New Directions For Optical Critical Dimension Metrology

18 APRIL 2024

Nigel Smith, Alex Boosalis, Mark Carr, Yiliang Liu, Pedro Vagos, Marieke Ordway, G. Andrew Antonelli

Onto Innovation Inc.
9025 NE Von Neumann Drive, Suite 100
Hillsboro, OR 97006

Motivation and Approach

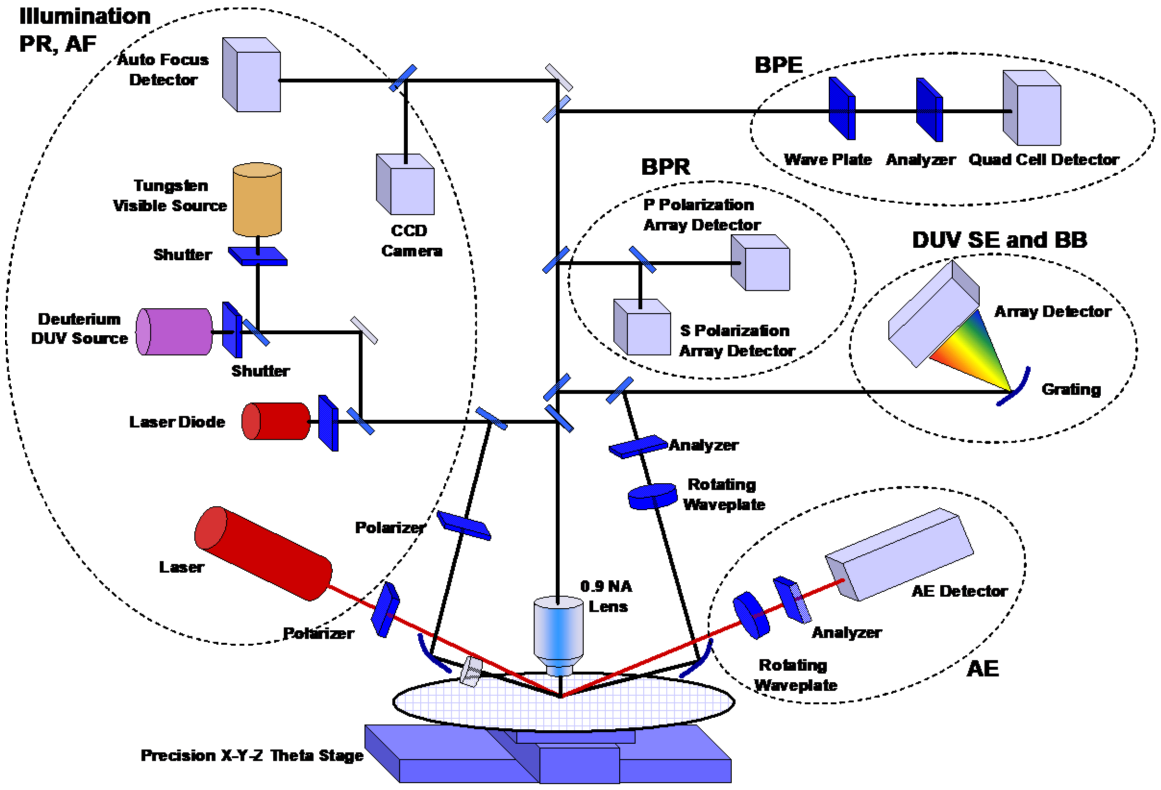
Address the limitations of existing OCD technologies

- Break parameter correlations by massive increase in number of incident angles measured.
  - Correlations occur when signal response to 2+ parameters are effectively the same.
  - When correlations occur, measurement capability is reduced.
- Reduce probe size
  - Some array sizes are below 10µm.
  - Probe within array variation, especially close to the edge of arrays.

Approach

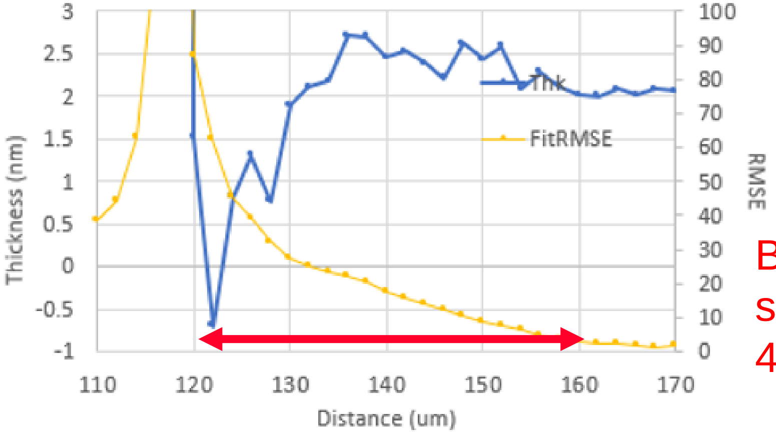
- Instrument based on microscopy with small field size: probe size limited by diffraction to ≈ 1µm or larger.

Non-interferometric Methods



"One can optimize reflection or transmission as a function of angle at a single wavelength or wavelength at a single angle, but not both simultaneously without degrading both" – Jon Opsal

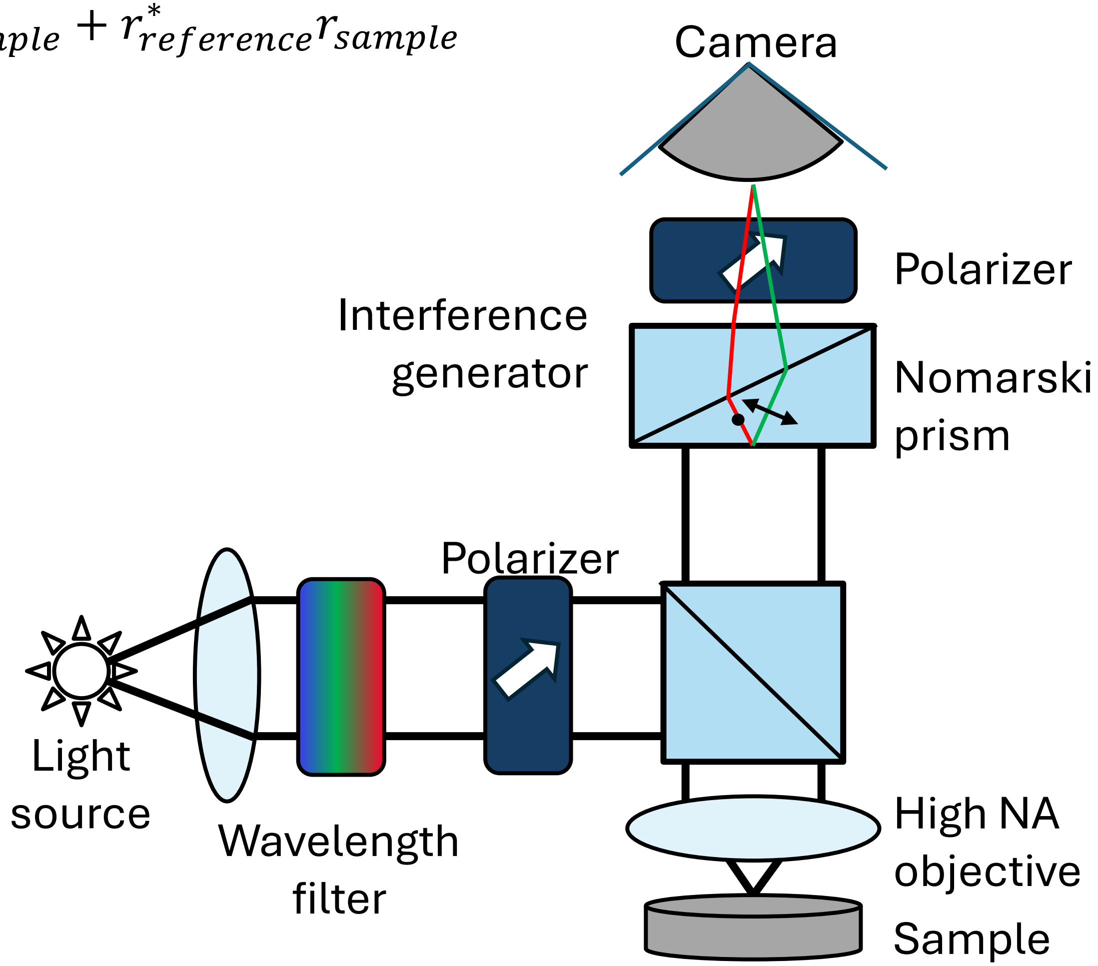
- BPR – Beam Profile Reflectometry (673 nm)
  - Pupil plane: multiple angles, one wavelength, one intensity.
- BPE – Beam Profile Ellipsometry (673 nm)
  - Pupil plane: multiple angles, one wavelength, full Mueller Matrix.
- AE – Absolute Ellipsometry (633nm)
- Vis – Visible Spectrometry (470-870 nm)
- BB – BroadBand Spectrometry (190-840 nm)
- RCSE – Rotating Compensator Spectroscopic Ellipsometry (190-840)
  - One angle, spectroscopic, full Mueller Matrix.



