Challenges and Opportunities in CMP Consumables – A Chemist's View

Yuzhuo Li
Center for Advanced Materials Processing
Department of Chemistry
Clarkson University
Potsdam, New York 13699
yuzhuoli@clarkson.edu

Some Interesting Questions

- At which technology node CMP will exhaust its usefulness?
- At which technology node copper will reach its limit as interconnect?
- What materials could/would be used to replace copper? Still need CMP?
- What practical considerations one must take for CMP consumables in 32 nm/450 mm processing?

Functions of CMP

- Planarization
  - Dielectric (PMD and IMD) CMP, NiP CMP, etc
- Formation of micro/nano structures
  - Cu CMP, W CMP, STI CMP, MEMS CMP
- Surface conditioning
  - NiP, Sapphire, MgO, etc

Which function(s) of the CMP is most sensitive to technology node?

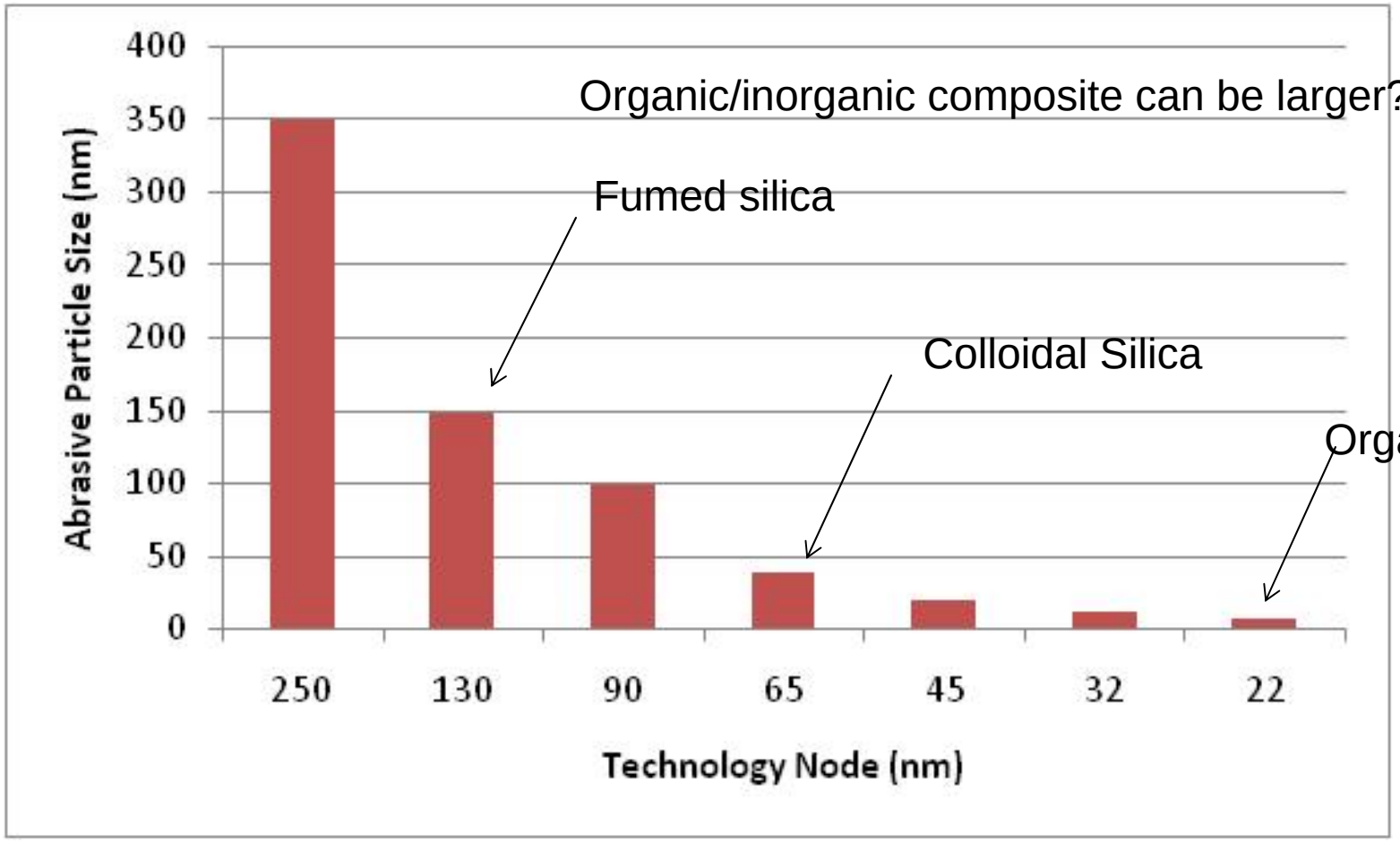
CMP Consumables

- ***Slurry***
- ***Pad***
- pCMP clean solution
- Pad conditioner
- Retainer ring
- Carrier film
- Filters
- etc

Cu CMP Slurry Formulation Strategy - in what order?

- Oxidizer
  - Peroxides, persulfates, periodates, etc
- pH
  - Consider stabilizer stability and Pourbaix diagram
- Complexing agent
  - Assist copper dissolution
- Passivating agent
  - Suppress isotropic copper dissolution
- ***"Abrasive" particles***
  - ***Softer the better? What are the roles of particles?***

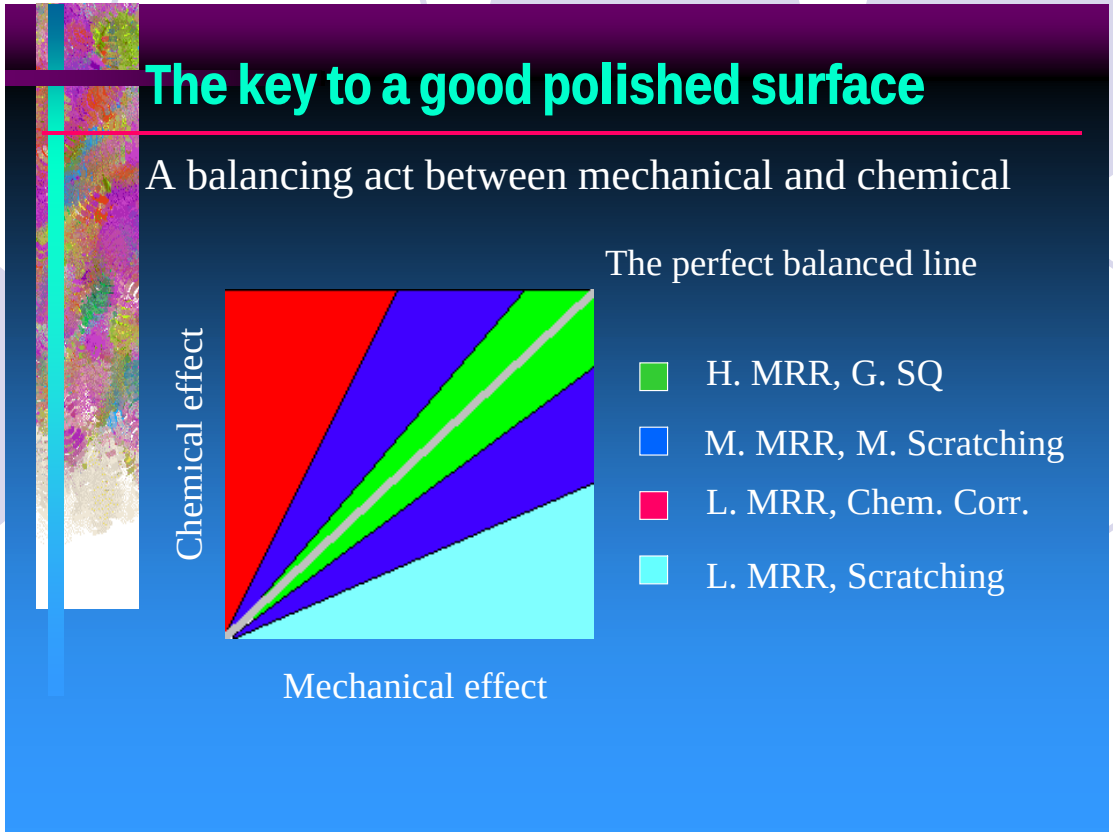
Technology Node and PS



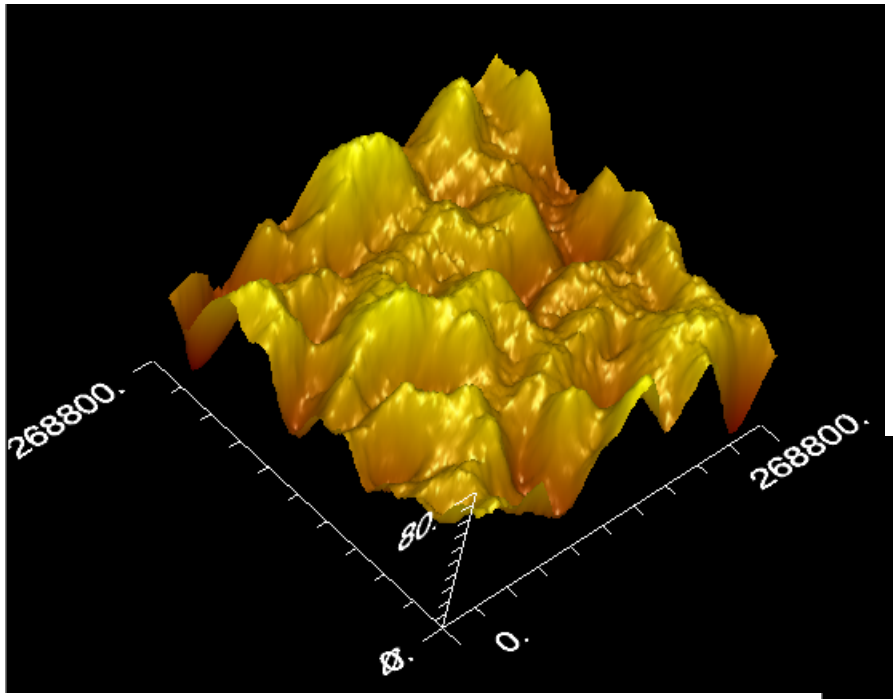
Removal rate
Step height reduction efficiency

