Comprehensive CD Uniformity Control in Lithography and Etch Process

Qiaolin Zhang[a], Cherry Tang[b],
Tony Hsieh[b], Nick Maccrae[b], Bhanwar Singh[b],
Kameshwar Poolla[a], Costas Spanos[a]

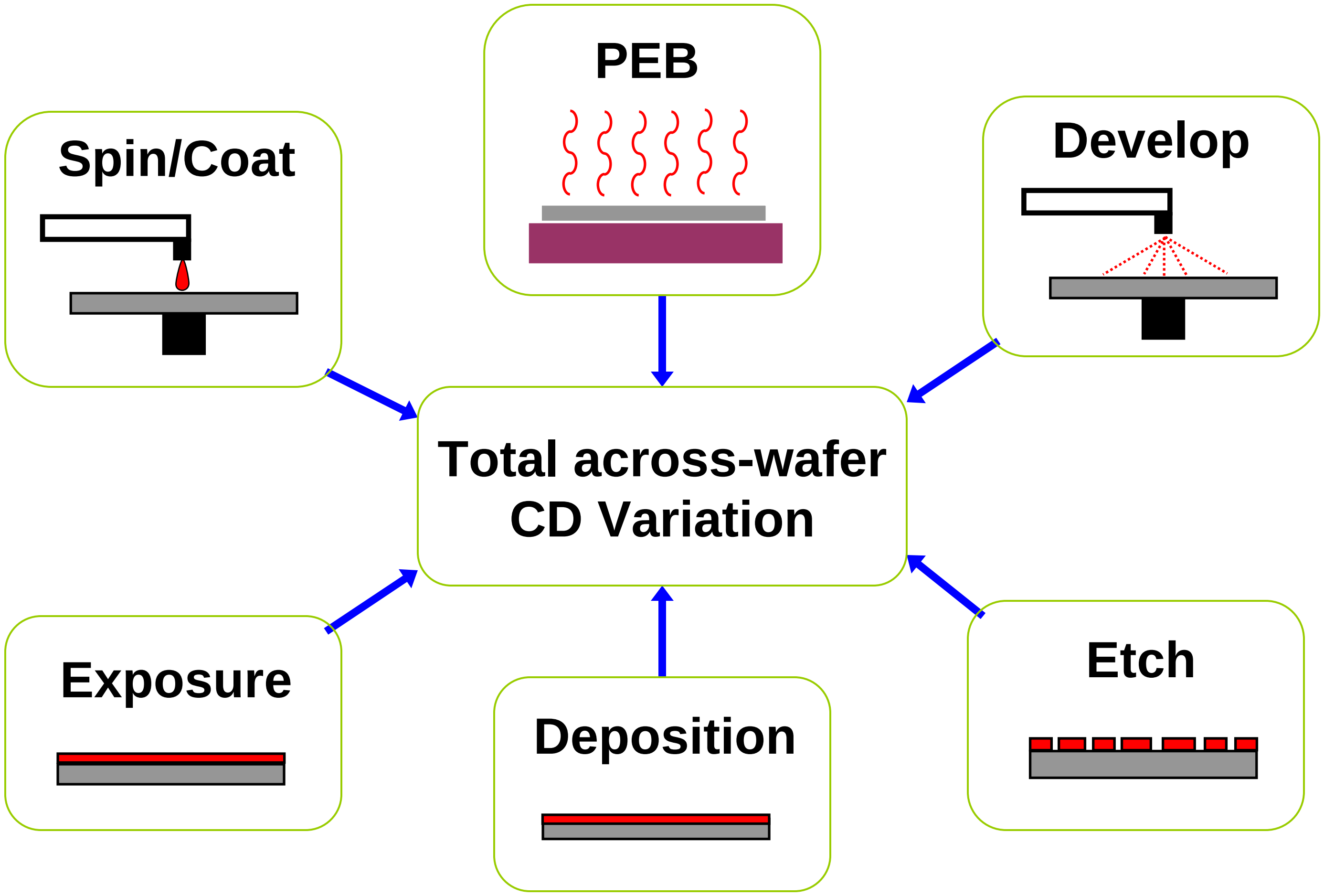
[a]Dept of EECS, UC Berkeley
[b]SDC, Advanced Micro Devices

March 3, 2005

Motivation

- Importance of across-wafer (AW) CD (gate-length) uniformity
  - Impacts IC performance spread and yield
  - Large AW CDV → large die-to-die performance variation → low yield
- How to cope with increasing AW CD variation?
  - Employ design tricks, ex. adaptive body biasing
    - Has limitations
  - Reduce AW CD variation during manufacturing
    - The most effective approach

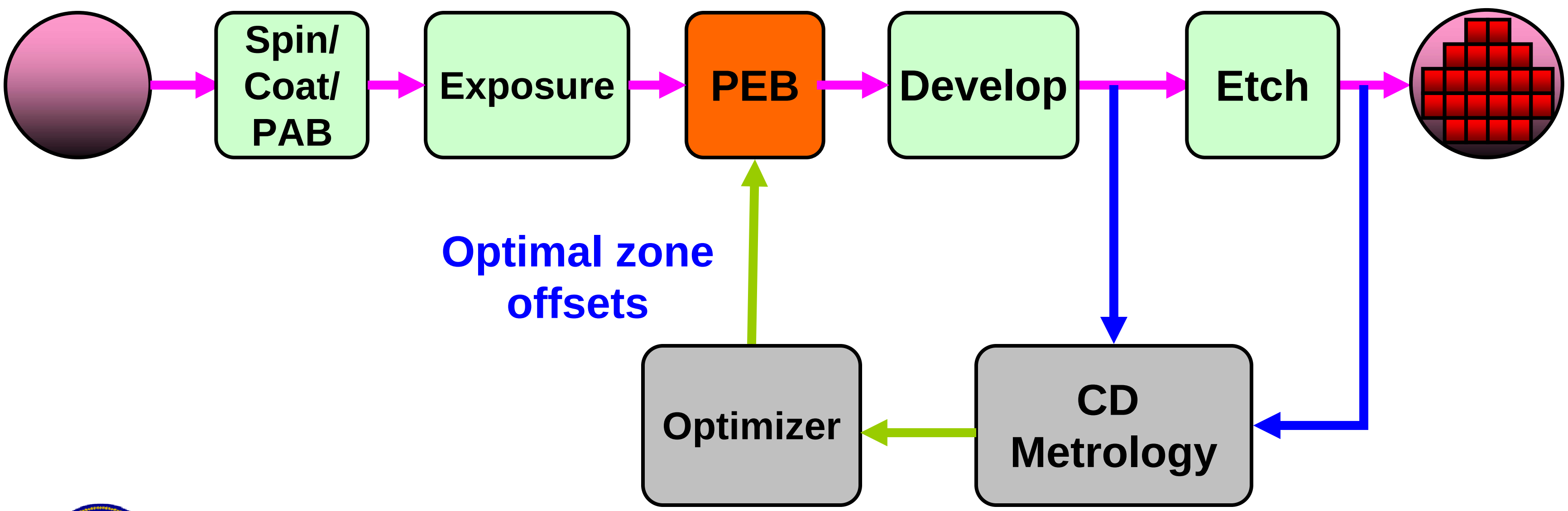
Across-wafer CD Variation Sources



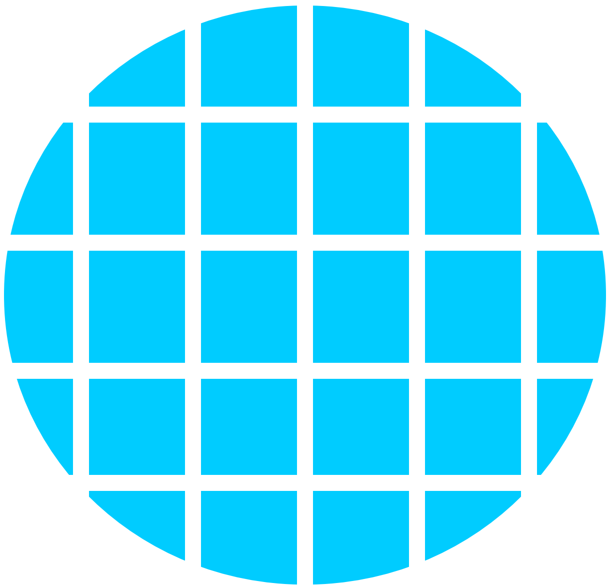
FLCC

CD Uniformity Control Approach

- Current litho clusters strive for uniform PEB profile of multi-zone bake plate and contemplate die-to-die exposure dose compensation to improve CDU.
- Our approach is to manipulate across-wafer PEB profiles to compensate for other systematic across-wafer poly CD variation sources

