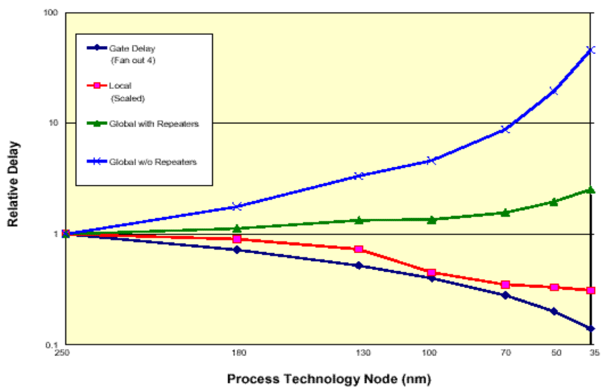
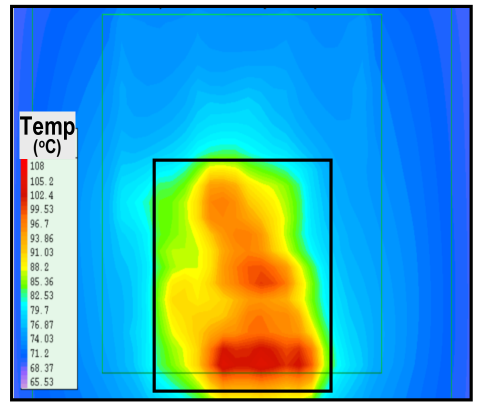
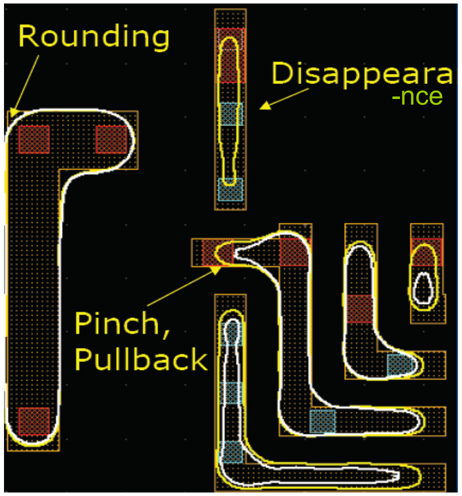
Design for Manufacturability and Reliability in Extreme CMOS Scaling and Beyond

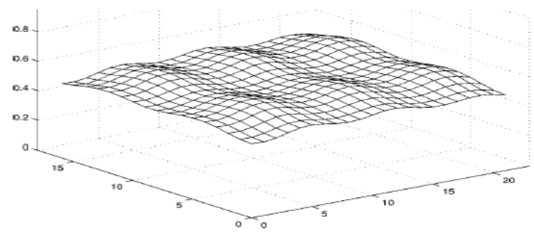
David Z. Pan
Dept. of Electrical and Computer Engineering
The University of Texas at Austin
http://www.cerc.utexas.edu/utda