BPE signal is sensitive to edges 40µm away

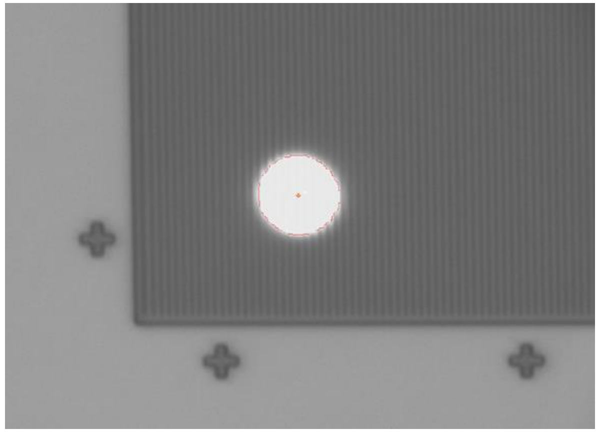
Interferometry

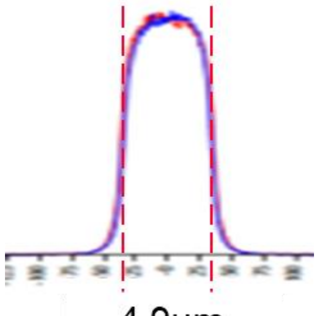
- $\frac{I_{out}}{I_{in}} = \left|r_{reference}\right|^2 + \left|r_{sample}\right|^2 + r_{reference}\, r_{sample}^* + r_{reference}^* r_{sample}$
- Interferometry allows measurement of the complex reflectance of the sample
- Examples: SIPE (J. Jung et al, *Proc. SPIE 11611, Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV*, 2021).



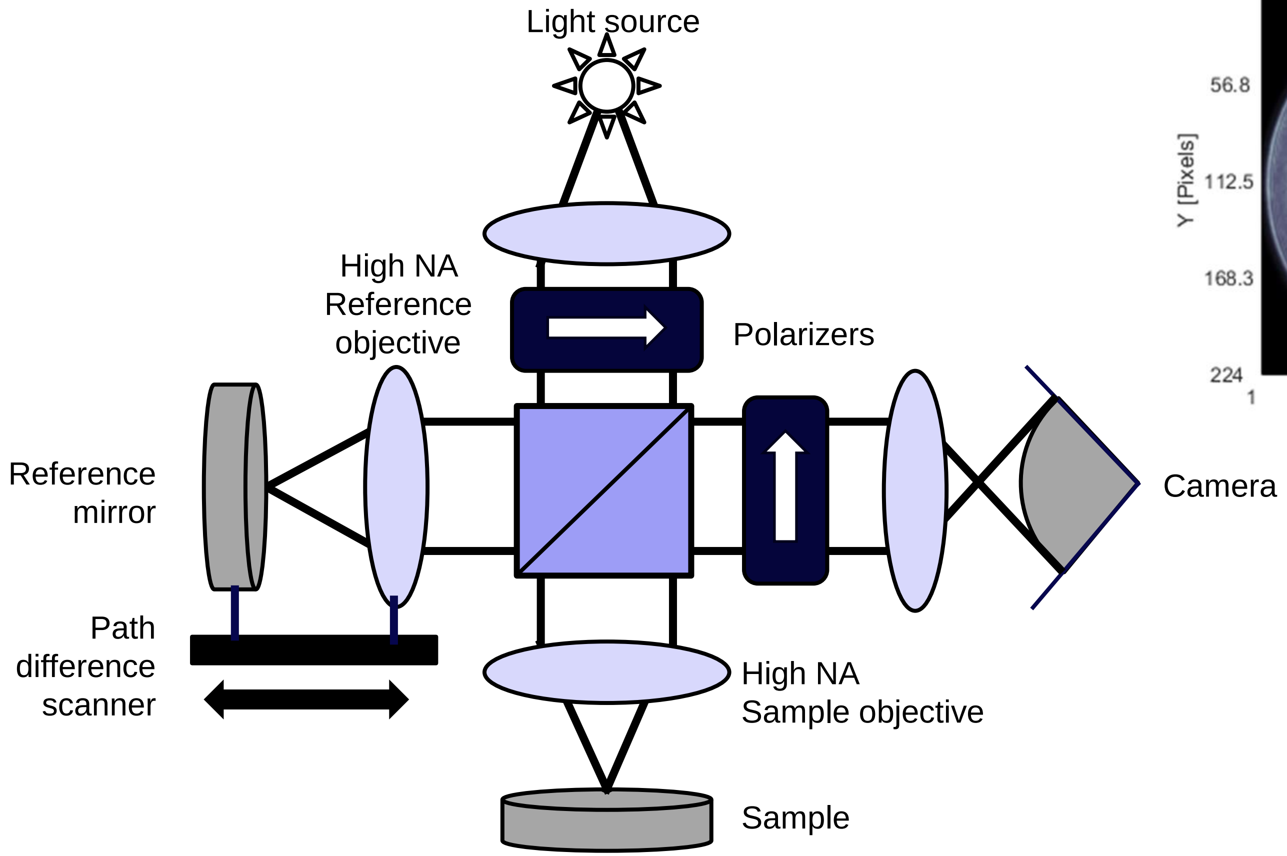
innovation™ onto

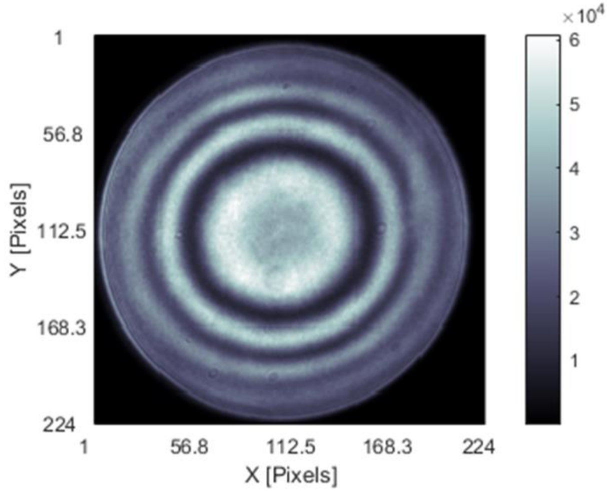
Pupil Plane Interferometry

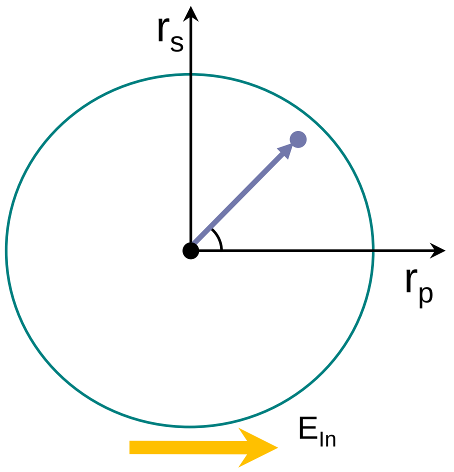




Spot size is controlled by field stop







Motivation and Approach

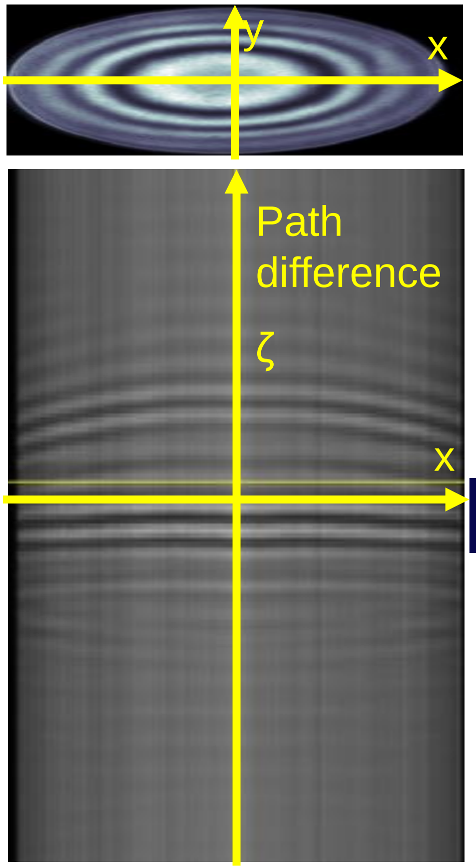
Address the limitations of existing OCD technologies

- Break parameter correlations by massive increase in number of incident angles measured.
  - Correlations occur when signal response to 2+ parameters are effectively the same.
  - When correlations occur, measurement capability is reduced.
- Reduce probe size
  - Some array sizes are below 10µm.
  - Probe within array variation, especially close to the edge of arrays.

