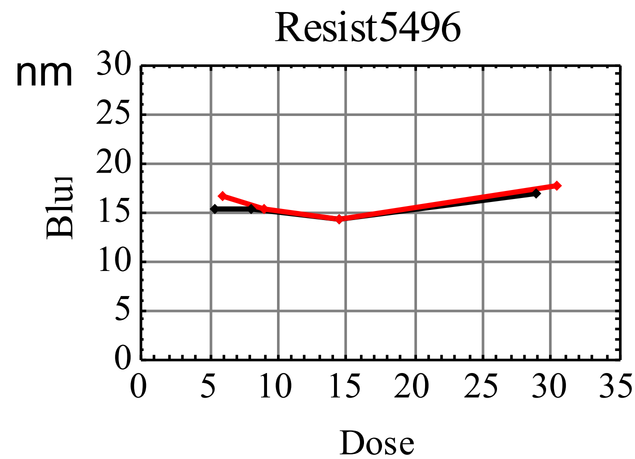


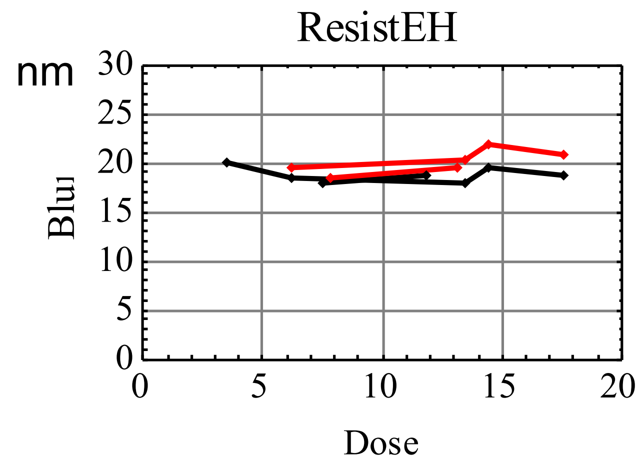


Blur, *R*, determined from data PSD's as a function of sizing dose







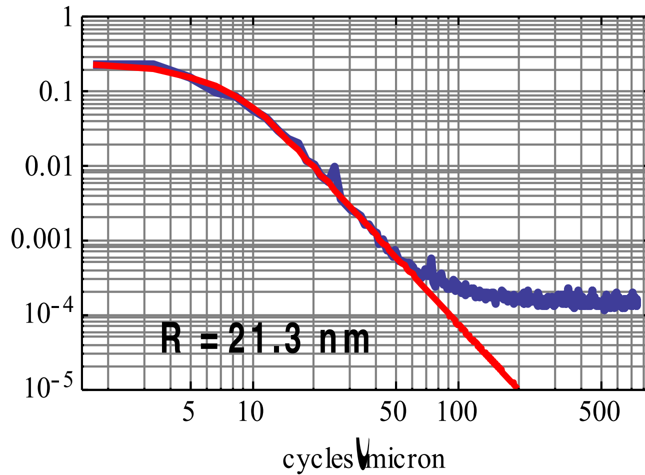


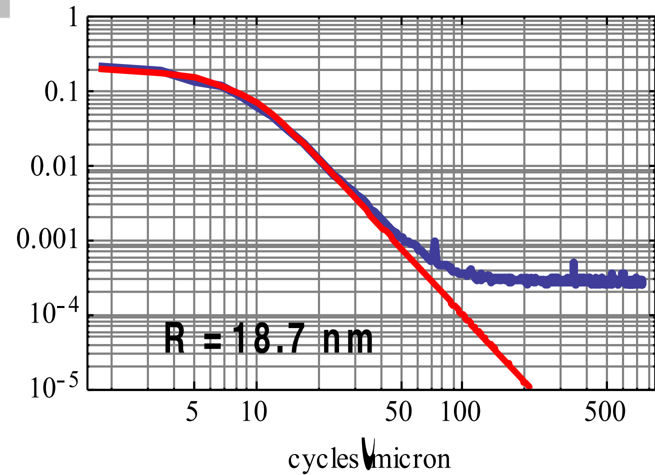
Lowest Dose

5435 EUV2D

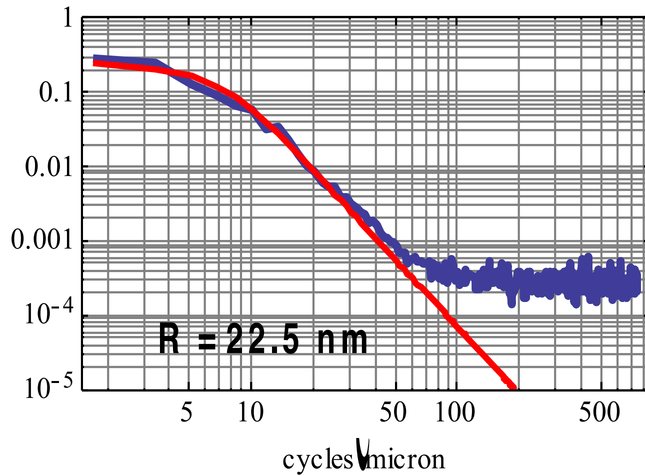
Highest Dose

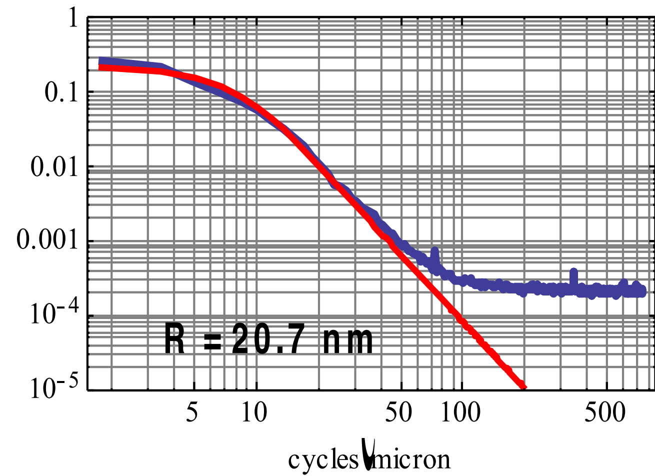
60 nm





50 nm



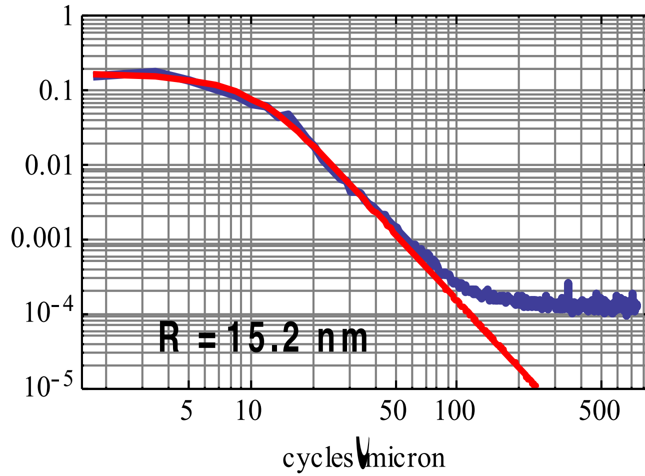


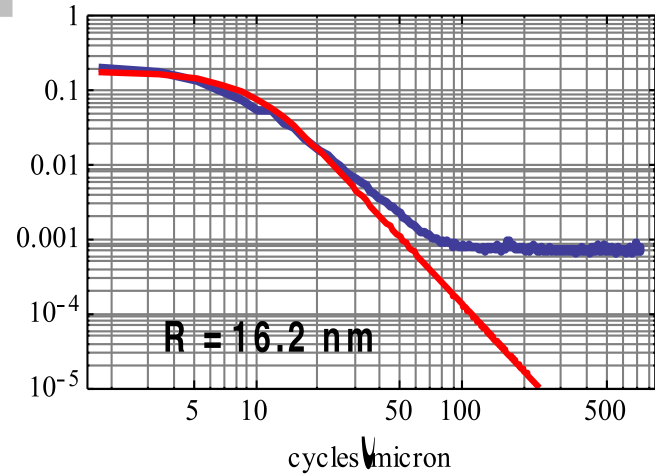
5271 MET2D

Lowest Dose

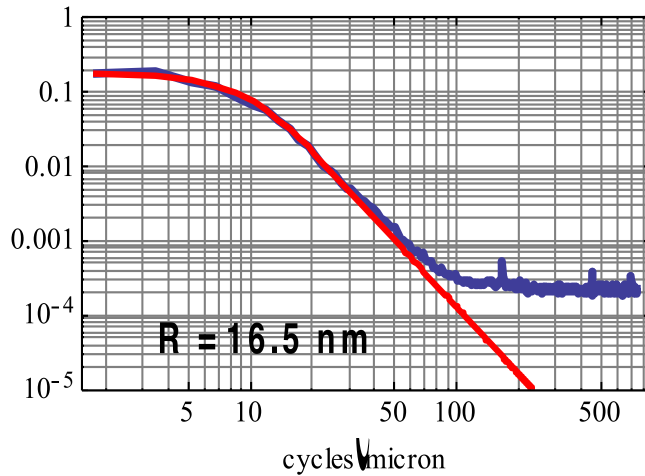
Highest Dose

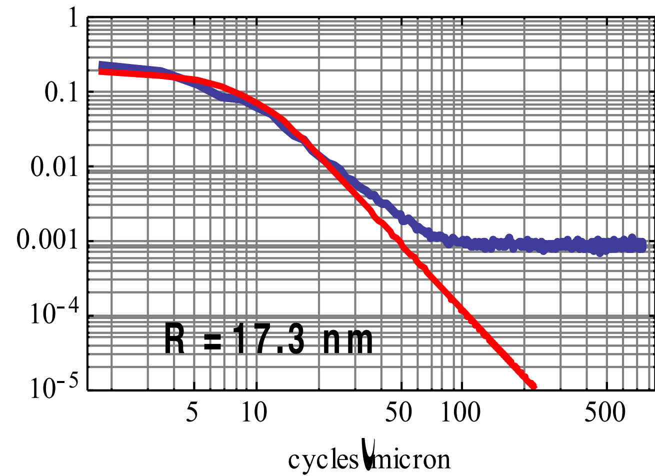
60 nm





50 nm



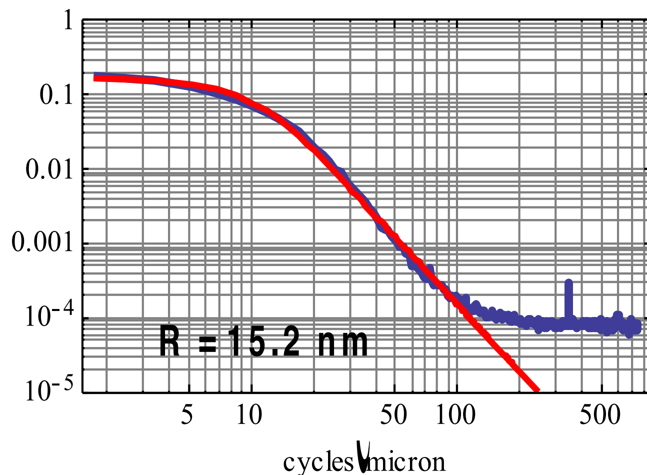