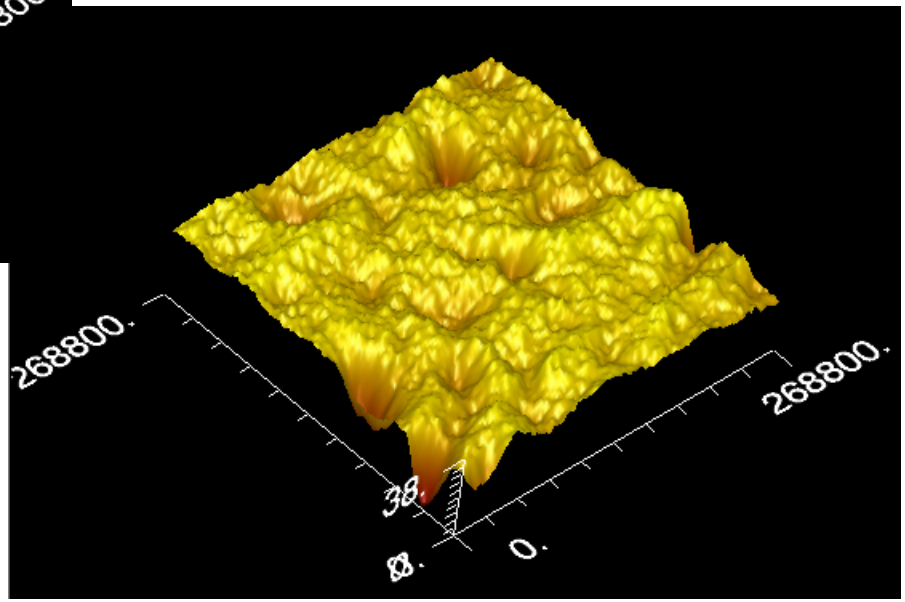
Abrasive Particle Size

Surface Quality
Defect count



Strong Chemical/Weak Mechanical





Weak Chemical/Mild Mechanical

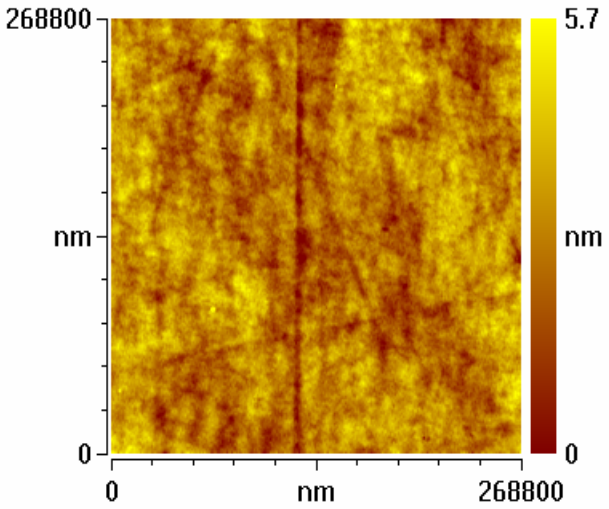
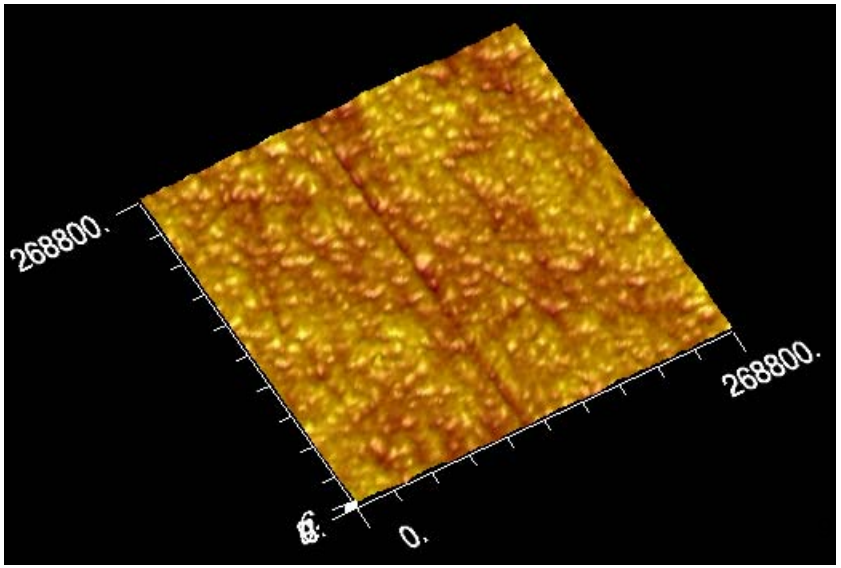


	Image
Mean	-0.00
Sq	0.95
Sa	0.76
Peak/Valley	12.41
Skewness	-0.15
Kurtosis	3.02



When there is a balance

Representative results from past studies

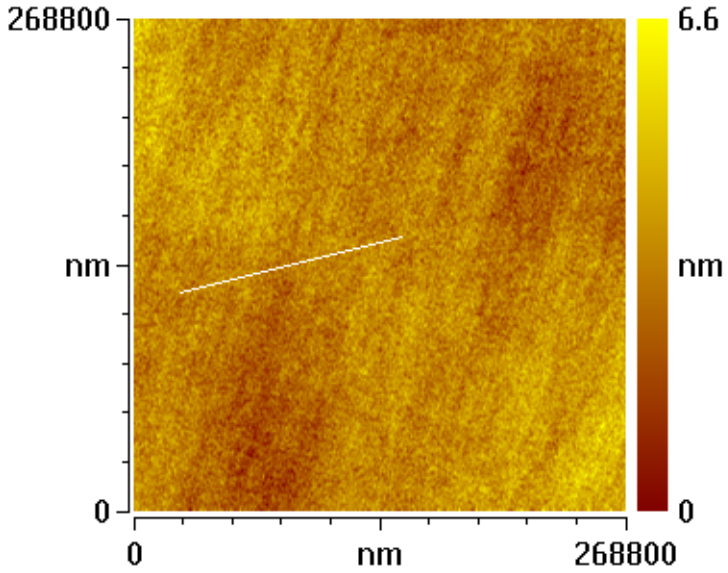
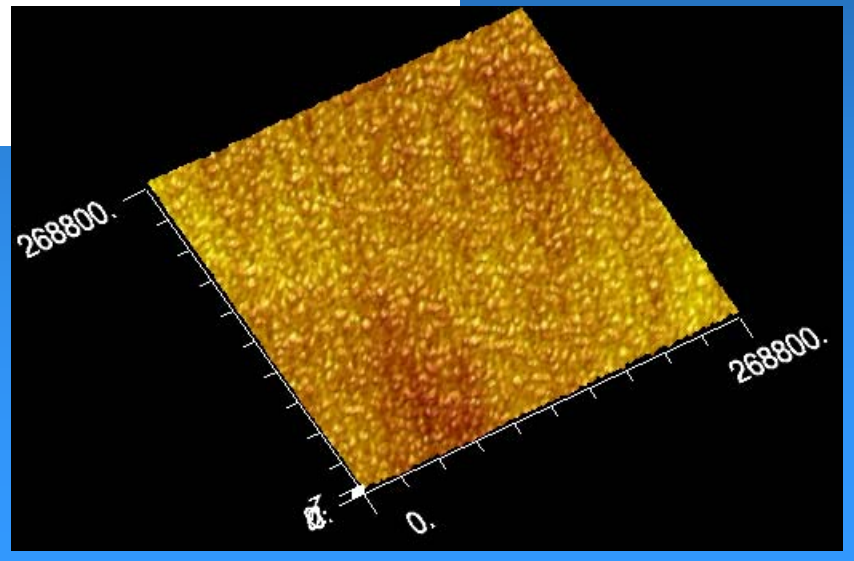


	Image
Mean	0.00
Sq	0.78
Sa	0.62
Peak/Valley	6.59
Skewness	-0.07
Kurtosis	3.15



Importance of Abrasive Particles

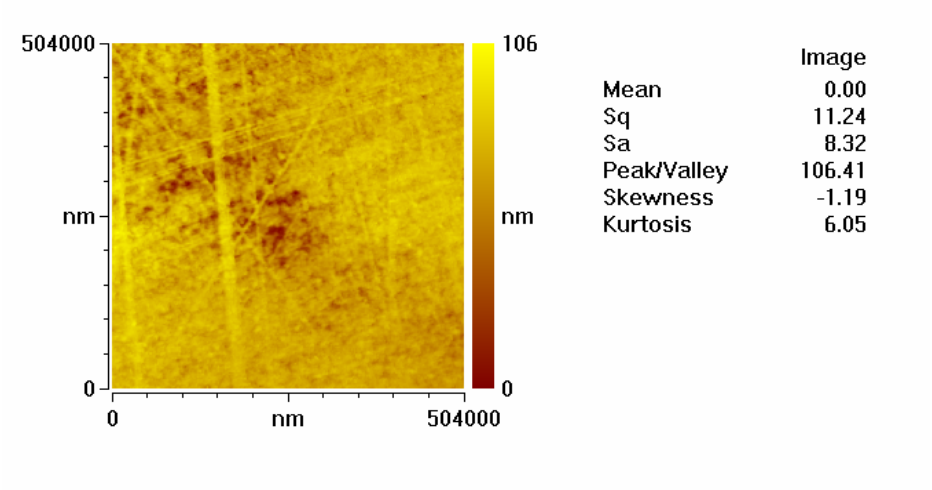


	Image
Mean	0.00
Sq	11.24
Sa	8.32
Peak/Valley	106.41
Skewness	-1.19
Kurtosis	6.05

No abrasive particles

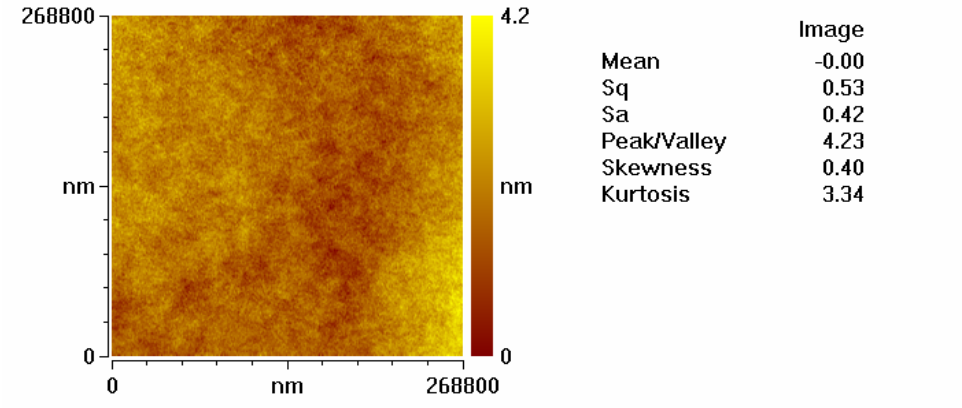


	Image
Mean	-0.00
Sq	0.53
Sa	0.42
Peak/Valley	4.23
Skewness	0.40
Kurtosis	3.34

3% 80 nm silica

Optimal Particle Size?