Approach

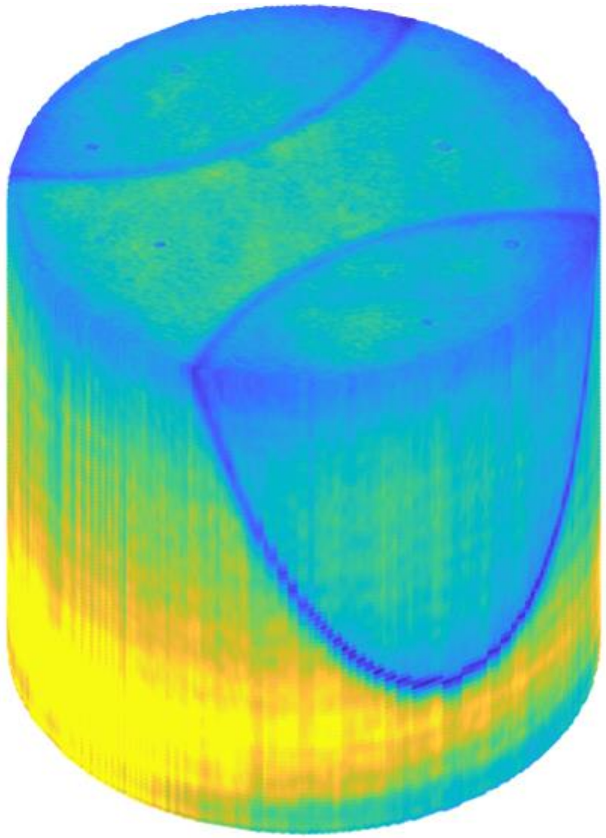
- Instrument based on microscopy with small field size: probe size limited by diffraction to ≈ 1µm or larger.
- Pupil plane detection allows measurement of sample reflection at many incident angles.
  - Unlike in the image plane, each pixel is at a unique angle, simplifying simulation of the signal formation.
- Use interferometry to extract complex reflectance.
- Use standard model fitting approaches to estimate sample parameters from measured data.

Signal Processing



Original data

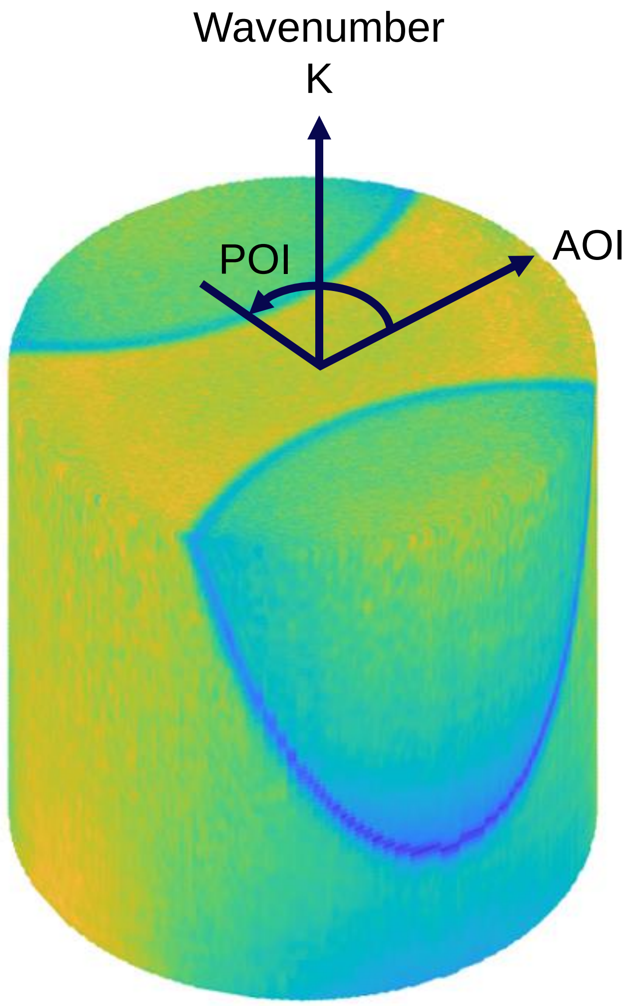
Multiple images as path difference changes



Take FFT at each pixel

$\zeta \rightarrow K$

FFT is **<u>complex</u>**

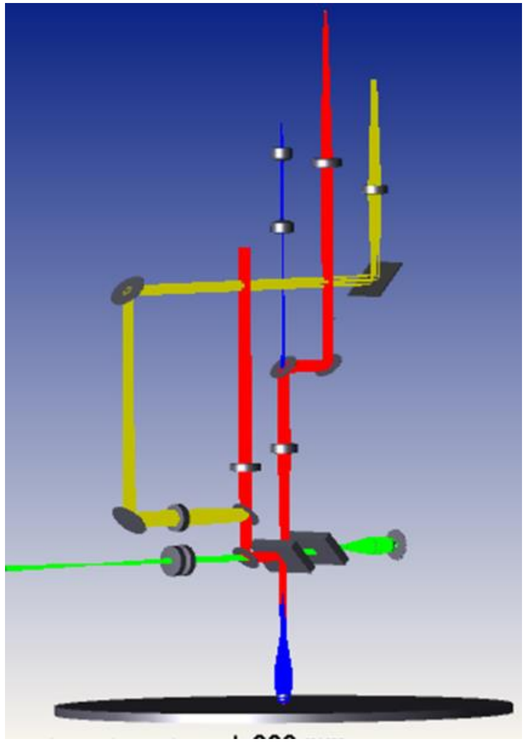


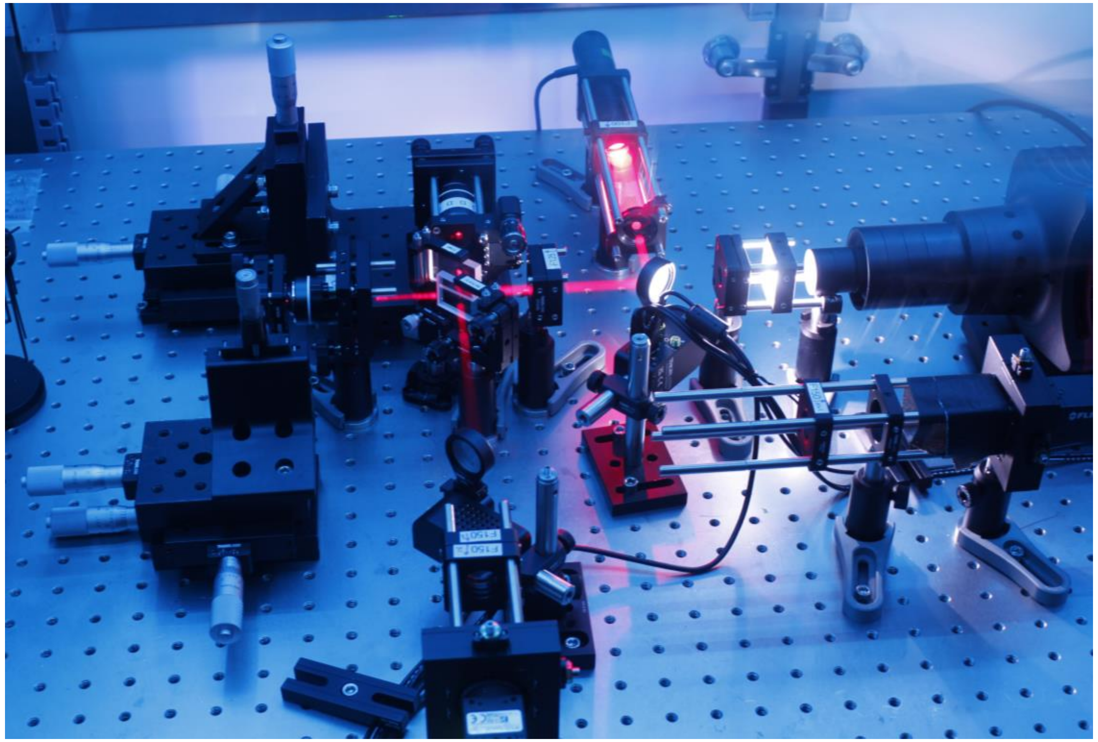
Remove pupil intensity variation and source spectrum