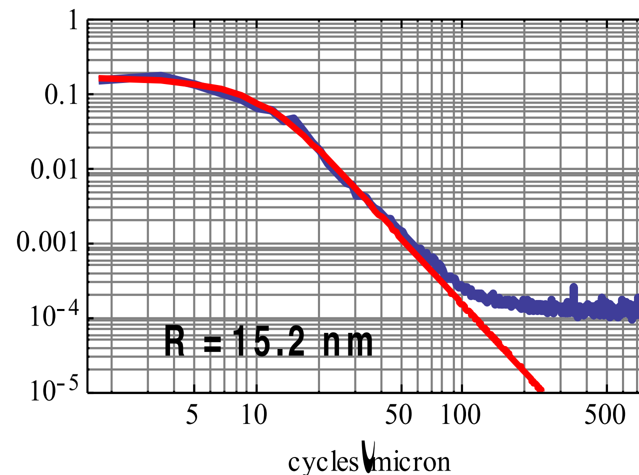
Lowest Dose

5496

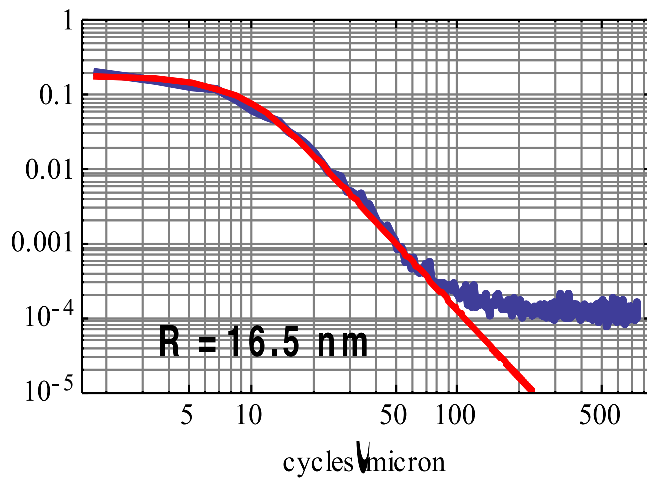
Highest Dose

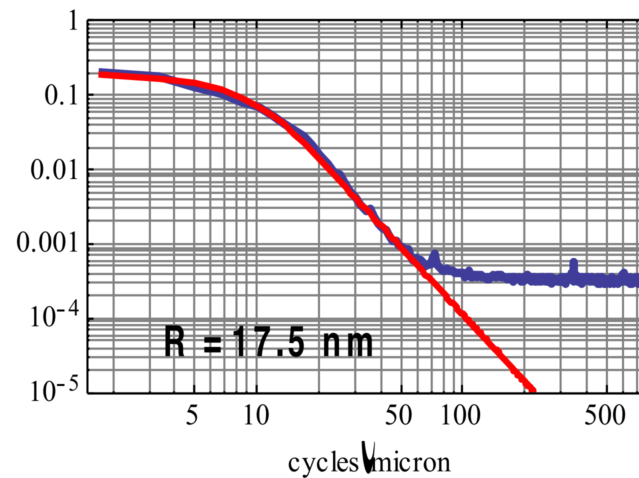
60 nm





50 nm



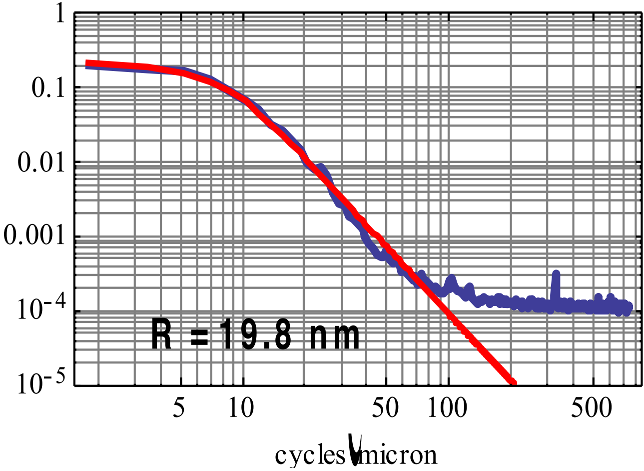


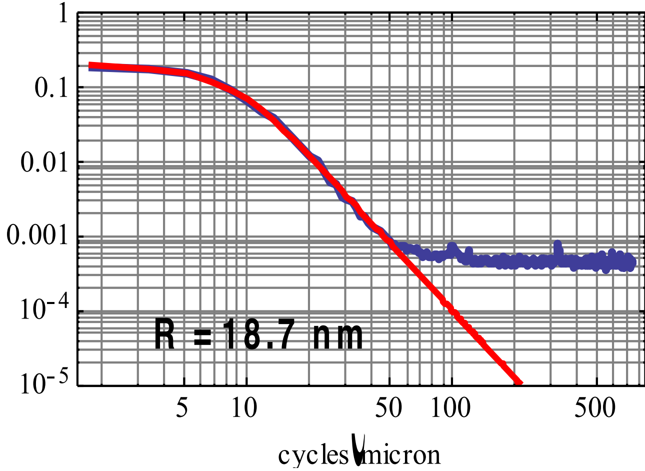
EH

Lowest Dose

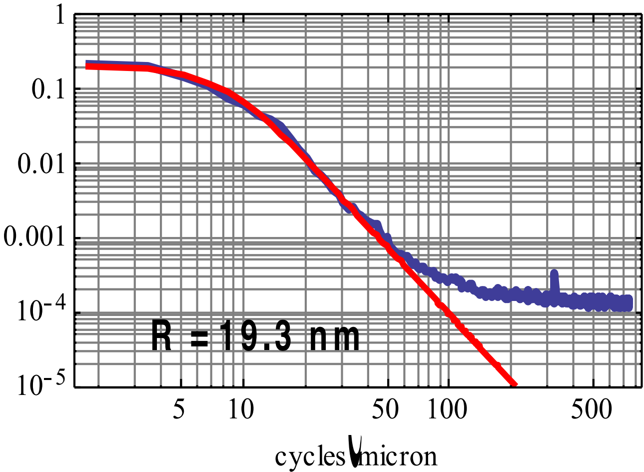
Highest Dose

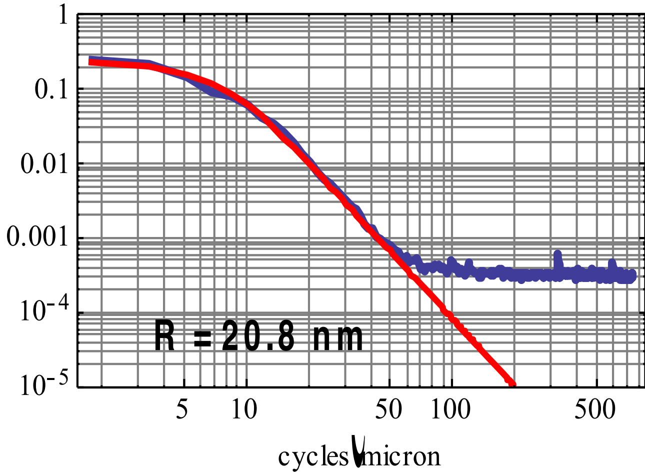
60 nm





50 nm





SPECKLE

There is no such thing as incoherent light

- All electromagnetic fields are perfectly coherent all the time.

- “Incoherence” is in the detection of the light not in the light itself

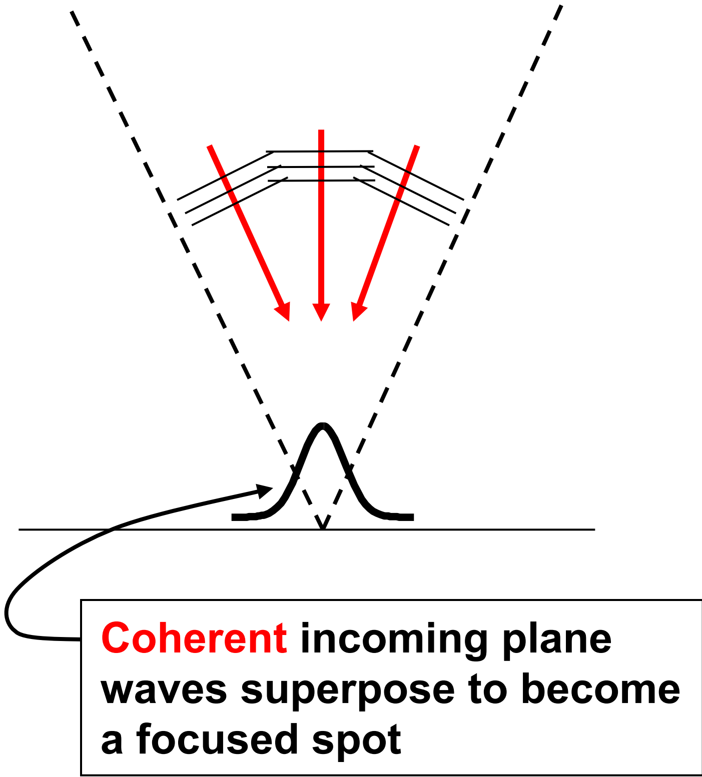