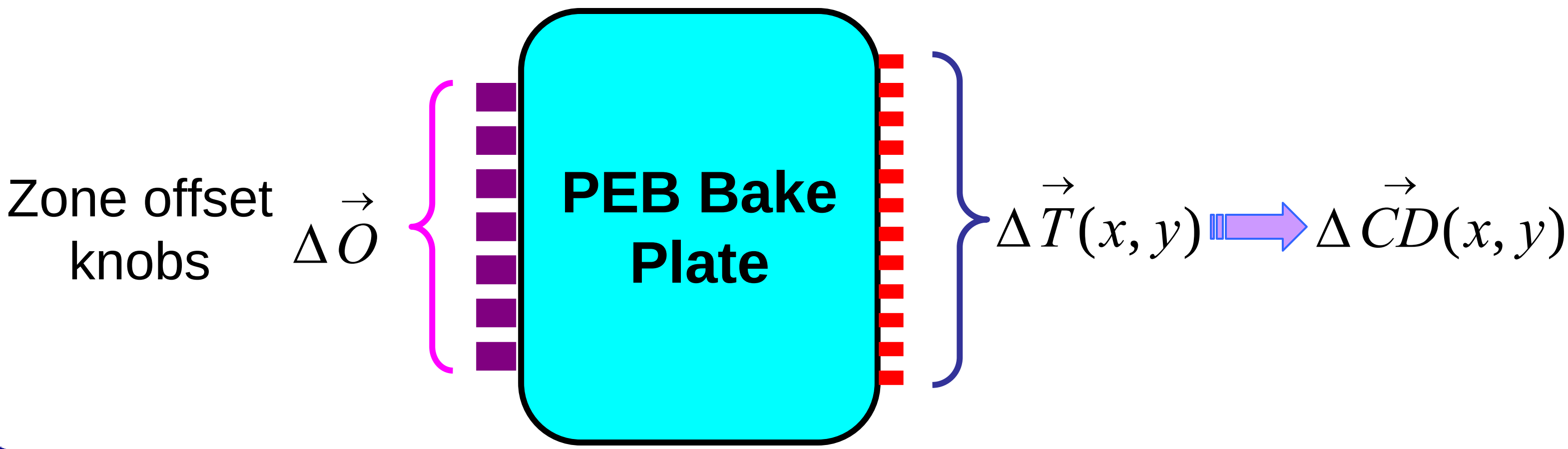


Multi-zone PEB Bake plate



General schematic setup of multi-zone bake plate

Each zone is given an individual steady state target temperature, by adjusting an *offset* value



Develop Inspection (DI) CDU Control Methodology

- The across-wafer DI CD is a function of zone offsets

$$\vec{T} = \begin{bmatrix} T_1 \\ \ldots \\ T_m \end{bmatrix} = \begin{bmatrix} g_1(O_1, O_2 \ldots O_7) \\ \ldots \\ g_m(O_1, O_2 \ldots O_7) \end{bmatrix}$$

$$\Delta \vec{T} = \vec{T} - \vec{T}_{baseline}$$

$$\vec{CD}_{DI} = \Delta \vec{T} S_{resist} + \vec{CD}_{baseline}$$

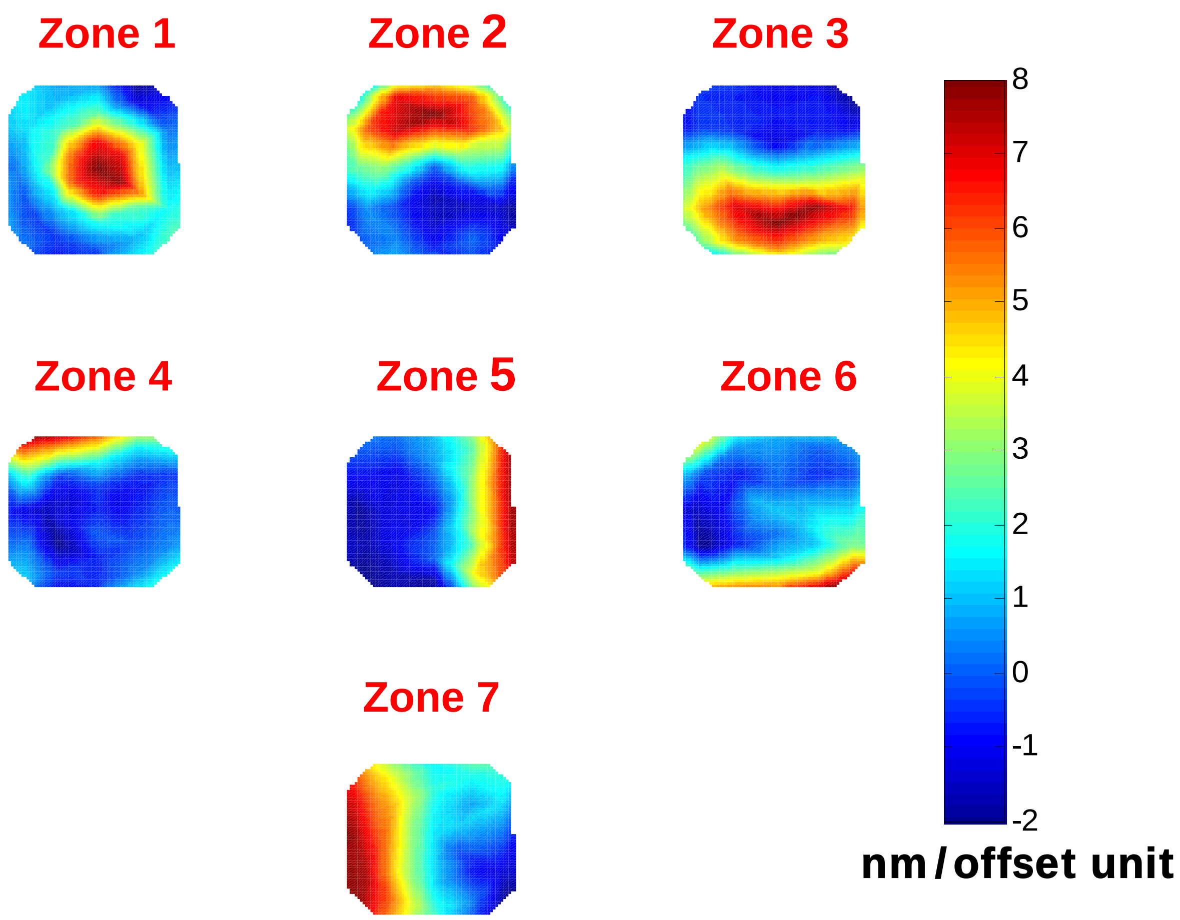
$$\vec{CD}_{DI} = \begin{bmatrix} CD_1 \\ \ldots \\ CD_n \end{bmatrix} = \begin{bmatrix} f_1(O_1, O_2 \ldots O_7) \\ \ldots \\ f_n(O_1, O_2 \ldots O_7) \end{bmatrix}$$

- Seen as a constrained nonlinear programming problem
- Minimize $\left(\vec{CD}_{DI} - \vec{CD}_{target}\right)^T \left(\vec{CD}_{DI} - \vec{CD}_{target}\right)$
- Subject to: $O^{Low} \leq O_i \leq O^{Up} \qquad i = 1,2 \ldots 7$

FLCC

Snapshot of Derived CD-to-Offset Model

- Empirically derived CD-to-offset model based on temperature-to-offset model and resist PEB sensitivity

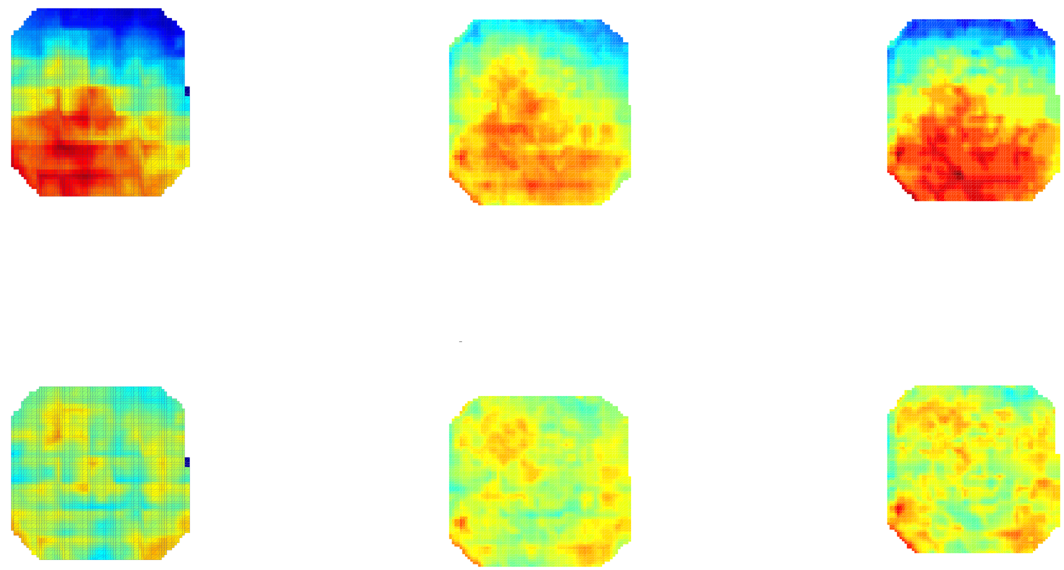


Simulation Results of DI CDU Control

Dense Line | **Semi-isolated Line** | **Isolated Line**

Experimentally extracted baseline CDU

Simulated optimal CDU after applying DI CDU control



	Dense Line	Semi-isolated Line	Isolated Line
CDU Improvement	72%	61%	69%

Final Inspection (FI) CDU Control Methodology

$$\vec{CD}_{DI} = \Delta \vec{T} S_{resist} + \vec{CD}_{baseline}$$

- Plasma etching induced AW CD Variation (signature)