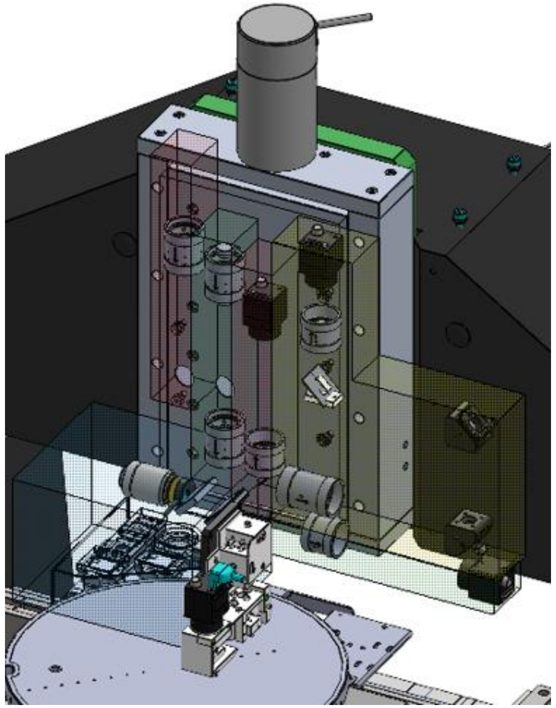
innovation™
onto

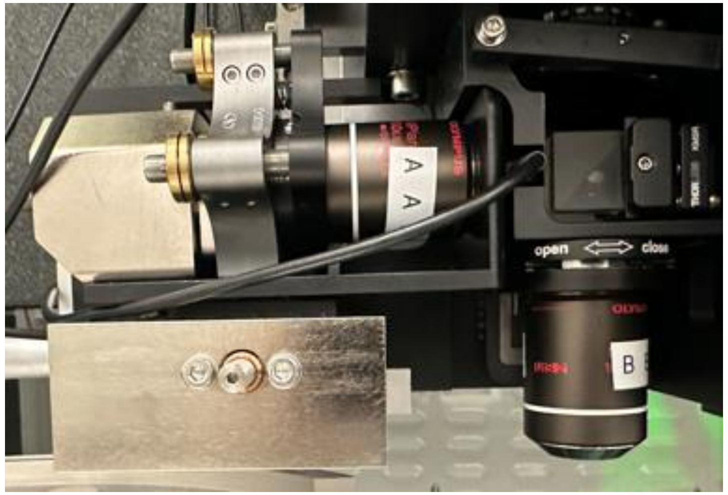
Instrument Realization

Adjustment vs. Stability

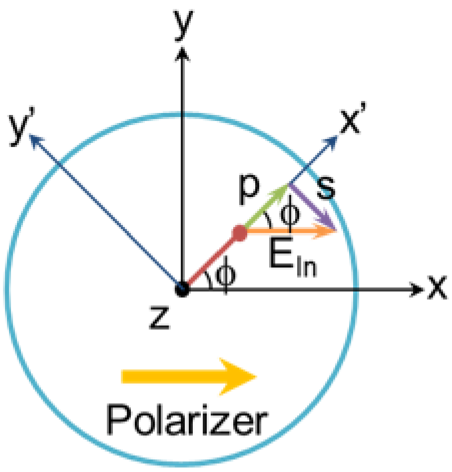


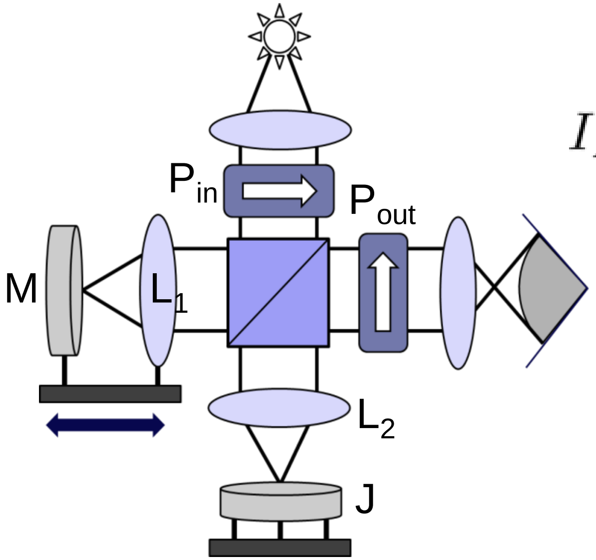






Simplified Analytical Model





$$\mathbf{E}_{out} = \mathbf{P}_{out}\mathbf{B}(\mathbf{L_2RJRL_2} + \mathbf{TL_1ML_1T})\mathbf{BP}_{in}\mathbf{E}_{in}$$

$$\mathbf{E}_{out} = \frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\left(\begin{bmatrix}\mathrm{c}^2 J_{pp} + \mathrm{cs}J_{sp} - \mathrm{cs}J_{ps} - \mathrm{s}^2 J_{ss} & -\mathrm{cs}J_{pp} + \mathrm{c}^2 J_{sp} + \mathrm{s}^2 J_{ps} - \mathrm{cs}J_{ss}\\ \mathrm{cs}J_{pp} + \mathrm{c}^2 J_{sp} + \mathrm{s}^2 J_{ps} + \mathrm{cs}J_{ss} & -\mathrm{s}^2 J_{pp} + \mathrm{cs}J_{sp} - \mathrm{cs}J_{ps} + \mathrm{c}^2 J_{ss}\end{bmatrix}\right.$$

$$\left. + \begin{bmatrix}e^{4\mathbf{i}\pi d/\lambda} & 0\\ 0 & -e^{4\mathbf{i}\pi d/\lambda}\end{bmatrix}\right)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\mathbf{E}_{in}$$

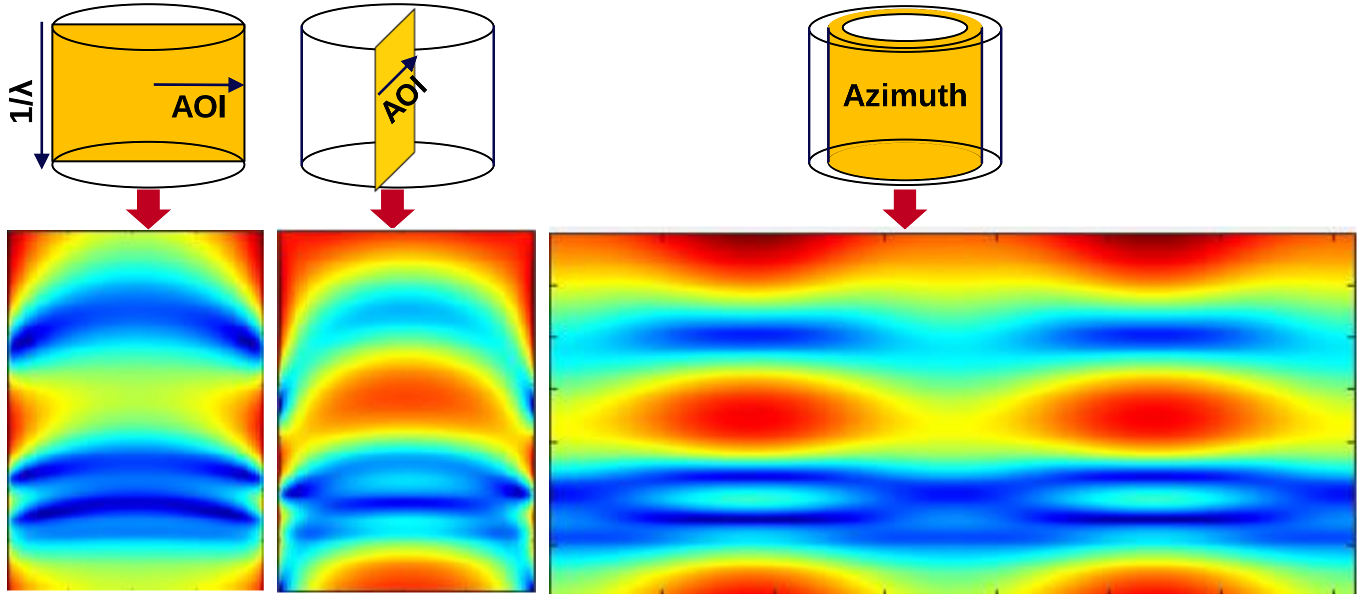
$$\begin{aligned} I_{AC} &= \mathbf{E}^*_{out}\mathbf{E}_{out}|_{AC} \\ &= I_0\cos^2(\phi)|J_{pp}|\cos(\Delta_{pp} - 4\pi d/\lambda) \\ &+ I_0\cos(\phi)\sin(\phi)|J_{ps}|\cos(\Delta_{ps} - 4\pi d/\lambda) \\ &- I_0\cos(\phi)\sin(\phi)|J_{sp}|\cos(\Delta_{sp} - 4\pi d/\lambda) \\ &- I_0\sin^2(\phi)|J_{ss}|\cos(\Delta_{ss} - 4\pi d/\lambda) \end{aligned}$$

- I_{AC} is modulation of interferogram.
- Full model:
  - 3D Jones Matrix.
  - Non-ideal system components.
- Use model parameters for sample to determine Jones matrix terms.
- Adjust Jones matrix terms for best fit to experimental signal.