  Consequence of the finite bandwidth of the detector.

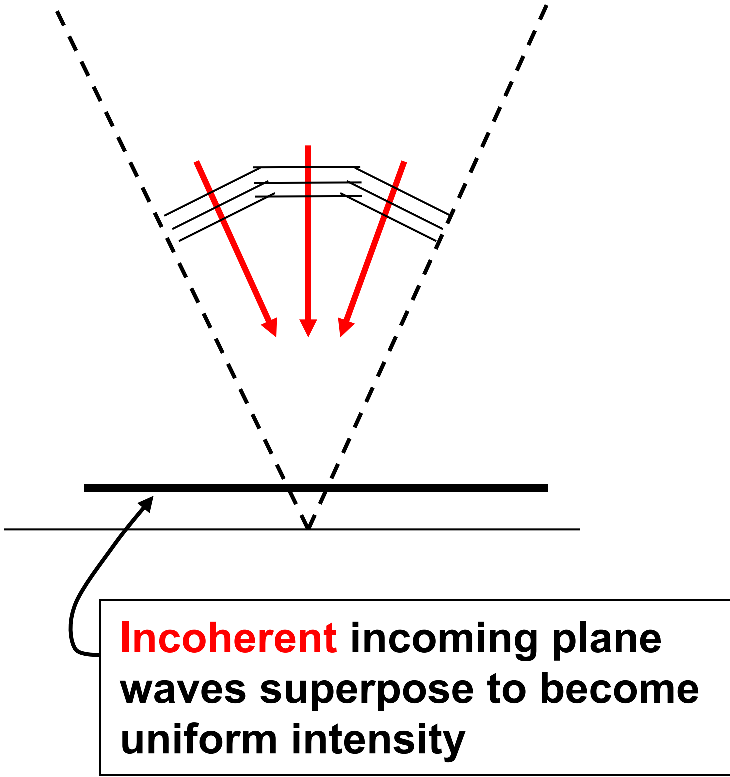
Coherent ➔ $\frac{\text{Detector Bandwidth}}{\text{Light Bandwidth}} >> 1$ **Incoherent** ➔ $\frac{\text{Detector Bandwidth}}{\text{Light Bandwidth}} << 1$

Well known in the speckle and statistical optics community
See, for example: Goodman, “Statistical Optics”
Loudon, “Quantum Theory of Light”

Implications for Lithography are relatively recent:
Rydberg, Bengtsson and Sandstrom, JM3, 2006,
Gallatin, 2002, Unpublished
Kritsun, et, al., 3Beams, 2008

Difference between a coherent and an incoherent sum of plane waves.