$$\Delta \vec{CD}_{p\_s} \stackrel{\Delta}{=} \vec{CD}_{FI} - \vec{CD}_{DI}$$

- Across-wafer FI CD is function of zone offsets

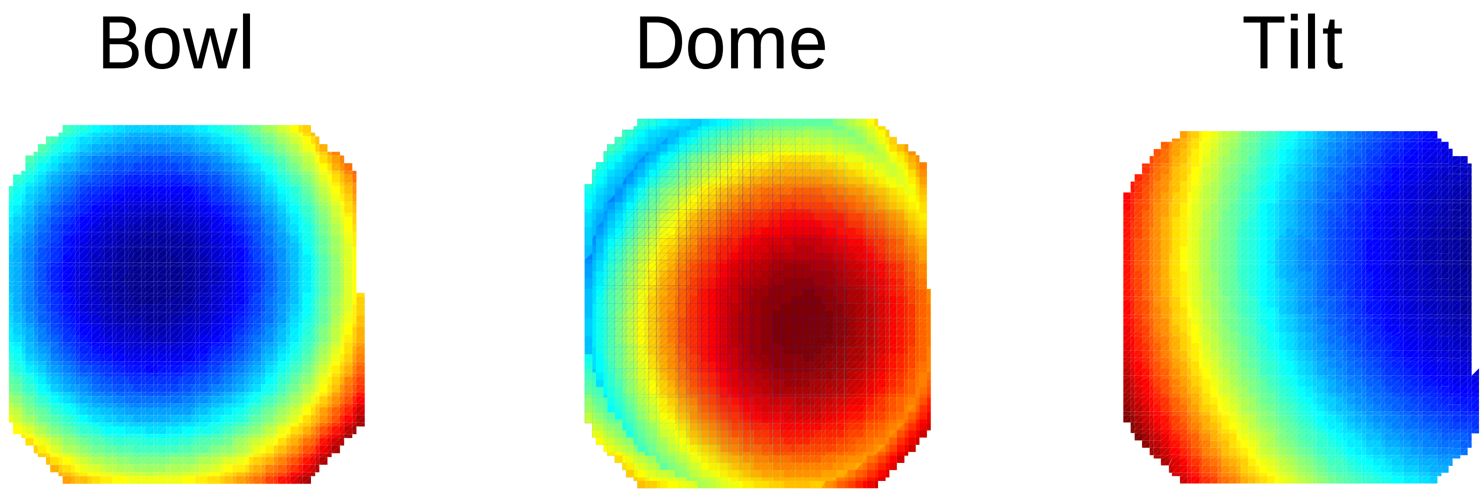
$$\vec{CD}_{FI} = \vec{CD}_{DI} + \Delta\vec{CD}_{p\_s} = \begin{bmatrix} g_1(O_1, O_2 \ldots O_7) \\ \ldots \\ g_n(O_1, O_2 \ldots O_7) \end{bmatrix}$$

- Now we minimize: $\left(\vec{CD}_{FI} - \vec{CD}_{target}\right)^T \left(\vec{CD}_{FI} - \vec{CD}_{target}\right)$
- Subject to: $O^{Low} \le O_i \le O^{Up} \qquad i = 1,2 \ldots 7$

Plasma Etching Induced AW CD Variation

- PEB-based DI control can be tuned to anticipate the plasma induced non-uniformity and cancel it.

- Use 3 plasma non-uniformity examples to simulate the proposed FI CDU control approach.



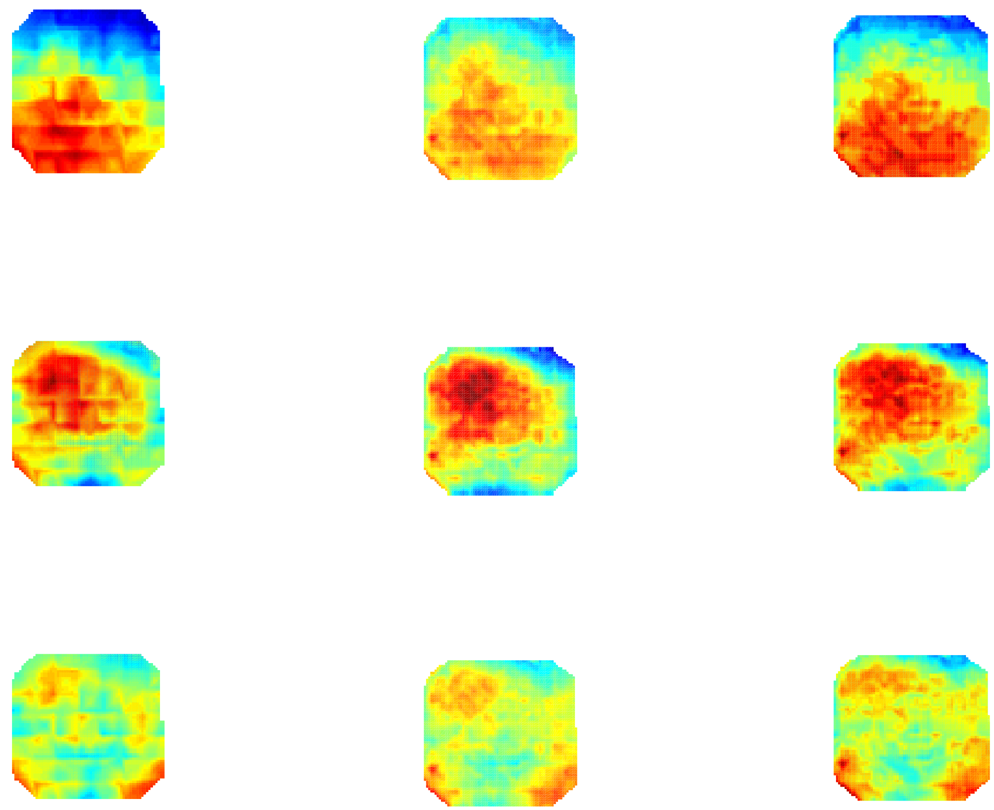
FI CDU Control Simulation - Bowl Plasma Signature

Dense | Semi-isolated | Isolated

Experimentally extracted baseline CDU

Simulated corrected DI CD after applying FI CDU control

Simulated optimal FI CD after applying FI CDU control



	Dense	Semi-isolated	Isolated
CDU Improvement	57%	37%	53%

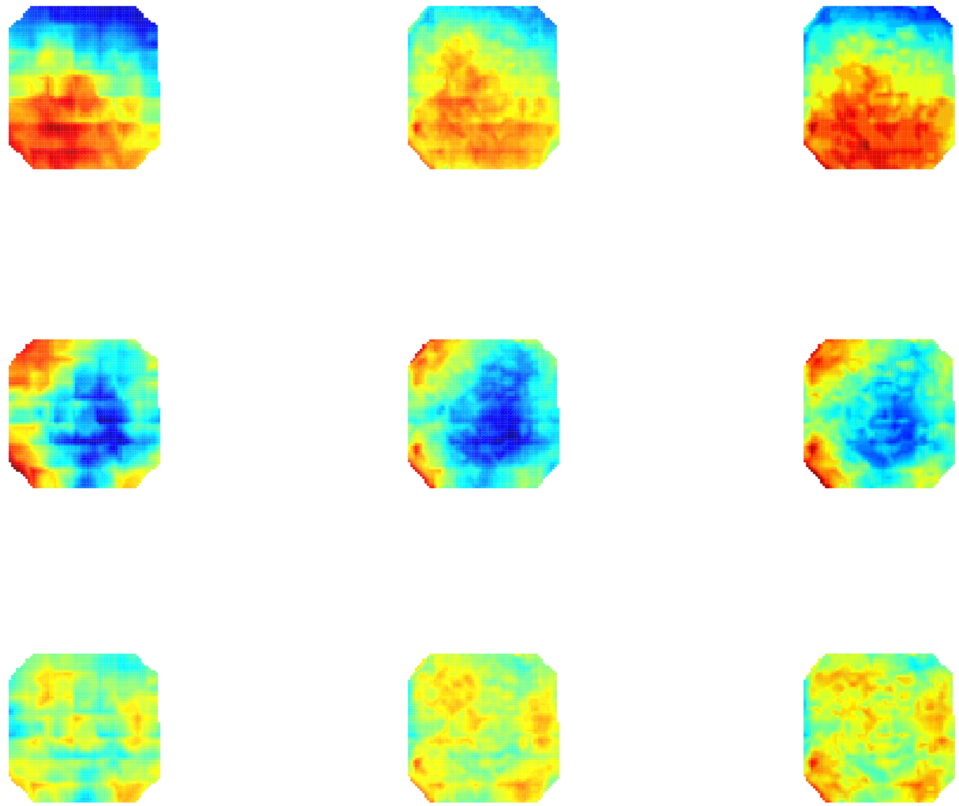
FI CDU Control Simulation - Dome Plasma Signature