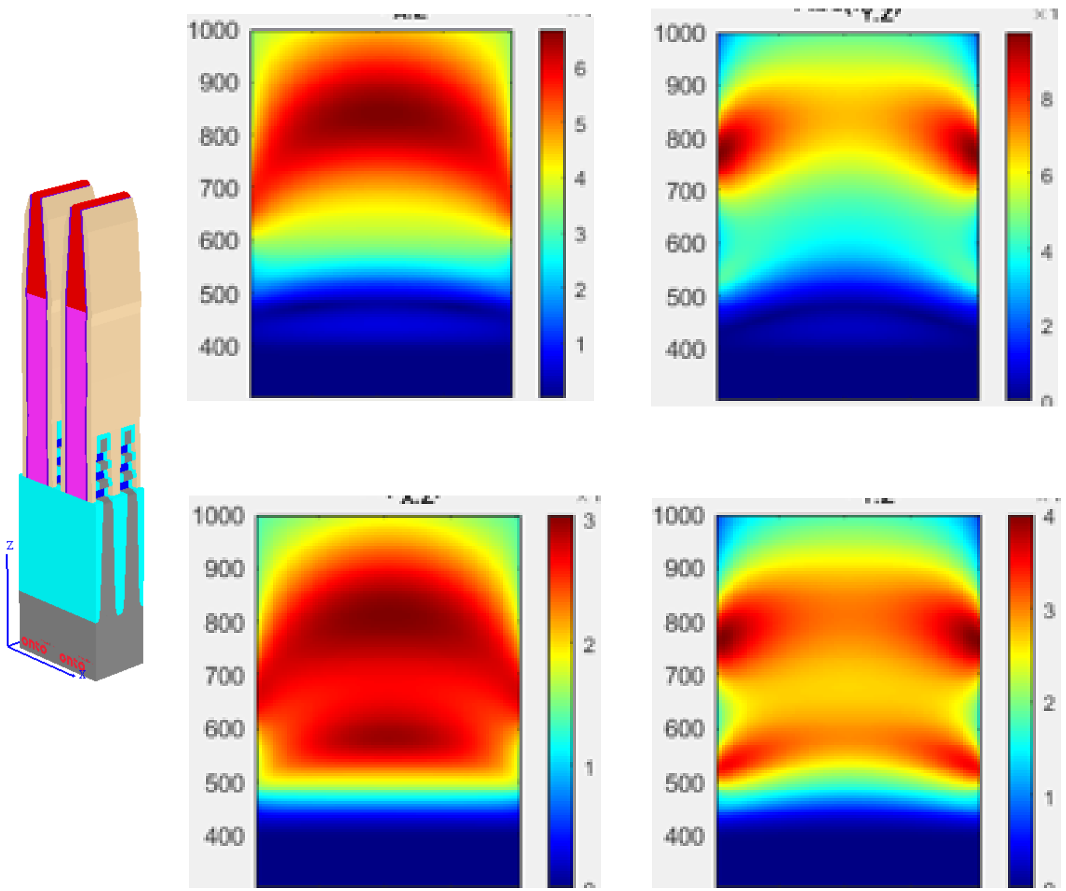
innovation™ onto

Simulations of Gate-All Around Transistor: CD Sensitivity

Signal intensity projections

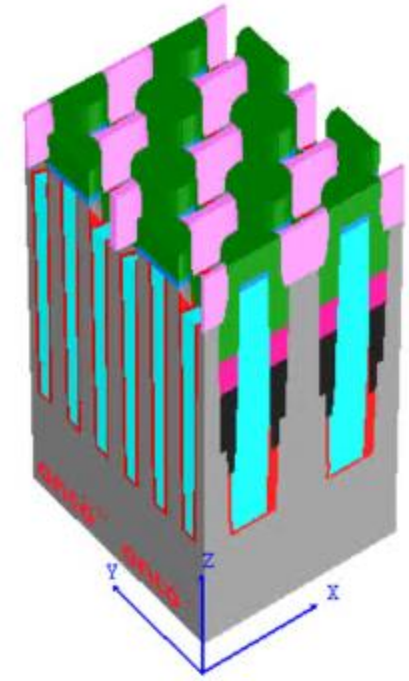


Sensitivity to 0.1nm process property change

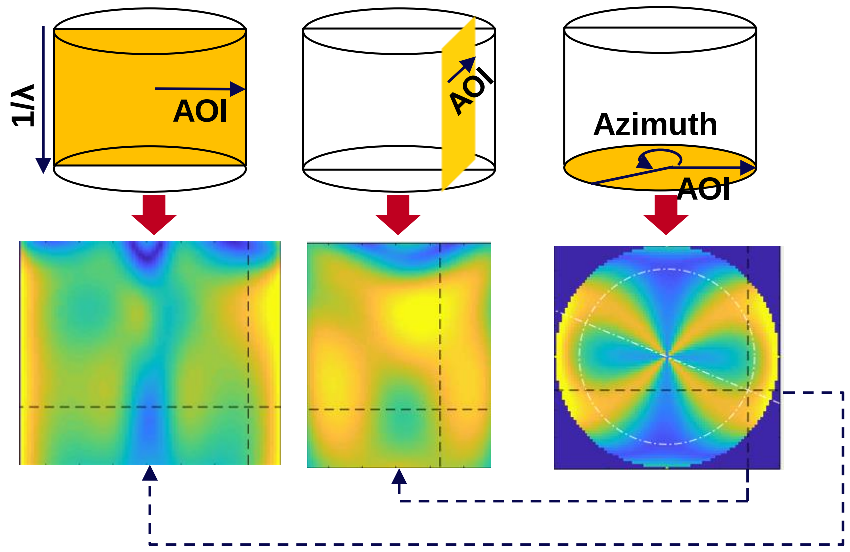


innovation
onto

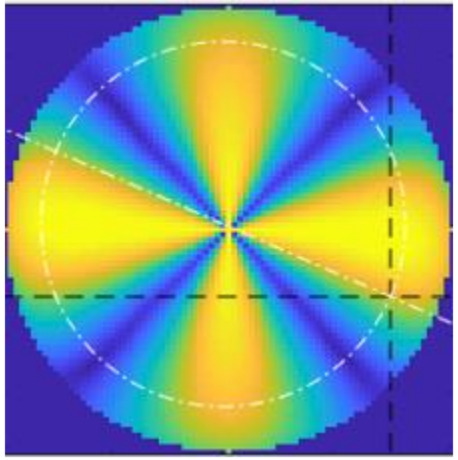
Simulations of DRAM Bit Line Gate Etch: Overlay Sensitivity



Signal intensity projections

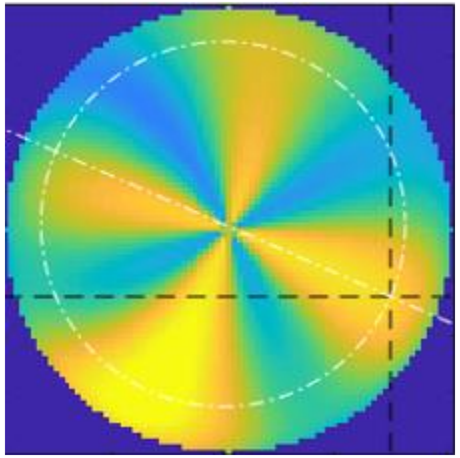


Sensitivity to property change



0.1nm CD change

Pattern symmetry is retained

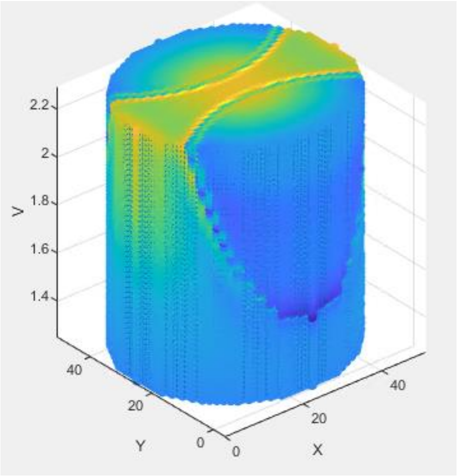


0.1nm overlay

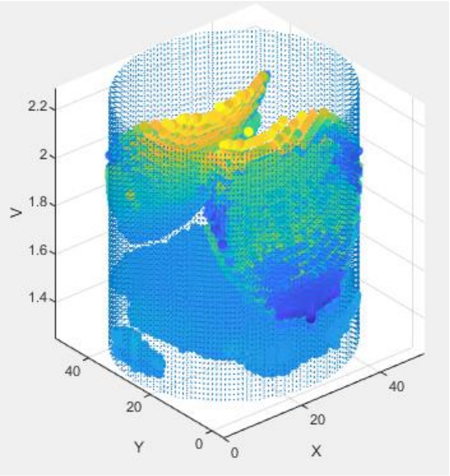
Pattern rotates and becomes asymmetric

Data Selection

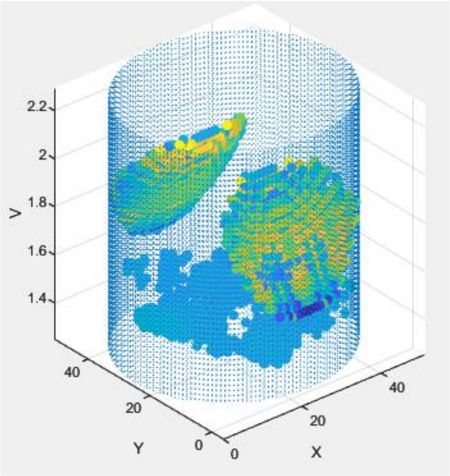
Reduce data volume to a selectable number of the most useful data points