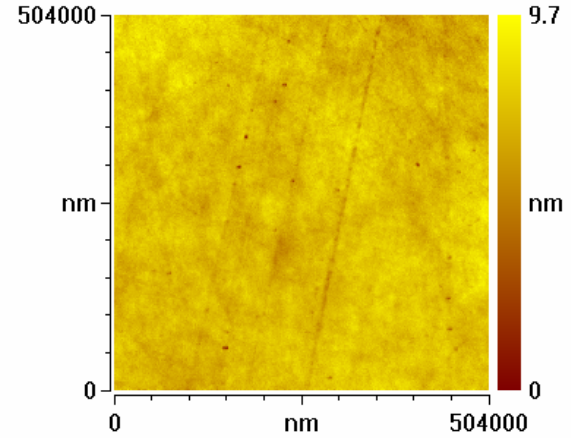


	Image
Mean	0.00
Sq	0.63
Sa	0.49
Peak/Valley	9.66
Skewness	-0.21
Kurtosis	4.38

3% 20 nm silica particles

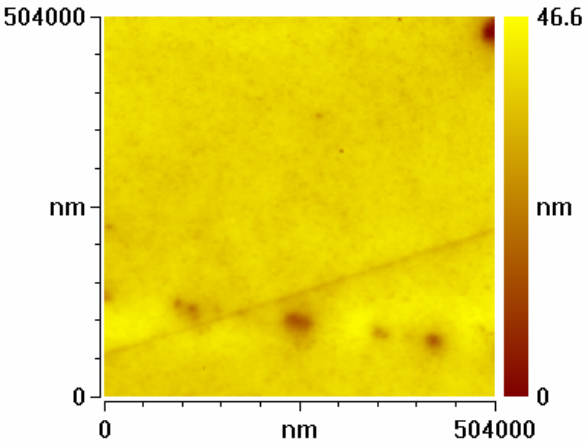


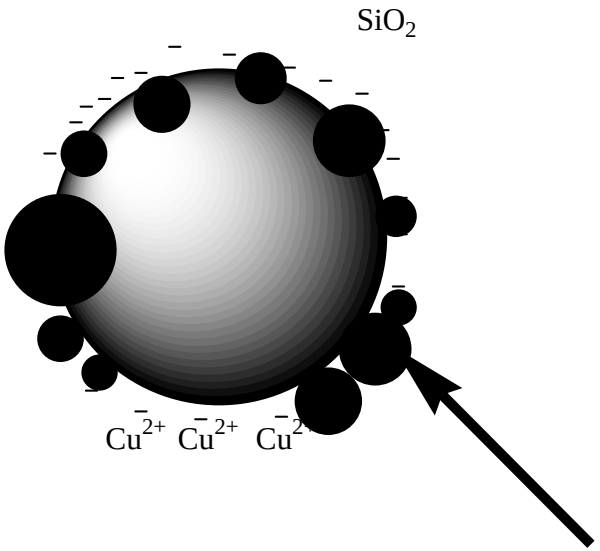
	Image
Mean	0.00
Sq	2.79
Sa	1.62
Peak/Valley	46.63
Skewness	-4.05
Kurtosis	39.32

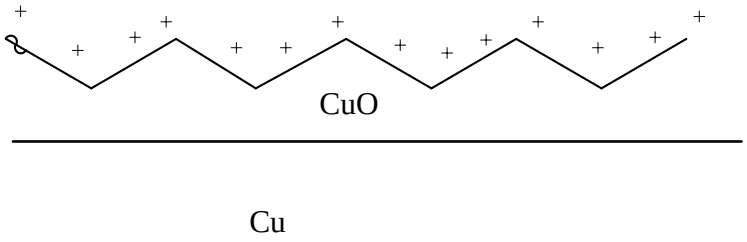
3% 130 nm silica particles

Functions of an abrasive

- To enhance the mechanical effect of a pad
  - Minimum requirement in hardness?
  - What about abrasive-free or abrasive-diet systems?
- To serve as a carrier
  - Chemically interact with pad, surface to be polished, and all slurry components
  - Physically remove the polishing debris away from the surface
- Serve as a particulate lubricant
  - Well-known in classic tribology
  - Special effects when the particles are in nanosize

Particulate adsorption



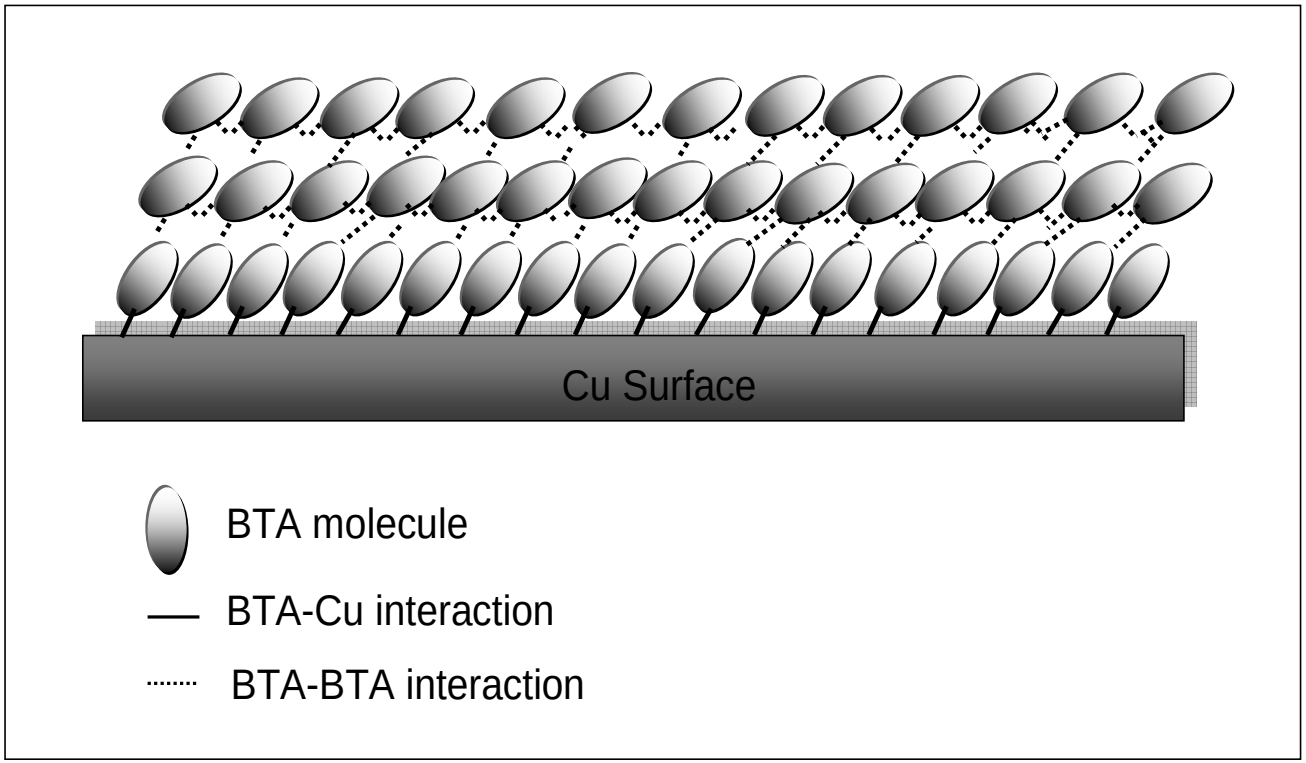


Notice the relative size ratio:

Abrasive particle/polishing debris

Passivation Film Formation

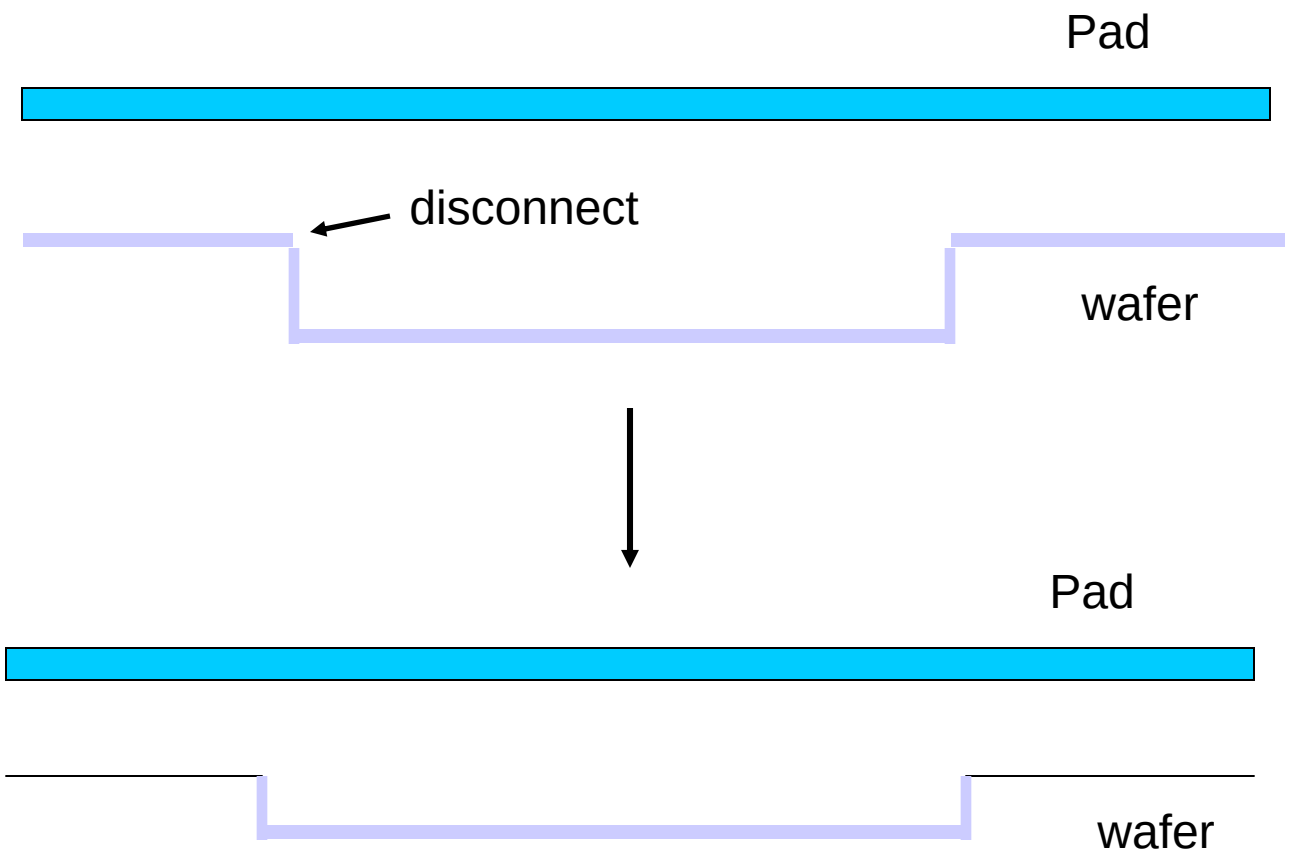
Where other things fit in here (complexing agent, surfactant, etc)