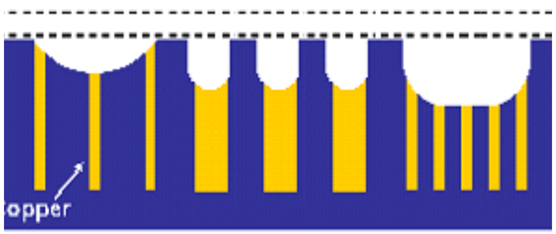
Nanometer Issues

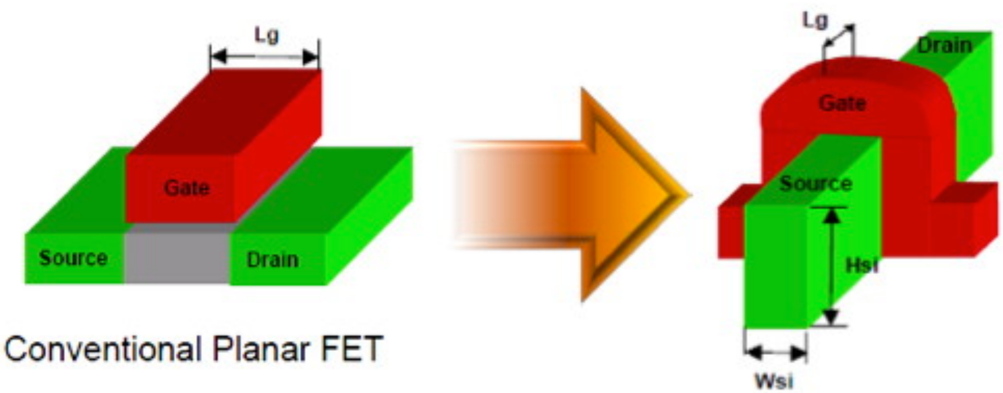






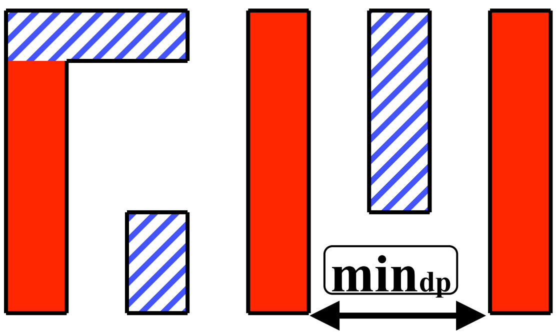




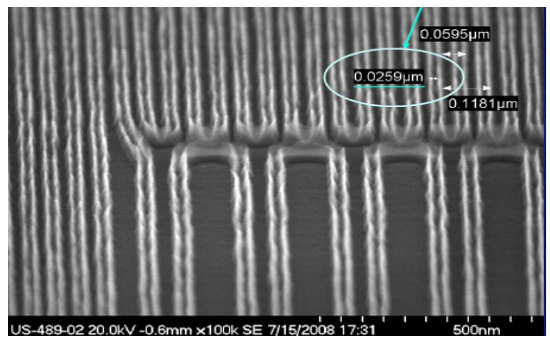


Emerging Lithography

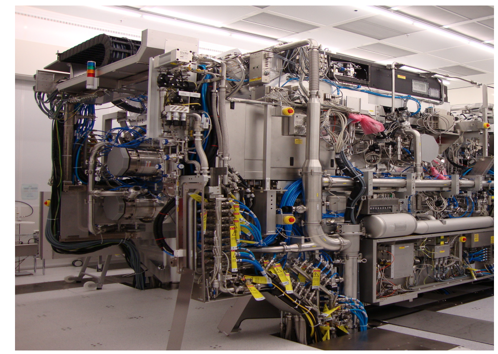
193i w/ DPL



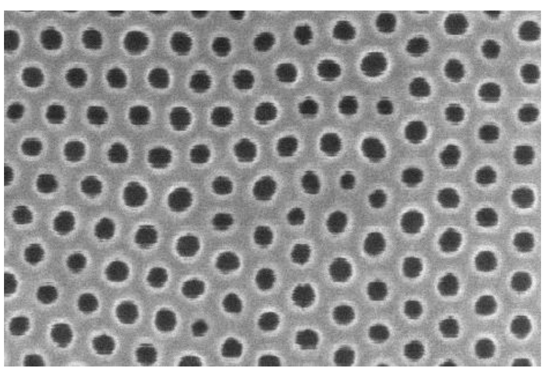
Quadruple patterning



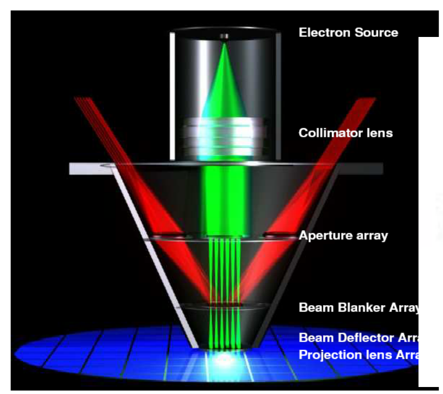
EUV



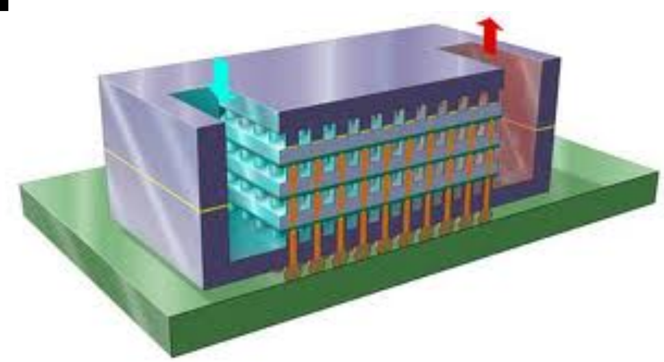
DSA



E-beam



