Modeling and Mapping of Nanotopography Interactions with CMP

D. Boning and B. Lee, MIT
W. Baylies, BayTech Group
N. Poduje, ADE Corp.
J. Valley, ADE Phase-Shift

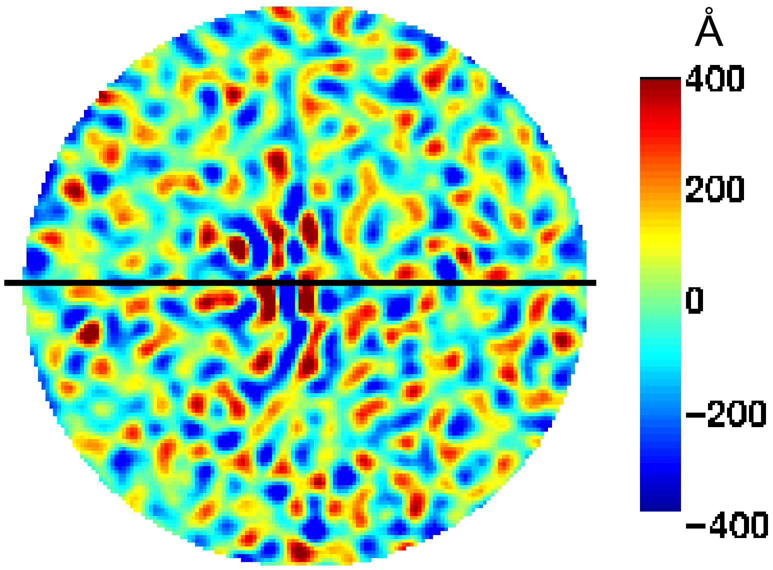
MRS Spring Meeting
Symposium on Chemical Mechanical Planarization
April 2002

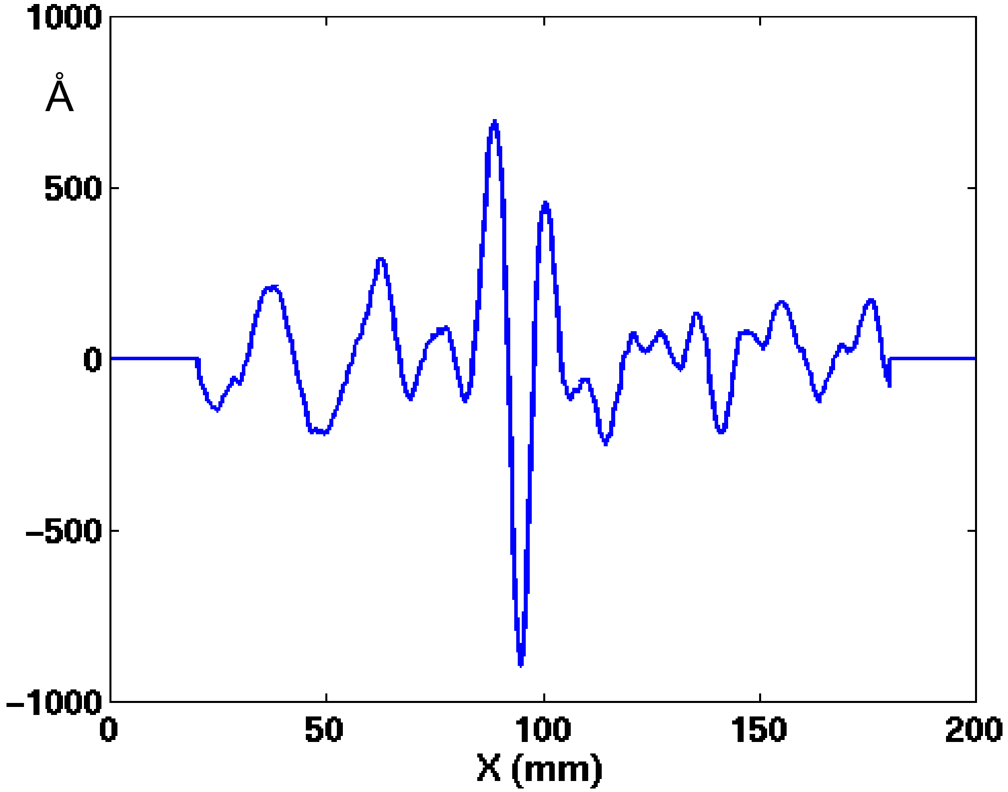
Outline

- Review: Nanotopography interaction with CMP
  - Experiments demonstrate oxide thinning over nanotopography
- Mapping nanotopography into oxide thickness deviation
  - Scaling
  - Linear filter
  - Contact wear
- Comparison of models
- Application of nanotopography/CMP model:
  - Prediction of STI clearing problem areas
  - Prediction of excessive nitride erosion problem areas
- Conclusions

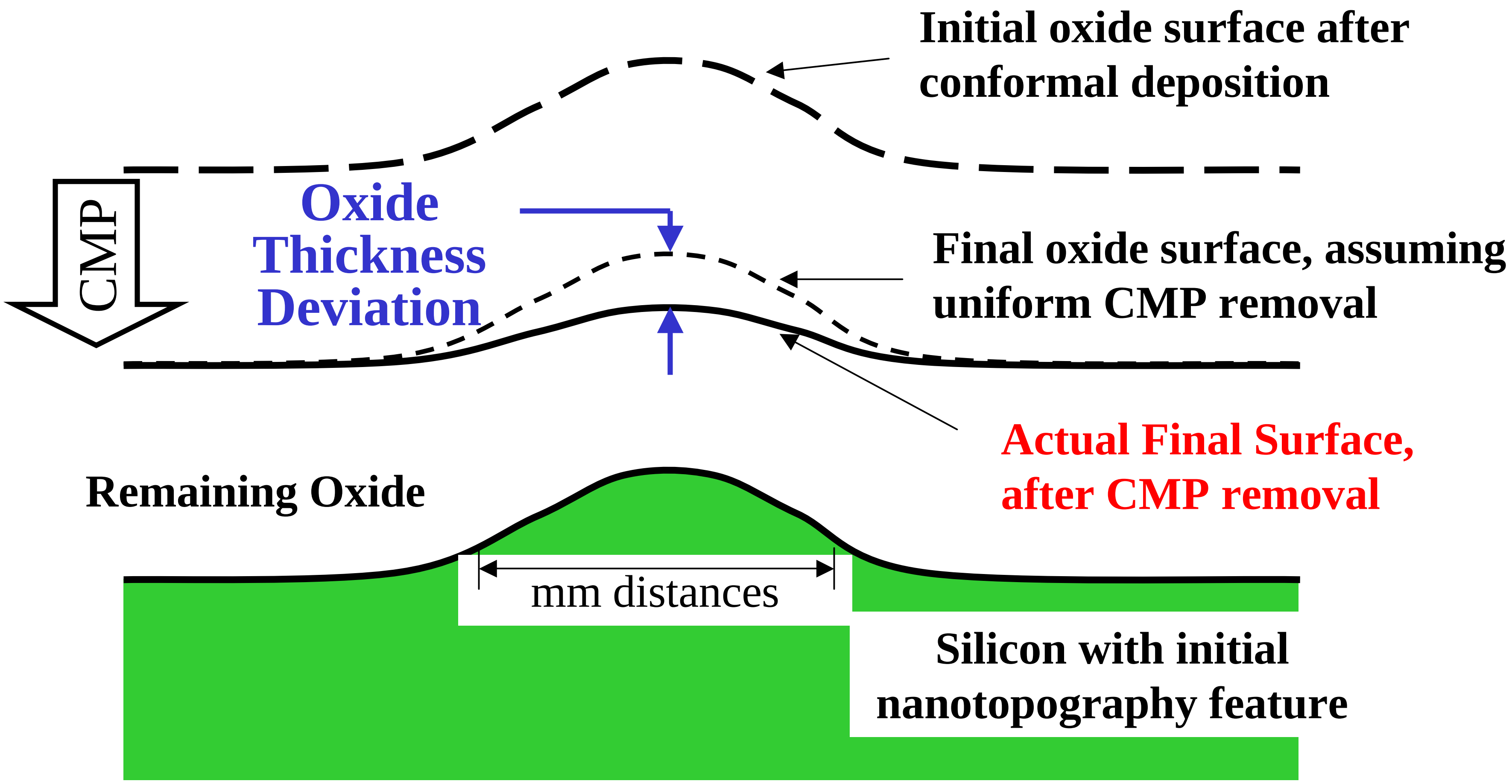
Review: Nanotopography

- Nanometer height variations occurring on millimeter lateral length scales in virgin silicon wafers