Interesting question

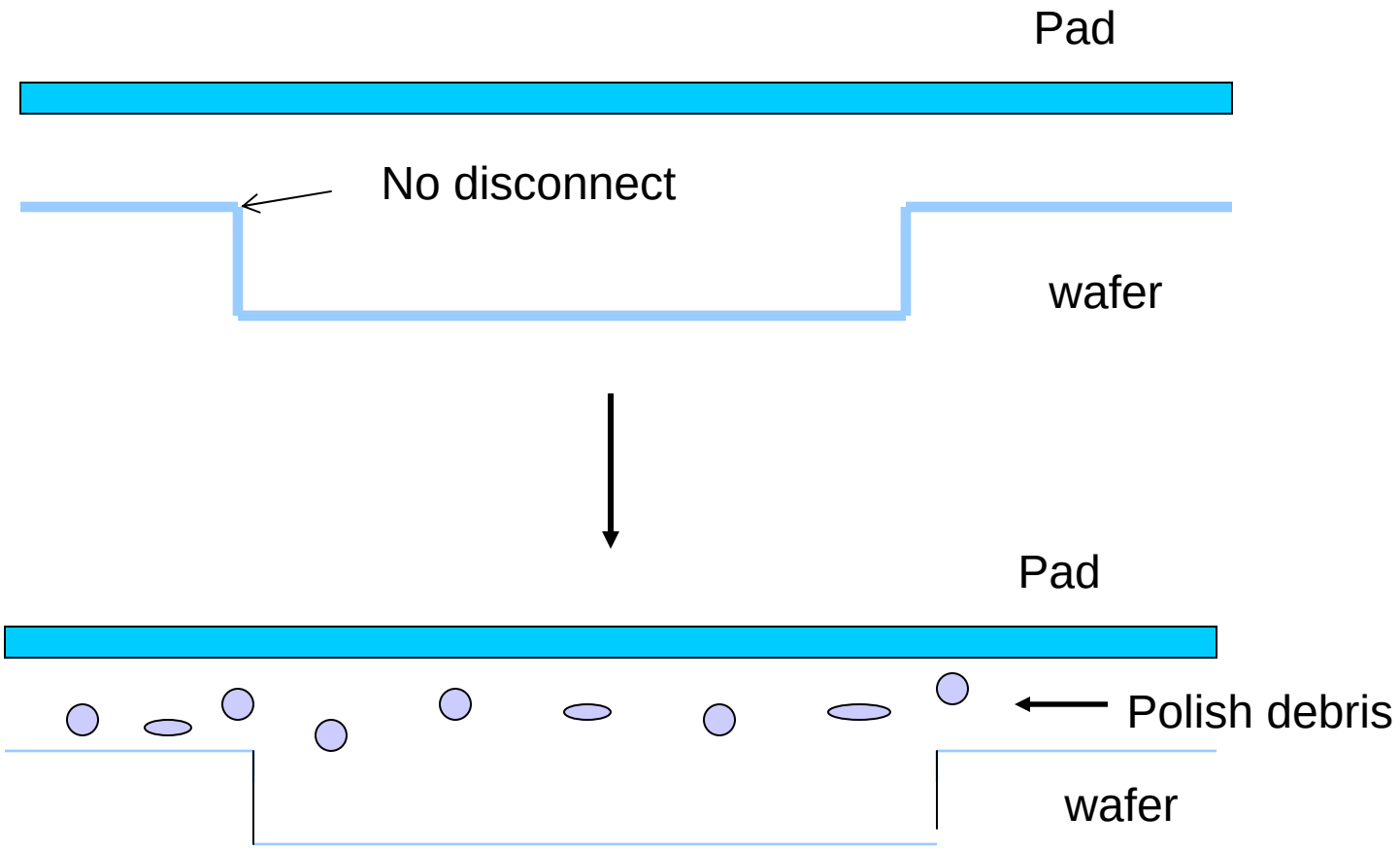
- Will higher MRR/SER ratio always translate to better SHRE?
  - Is it true that tougher the passivation film the better the SHRE?
- Case 1: non-BTA based passivating film, zero static etch rate, > 5000A/min MRR, no step height reduction efficiency
- Case 2: surfactant based passivating film, very low static etch rate (<50A/min), high removal rate (>5000A/min), SHRE < 30%

Kaufman Model



Step height reduced
Trench width kept the same

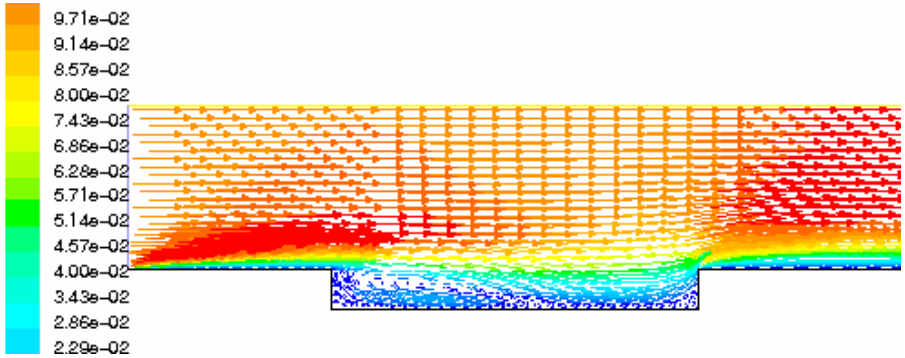
Delamination model



No step height reduction or
Increased step height
“Trench” width slightly increased

Slurry with Higher Viscosity Gives Lower SHRE

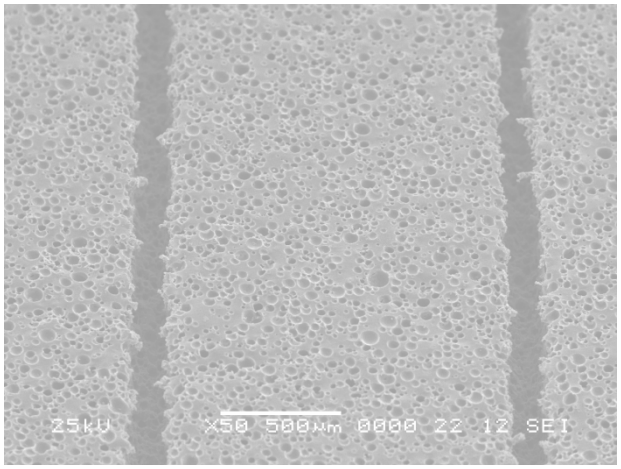
slurry	Relative viscosity	SHRE (%)	Dishing at 100/100 um lines (A)
original	1.00	89	600
Original plus IPA	1.35	65	1200

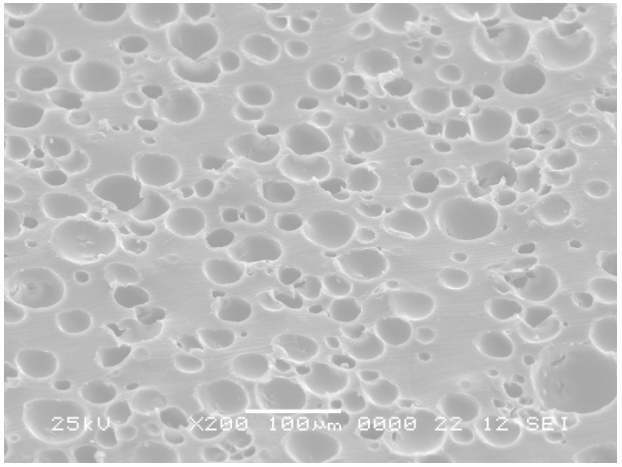


Simulated linear velocity difference between the fluid in a recessed area and near the pad. The reduced flow, to certain extent, helps the preservation of the passivating film in the recessed area.