Dense | **Semi-isolated** | **Isolated**

Experimentally extracted baseline CDU

Simulated corrected DI CD after applying FI CDU control

Simulated optimal FI CD after applying FI CDU control



	Dense	Semi-isolated	Isolated
CDU Improvement	69%	56%	65%

FLCC

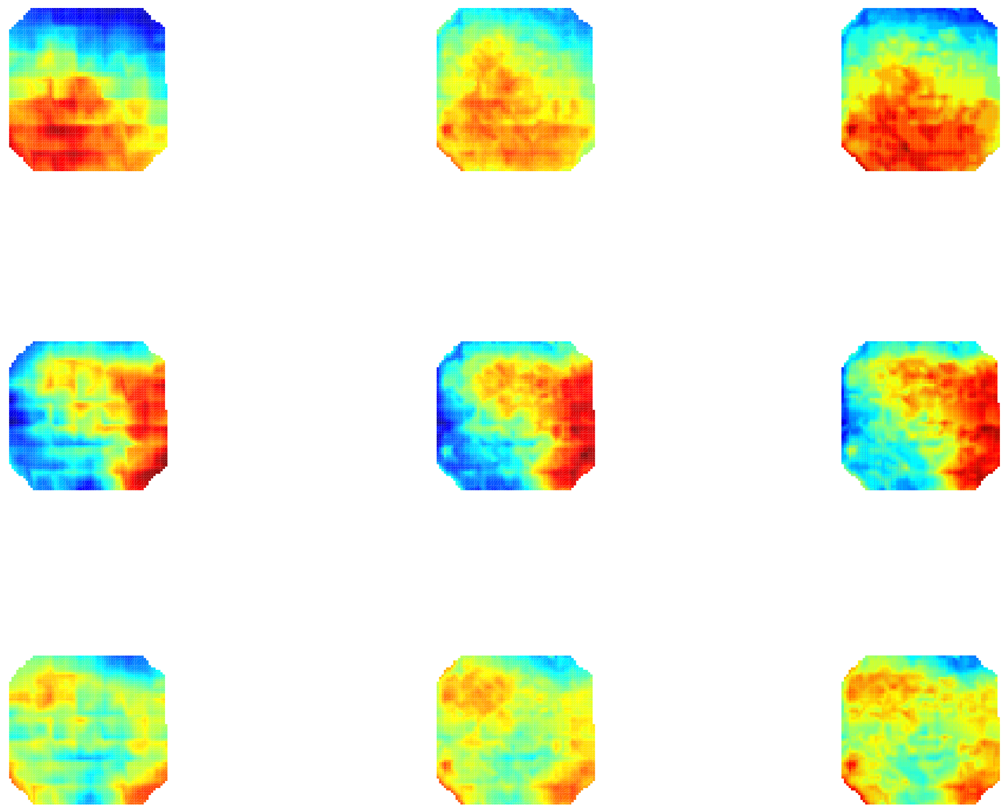
FI CDU Control Simulation - Tilted Plasma Signature

	Dense	Semi-isolated	Isolated

Experimentally extracted baseline CDU

Simulated corrected DI CD after applying FI CDU control

Simulated optimal FI CD after applying FI CDU control



	Dense	Semi-isolated	Isolated
CDU Improvement	56%	34%	52%

FLCC

Simultaneous CDU Control for Multiple CD Targets

- Multi-objective optimization of CDU for multiple targets
- Minimize the weighted sum of deviation of each target

$$J = \sum_{i=1}^{n} W_i \left\| \overrightarrow{CD_i} - \overrightarrow{CD\_T_i} \right\|^2$$

  - Subject to:

$$\begin{cases} 0 \le W_i \le 1 & 1 \le i \le n \\ \sum_{i=1}^{n} W_i = 1 & \end{cases}$$

  - Optimal zone offsets:

$$\vec{O}_{opt} = \arg\min_{\vec{O}} \left(\sum_{i=1}^{n} W_i \left\| \overrightarrow{CD_i} - \overrightarrow{CD\_T_i} \right\|^2 \right)$$

  - The relative magnitude of the weighting factor indicates the importance of meeting the corresponding CD target

Simultaneous CDU Control for Multiple CD Targets

- What is the best improvement possible for multiple targets?
- How can we *automatically* find the corresponding weighting factors and optimal zone offsets?
- Minimax optimization
  - Weighting factors of the j[th] iteration along the optimal searching trajectory:

$$\vec{W}_j = \begin{bmatrix} W_{1,j} & \dots & W_{n,j} \end{bmatrix}^T$$

$$\vec{O}_j = \arg\min_{\vec{O}} \left(\sum_{i=1}^{n} W_{i,j} \left\| \overrightarrow{CD_i} - \overrightarrow{CD\_T_i} \right\|^2 \right)$$

  - Minimax to find optimal weighting factors and offsets

$$\vec{W}_{opt} = \begin{bmatrix} W_{1,opt} & \dots & W_{n,opt} \end{bmatrix}^T = \arg\min_{\vec{W}_j} \left(\max(\sigma_{1,j} \quad \dots \quad \sigma_{n,j}) \right)$$

$$\vec{O}_{opt} = \arg\min_{\vec{O}} \left(\sum_{i=1}^{n} W_{i,opt} \left\| \overrightarrow{CD_i} - \overrightarrow{CD\_T_i} \right\|^2 \right)$$

Simultaneous CDU Control for Multiple CD Targets

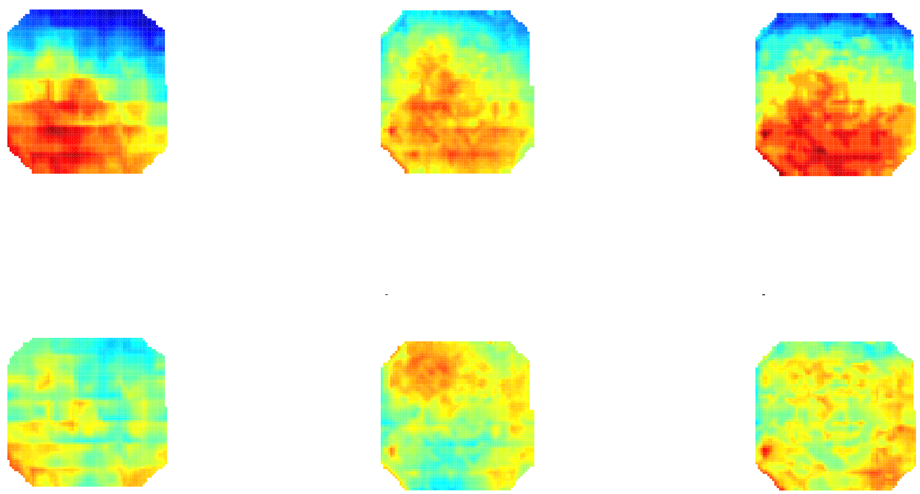
Simulation of simultaneous CDU control for *dense, semi-iso and iso lines*

	Dense Line	Semi-iso Line	Iso Line
Wd =0.36; Ws =0.33 ; Wi =0.31	**64.9%**	**40.7%**	**66.4%**
Wd = 0.90; Ws =0.05 ; Wi =0.05	71.8%	15.9%	61.8%
Wd = 0.05; Ws =0.90 ; Wi =0.05	48.2%	60.7%	54.1%
Wd = 0.05; Ws =0.05 ; Wi =0.90	64.1%	32.4%	68.6%

Dense **Semi-isolated** **Isolated**

Experimentally extracted baseline CDU

Simulated optimal FI CD after applying simultaneous CDU control



Summary and Conclusions

- Extracted CDU signatures of dense, iso and semi-iso
- CD-to-offset model enables DI & FI CDU control
  - The derived CD-to-offset model is based on temperature-to-offset model and resist PEB sensitivity
  - Offers better fidelity than the old CD-to-offset model purely based on CD measurement
  - Simulation indicates promise of DI & FI CDU control
- Multi-objective & minimax optimization schemes enable simultaneous CDU control for multiple CD targets
- Work in SDC at AMD are under way to validate this approach experimentally