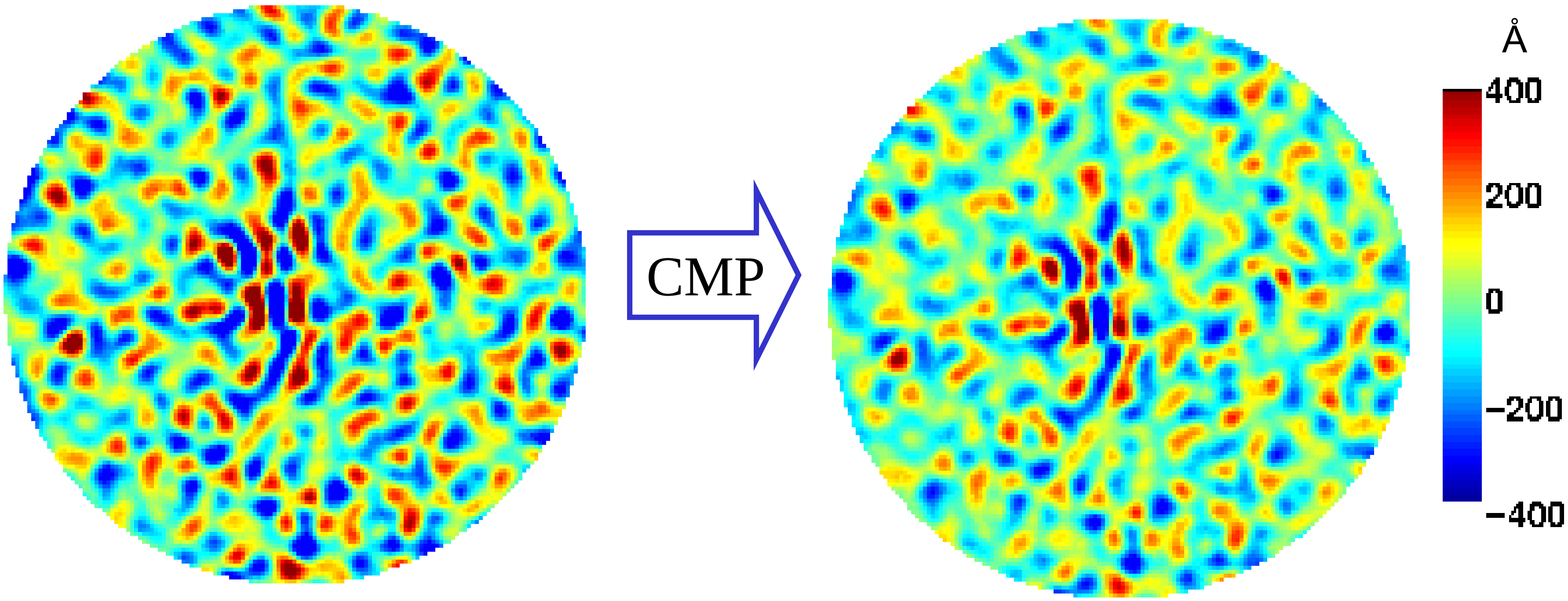
Problem: Oxide Thickness Deviation (or Oxide Thinning) During CMP



Nanotopography & Oxide Thinning

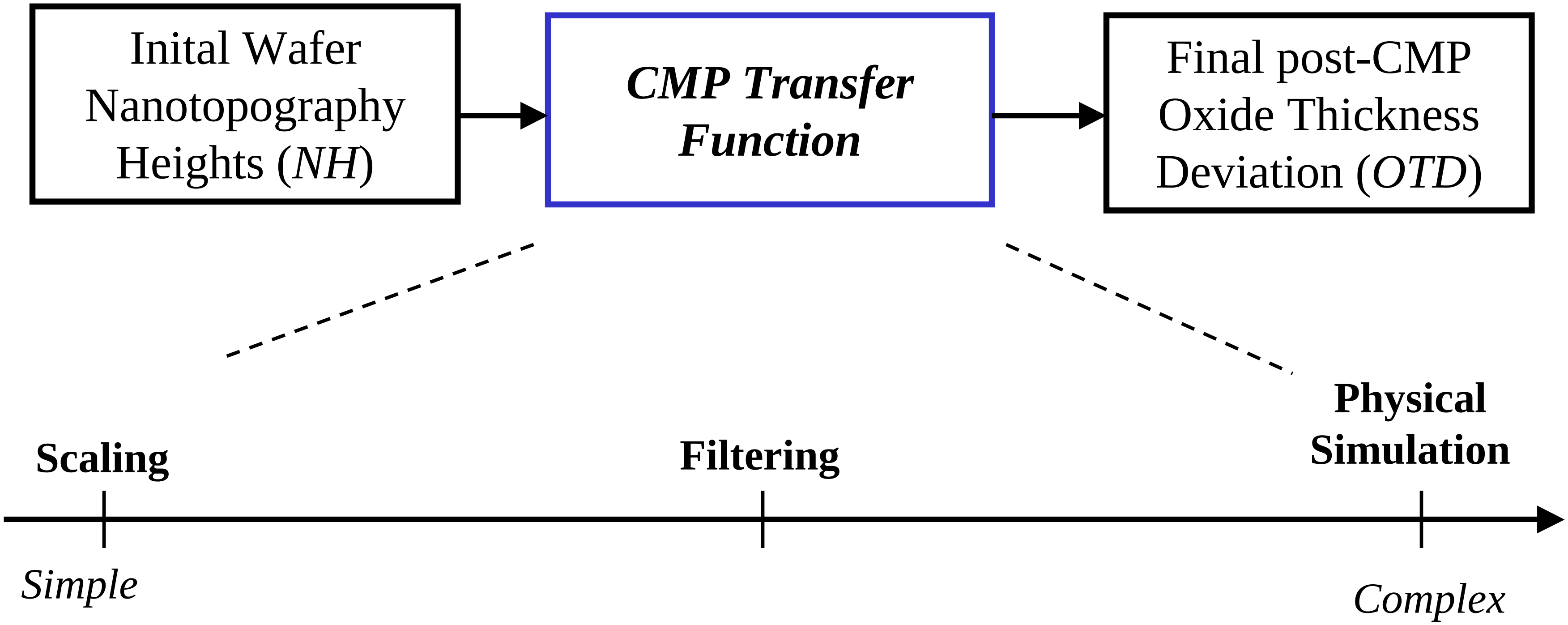
Nanotopography
filtered wafer height map