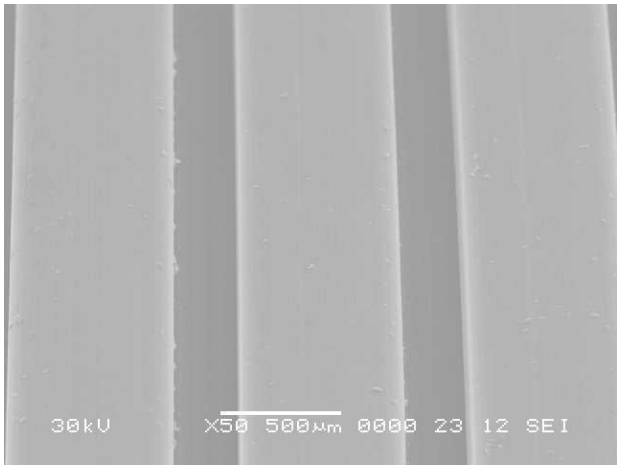
Porous vs. Non-Cell Pads

Porous



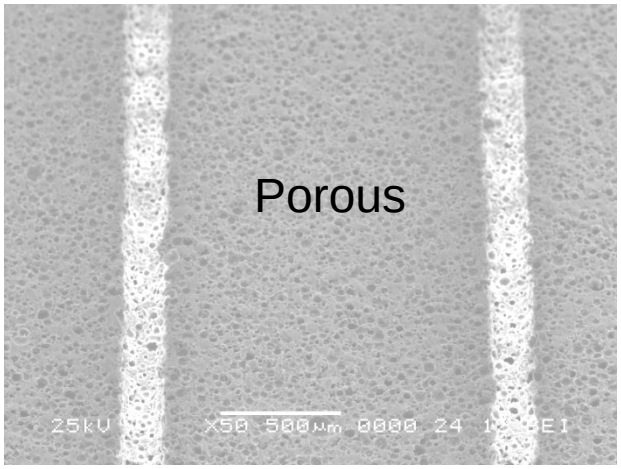


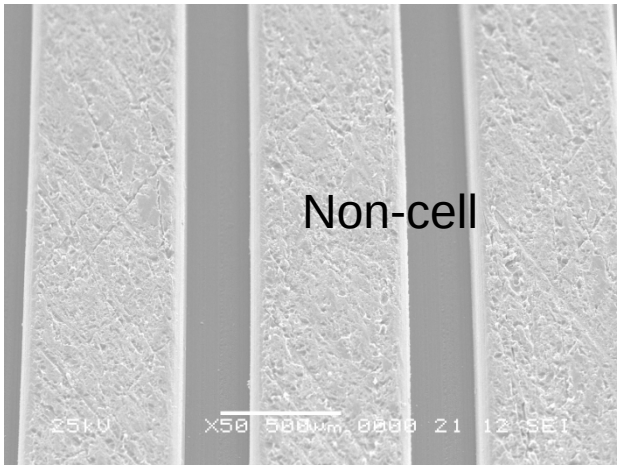
Non-cell

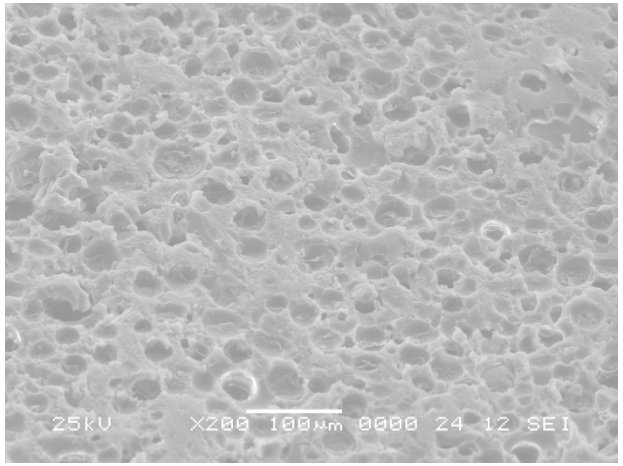


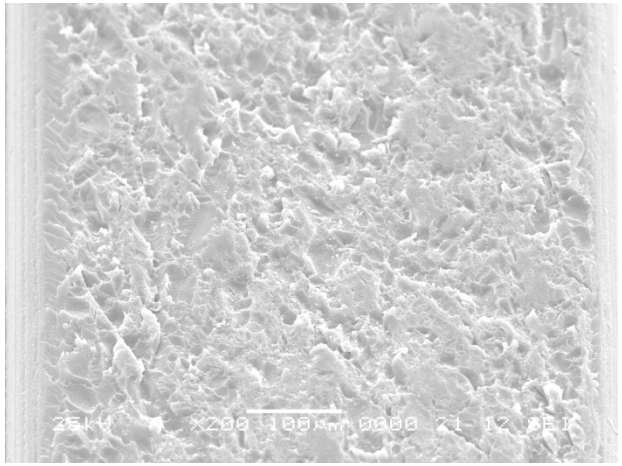
Porous vs. Non-cell Pads

After breaking-in and conditioning

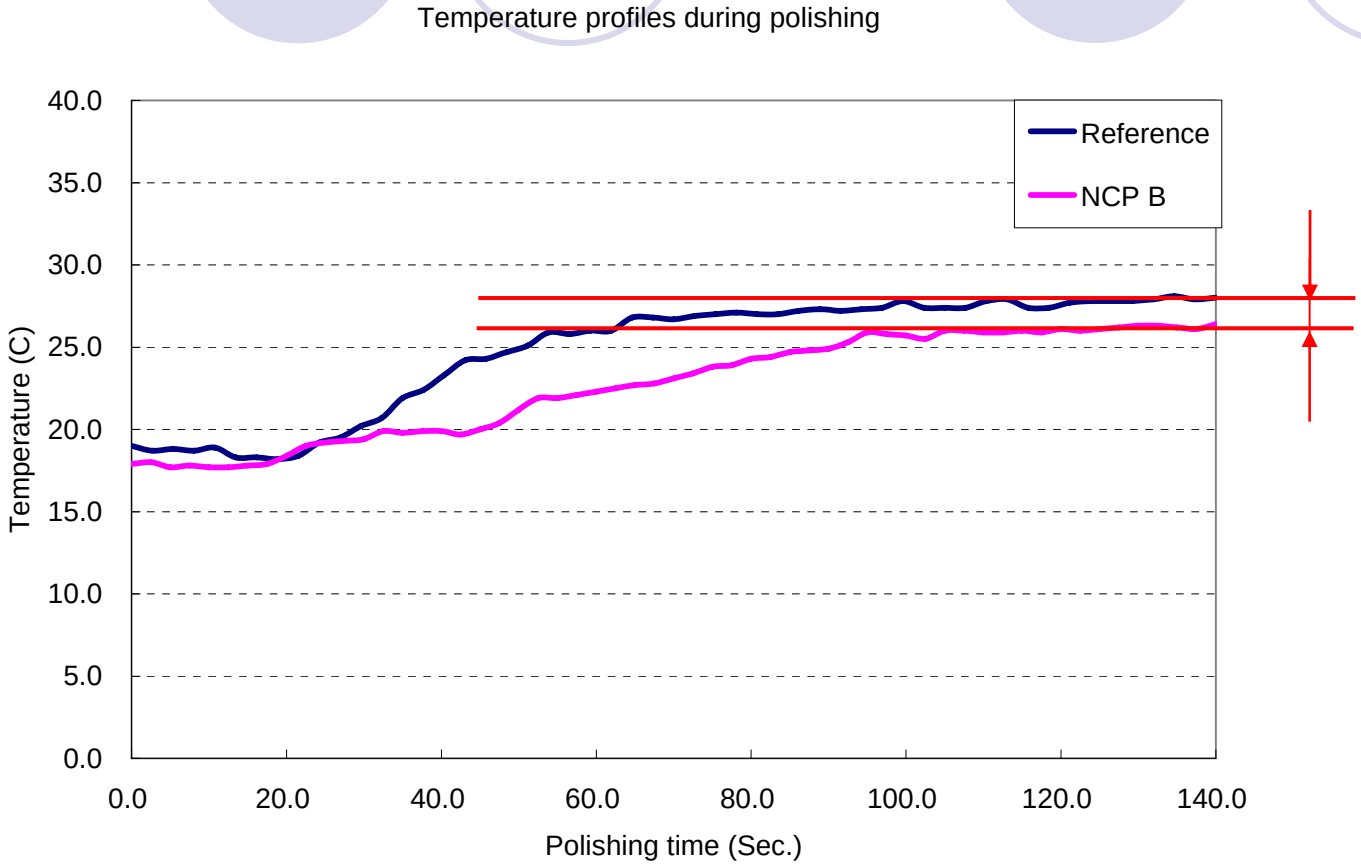








Pad Surface Temperature

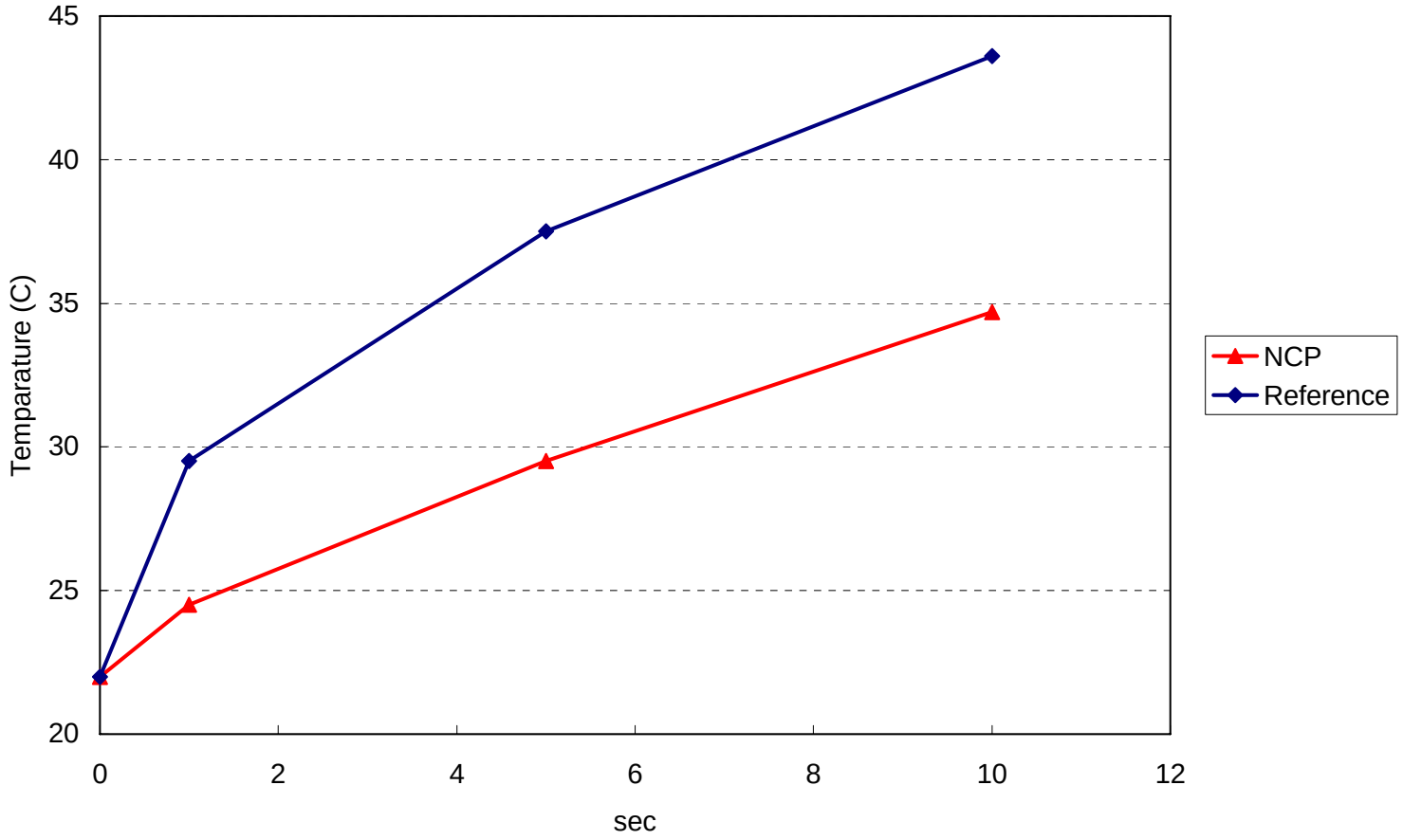


NCP pad temperature is lower during the polishing.

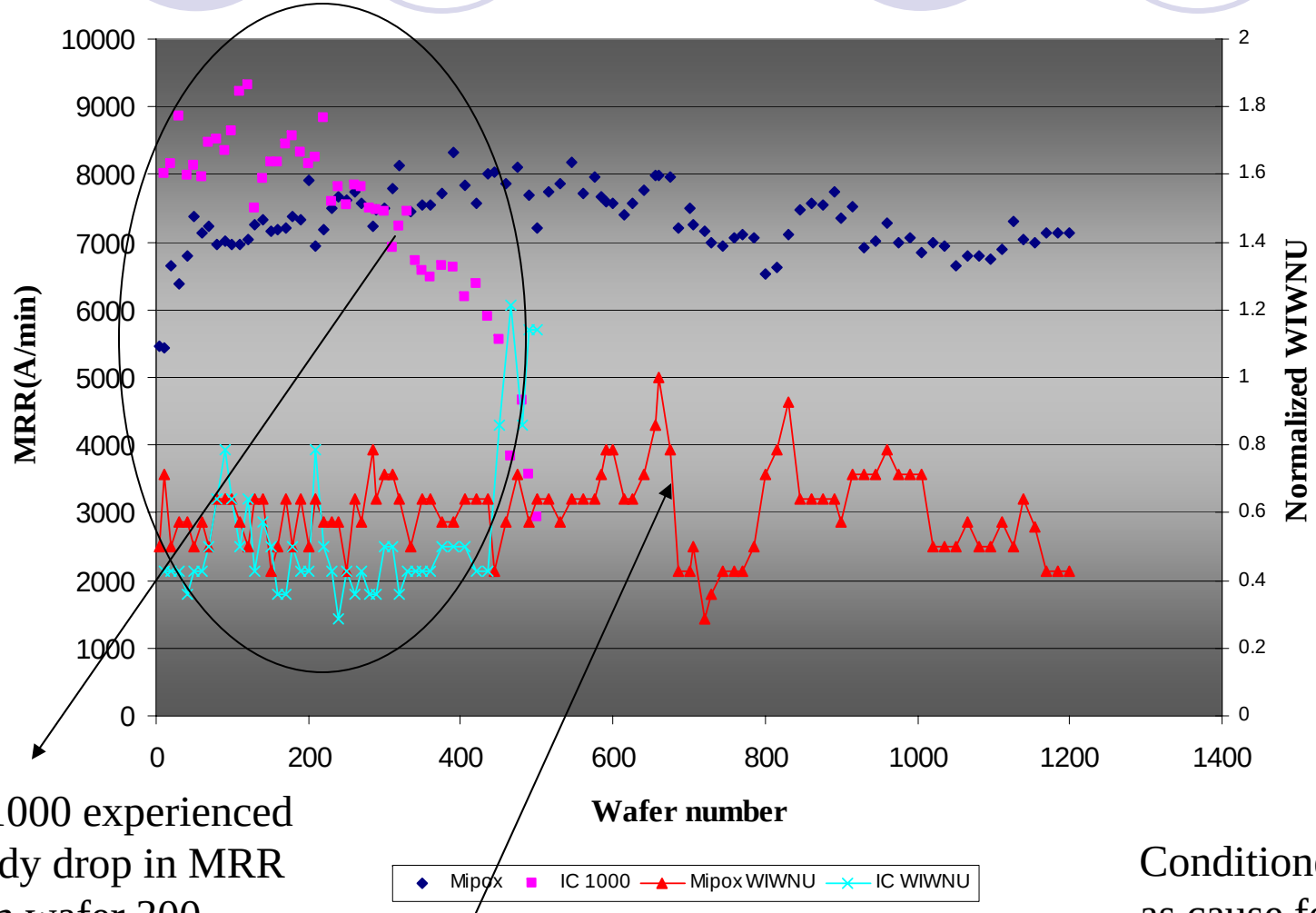
Lower local temperature

Slower to develop glazing effect?

Due to Greater Heat Capacity?