Sum of plane waves

Frequencies and Phases the same $\Rightarrow$ Intensity = Time Independent focused "spot"

$$\text{Intensity} = \left| \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y - i\omega_n t + i\phi_n} \right|^2 = \left| e^{-i\omega t + i\phi} \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y} \right|^2 = \left| \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y} \right|^2$$

Same frequency Different (random) phases $\Rightarrow$ Intensity = Time Independent speckle pattern

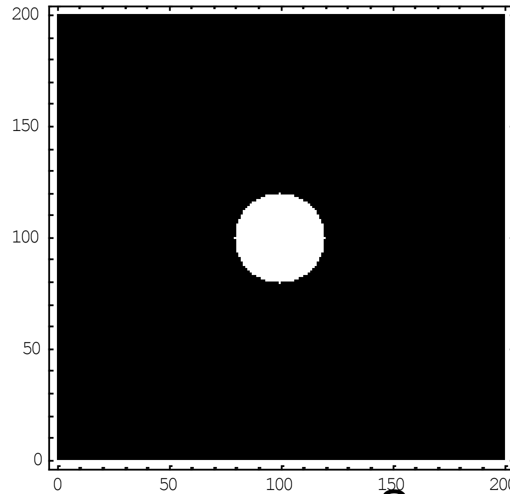
$$\text{Intensity} = \left| \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y - i\omega_n t + i\phi_n} \right|^2 = \left| e^{-i\omega t} \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y + i\phi_n} \right|^2 = \left| \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y + i\phi_n} \right|^2$$

$$\text{Intensity} = \left| \sum_n a_n e^{i\beta_{nx}x + i\beta_{ny}y - i\omega_n t + i\phi_n} \right|^2$$

Different frequencies Different (random) phases $\Rightarrow$ Intensity = Time Dependent speckle pattern

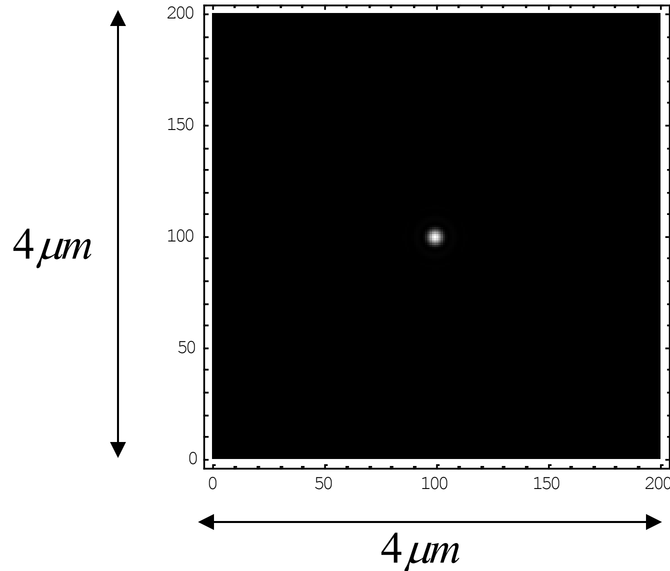
Pupil
NA = 1
= 20pt radius

$\lambda = 193nm$



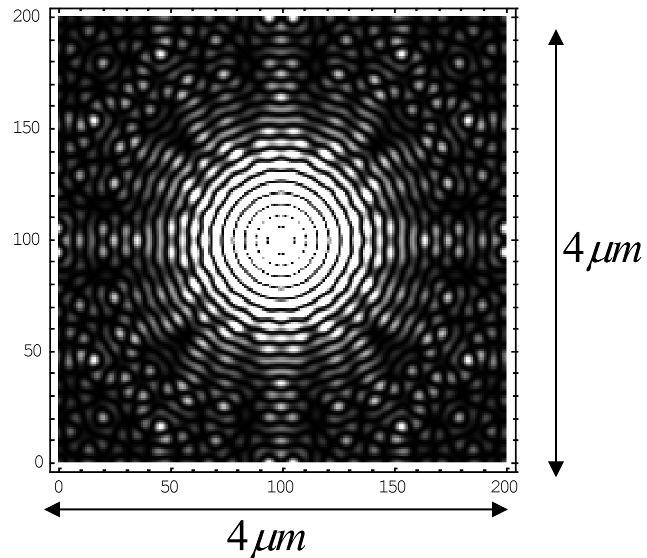
Grayscale covers all values.(full range)

Same data rescaled to show details at low intensity



$4\mu m$

$4\mu m$



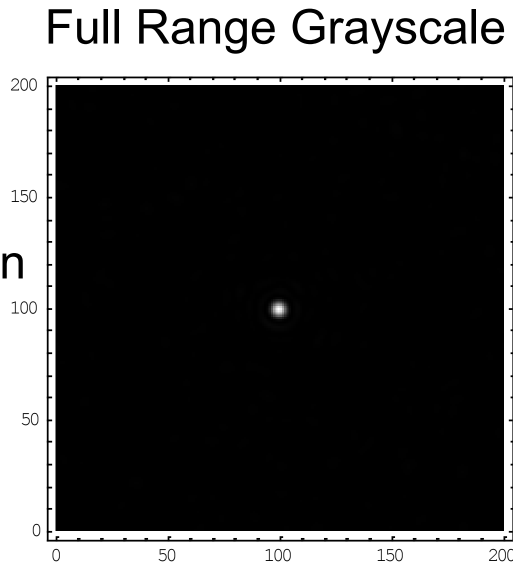
$4\mu m$

$4\mu m$

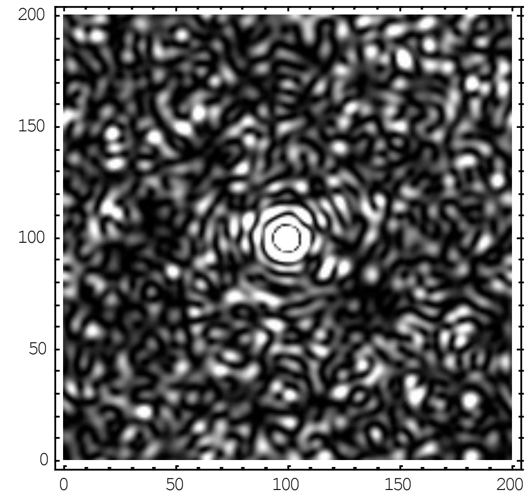
Full Range Grayscale

Low intensity grayscale

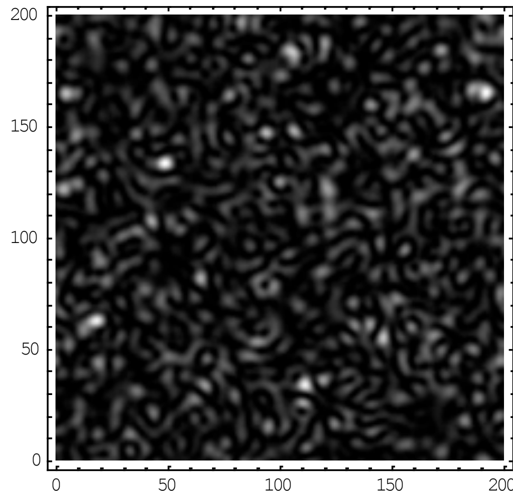
0.5 wave
Uniform distribution of random wavefront error