Inverted oxide thickness
with zero mean



Single-Sided Polish Wafer Data @ 20mm EE

Goal: Predictive Mapping of Nanotogography to Oxide Thickness Deviation



Inital Wafer Nanotopography Heights (*NH*)

CMP Transfer Function

Final post-CMP Oxide Thickness Deviation (*OTD*)

Scaling

Filtering

Physical Simulation

Simple

Complex

Scaling Model

- Assume CMP reduces nanotopography height by a constant scaling factor, on a point by point basis:

$$OTD(x,y) = \alpha \cdot NH(x,y)$$

- Model parameter extraction
  - α: scaling factor
  - Use pre-CMP measured NH data for a given wafer and post-CMP film thickness data for same wafer
  - Extract model parameter; assume it applies for any pairing of that CMP process (pad, slurry, tool, settings, ...) and wafer type

After Schmolke et al., *JECS,* 2002

Filtering Model

- Approximate interaction of nearby topography by treating the CMP process as a 2D linear filter

$$OTD(x, y) = s \cdot NH(x, y) \otimes h(r)$$

- Filtering efficiently performed in the frequency domain:

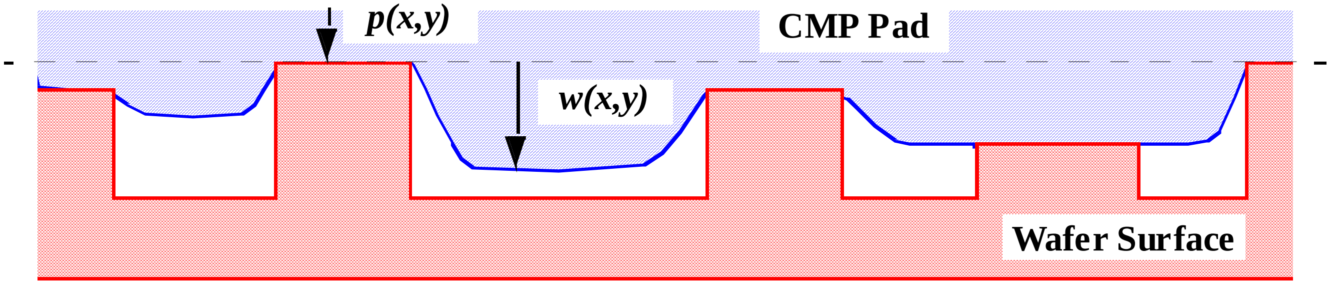
$$H(\omega) = 2e^{-k\left(\frac{\omega}{\omega_{cp}}\right)^2} - e^{-2k\left(\frac{\omega}{\omega_{cp}}\right)^2}$$

$$\omega_{cp} = \omega_c \big/ 1.331, \quad \omega_c = 2G / L_C$$

- Double Gaussian filter structure shown above is used
- Model parameter extraction (using experimental data)
  - L_C: filter cutoff length
  - s: filter scaling factor

Contact Wear Model

- Explicitly model the pressure distribution *p(x,y)* across wafer surface as pad bends around nanotopography



$$w(x,y) = \frac{(1-\upsilon^2)}{\pi E} \int_A \frac{p(\xi,\eta)}{\sqrt{(x-\xi)^2 + (y-\eta)^2}}\, d\xi\, d\eta$$

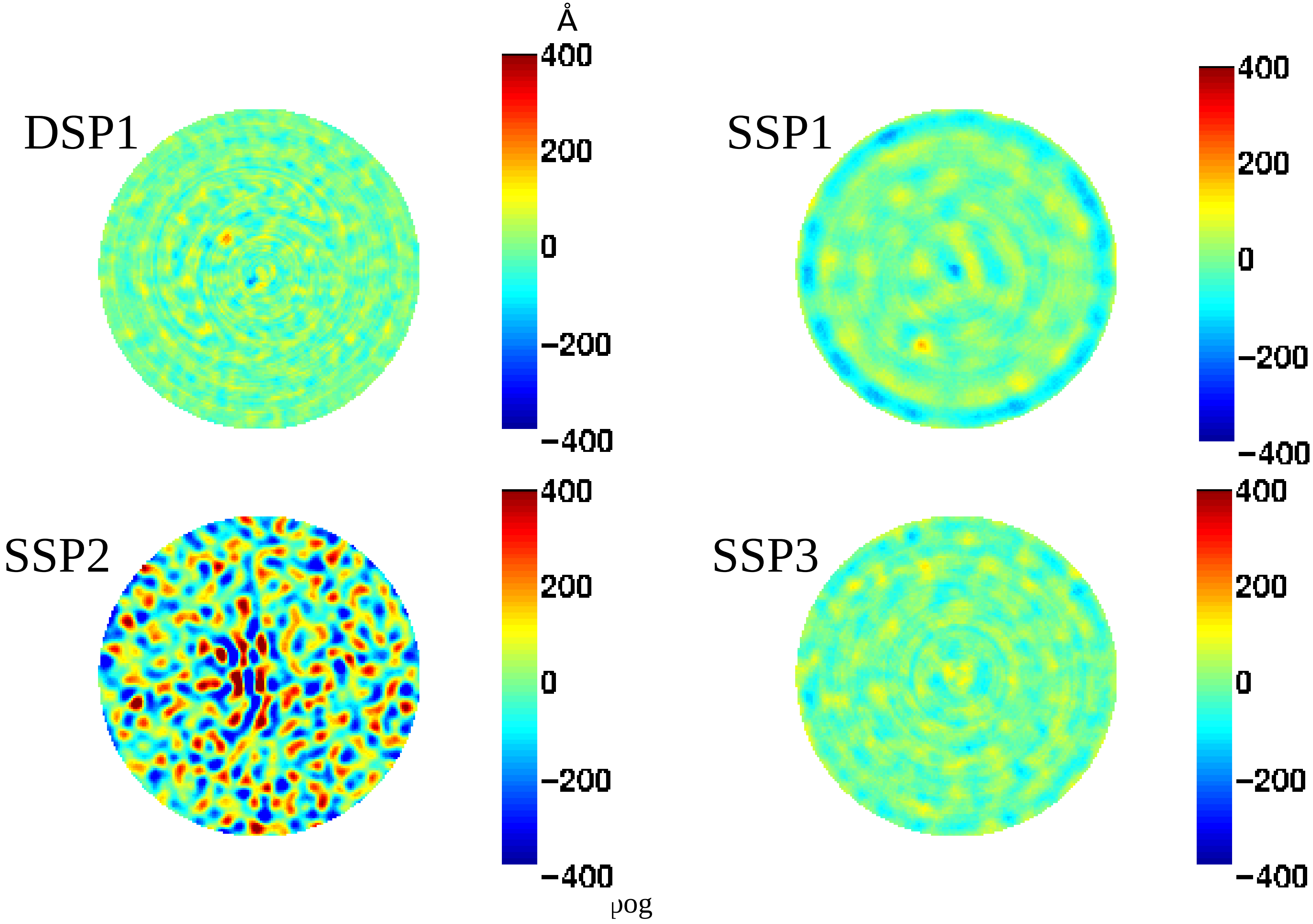
Chekina, *JECS* 1998;
Yoshida, *Proc. ECS*, 1999

- Under assumption of full pad contact, displacement *w(x,y)* is known at each time step and *p(x,y)* can be found using FFT
- Model parameter extraction (using experimental data)
  - *E*: Effective Young's modulus for the pad