Results and Discussion
Blanket wafer data as a function of pad life



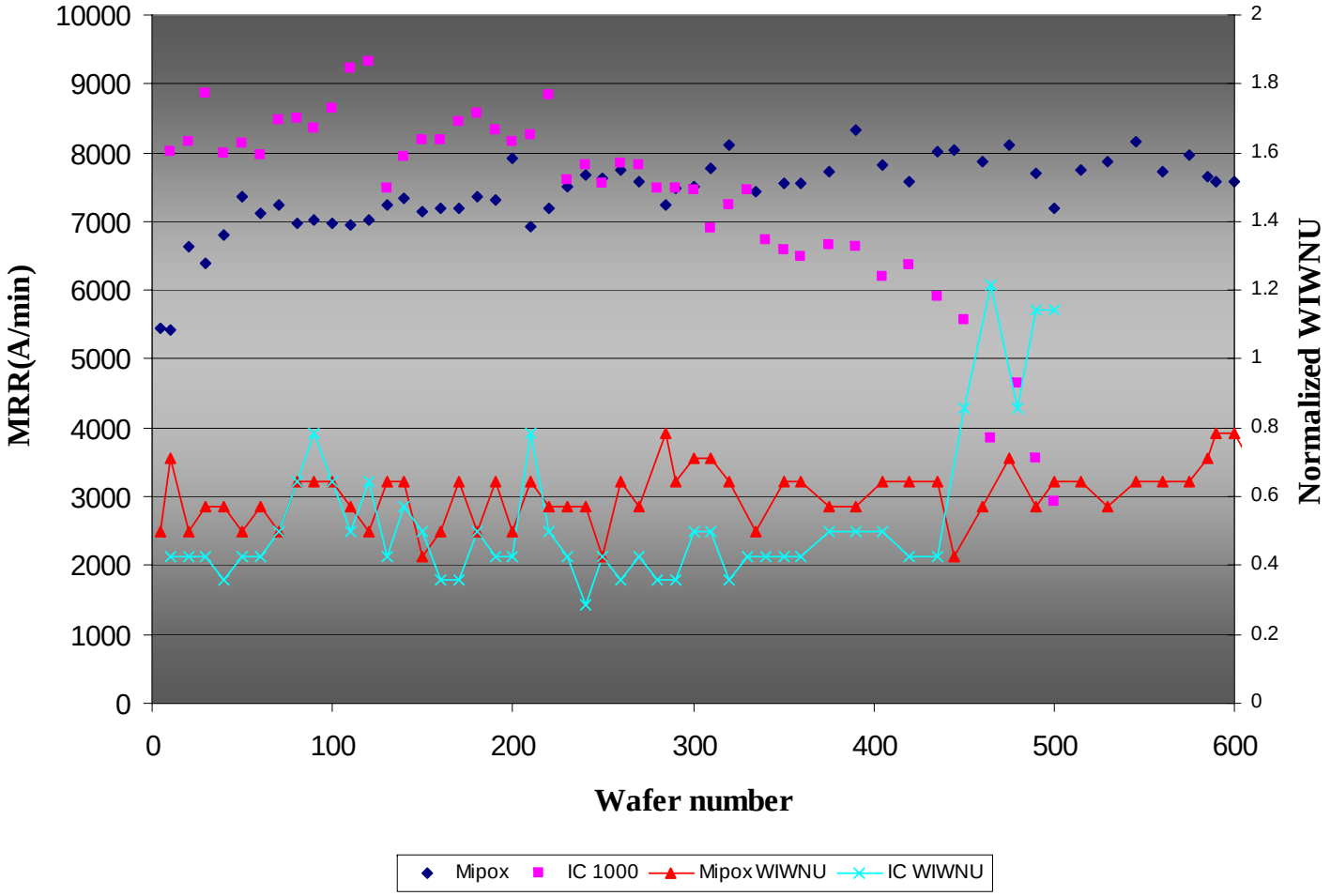
Mipox pad performance stable even after 1200 wafers

IC 1000 experienced steady drop in MRR from wafer 300

Carrier film changed

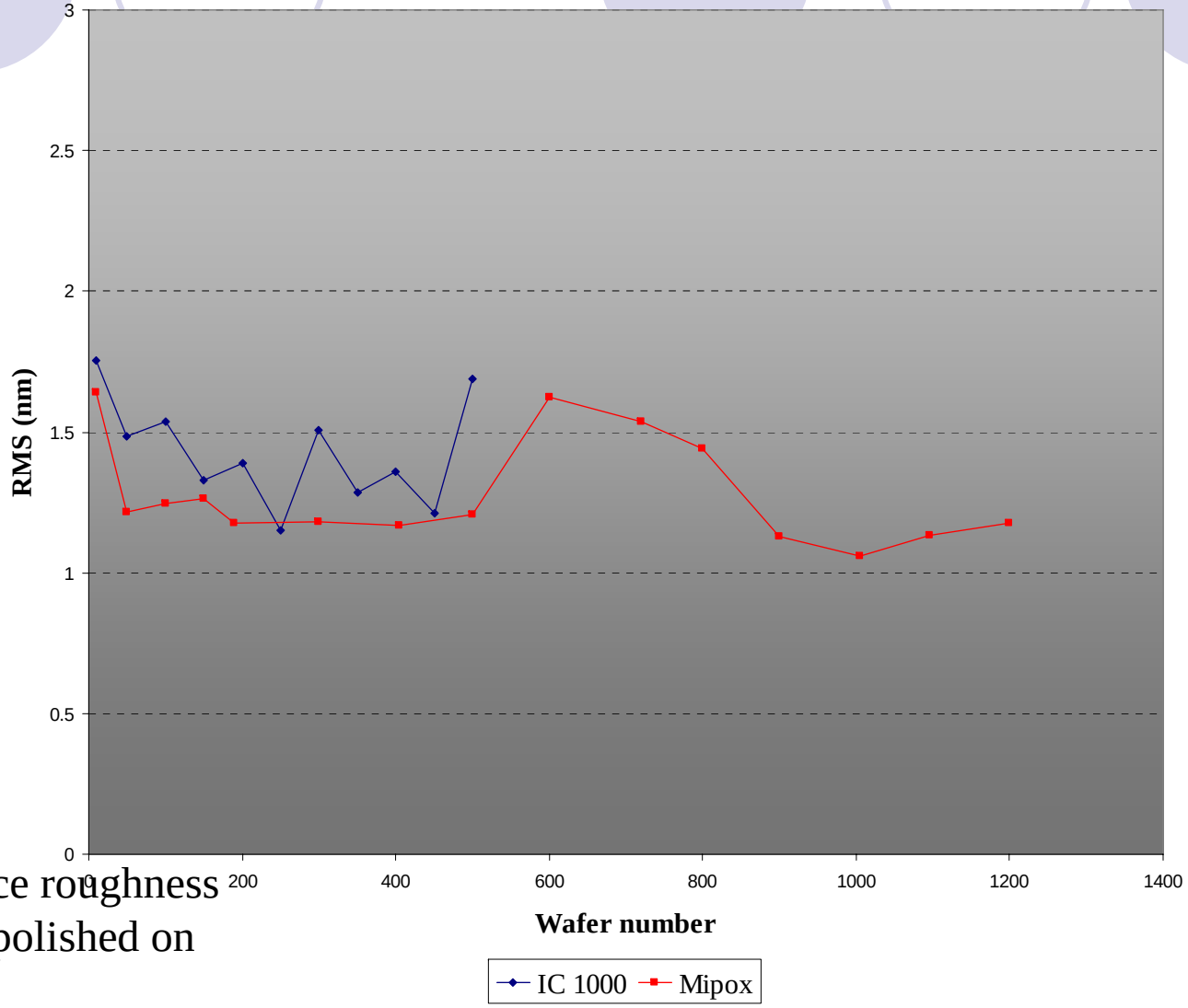
Conditioner ruled out as cause for drop in removal rate

Blanket wafer data as a function of pad life (contd.)



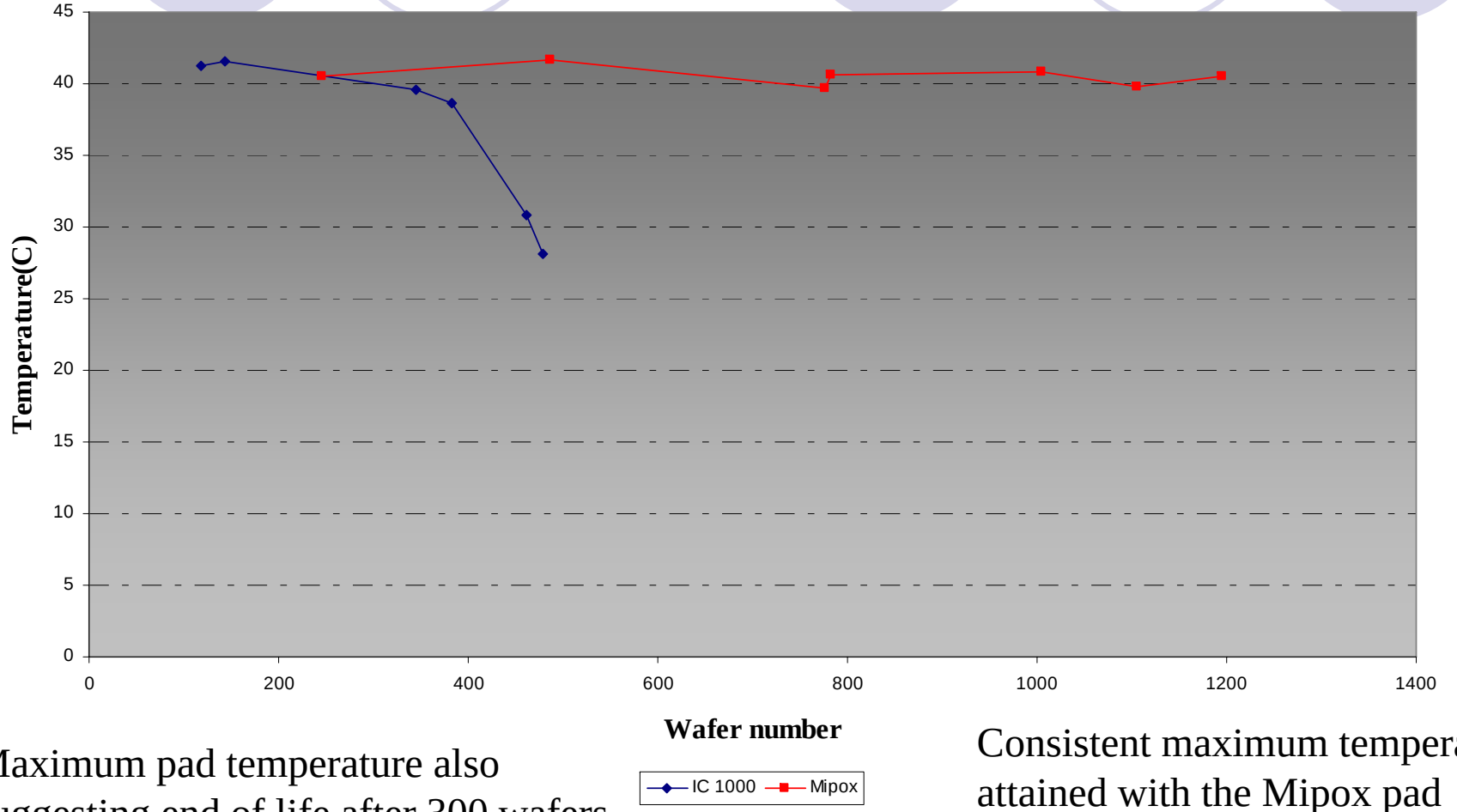
MRR starts to decrease after 300 wafers in case of IC-1000