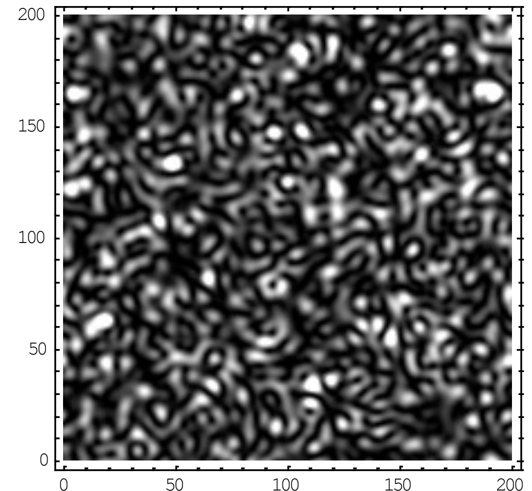
Same data



1.0 wave
Uniform distribution of random wavefront error

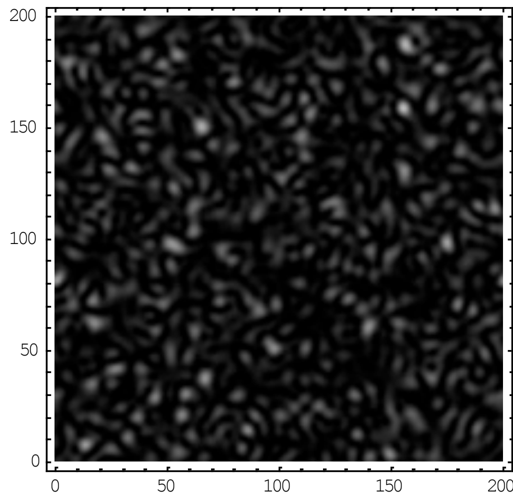


Same data



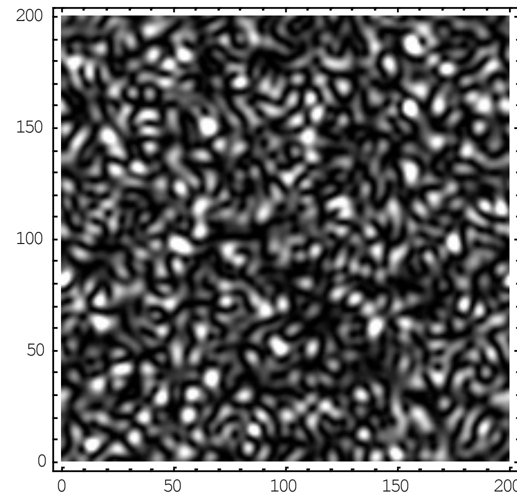
1 wave uniform phase distribution
Uniform amplitude distribution (0 to 1)