Experiments

- Experiments performed to study the interaction between nanotopography and different CMP processes
  - Pre-CMP measurements of *NH* on virgin wafers
  - Thermal oxidation to grow conformal oxide on surface
  - Polish and Post-CMP measurement of *OTD*
- Four nanotopography wafer types
  - Three types of Single-sided polish (SSP) epi wafers
  - One type of Double-sided polish (DSP) epi wafers
- Three CMP processes
  - Process B (soft pad): PL = 3.4 mm
  - Process G (medium pad): PL = 6.4 mm
  - Process F (stiff pad): PL = 9.7 mm
- Two wafer replicates examined for each process/wafer combinations

Wafer Nanotopography Signatures



Scaling Model - Extracted Coefficients

Process	Planarization Length (mm)	Wafer Type	α	
			Wafer 1	Wafer 2
B	3.4	SSP1	0.07	0.08
		SSP2	**0.10**	0.14
		SSP3	0.13	0.11
G	6.4	DSP1	0.38	**0.44**
		SSP1	0.19	**0.17**
		SSP2	**0.28**	**0.27**
		SSP3	0.27	**0.25**
F	9.7	DSP1	0.72	0.73
		SSP1	0.84	0.85
		SSP2	**0.74**	0.72
		SSP3	0.71	0.69

- α increases with increasing planarization length
- α seems to depend on both process and wafer type

Filter Model - Extracted Coefficients

Process	Planarization Length (mm)	Wafer Type	Filter length: L_C (mm)		Scaling factor: s	
			Wafer 1	Wafer 2		
B	3.4	SSP1	8	9	0.45	0.42
		SSP2	**12**	9	0.20	0.35
		SSP3	7	100	0.48	0.10
G	6.4	DSP1	11	11	0.66	0.66
		SSP1	18	16	0.43	0.44
		SSP2	**14**	14	0.48	0.46
		SSP3	12	13	0.56	0.51
F	9.7	DSP1	**100**	**100**	**0.72**	**0.73**
		SSP1	**92**	**100**	**0.85**	**0.87**
		SSP2	**100**	**100**	**0.74**	**0.72**
		SSP3	**100**	**61**	**0.71**	**0.71**

- L_C increases with increasing planarization length
- In stiff pad case, L_C ~ full window size, in which case the filter based model $\cong$ simple scaling model $\Rightarrow s = \alpha$

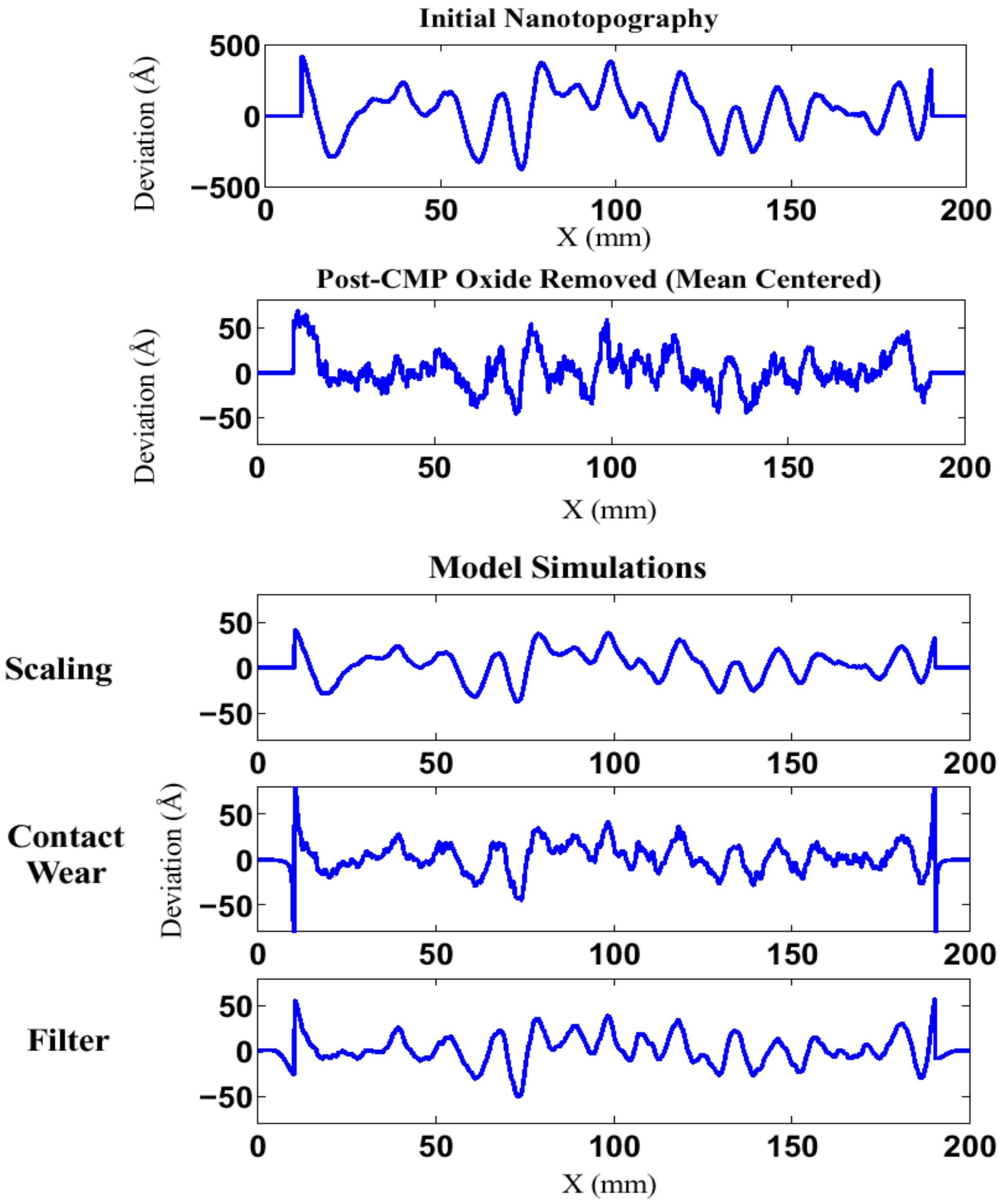
Contact Wear Model - Extracted Coefficients

Process	Planarization Length (mm)	Wafer Type	Effective Pad Modulus: *E* (MPa)	
			Wafer 1	Wafer 2
B	3.4	SSP1	20	25
		SSP2	**20**	30
		SSP3	30	25
G	6.4	DSP1	75	75
		SSP1	60	50
		SSP2	**75**	60
		SSP3	60	60
F	9.7	DSP1	250	**260**
		SSP1	900	**800**
		SSP2	**300**	**280**
		SSP3	375	**360**

- *E* increases with increasing planarization length
- Consistent extractions of *E* for processes B and G; process F has large *E* but varies substantially depending on wafer

Model Comparison

- Simulate 100 mm central region of wafer (to avoid edge effects) to predict OTD
- Case shown:
    - Process B, SSP2, wafer 1
- Observations:
    - all models capture trends
    - filter and contact wear models handle interactions with nearby nanotopography (including in Y direction not pictured)



Model Error Comparisons

Wafer Data					Prediction Error - RMS (nm)					
Process	PL (mm)	Wafer Type	σ_{OTD} (nm)		Scaling		Filter		Contact Wear	
			W1	W2	W1	W2	W1	W2	W1	W2
B	3.4	SSP1	1.20	1.70	1.18	1.67	1.13	1.65	1.15	1.65
		SSP2	1.87	3.62	1.27	2.30	1.16	1.57	1.13	1.67
		SSP3	1.16	1.08	1.08	1.00	1.00	1.01	1.01	0.94
G	6.4	DSP1	2.01	1.70	1.23	1.03	0.90	0.82	0.97	0.88
		SSP1	1.05	1.04	0.81	0.85	0.77	0.81	0.74	0.79
		SSP2	5.12	4.30	2.11	2.01	1.50	1.42	1.47	1.45
		SSP3	1.38	1.36	0.94	0.96	0.83	0.90	0.84	0.87
F	9.7	DSP1	3.51	3.21	1.91	1.95	1.91	1.96	1.98	2.04
		SSP1	3.42	3.38	1.98	1.39	1.98	1.40	2.04	1.44
		SSP2	13.12	11.11	2.17	2.00	2.18	2.01	2.79	2.64
		SSP3	2.82	2.90	1.49	1.45	1.49	1.43	1.58	1.55