Technology/Circuit Co-Design: Impact of Spatial Correlation

Paul Friedberg

Department of Electrical Engineering and Computer Sciences

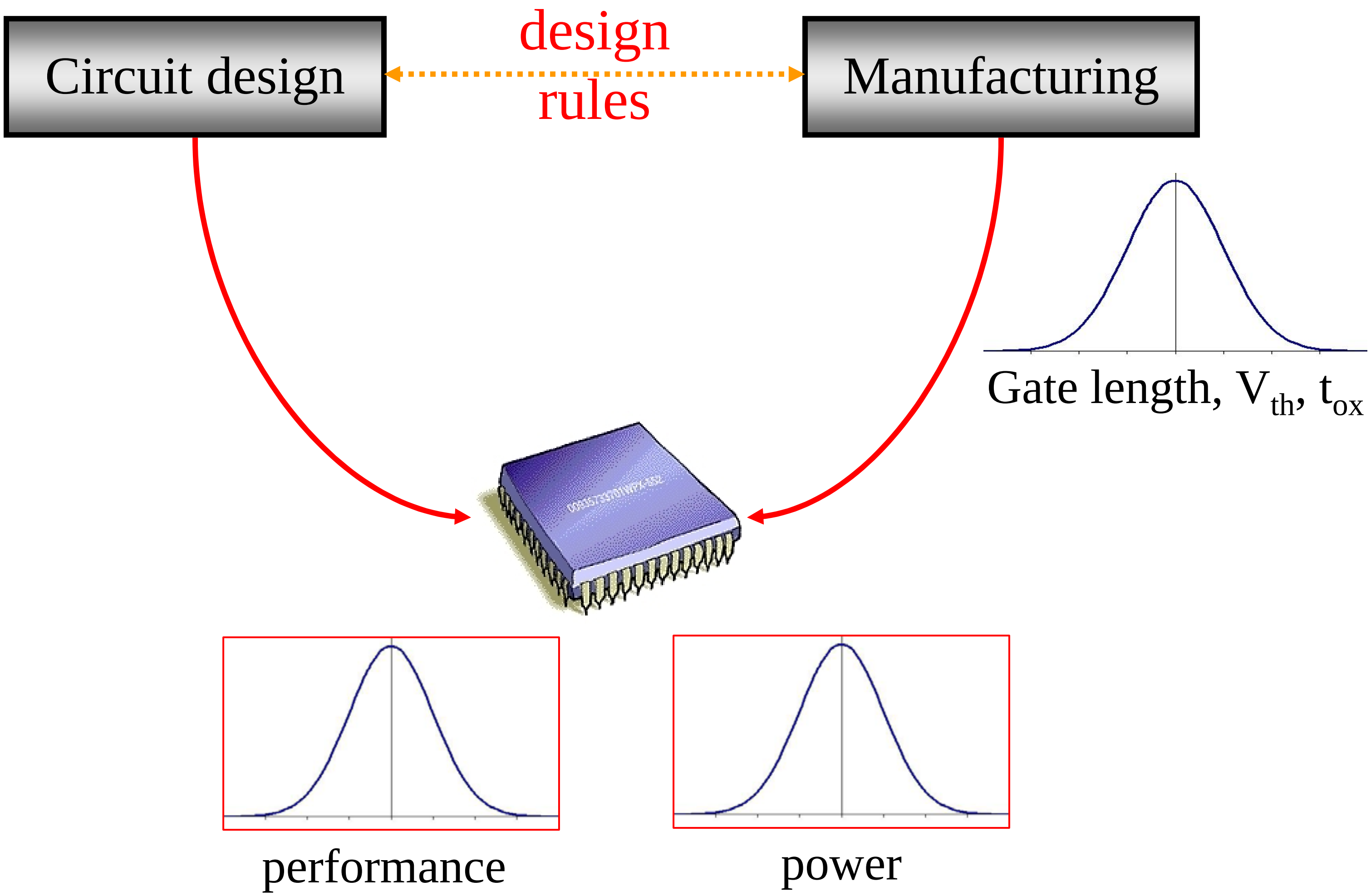
University of California, Berkeley

Feb. 14, 2005

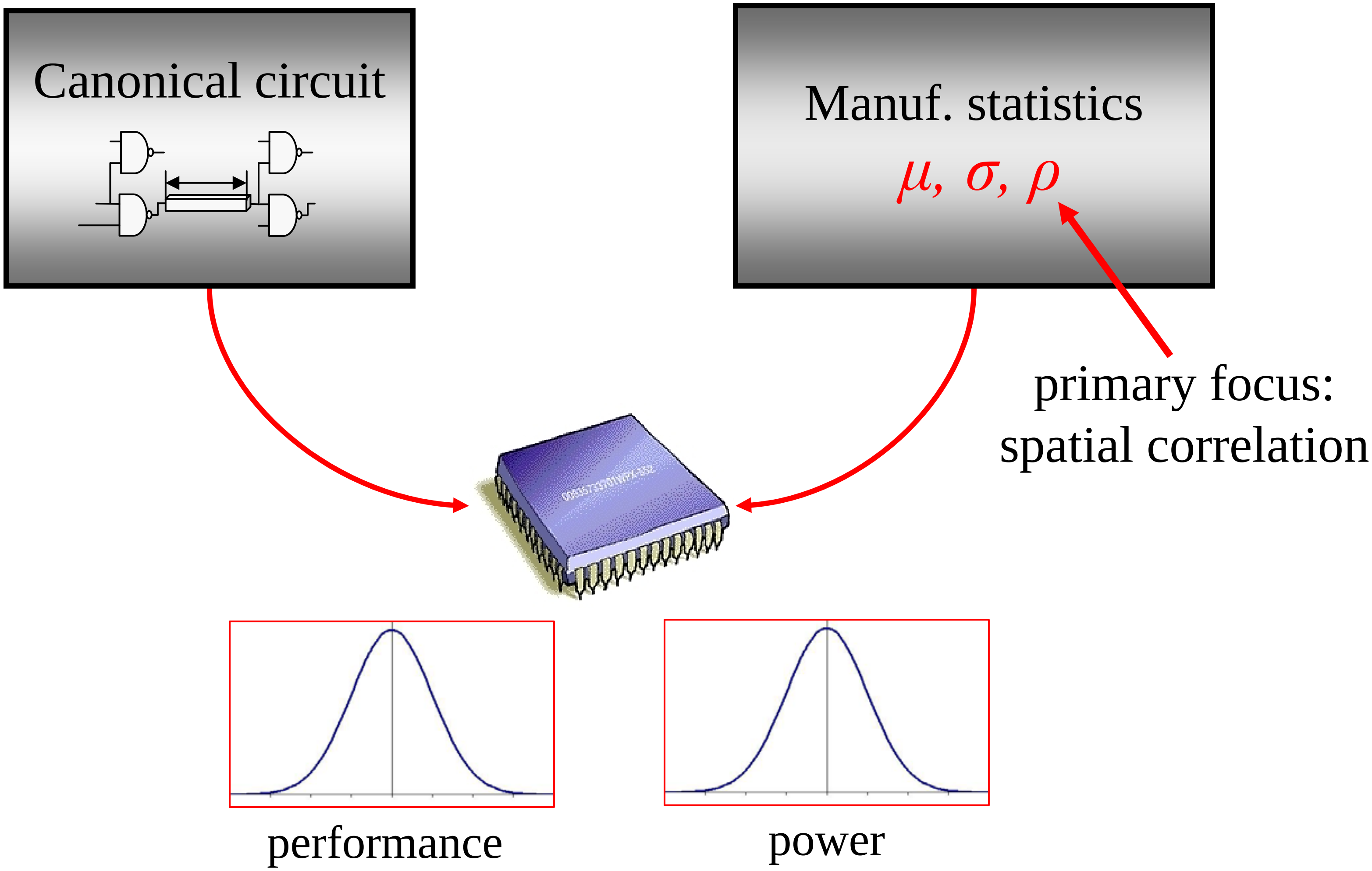
Outline

- Motivation
- Spatial Correlation Extraction
- Impact of Spatial Correlation on Circuit Performance
- How does process control impact spatial correlation?
- Conclusions/Future Plans

Motivation: reality



Motivation: simulation



Spatial Correlation Analysis

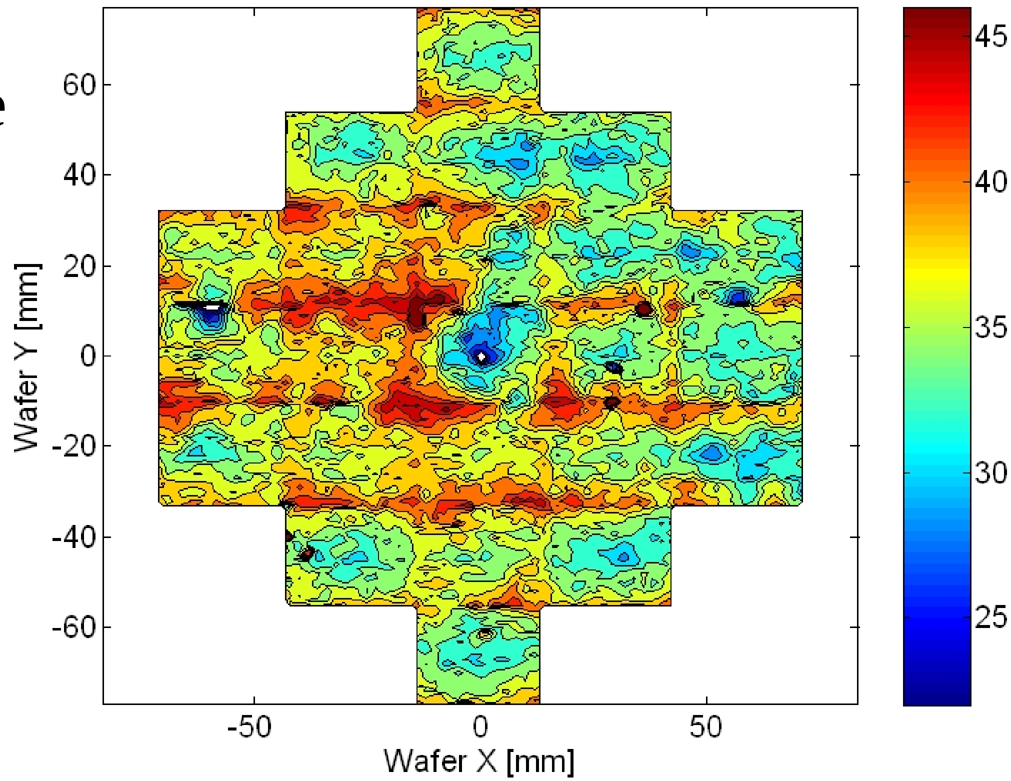
- Exhaustive ELM poly-CD measurements (280/field):

- Z-score each CD point, using wafer-wide distribution:

$$z_{ij} = \left(x_{ij} - \overline{x}_j\right) / \sigma_j$$

- For each spatial separation, calculate correlation ρ among all within-field pairs:

$$\rho_{jk} = \left(\sum z_{ij} * z_{ik}\right) / n$$

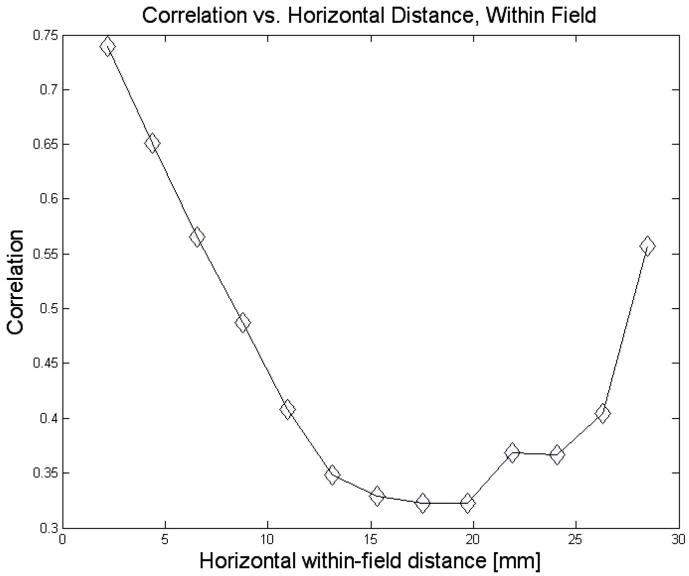


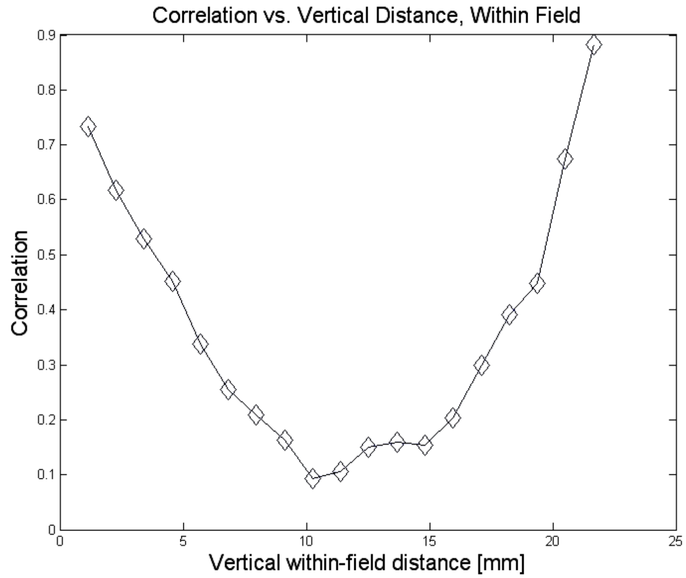
(CD data courtesy of Jason Cain)