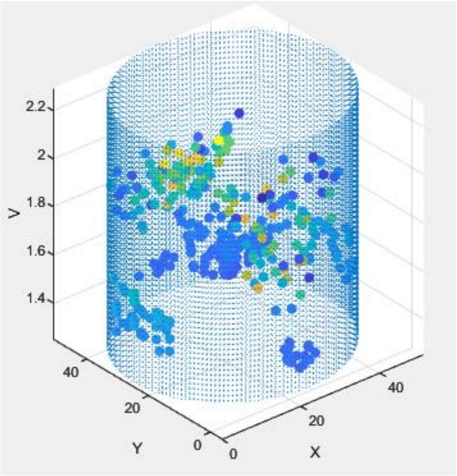
Original data cylinder



Filter 1



Filter 2

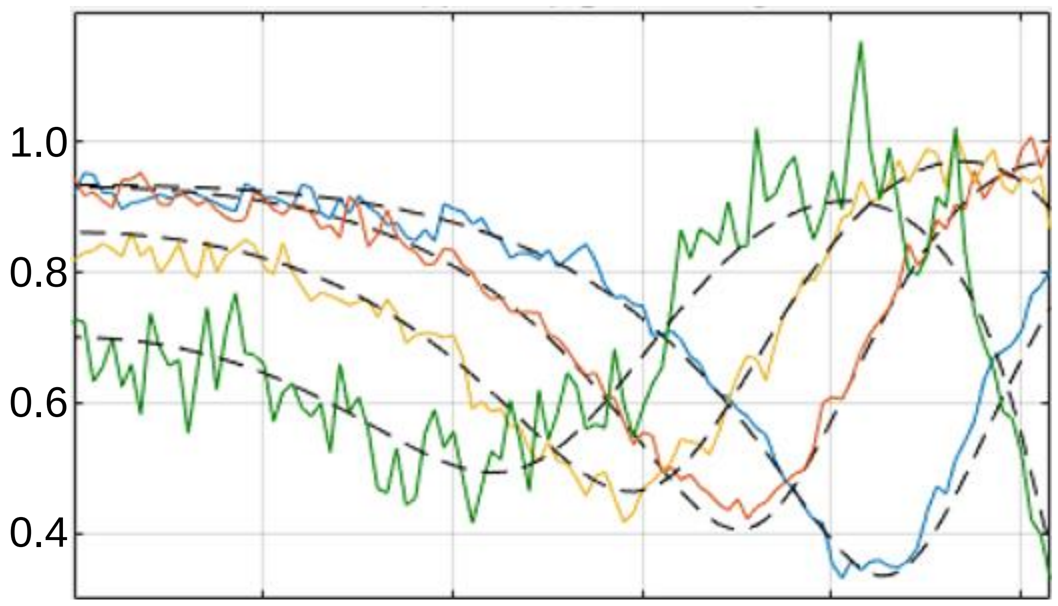


Filter 3

Thin Film Signal and Simulation

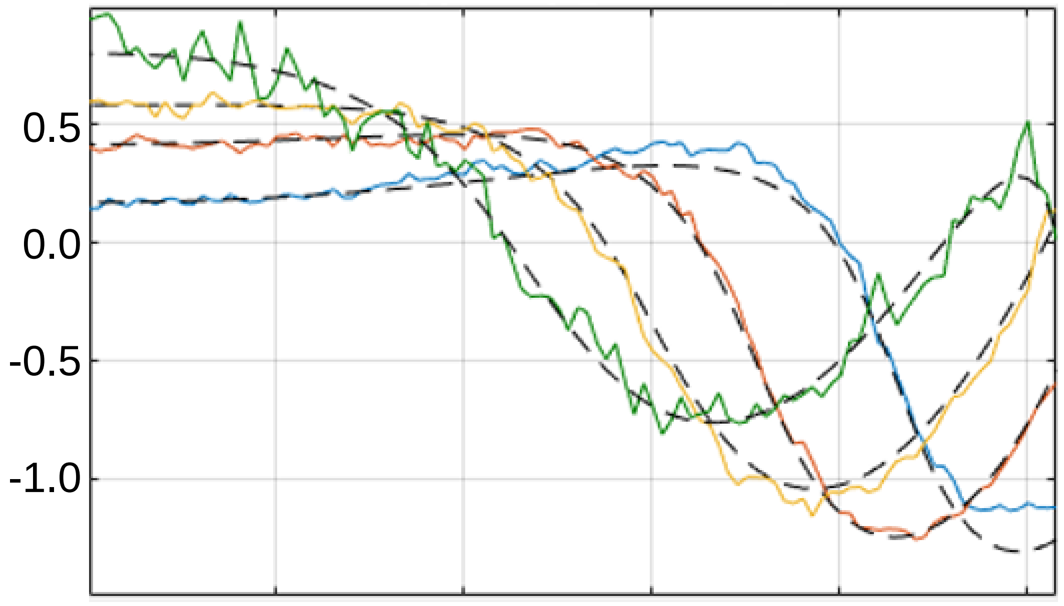
1045nm SiO_2

Signal Magnitude vs. AOI



769.2 nm
625.0 nm
526.3 nm
454.5 nm

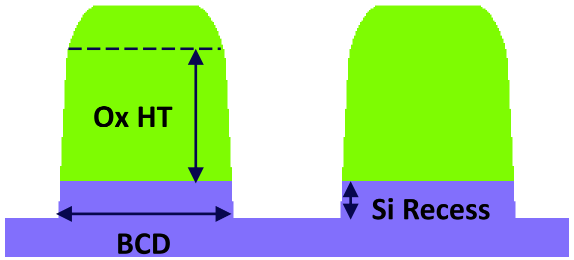
Signal Phase (radian) vs. AOI

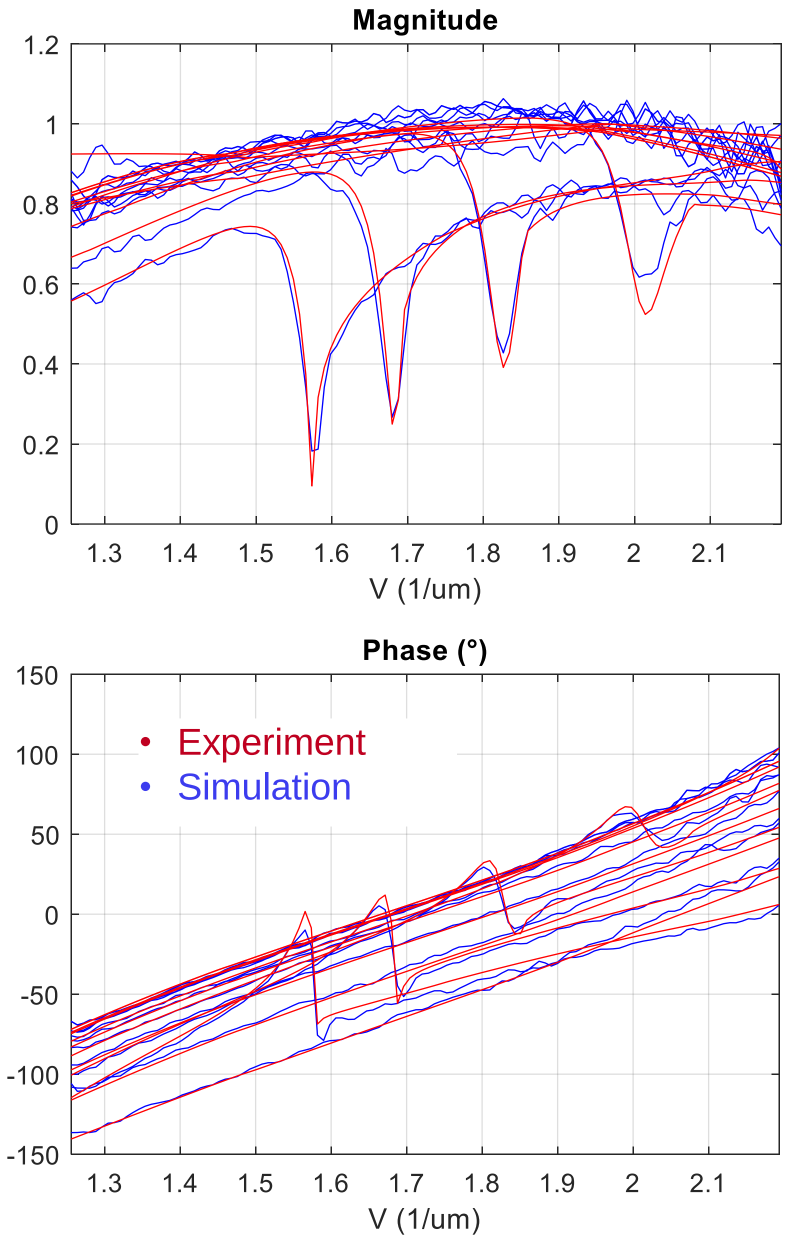


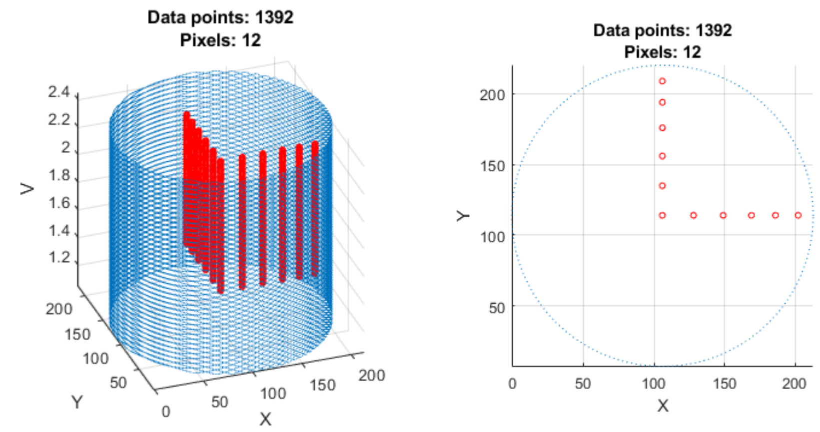
- Good use of phase data requires a common "zero phase" for all data points.
- Platform drift changes phase.
- The tool must be stable throughout the measurement.