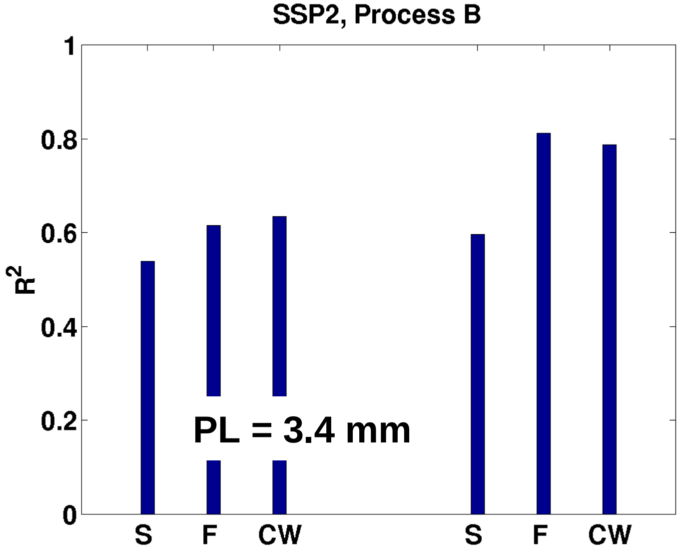
- Some noise floor (random or other unexplained variation in OTD) that is not predicted by any model: ~ 1-2 nm

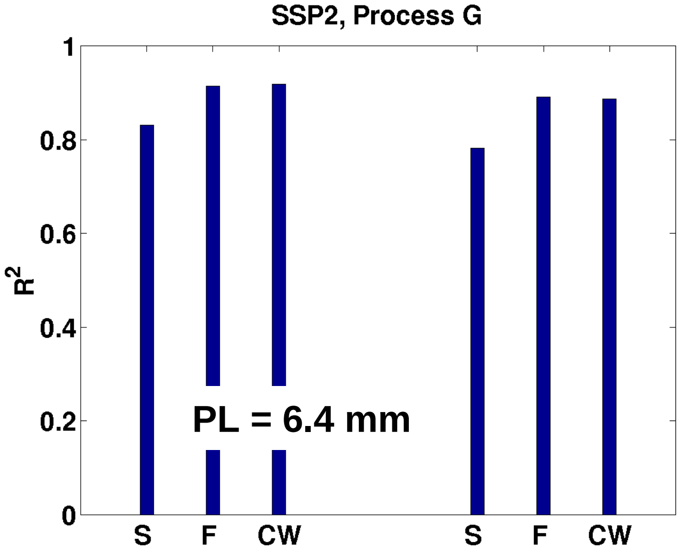
Model Comparisons

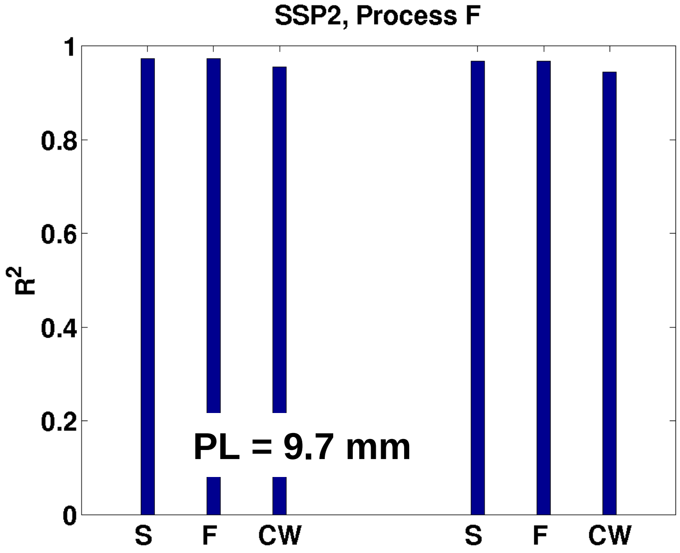
- R^2 for SSP2 case (2 wafers each process)
  - fraction OTD variance captured by each model:

$$R^2 = \frac{\sigma^2_{\text{model}}}{\sigma^2_{\text{OTD}}} = 1 - \frac{\sigma^2_{\text{err}}}{\sigma^2_{\text{OTD}}}$$

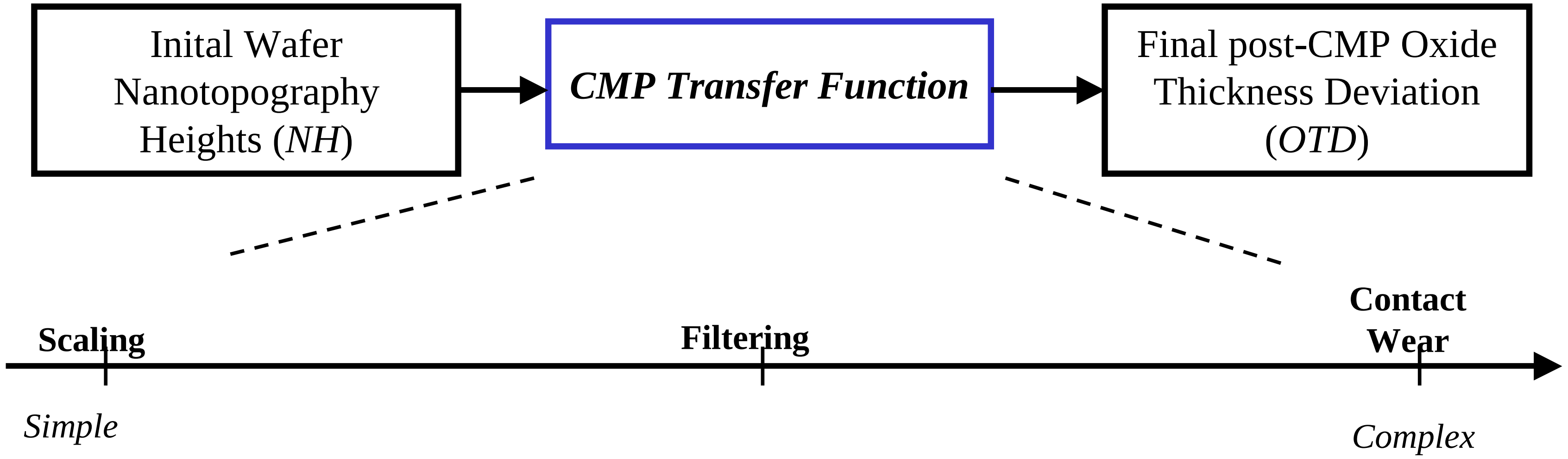
S - scaling
F - filter
CW - contact wear







Model Comparisons



Inital Wafer Nanotopography Heights (*NH*)

CMP Transfer Function

Final post-CMP Oxide Thickness Deviation (*OTD*)