3D-IC

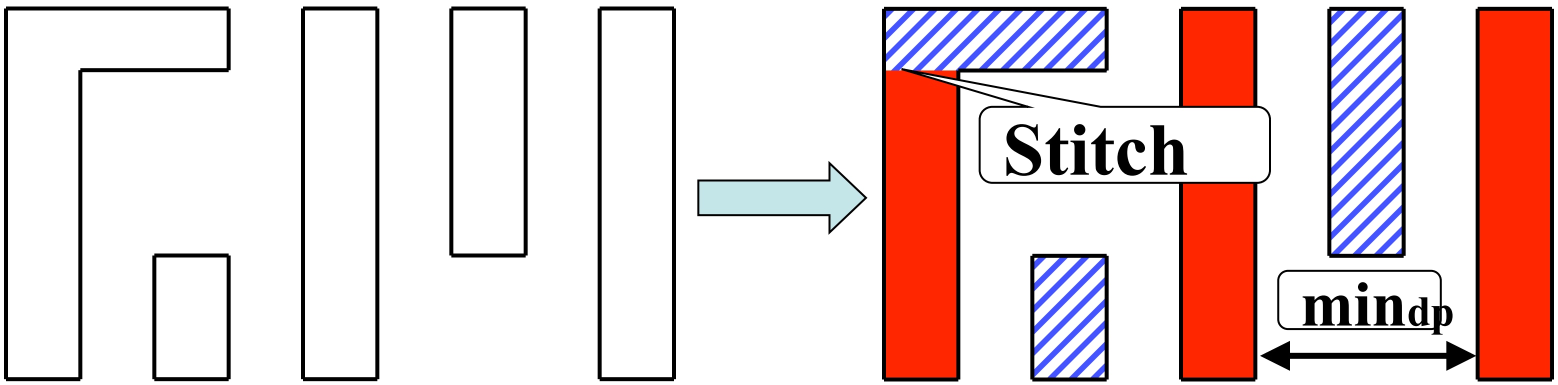


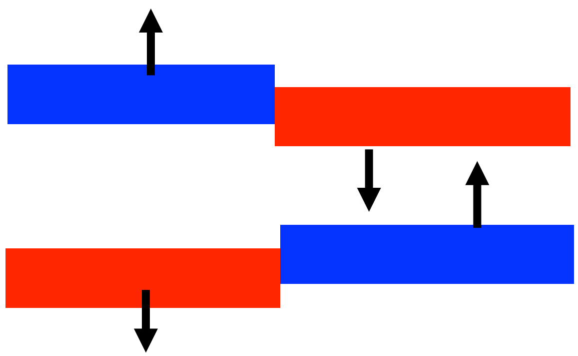
The "Moore", the Merrier!

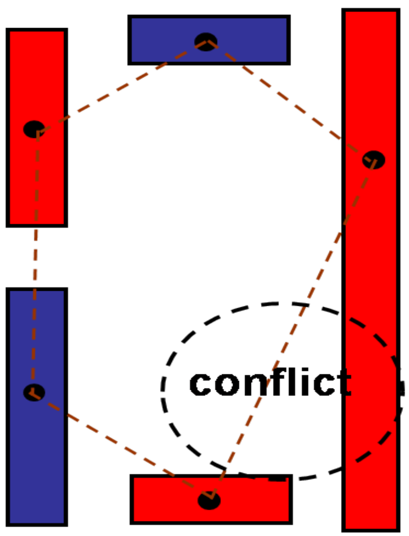
- **More Moore**
  - Nano-Patterning for Extreme Scaling
  - Lithography Aware Physical Design
- **A different kind of "Moore"**
  - 3D Integration
  - New devices/material/…

➔ Need synergistic design and technology co-optimization for cross-layer resilience

What is Double Patterning?