Surface quality of polished wafers



Comparable surface roughness values for wafers polished on both the pads

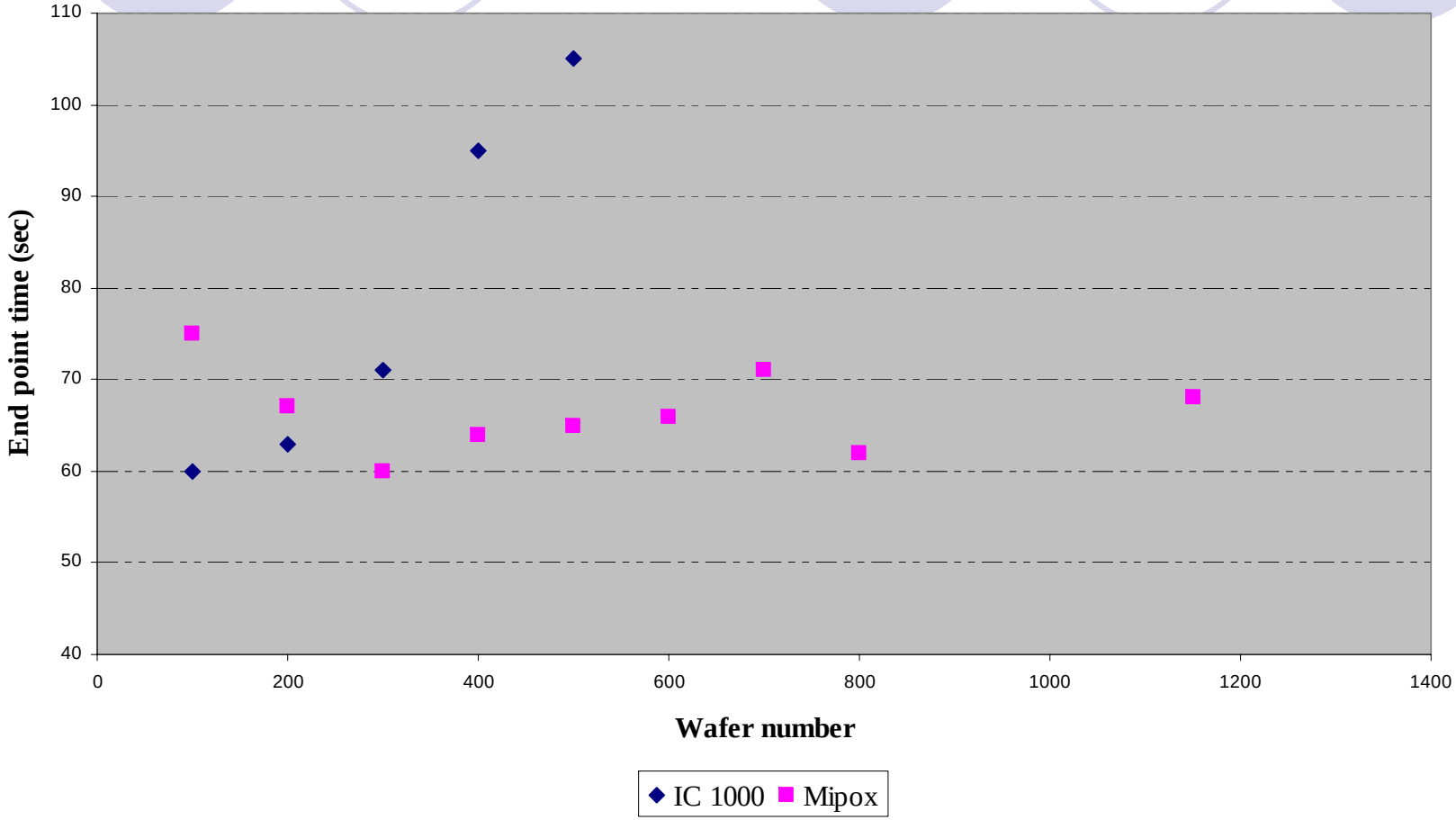
Pad temperature measurement



Maximum pad temperature also suggesting end of life after 300 wafers for the IC 1000 pad

Consistent maximum temperatures attained with the Mipox pad

End-point times as a function of wafer number

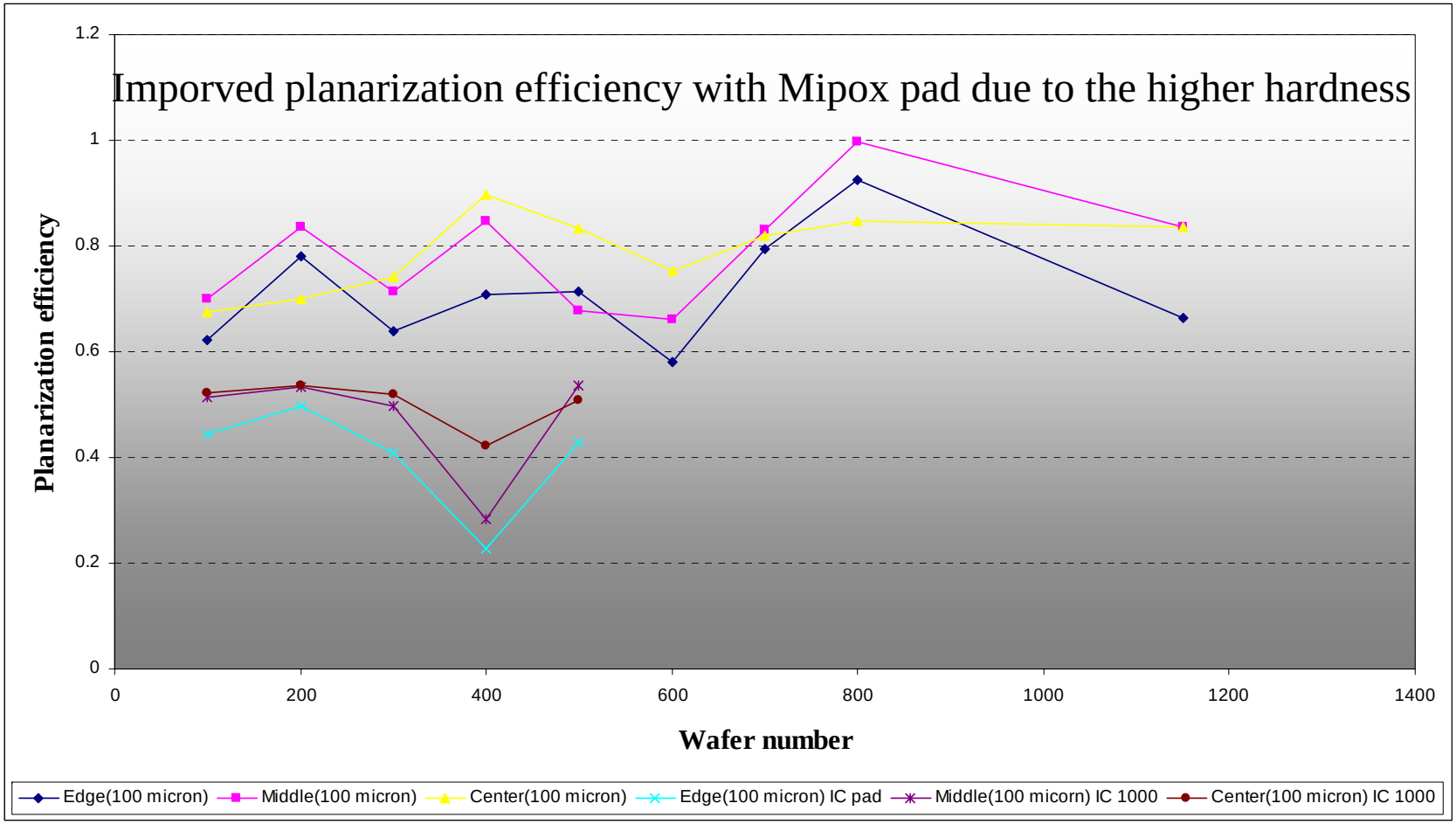


Consistent end-point times with the Mipox pad

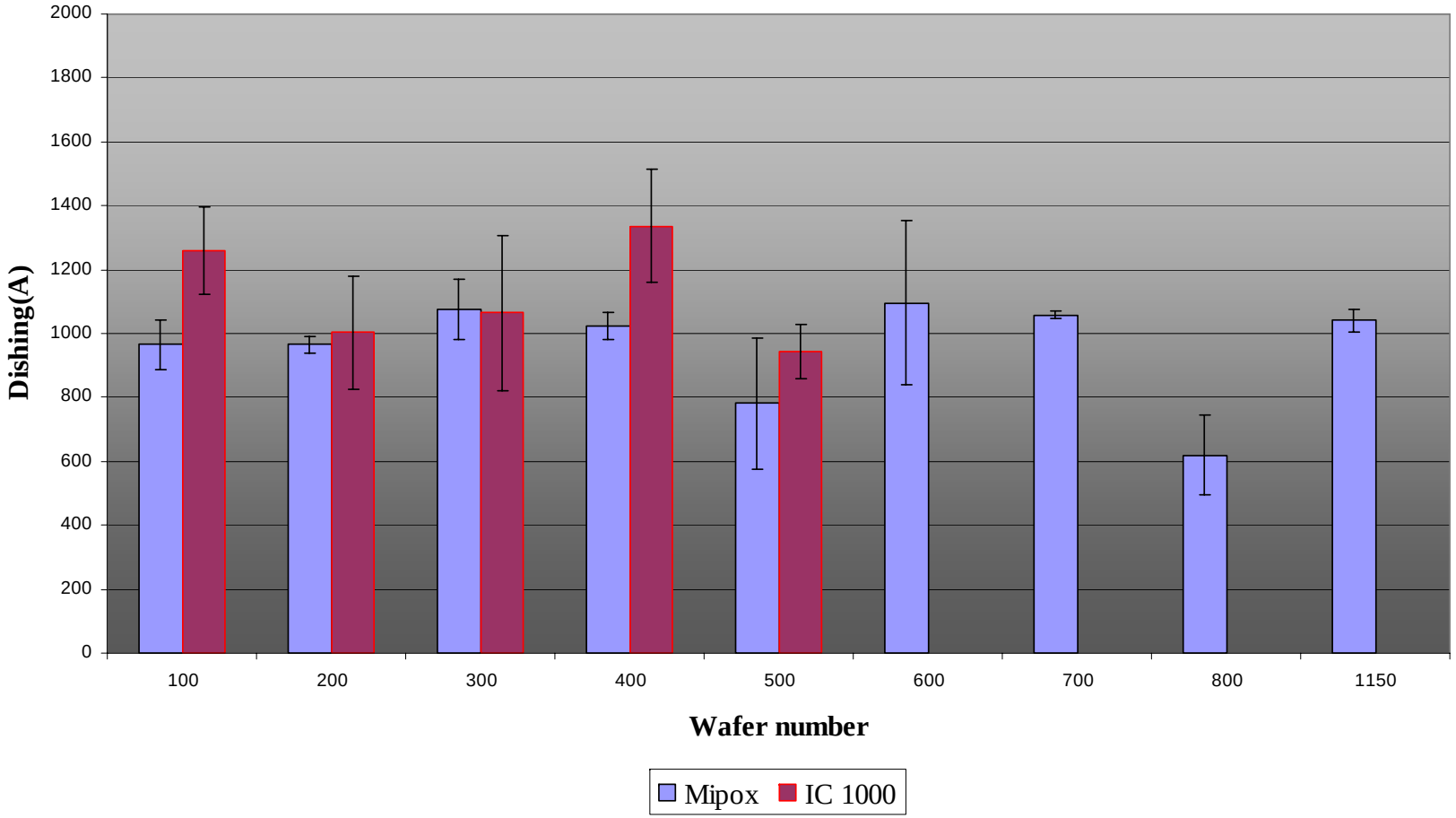
Steady increase in the end-point times with the IC-1000 pad after wafer 300

Patterned wafer results

Planarization efficiency as a function of wafer number



Dishing as a function of wafer number



Lesser dishing in most of the runs with the Mipox pad due to the higher stiffness in comparison to the IC-1000 pad