Scaling

Filtering

Contact Wear

Simple

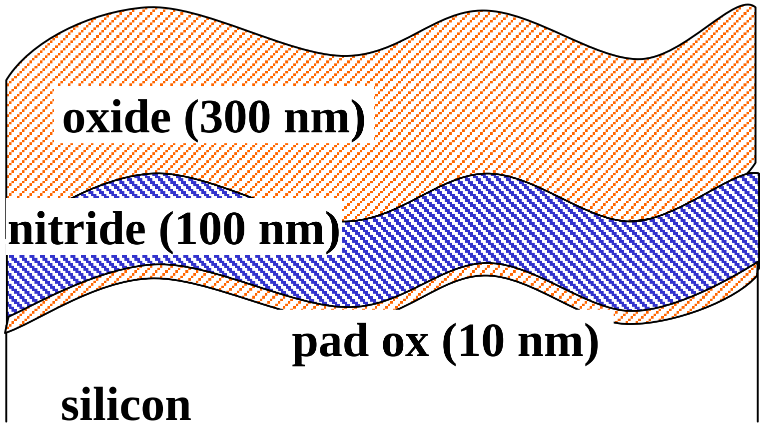
Complex

- Filter and contact wear models:
  - perform equally well
  - outperform scaling model in short and medium PL processes
- Scaling and filter models:
  - Major limitation: not clear how to apply to CMP of multiple films as occur in shallow trench isolation (STI)

Outline

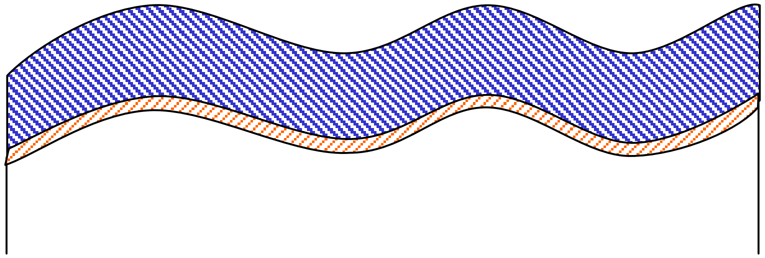
- Review: Nanotopography interaction with CMP
- Mapping nanotopography into oxide thickness deviation
  - Scaling
  - Linear filter
  - Contact wear
- Comparison of models
- ➢ Application of nanotopography/CMP model:
  - Prediction of STI clearing problem areas
  - Prediction of excessive nitride erosion problem areas
- Conclusions

Model Application: Predict Effect of Nanotopography on CMP of STI Stack



“Ideal” result:

- Complete removal of oxide
- No removal of nitride



STI stack over nanotopography: **pre-CMP**

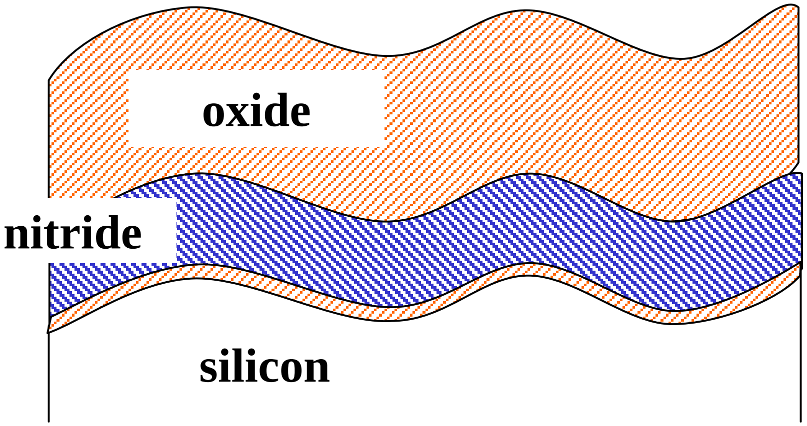
CMP

STI stack over nanotopography: **post-CMP**

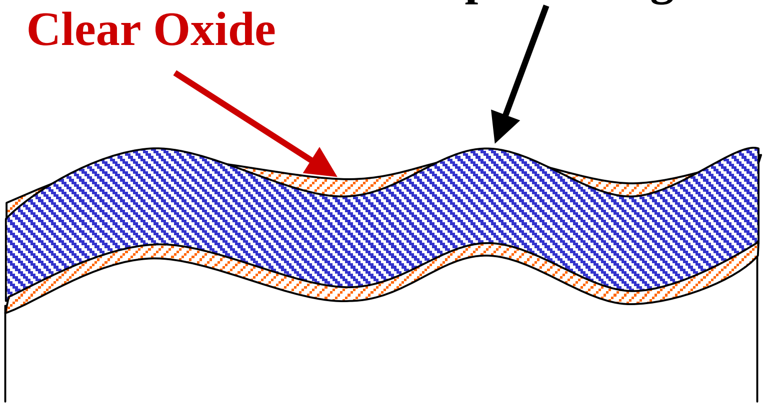
- **5:1 oxide:nitride selectivity**
- **Medium effective pad modulus: 147 MPa**

Problem #1: Failure to Clear Oxide

- Polish time is determined by initial clearing (initial endpoint) plus overpolish time
  - Polish to initial clear: 84 seconds for this wafer/stack
  - Assume fixed overpolish time: 14 seconds
- Examine regions where oxide remains
  - These result in device failure