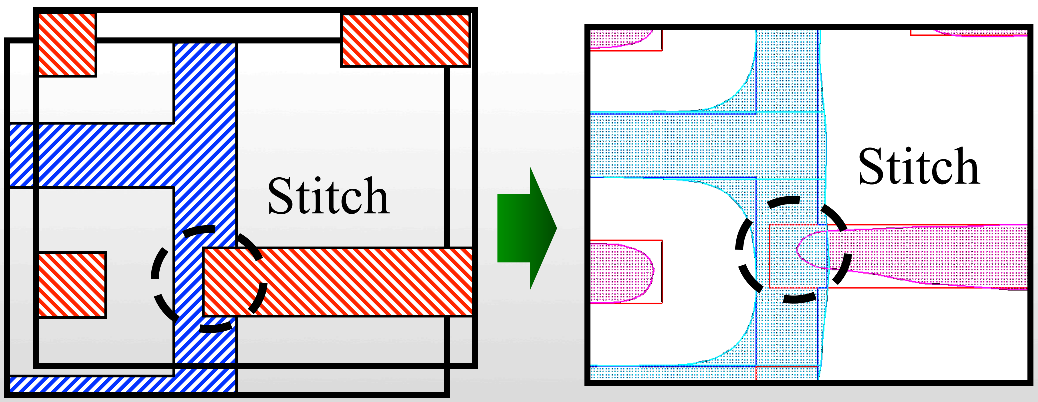






Dealing with Overlay in LELE

Minimum Stitch Insertion

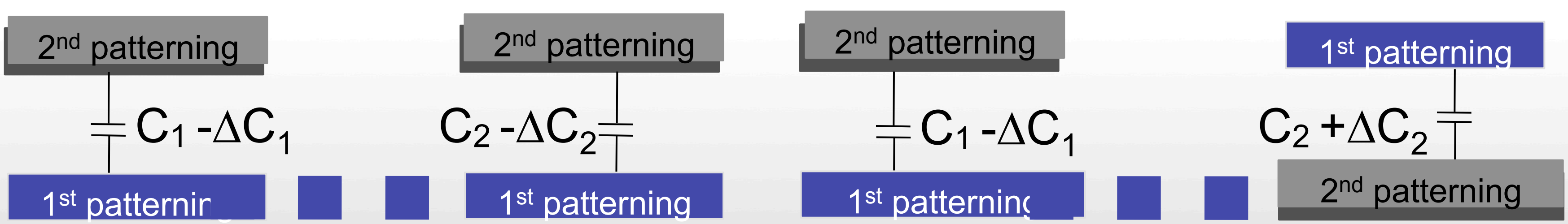


[Lucas SPIE'08]

1) **Minimize stitch #**
2) **A bit more overlap margin for stitch, but area increases**

Overlay Compensation

[Yang+, ASPDAC10]

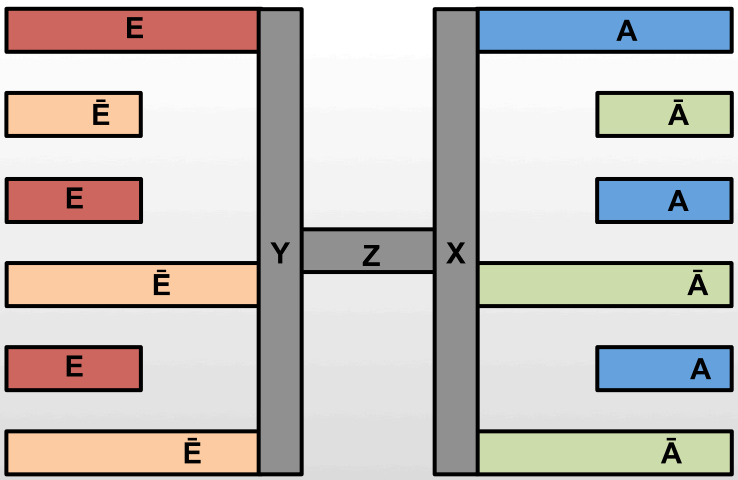


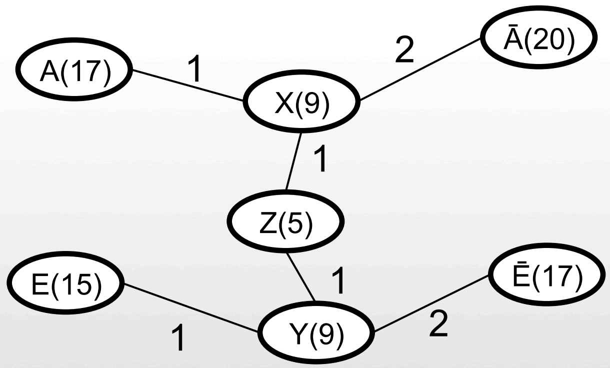
Without Overlay Compensation **With Overlay Compensation**

A Graph-Partitioning Based, Multi-Objective Decomposer

Decomposition Graph Construction

[Yang+, ASPDAC10]