FLCC

Spatial Correlation Dependence

- Within-field correlation vs. horizontal/vertical distance, evaluated for entire wafer:



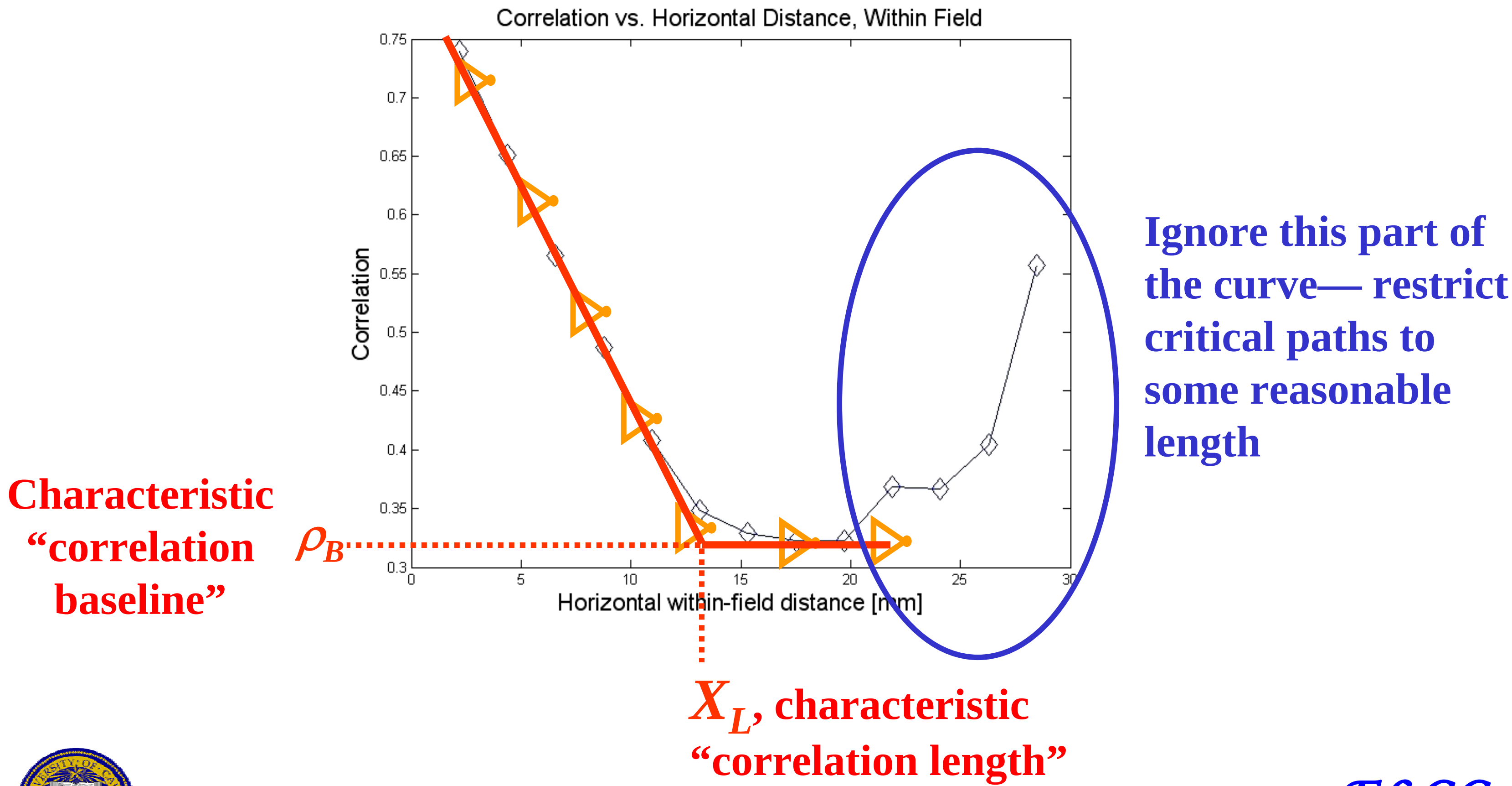


- Statistical assumptions are violated (distribution is not stationary): we will address this later

FLCC

Spatial Correlation Model

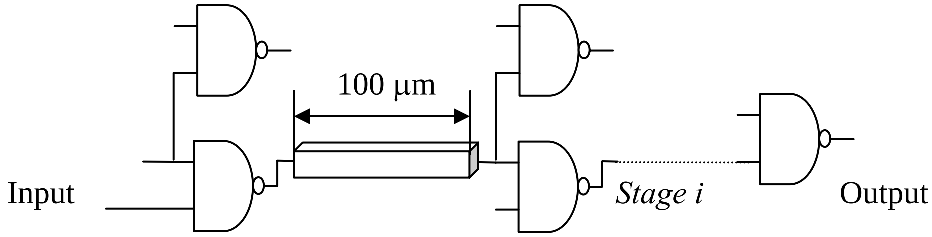
- Fit rudimentary linear model to spatial correlation curve extracted from empirical data:



FLCC

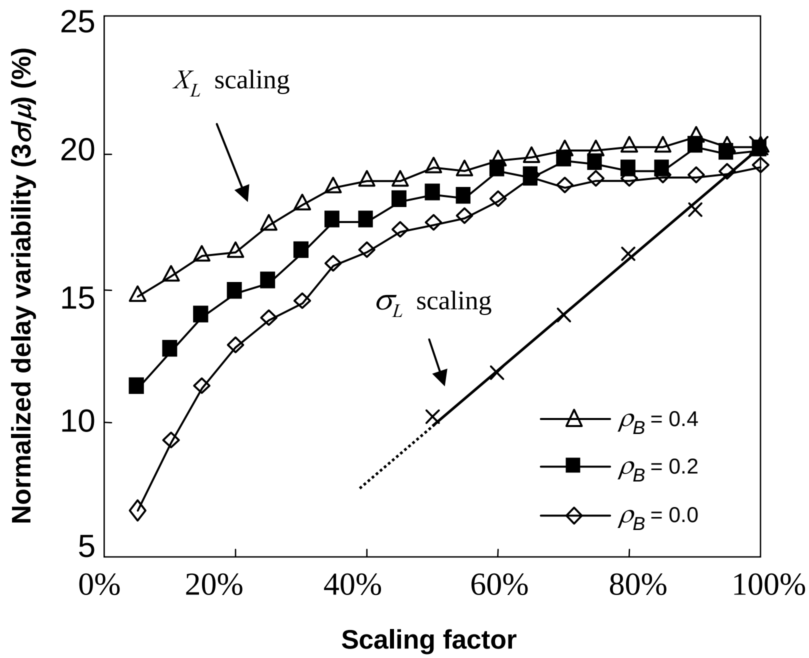
Monte Carlo Simulations

- Use canonical circuit of FO2 NAND-chain w/ stages separated by 100μm local interconnect, ST 90nm model:



- Perform several hundred Monte Carlo simulations for various combinations of X_L, ρ_B, and σ/μ (gate length variation)
- Measure resulting circuit delays, extract normalized delay variation ($3\sigma/\mu$)

Delay Variability vs. X_L, ρ_B, σ/μ



- Scaling gate length variation directly: most impactful
- Reducing spatial correlation also reduces variability, increasingly so as ρ decreases

FLCC

Origin of Spatial Correlation Dependence

- CD variation can be thought of as nested systematic variations about a true mean:

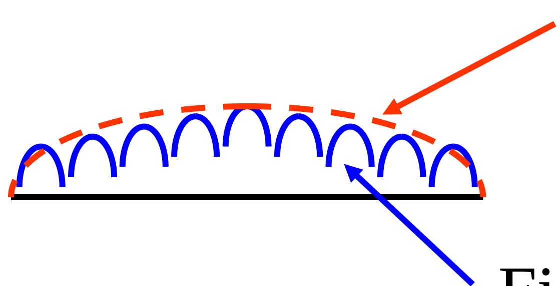
$$CD_{ij} = \mu + mask + f_i + w_j + \varepsilon_{ij}$$

Spatial components ($f_i + w_j$)