Full Range Grayscale

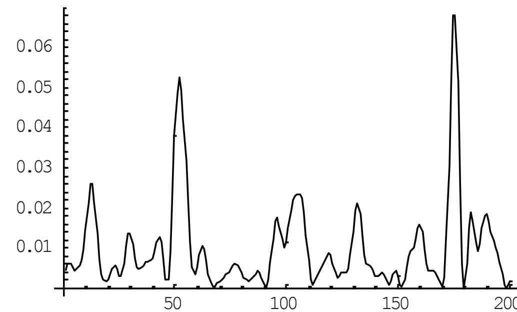


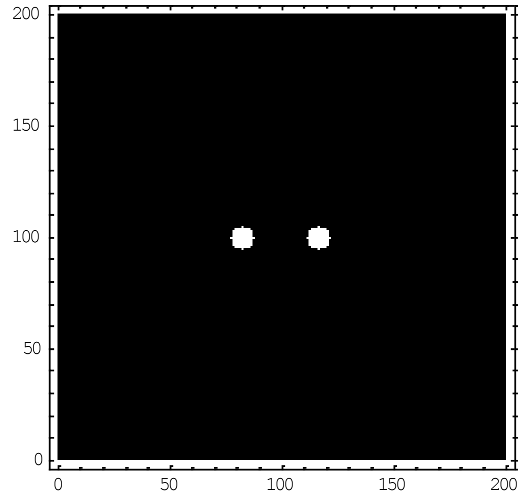
Same Data

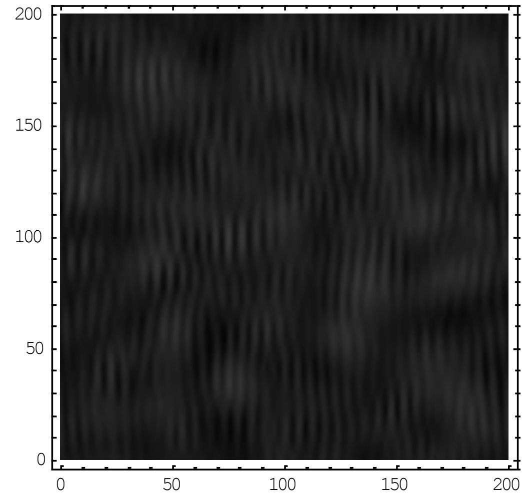
Low intensity grayscale



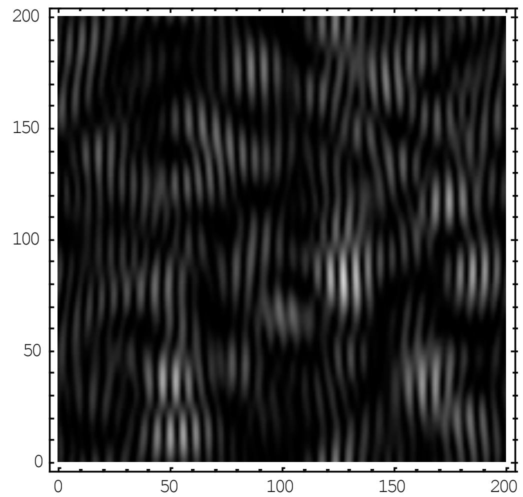
Single speckle pattern is 100% contrast

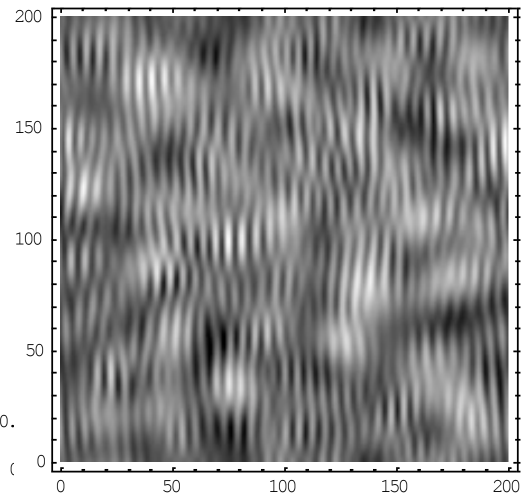


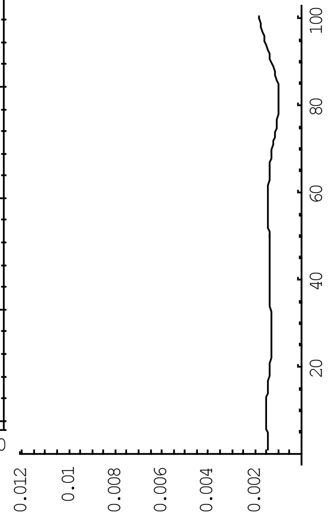


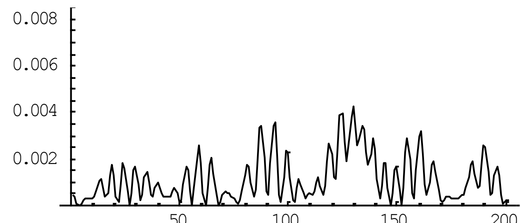


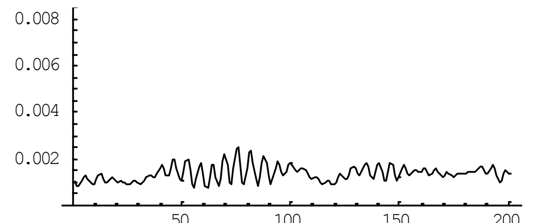
20 pattern average

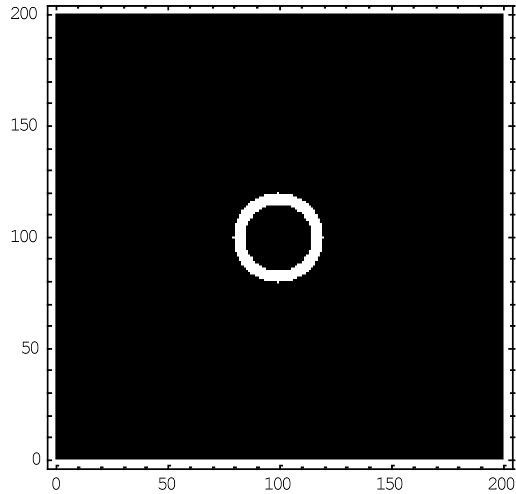


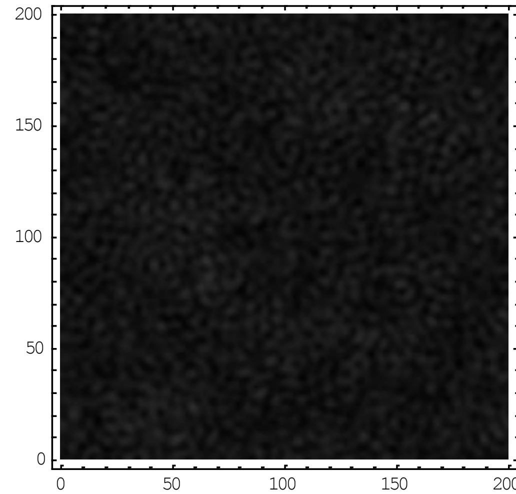




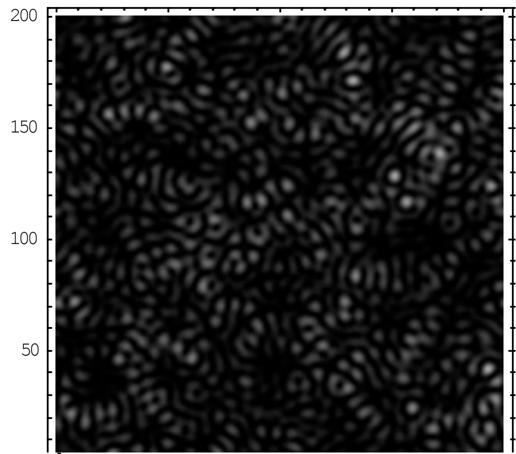


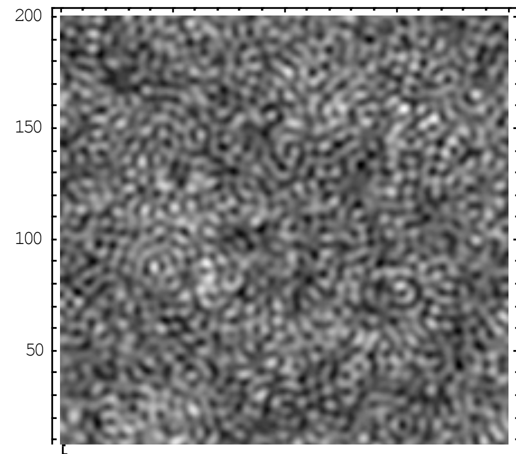


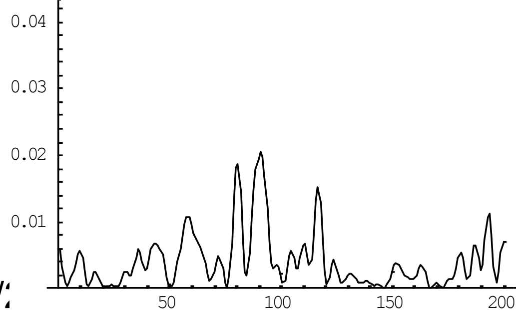


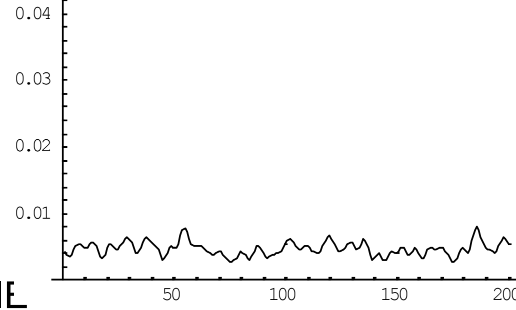


20 pattern average









- Any single speckle pattern has 100% contrast ➔ 100% intensity variation
- Sum of N independent speckle patterns has residual contrast (residual intensity variation)

$$\sigma \approx \frac{1}{\sqrt{N}} \times 100\%$$

- For time dependent speckle: N ~ Bandwidth of the Light X Detector Integration Time

Implication for Lithography:

Detector Integration Time = Excimer pulse length X Number of pulses

~ 20nsec X 50 pulses = 10^{-6} seconds

Illumination Bandwidth = Excimer bandwidth ~ 1pm = 10^{-3}nm ➔ ~10^{10}Hz

$$\Rightarrow \quad N \sim 10^4 \quad \Rightarrow \quad \frac{100\%}{\sqrt{N}} \sim 1\% = \text{Residual "speckle"}$$

Illumination is not, and cannot be, perfectly uniform!
Nonuniformity completely consumes Dose Uniformity Spec of < 1%

Impact of Line Width Roughness on Intel's 65 nm process devices

Manish Chandhok, Suman Datta, Daniel Lionberger, and Scott Vesecky

Intel Corporation, Hillsboro, OR 97124

ABSTRACT

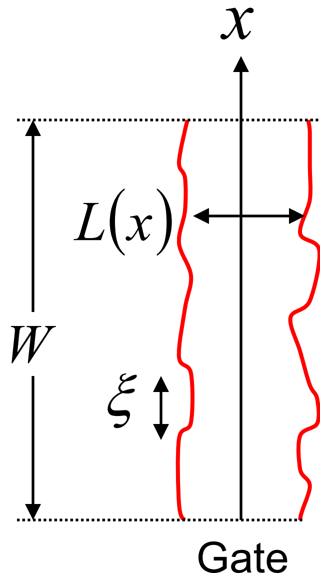
Line Width Roughness (LWR) is the random variation of MOS gate length along the gate width. LWR is undesirable because it degrades drive current (Ion), increases off-current (Ioff), and causes a random variation of device parameters across a die. Previously, [1] it was determined that LWR did not impact Intel's 130 nm process devices. As device sizes shrink, the sensitivity to LWR increases, so the amount of LWR that can be tolerated in future generations needs to be re-assessed. In this paper we will present the experimental results of the effects of LWR on Intel's 65 nm process.
It was found that both nominal drive current and its variation degrade with increased LWR. Additionally, Ioff increased exponentially with increased LWR. In order to maintain less than 2% degradation in Ion from LWR, the 3-Sigma % LWR should be less than 10% of the nominal final check critical dimension (FCCD). Thus, for future generations, LWR needs to scale as gate lengths decrease or else any potential benefits in increased drive current would be offset by large amounts of leakage.