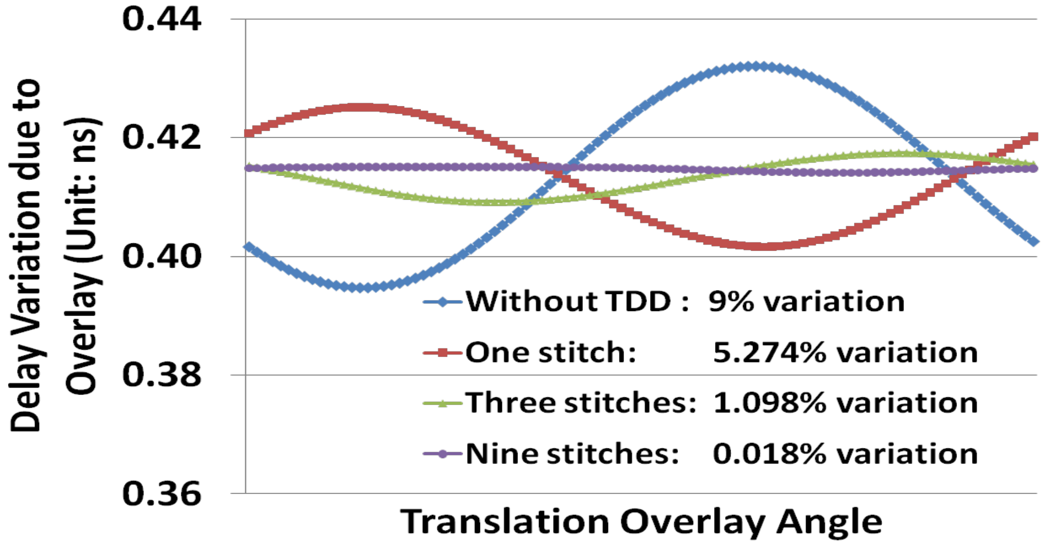
Constraint:
(A, Ā) and (E, Ē) are repulsive pairs.

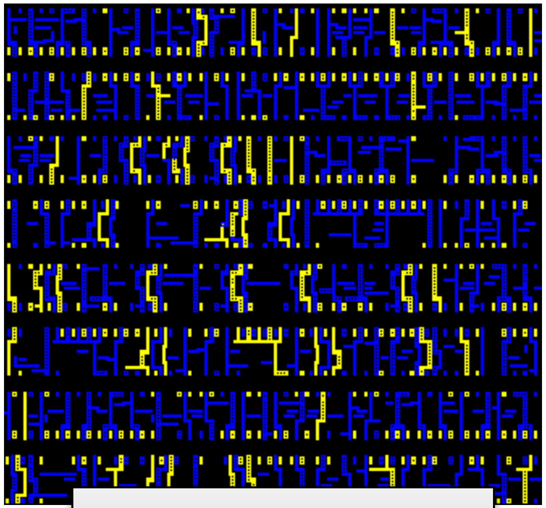
Theorem : Stitch minimization problem is equivalent to the min-cut partitioning of the decomposition graph

Extensions of the framework: to incorporate other constraints and costs into graph partitioning, e.g., balanced density, overlay compensation, and so on

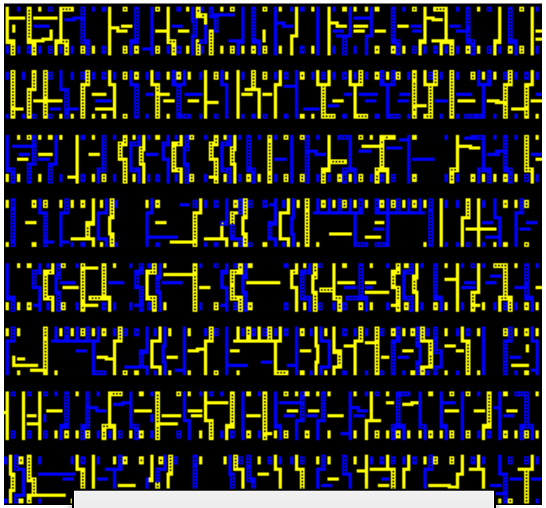
Overlay Compensation & Density Balancing

[Yang+, ASPDAC10]