μ: True mean

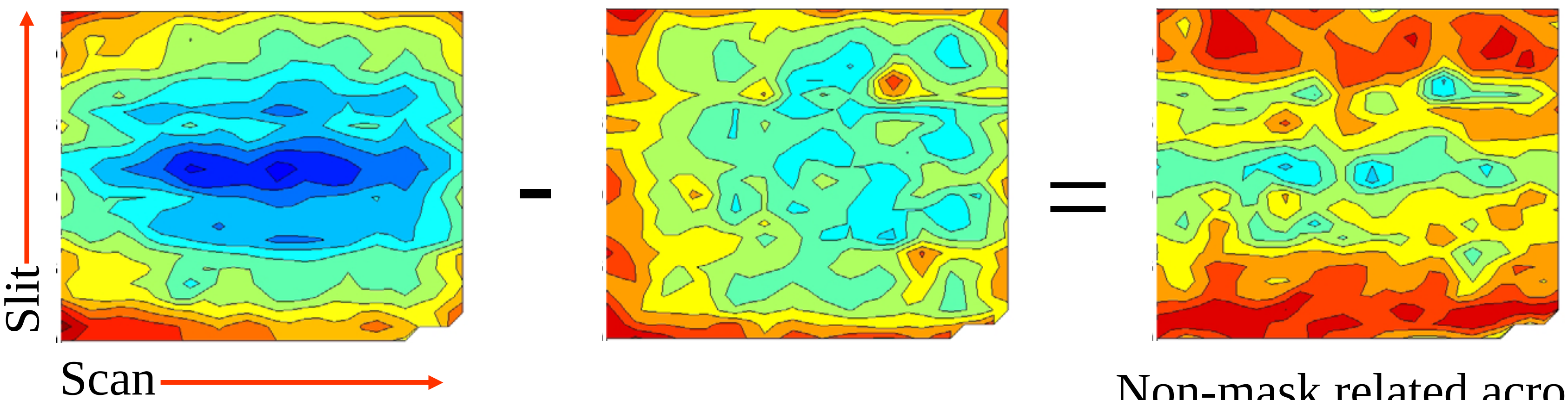
f_i: Across-field

w_j: Across-wafer



Origin of Spatial Correlation Dependence

- Within-die variation:

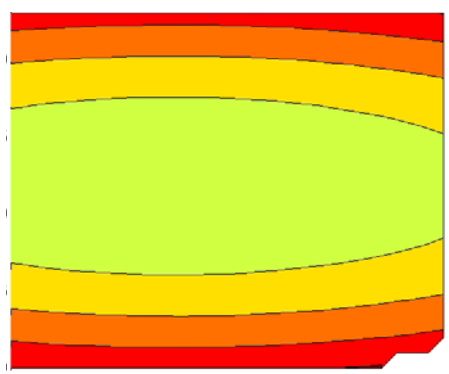


Average Field

Scaled Mask Errors

Non-mask related across-field systematic variation

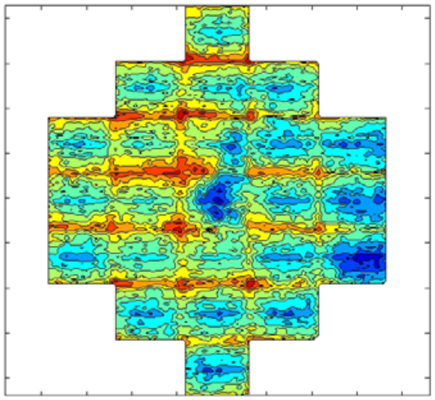
Polynomial model of across-field systematic variation



Removing this component of variation will simulate *WID* process control

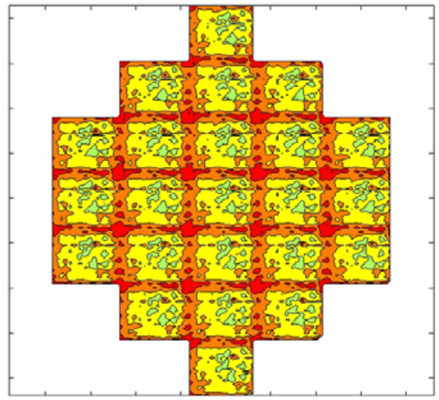
Origin of Spatial Correlation Dependence

- Across-wafer variation extraction:



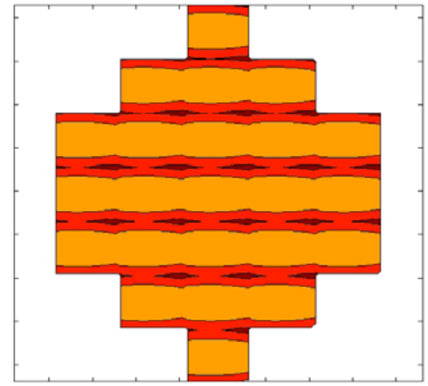
Average Wafer

\-



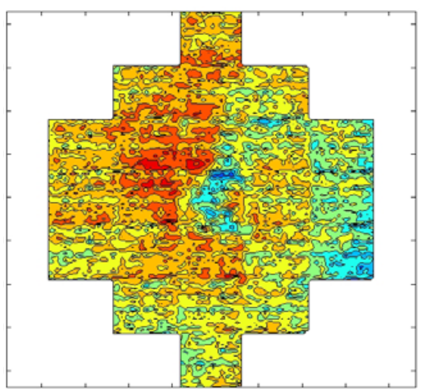
Scaled Mask Errors

\-



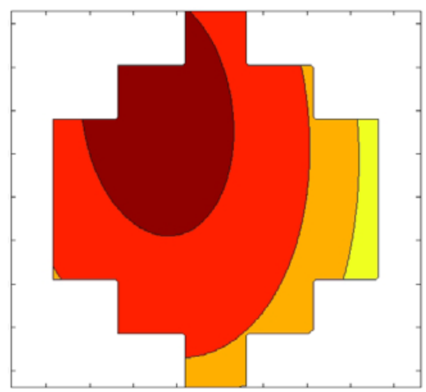
Across-Field Systematic Variation

=



Across-Wafer Systematic Variation

→

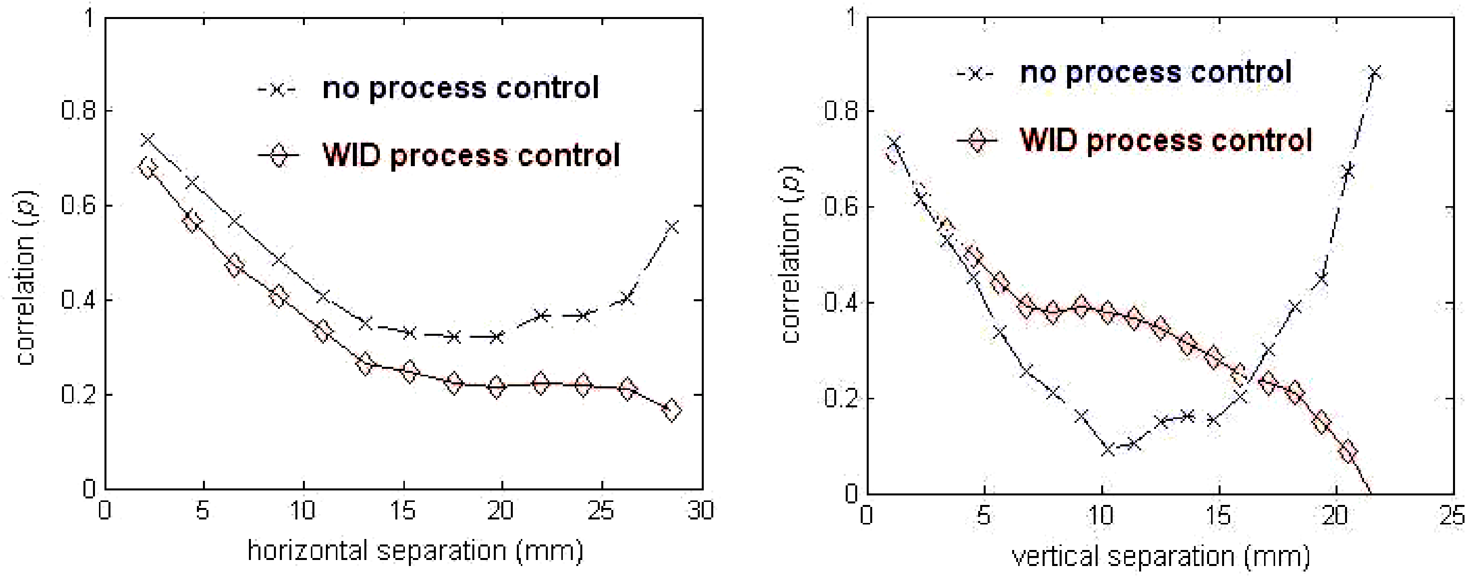


Polynomial Model

Removing this component of variation will simulate *AW* process control

Artificial *WID* Process Control

- By removing the within-field component of variation, we get distinctly different correlation curves:

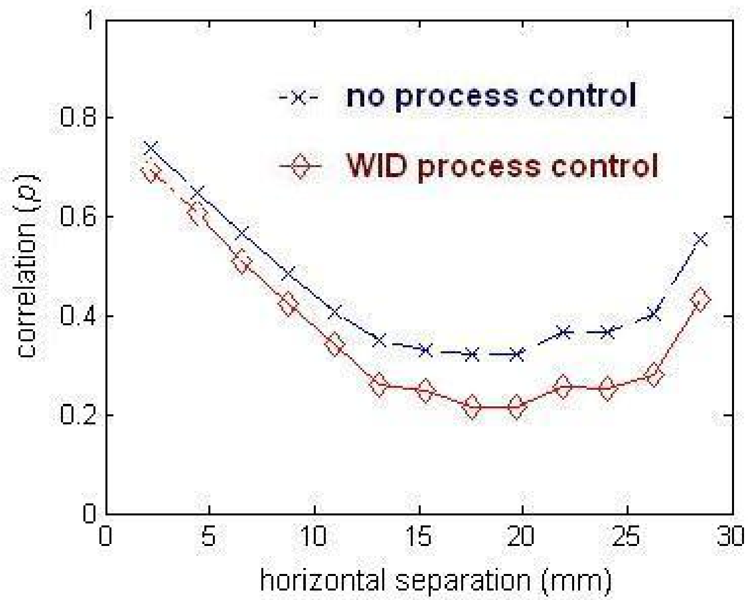


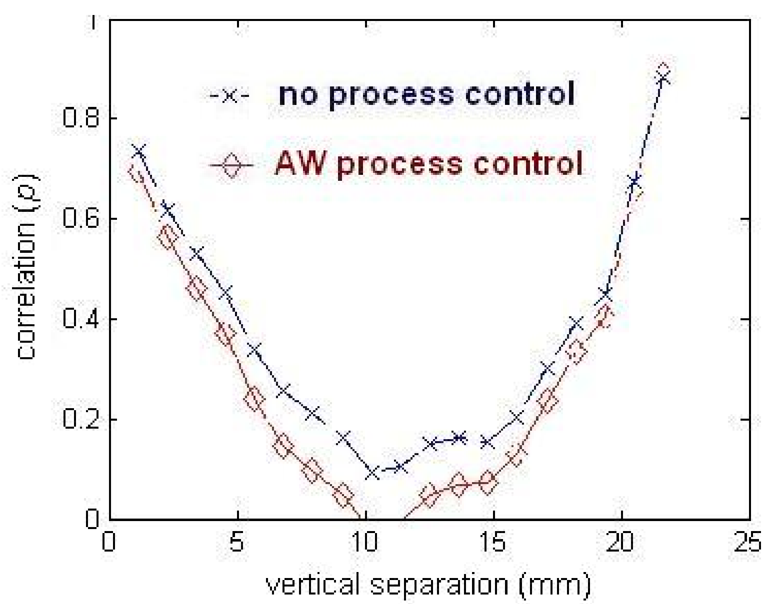
- Shape of curve changes; correlation decreases for horizontal, but *increases* for vertical

FLCC

Artificial *AW* Process Control

- Removing the across-wafer component only:

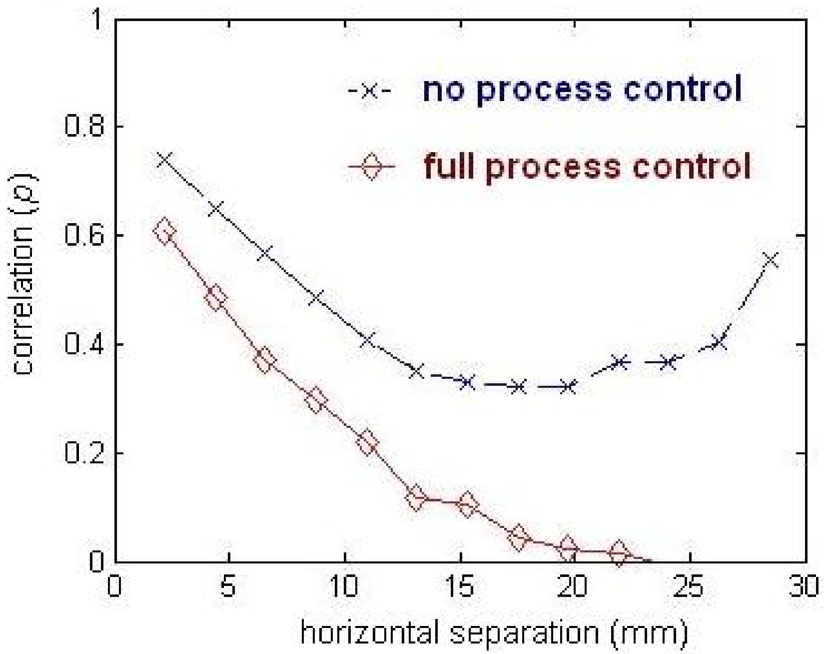


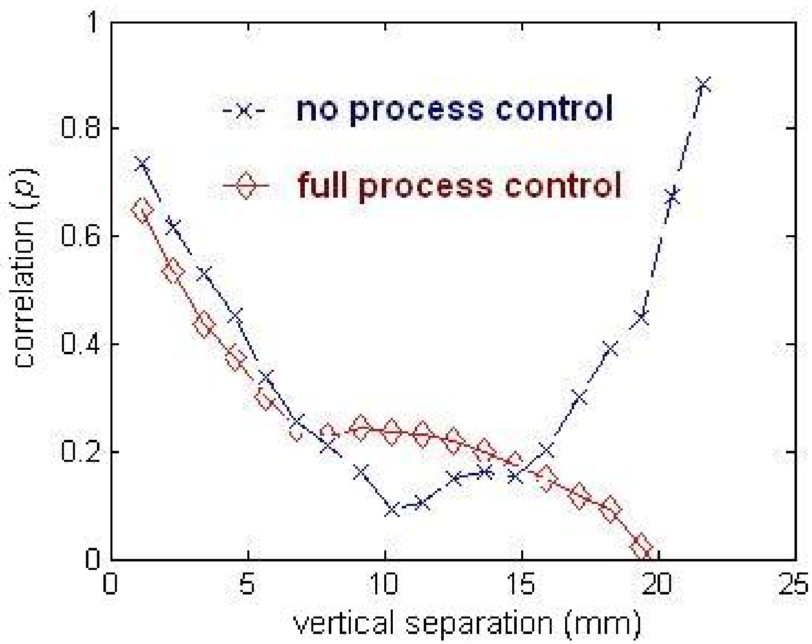


- Shape stays roughly the same; correlation decreases across the board

Artificial *AW+WID* Process Control

- Removing both *AW* and *WID* components, get a cumulative effect larger than the sum of the parts:

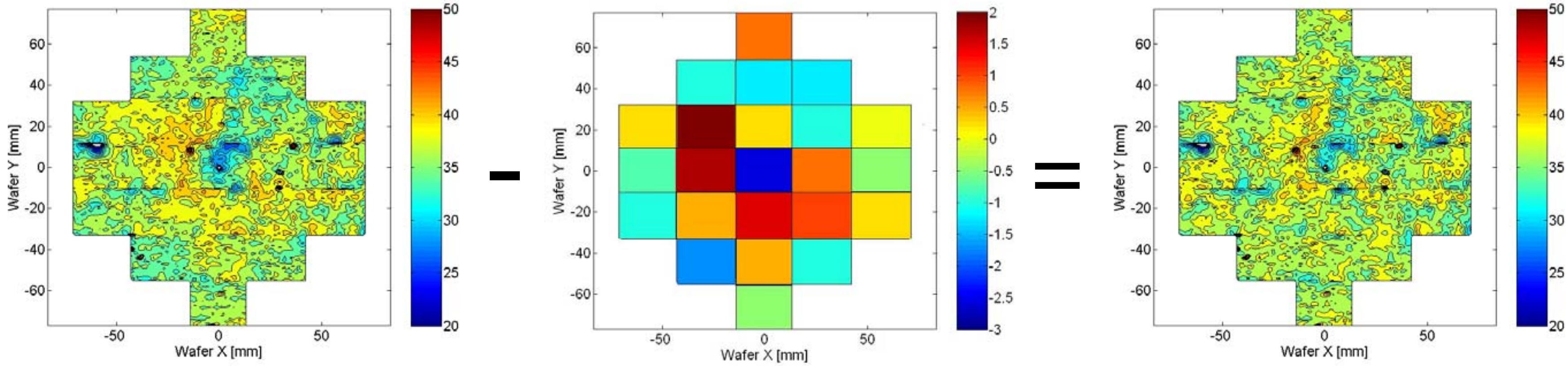


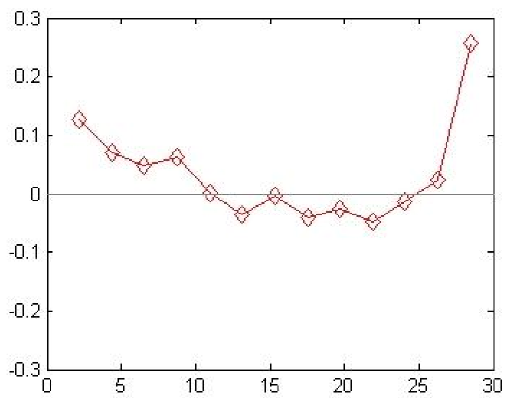


FLCC

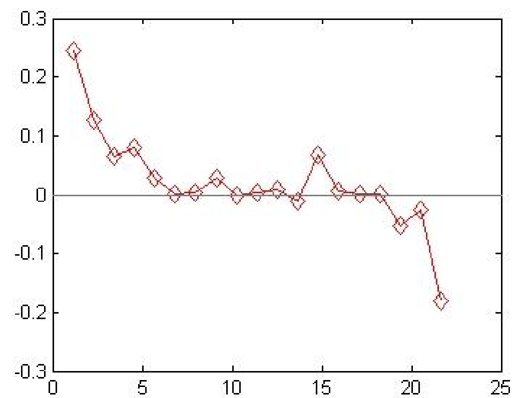
Additional process control

- One more round of control: die-to-die dose control





Horizontal separation



Vertical separation

FLCC

Conclusions

- Correlation effects are significant: should definitely be included in MC simulation frameworks
- Spatial correlation virtually *entirely* accounted for by systematic variation

  → Complete process control can almost completely reconcile correlation
- As process control is implemented, σ and ρ are simultaneously reduced: a double-win
- The closer to complete control, the greater the impact of additional control on correlation
  - Last "little bit" of systematic variance in the distribution causes substantial correlation