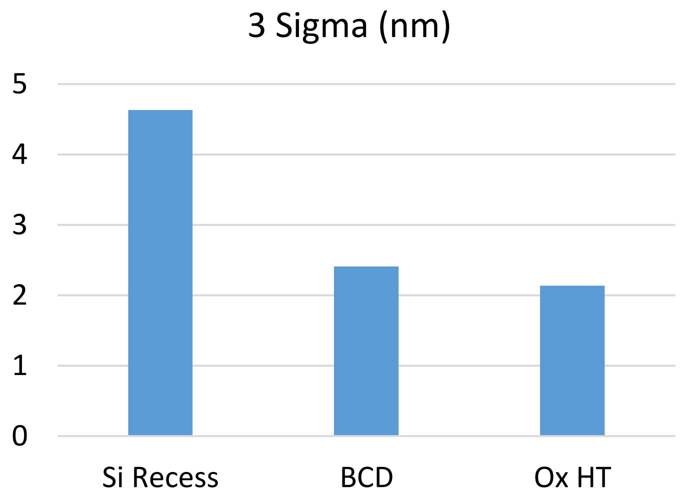
Measurement Case 1: 360nm pitch 1:1 Grating

Sampling plan





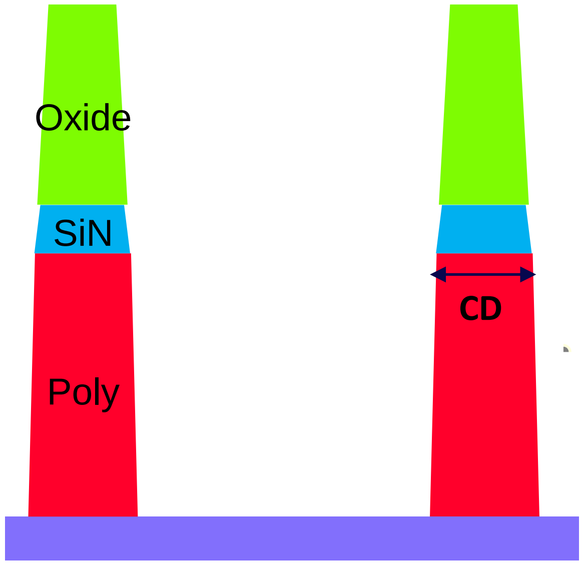


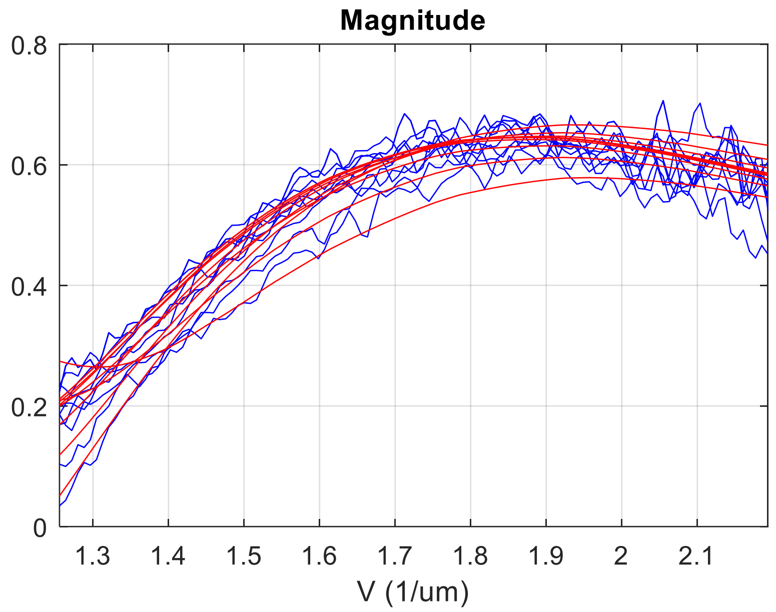


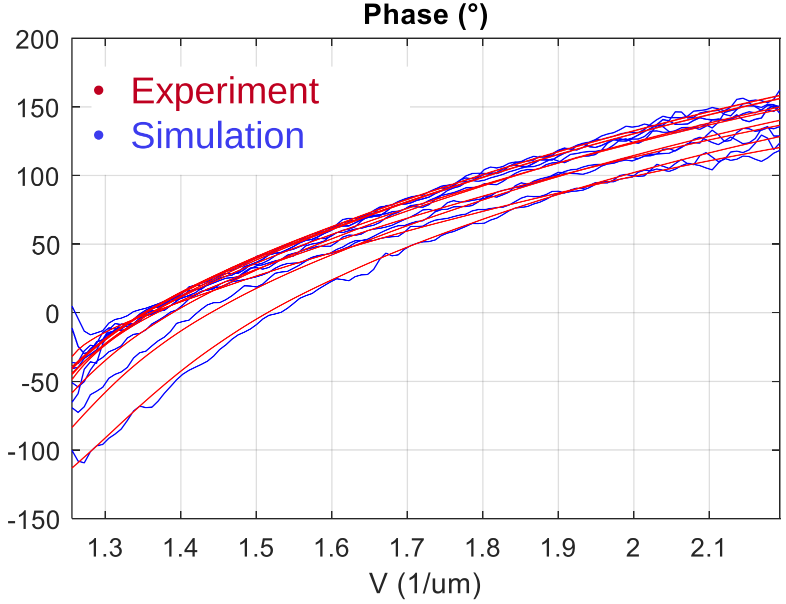
- Good use of phase requires wavelength-to-wavelength continuity.
- Platform stability is vital.

innovation onto

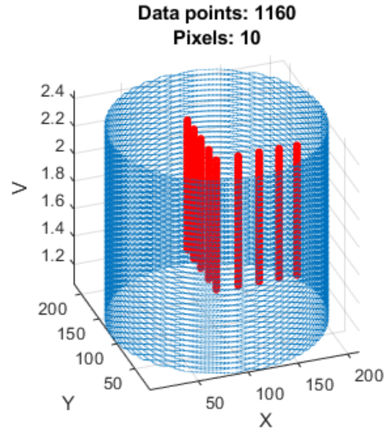
Measurement Case 2: Poly CD

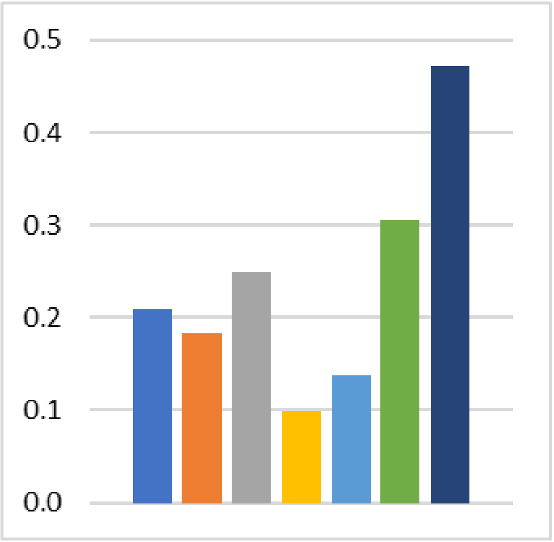






CD precision





- Good use of phase requires wavelength-to-wavelength continuity.
- Platform stability is vital.

innovation onto

Vibration

$$I(z) = I_{DC} + I_{AC}(z) = S(k, AOI, Azimuth, t)(1 + C(k, AOI, Azimuth, z)e^{ikz})$$

Source intensity, tool

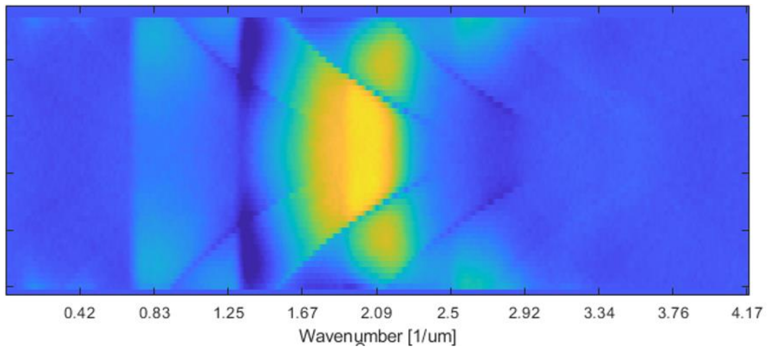
Signal envelope: sample

Modulation

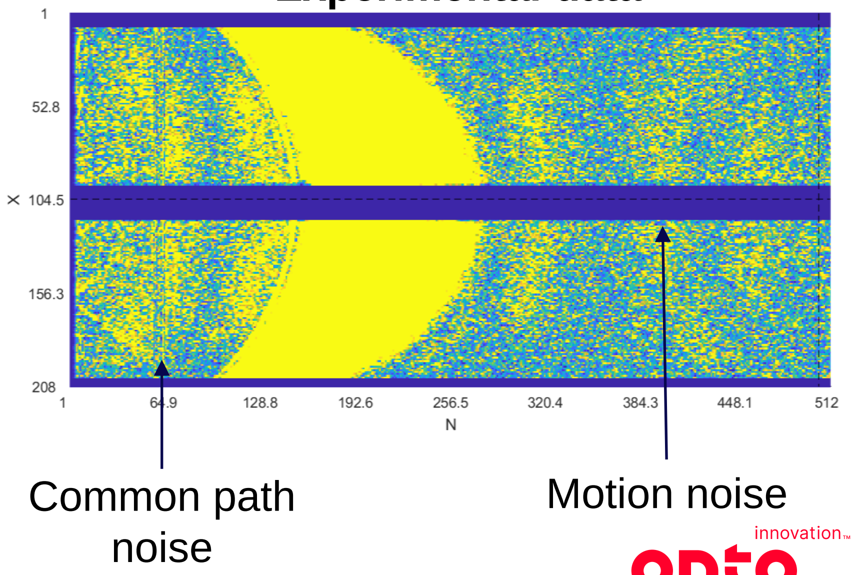
Single frequency noise: $z(t) = Vt + A\sin 2\pi f_z t$