STI stack: pre-CMP



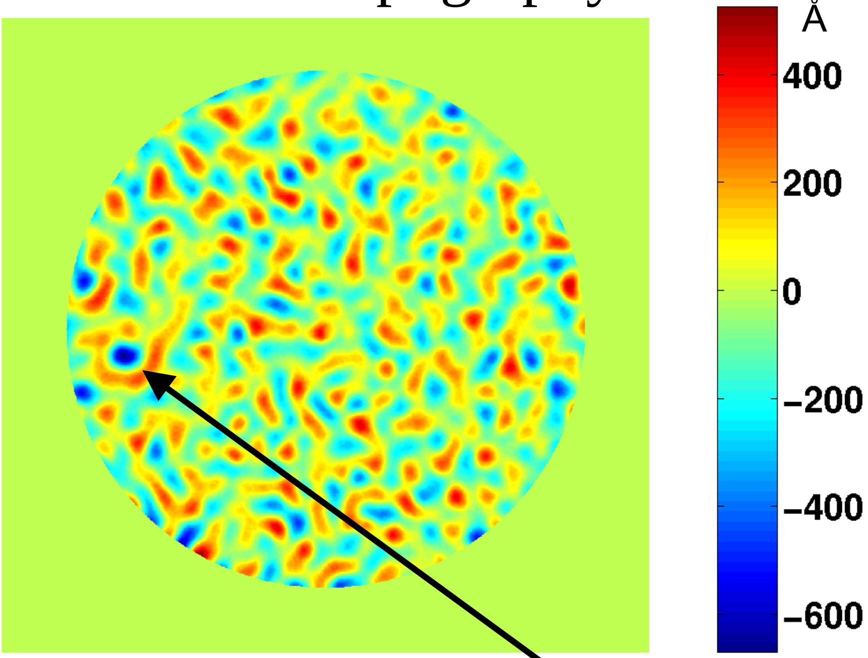
Failure to Clear Oxide

Initial clearing or "Endpoint" is detected when top points begin to clear

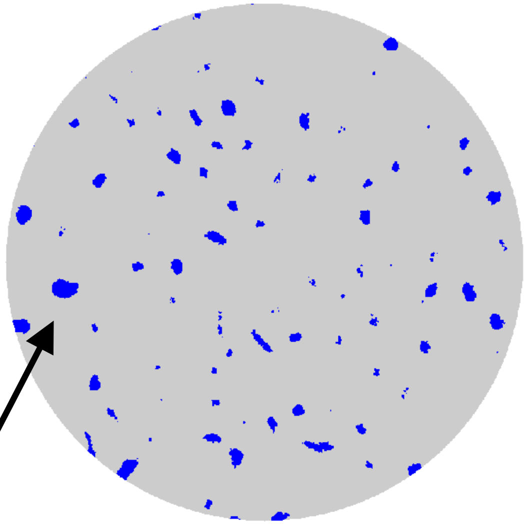
Post-CMP with too little overpolish

Simulated Result: Oxide Clearing Map

Initial Nanotopography



3% of wafer does not clear



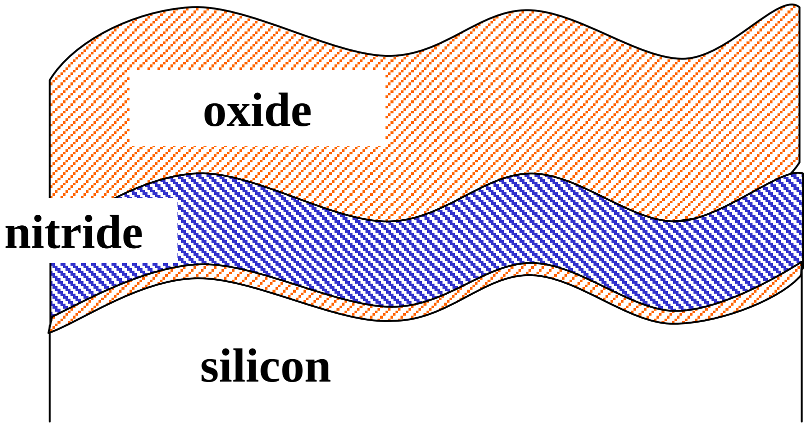
Blue indicates uncleared areas

Initial Nanotopography

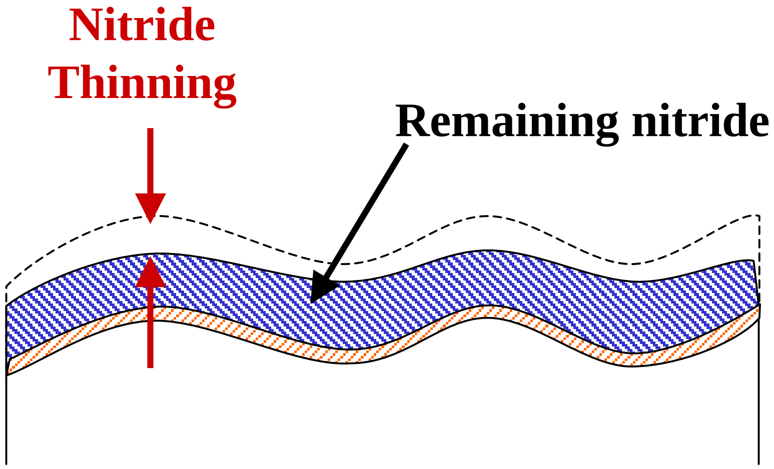
Note visual correlation with initial nanotopography “low spots”

Problem #2: Excessive Nitride Loss

- Failure to clear - causes incomplete transistor formation
  - Alternative: Increase overpolish time to ensure complete clearing of oxide in all nanotopography valleys everywhere on wafer
- Nanotopography thus forces *additional* overpolish time!
  - In addition to overpolish due to wafer level or chip pattern effects
- Resulting problem: excessive nitride loss - causes transistor performance degradation

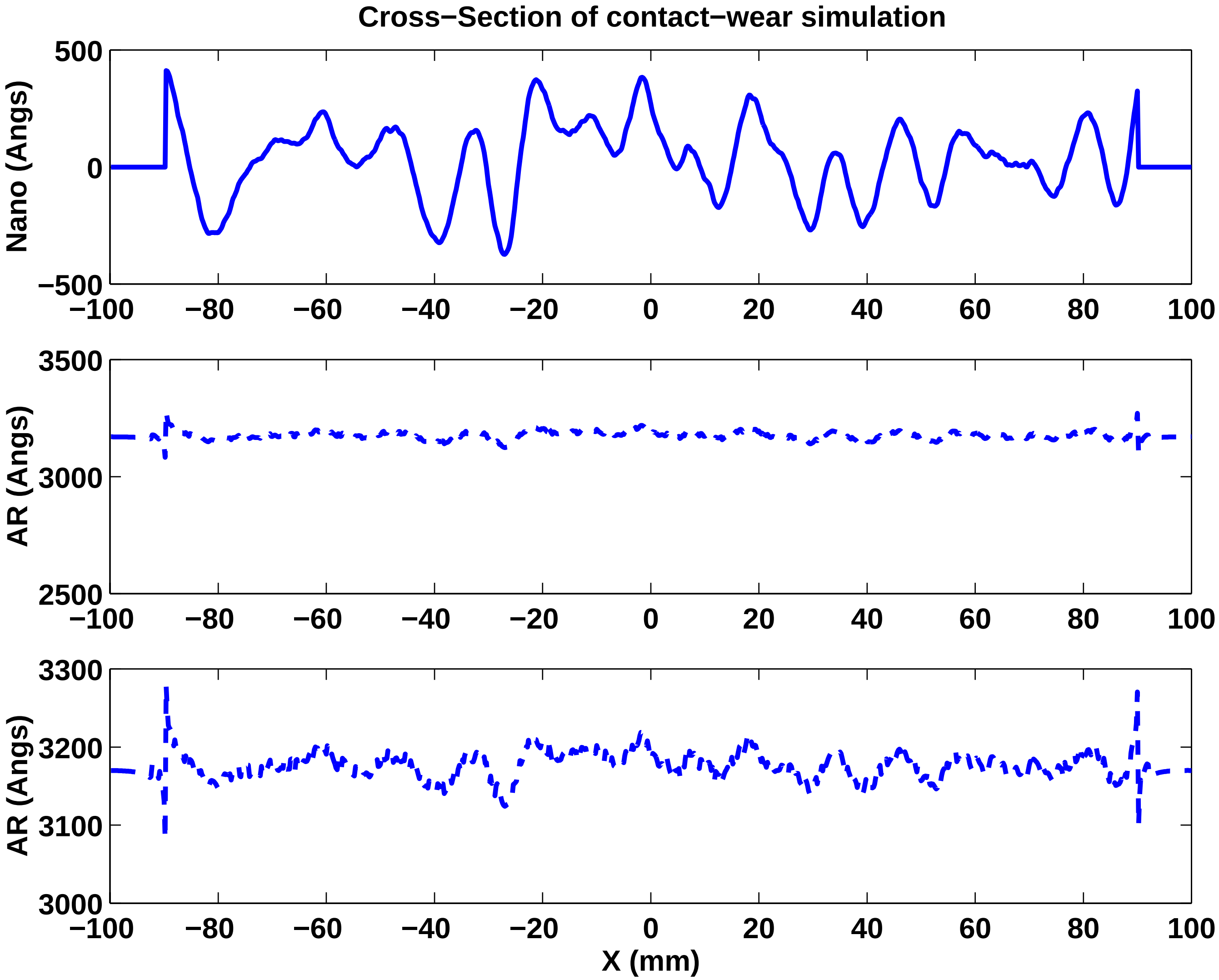


STI stack: pre-CMP



STI stack: post-CMP

Total Removal (After Just Clearing Everywhere)