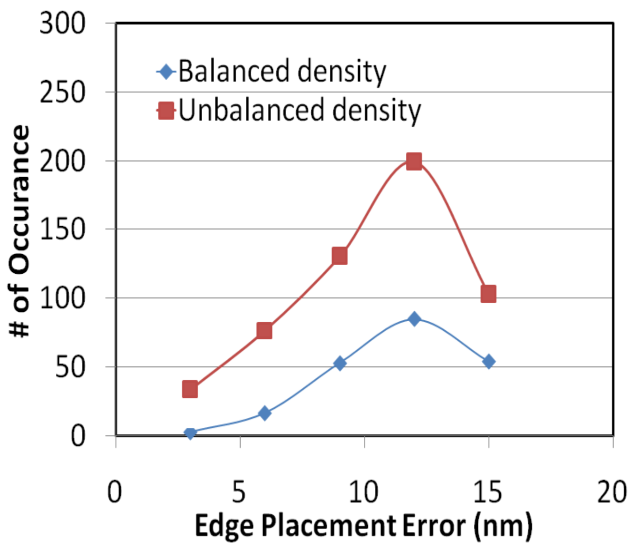




C432:27% and 73%
(7 stitches)

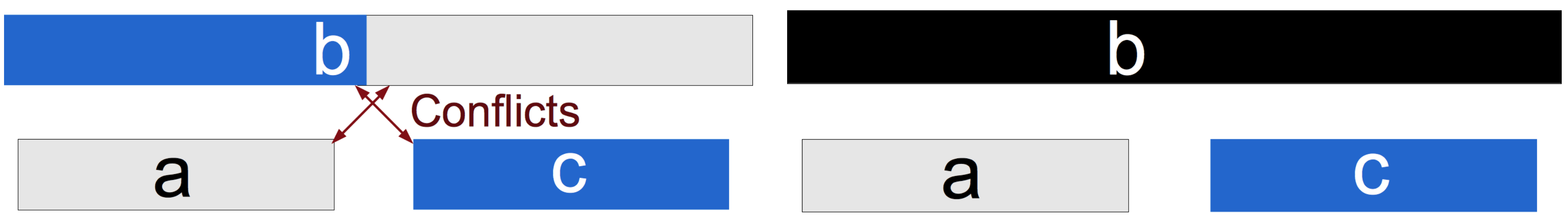


C432:50% and 50%
(17 stitches)



Triple Patterning

- What's triple patterning lithography (TPL)?
  - Extension of double patterning concept
  - Original layout is divided into three masks
  - Triple effective pitch
- Why TPL?
  - Delay of next generation lithography (EUV, E-beam)
  - Resolve conflicts of DPL
  - Achieve further feature-size scaling (14nm, 11nm)



LAPD

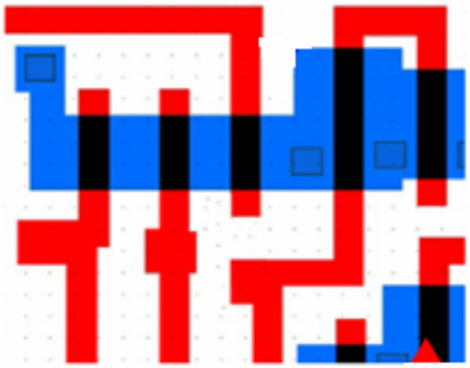
Another LAPD

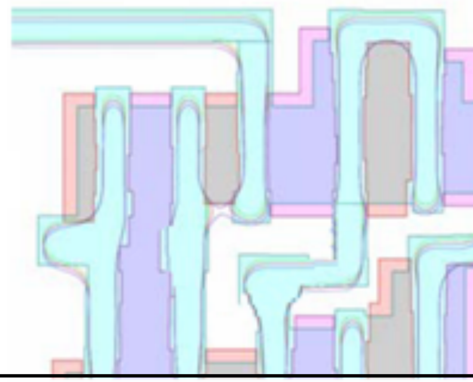
- Double/multiple patterning layout compliance/decomposition
- Still something could go wrong!
- Lithography Aware Physical Design (LAPD) ➔
- Litho Hotspot Detection
- Litho Friendly Design
  - Hotspot Avoiding/Correction
  - Correct by Construction/Prescription