SURFACE IMAGE

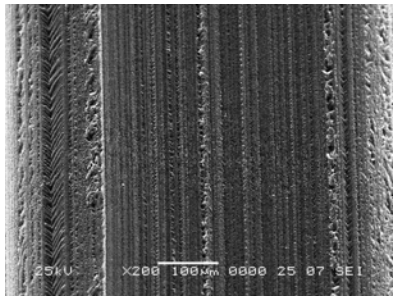
Reference: New unconditioned pad

Magnification : X200

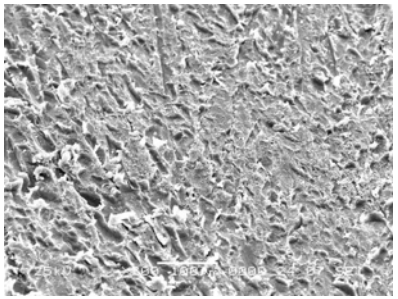
Angle : 45degrees

After 1200 wafers

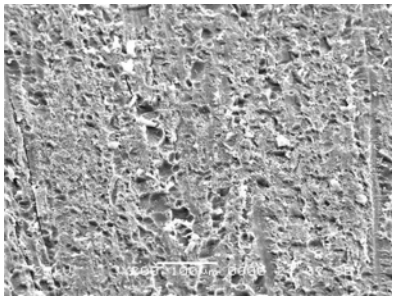
Mipox pad



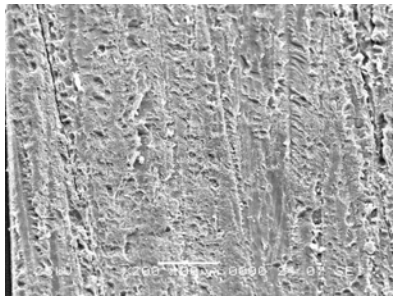
Reference



ID



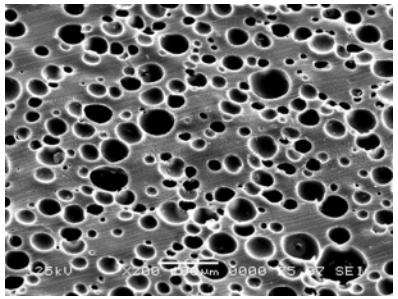
MD



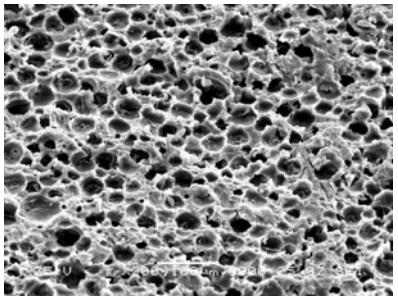
OD

IC 1000 pad

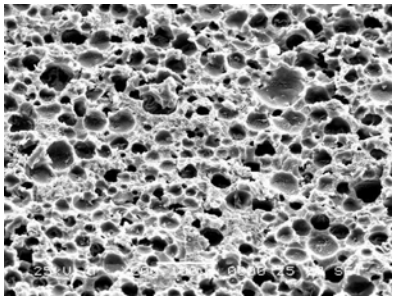
After 500 wafers



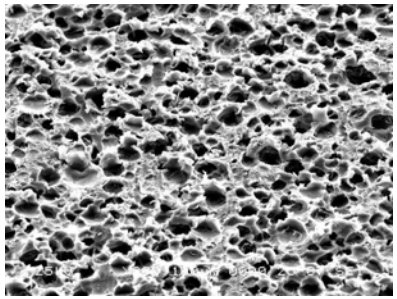
Reference



ID



MD



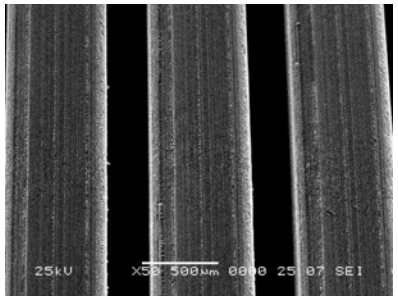
OD

SURFACE IMAGE

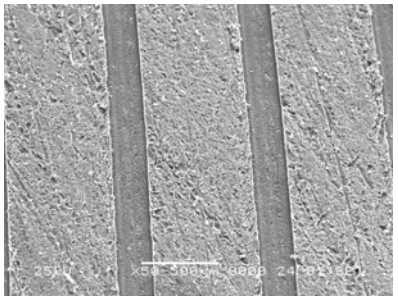
Magnification : X50

Angle : 45degrees

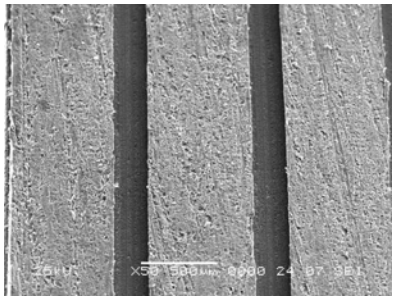
Mipox pad



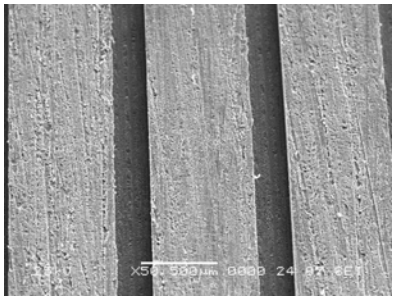
Reference



ID

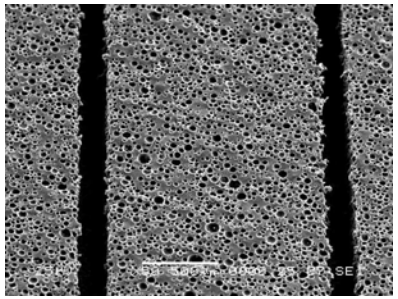


MD

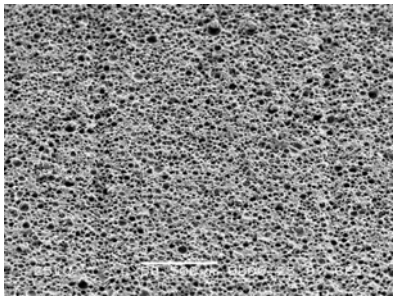


OD

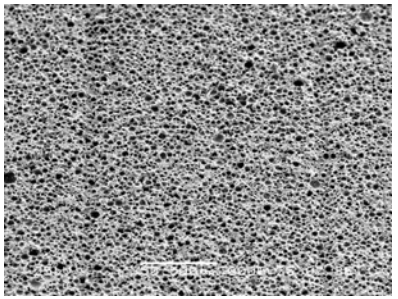
IC 1000 pad



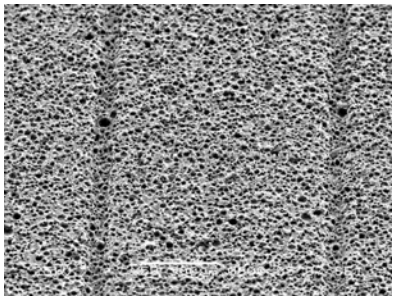
Reference



ID

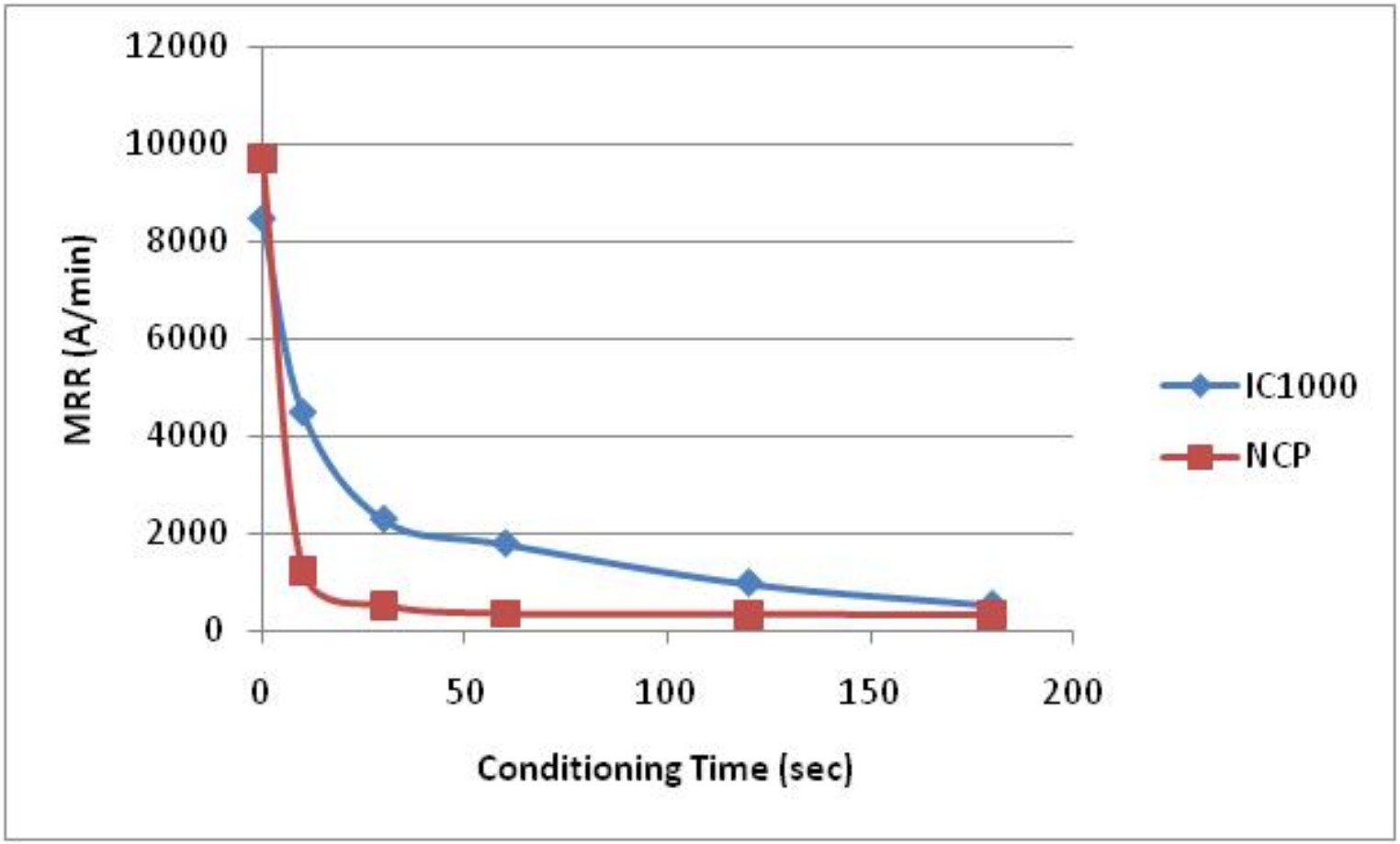


MD



OD

Lower Retention of Slurry Chemistry on NCP Pad



Potential advantages

- Higher hardness and modulus values
  - Improved pad-life
  - Higher planarization efficiency and lower dishing
- Non-porous nature
  - Feasibility of process development at lower flow-rates due to the non-porous nature
  - Feasibility of rapidly changing slurry chemistry during Cu CMP process
- For more information, see Mipox poster

Summary

- For technology node sensitive CMP processes, abrasive particle size will continue to reduce
  - Smaller the particle size, the true abrasiveness is reduced.
  - Larger the surface area, the surface adsorption property becomes more important
  - The abrasiveness of these particles will be transmitted or express via the tips of the pads
  - Non-cell type pads may offer several potential advantages including longer life time