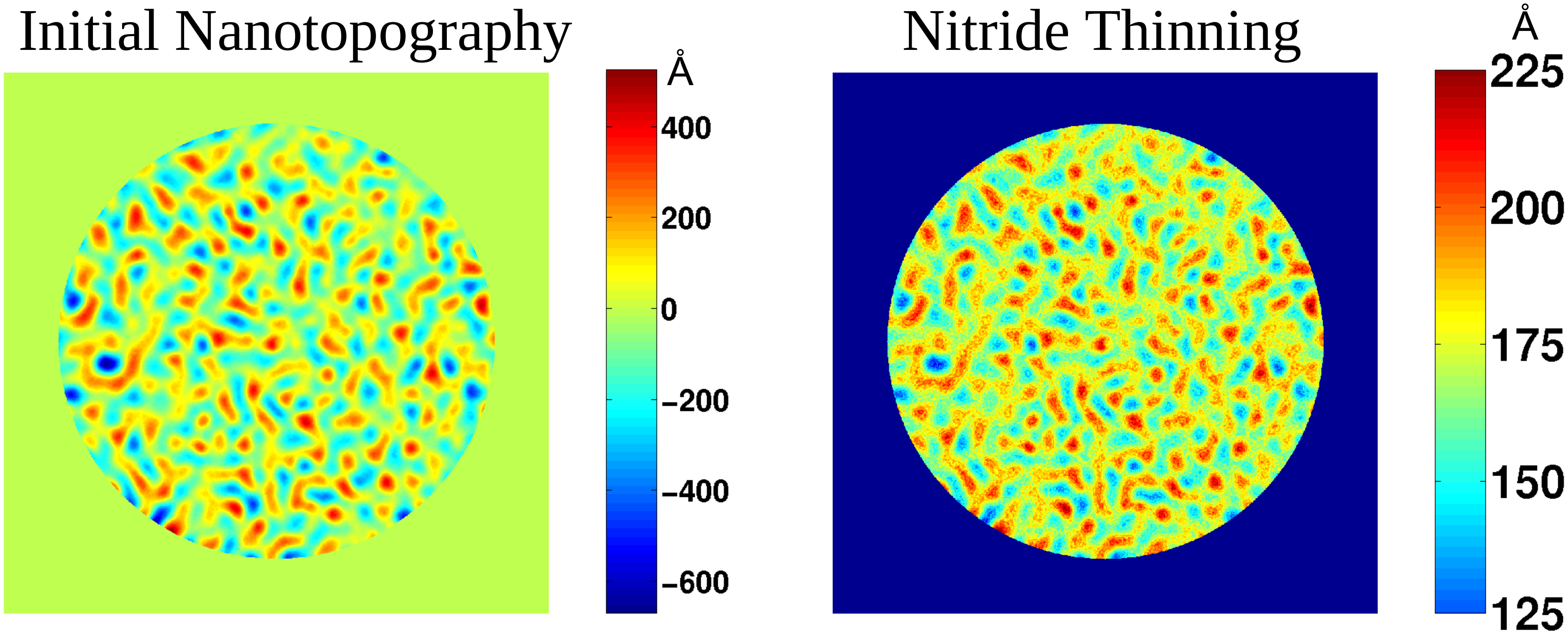


Initial Nanotopography

Total Removed

Zoom on Total Removed

Thinning of Nitride Layer (Under Oxide)

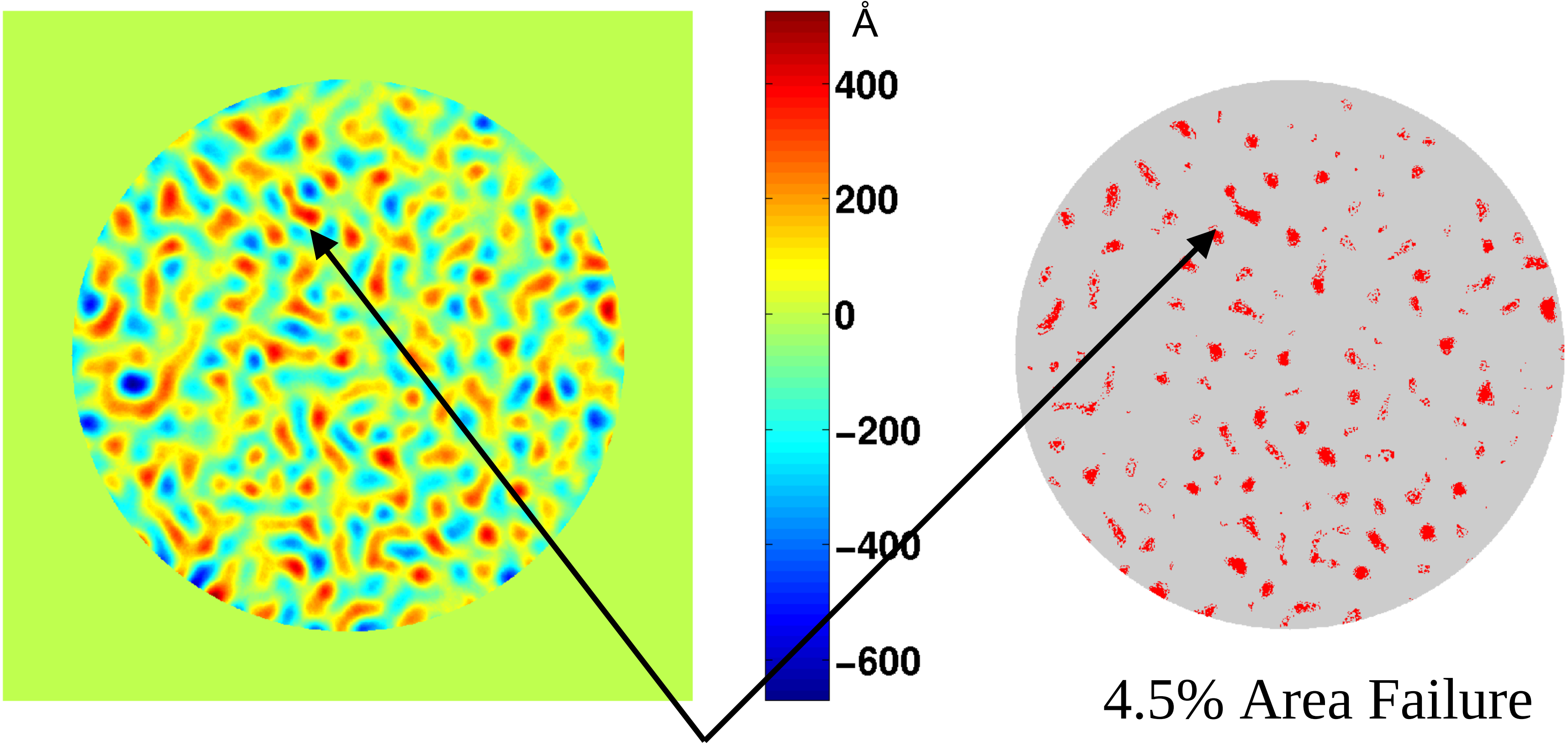


E= 147 MPa

Nitride Thinning – Device Failure Map

Initial Nanotopography

Potential Failure Locations



4.5% Area Failure

Note visual correlation with nanotopography "high spots"

E= 147 MPa

Red indicates excessive nitride thinning – greater than 20 nm budget (20% of 100 nm)

Conclusions

- Nanotopography interacts with CMP to cause localized thinning of surface films
- Modeling approaches
  - Scaling: does not capture localized spatial interactions
  - Linear filter
  - Contact wear: good results; flexible application
- STI yield concerns can be predicted and yield risk maps produced from nanotopography maps using contact wear simulations
  - Problem #1: failure to clear oxide
  - Problem #2: excessive nitride thinning