Chandhok, et. al., SPIE 2007

LER impact on Leakage Current

LER → LWR → Gate Length Roughness (GLR)

Goldfarb EIPBN 04



Device Parameter

$$I = \int_0^W J(L(x))dx$$

Croon IEDM 02
Croon ESSDERC 03

Gate Length fluctuates about average

$$L(x) = \bar{L} + \delta L(x)$$

Relative fluctuation

$$\frac{\sigma_I}{\langle I \rangle} \sim \left(\frac{\partial J / \partial \bar{L}}{J(\bar{L})} \right) \left(\frac{\sigma_{GLR}}{\sqrt{W / 2\xi}} \right)$$

LER weighting for CD variation $\sim \dfrac{\sigma_{LER}}{\sqrt{W / 2\xi}}$

Note: The analysis here retains only the lowest order linear term. If J is highly nonlinear and/or $\delta L/L$ is not small then can get highly nonlinear dependence of leakage current on LER

Can resist LER be fixed??

i.e.,

Can we “break” the RLS tradeoff and get what we want??

Two approaches considered here

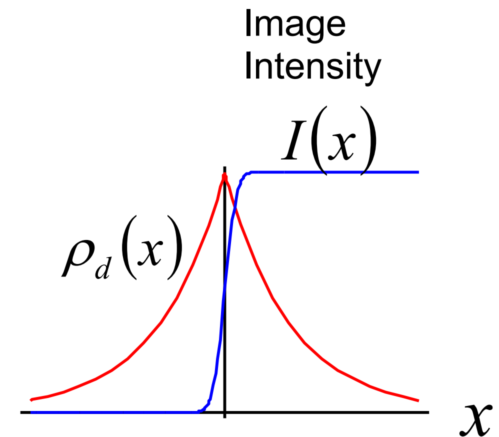
1. Anisotropic resist blur
    - Only a “What if” analysis…. Feasibility is not determined
2. Higher PAG loading and/or higher Q

Anisotropic resist "blur"

Simplification…..Assume the following

- Ignore resist absorption ➔ $I(\vec{r}) = I(x,y,z) \to I(x,y)$
- Resist line edge oriented along y ➔ $I(x,y) \to I(x)$
- Image has much higher resolution than the resist

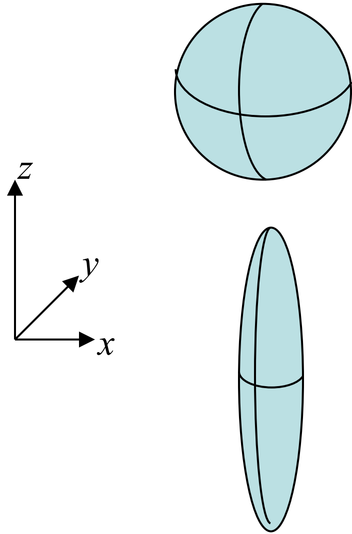
$$\frac{\partial I(x)}{\partial x} \approx \delta(x) \quad \text{relative to} \quad \rho_D(\vec{r})$$



- Assume Gaussian deprotection blur

$$\rho_d(\vec{r}) \approx \exp\left[-\left(\frac{x^2}{2\sigma_x^{\ 2}} + \frac{y^2}{2\sigma_y^{\ 2}} + \frac{z^2}{2\sigma_z^{\ 2}}\right)\right]$$

Deprotection "blur" is NOT Gaussian.
Gaussian form is used here only to simplify the analysis.
➔ The integrals can be done analytically for a Gaussian.



➔ Spherical Deprotection Blur: $R_x = R_y = R_z$

$$\text{Volume} \sim R_x R_y R_z$$

➔ Anisotropic Deprotection Blur:

Shrink in x and y | Expand in z

$$R_x \rightarrow \frac{R_x}{s} \qquad R_y \rightarrow \frac{R_y}{s} \qquad R_z \rightarrow s^2 R_z$$