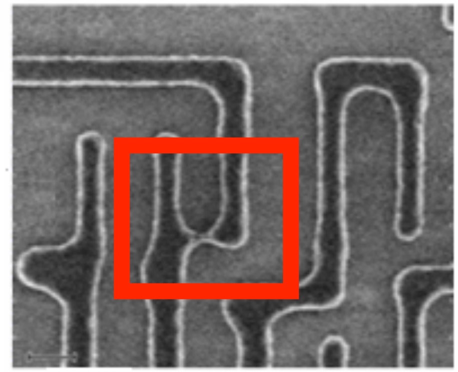
Detection

Correction

Lithography Hotspot Detection


Layout

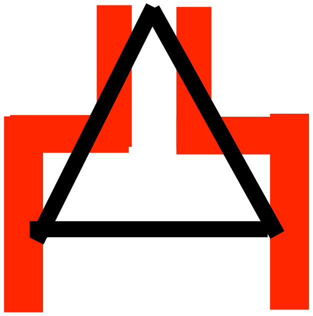

Litho simulations


Chip

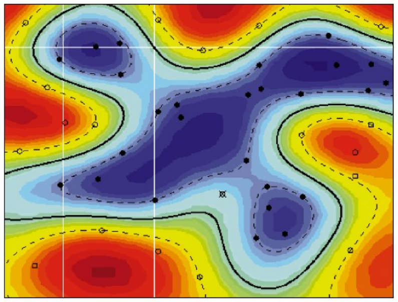
- **Lithographic hotspots**
  - What you see (at design) is NOT what you get (at fab)
  - Hotspots mean poor printability
  - Highly dependent on manufacturing conditions
  - Exist after resolution enhancement techniques
- **Litho-simulations are extremely CPU intensive**
  - Full-blown OPC could take a week
  - Impossible to be used in inner design loop

Various Approaches

[Xu+ ICCAD07]
[Yao+ ICCAD08,
[Khang SPIE06],
etc.



Pattern/Graph Matching



SVM [J. Wuu+ SPIE09]
[Drmanac+ DAC09]
Neural Network Model
[Norimasa+ SPIE07][Ding + ICICDT09]
Regression Model
[Torres+ SPIE09]

Data Mining/Machine Learning

- **Pros and cons**
  - Accurate and fast for known patterns
  - But too many possible patterns to enumerate
  - Sensitive to changing manufacturing conditions
  - High false-alarms

- **Pros and cons**
  - Good to detect unknown or unseen hotspots
  - Accuracy may not be good for "seen" patterns (cf. PM)
  - Hard to trade-off accuracy and false alarms

A New Meta-Classification Paradigm

Pattern Matching Methods
Good **for detecting previously known types of hotspots**

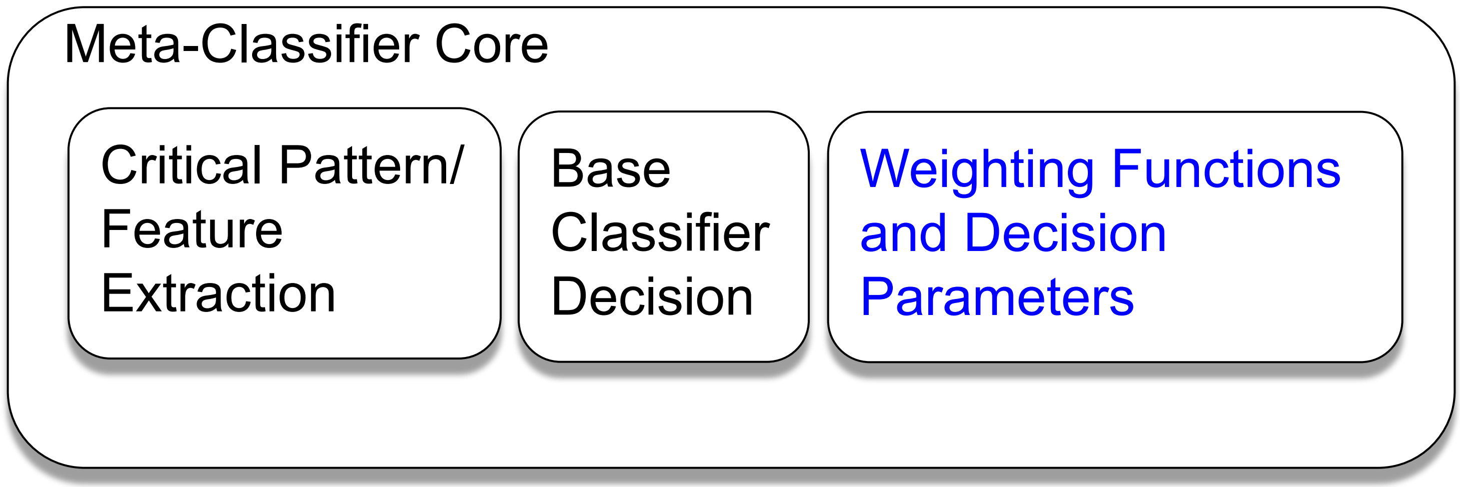
Machine Learning Methods
Good for detecting **new/previously unknown hotspots**