Noise in signal: $\mathcal{FT}\left\{\frac{\partial C}{\partial z} dZ\right\}(k) = ikA\mathcal{FT}\{C\}(k \pm 2\pi f_z/V)$

Simulation: 7nm noise at 18Hz



Experimental data



- Common path noise gives rise to single frequency bands in the spectrum
- Motion noise inside the interferometer gives rise to side bands in the spectrum
- The noise contribution is independent of AOI
- The noise (f) frequency / side band is at $f_z/F\Delta$. (Frame rate F, sample spacing Δ; FΔ = V = scan speed).
- Noise is important if it appears inside the measured spectral range…
- … e.g. if $f_z = 0$! **Low frequency noise is important.**

innovation™ onto

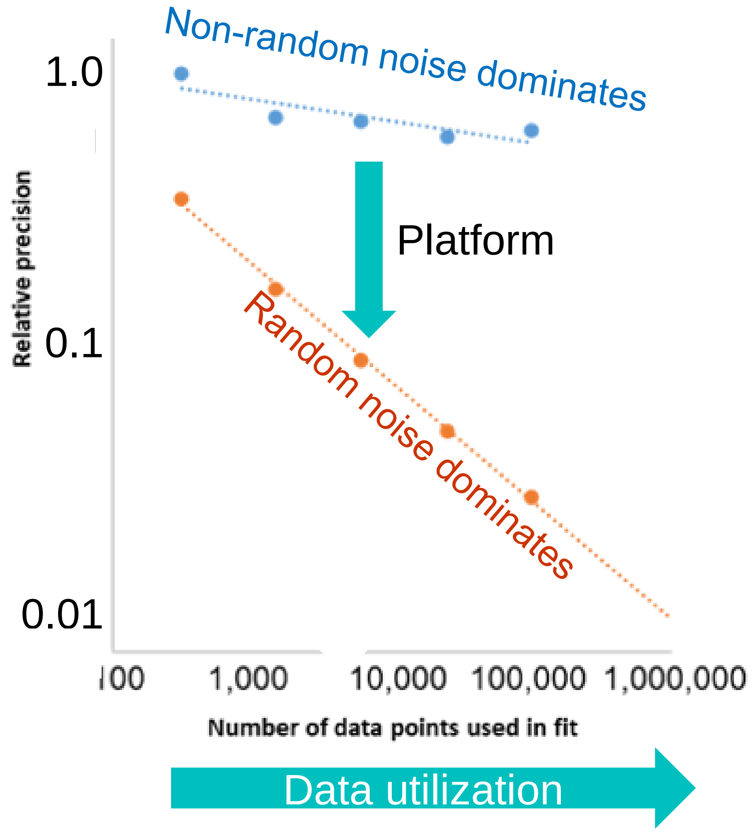
The Future #1: Increasing Data Utilization

- Dominant non-random noise:
  - low frequency vibration
  - focus error (Z stage capability).
- Random noise
  - Shot noise in camera
    - camera well depth,
    - light intensity
    - exposure time
  - High frequency vibration.
- If dynamic noise dominates then precision improves only slowly with increasing number of data points.

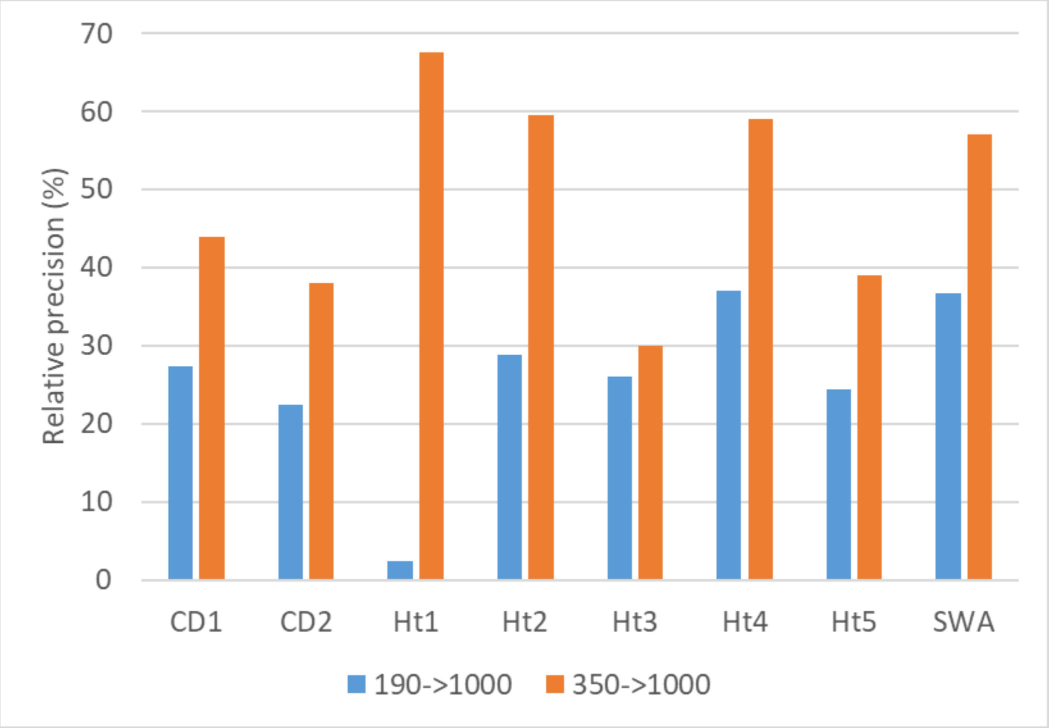
$$Precision \propto \sqrt{(dynamic\ noise)^2 + \frac{1}{N}(random\ noise)^2}$$

- If random noise dominates then for N data points:

$$Precision \propto \frac{1}{\sqrt{N}}$$

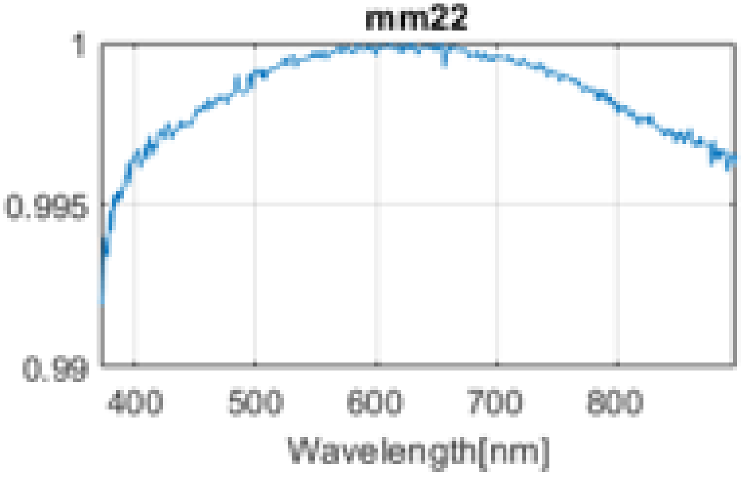


The Future #2: Shorter Wavelengths



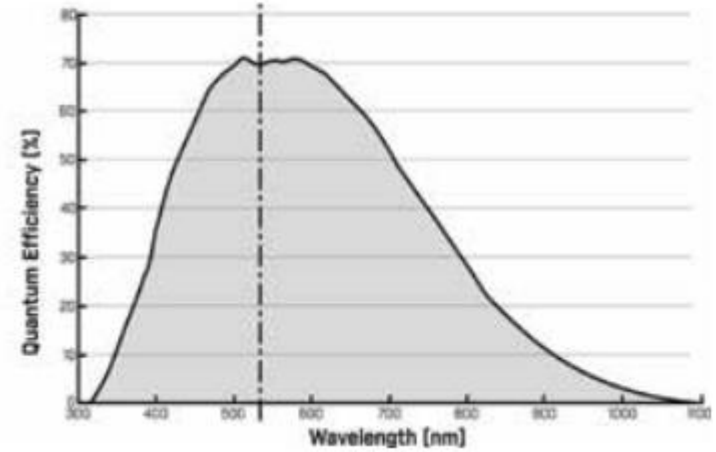
Precision using 190-1000 and 350-1000nm wavelength ranges relative to prevision using 450-1000nm

DRAM example.



Objective

Measured response



Sensor response

(manufacturer data)

Select components for shorter wavelengths.
Reflecting optics might be necessary.

innovation

onto

Conclusions

- Pupil plane interferometry is a viable OCD technique
  - Separation of incident angles enables a viable signal formation model.
  - Then use familiar model fitting methods to extract OCD parameters.
  - Requires reduction of the data to a manageable number of highly effective data points.
  - Phase data should be continuous throughout the data set.
    - The tool must be very stable during data acquisition.
- Probe size is much smaller than other methods.
- Significant improvement in measurement capability is feasible
  - By utilizing as much of possible of the acquired data cylinder.
  - Implies a change of strategy from simulation and fitting towards machine learning methods.

innovation™ onto

Thank You

谢谢 | 謝謝

ありがとう

Obrigado

Danke

감사합니다

Merci

info@ontoinnovation.com
www.ontoinnovation.com

onto innovation™