s = Scaling Factor
$s \geq 1$

$$\text{Volume} \sim \frac{R_x}{s}\frac{R_y}{s} s^2 R_z = R_x R_y R_z$$

…do the math…

$$LER \sim 1/s$$

$$\text{Horizontal Blur} \sim 1/s$$

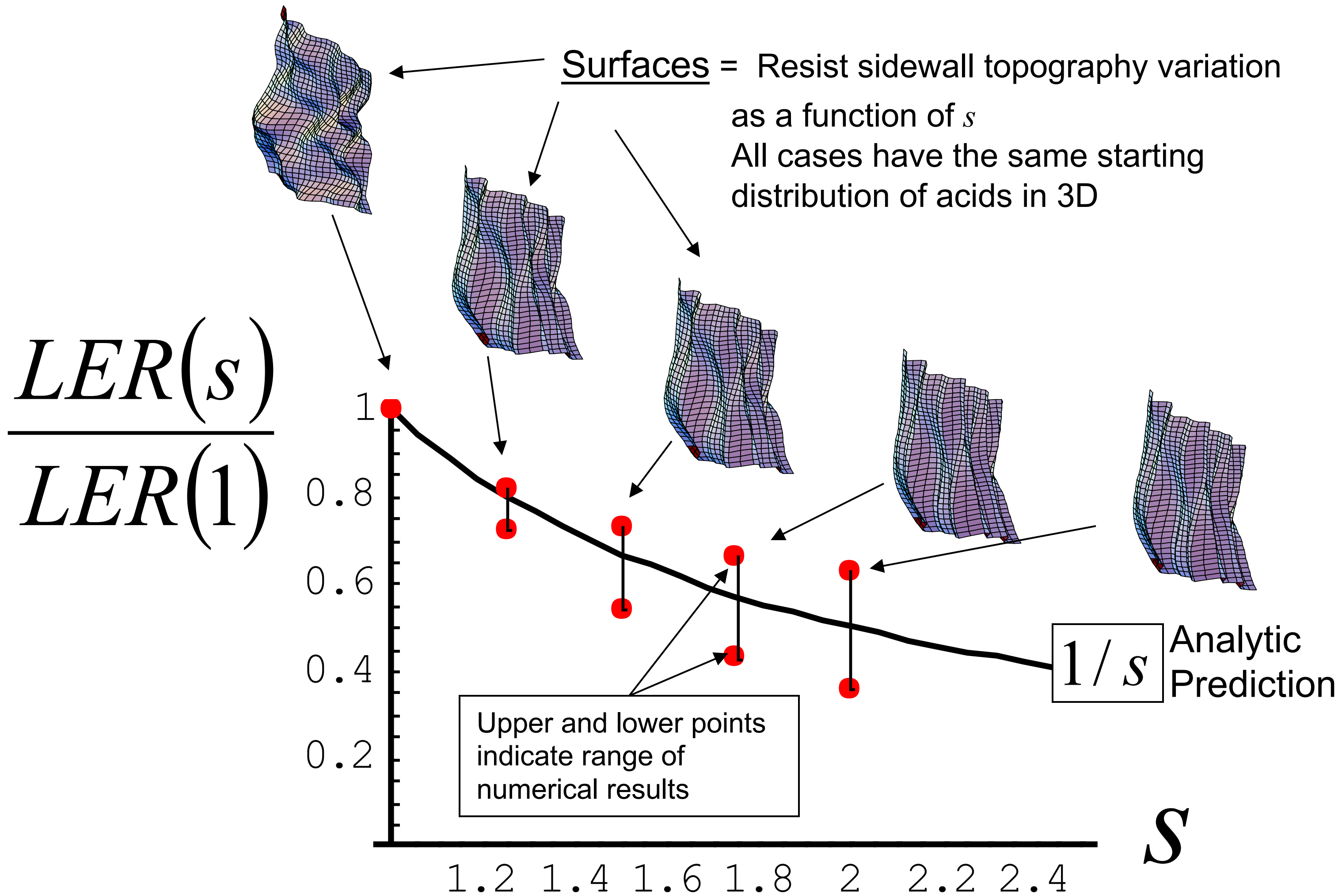
$$\text{Dose} = \text{Fixed}$$

Improved LER and resolution

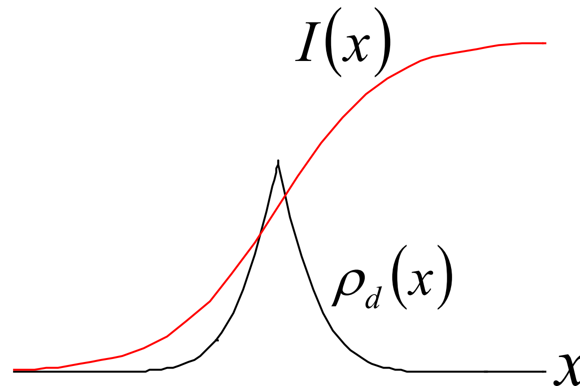
Dose unchanged

Anisotropic diffusion explicitly breaks the RLS tradeoff by improving resolution and LER at fixed Dose

Anisotropic resist "blur" Numerical results bracket the analytic prediction



Repeat the analysis for the regime where the resist has better resolution than the image



…do the math…

$$LER \sim \text{Fixed}$$

$$\text{Horizontal Blur} \sim 1/s$$

$$\text{Dose} = \text{Fixed}$$

LER unchanged

Improved resolution

Dose unchanged

Anisotropic diffusion breaks the RLS tradeoff by improving the resolution at fixed LER and Dose

Using Quantum Yield Q to improve LER

248 nm and 193 nm photons release acids ➔ Energy required to release an acid is at most ~ 5 to 6 eV.

EUV photon has 92 eV of energy ➔ Each EUV photon has enough energy to release at least 92 eV/5 eV~18 acids

"acid bottleneck" Neureuther, et al., JVST B 2006

DATA: $Q_{DUV} \sim 1/3$ ➔ 1 in 3 absorptions results in acid release

Brainard, et al., SPIE 2004

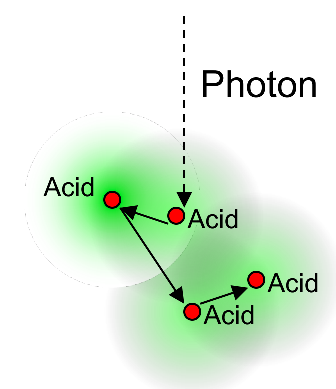
$Q_{EUV} \sim 2$ ➔ EUV is not close to using its energy efficiently

BUT Not all acids can be released in the same position ➔ adds blur

- Assume acids are released along random walk path with steps spaced by $\rho^{-1/3}$

Q released acids ➔ Extra "exposure" blur $r \sim (Q-1)^{1/2} \rho^{-1/3}$

Combine deprotection "blur" from each acid with the random walk distribution of released acids → "Net Blur Radius"



$$Net\ Blur\ Radius \sim \sqrt{R^2 + r^2} = \sqrt{R^2 + (Q-1)\rho^{-2/3}}$$

R = Deprotection blur radius = FWHM/2 ~ 15nm

Spatial distribution of Q released acids

Added Blur (nm)

Net Blur Radius (nm)

Q =16

Q = 8

Q = 4

Q = 2

Q = 1

Clearly want "high" PAG density to avoid significantly increasing blur

$\bar{\rho}_{PAG}\ (\#/nm^3)$

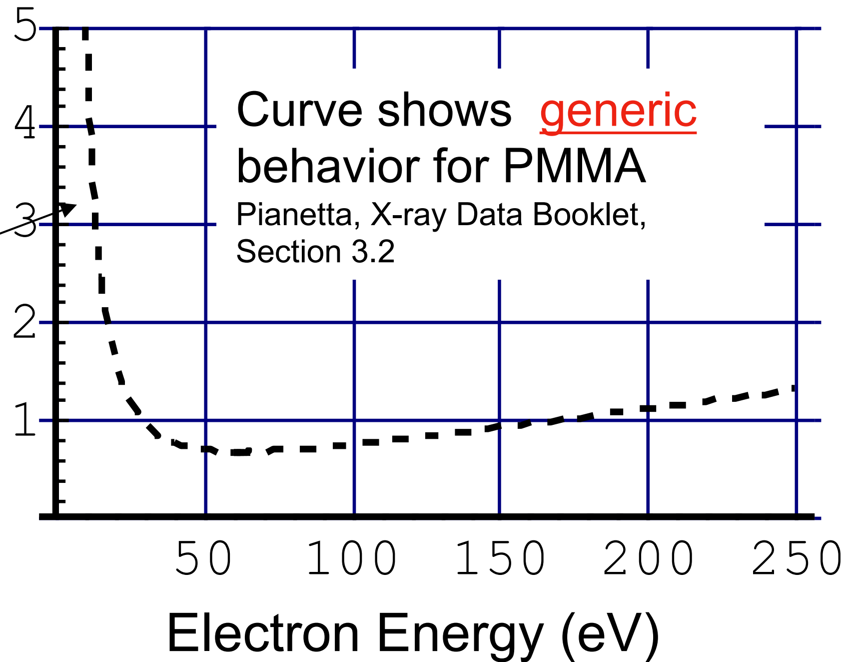
Mechanism for releasing multiple acids per photon:

Expect the dominant pathway for spreading energy out to multiple PAG molecules to be by secondary (Auger) electrons.

Han, et al., JVST B '02
Drygiannakis, et al., Poster 6519-36

Inelastic Mean Free Path of electrons In materials

Electrons with energy less than ~25 eV can "travel" far enough to interact with multiple PAGs.



$$1\sigma \text{ LER} \approx \frac{1}{ILS}\sqrt{\frac{1}{\bar{\rho}_{\text{PAG}}\alpha\, Q\, v\, E_{size} R^3}}$$

1. Increasing Q and/or $\bar{\rho}_{PAG}$ with everything else fixed

   LER decreases with Dose and blur fixed

   ➔ “breaks” RLS tradeoff

2. Dose E nominally decreases with increasing Q and $\bar{\rho}_{PAG}$

   LER ~ constant but with decreasing Dose and fixed blur

   ➔ “breaks” RLS tradeoff

<u>DATA</u> ➔ Case 1: Both LER and Dose decrease with increasing PAG loading

Choi, et al., SPIE 2007
Leunissen, et al., MNE 2005.

Question: “Does resist LER really matter???”

- Many process steps between resist patterning and the final device
- These process steps can smooth the edge roughness

See for example:

Palmateer, Sematech report, 1998
Jang, et. al., JVSTB 2004
Hwang, et. al., JVSTB 2004
Goldfarb, et. al., SPIE 2004
Prabhu, et. al., SPIE 2004
Mahorowala, et. al., SPIE 2005

➔ So maybe we don’t need to worry about this afterall

Thanks for your attention

The End