A New Unified Formulation (EPIC)
Good for detecting all types of hotspots with advantageous accuracy/false-alarm (**Meta-Classifier**)

- Meta-Classification combines the strength of different types of hotspot detection techniques

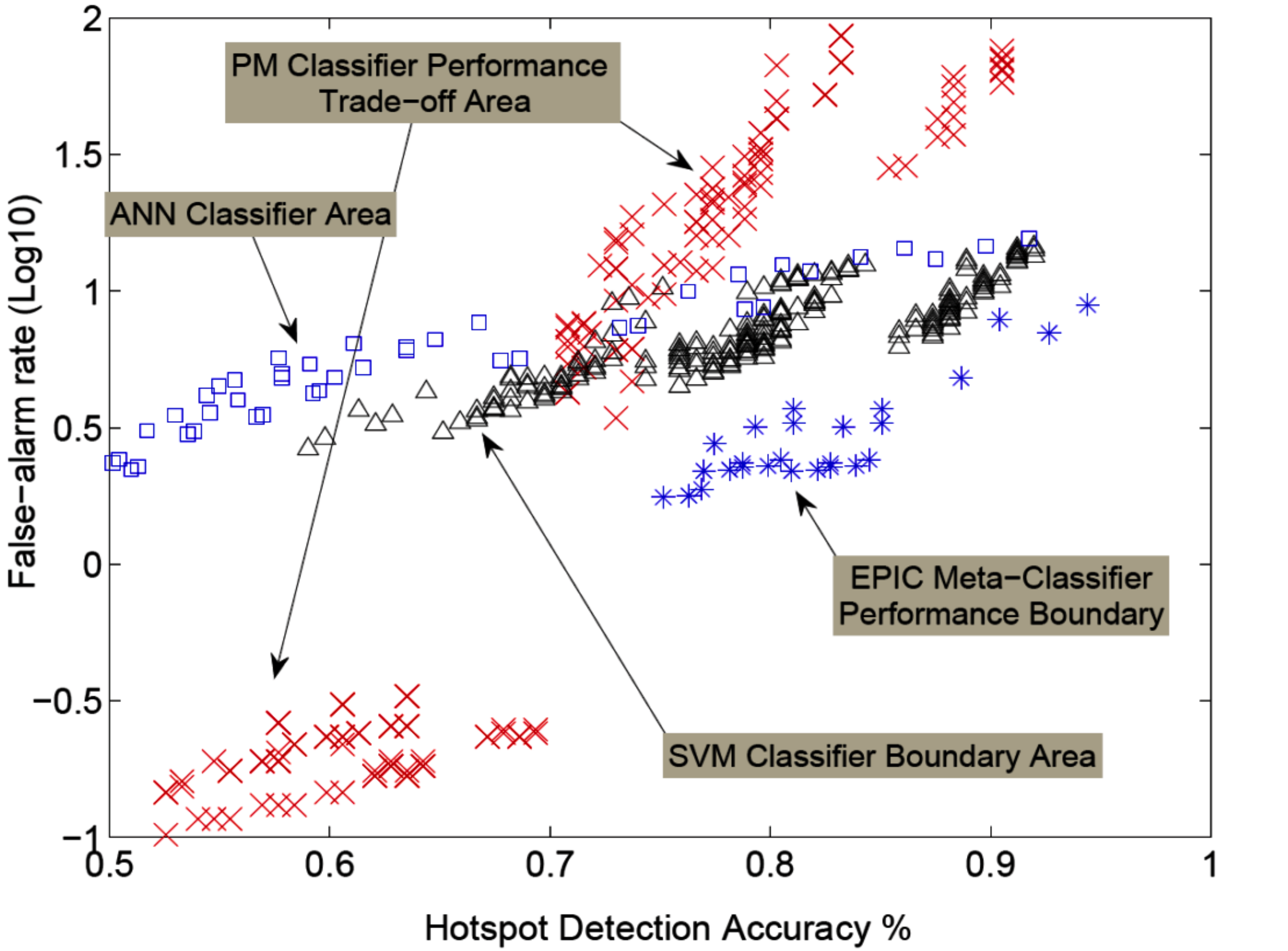
[Ding et al, ASPDAC 2012]

Components of Meta-Classifier Core

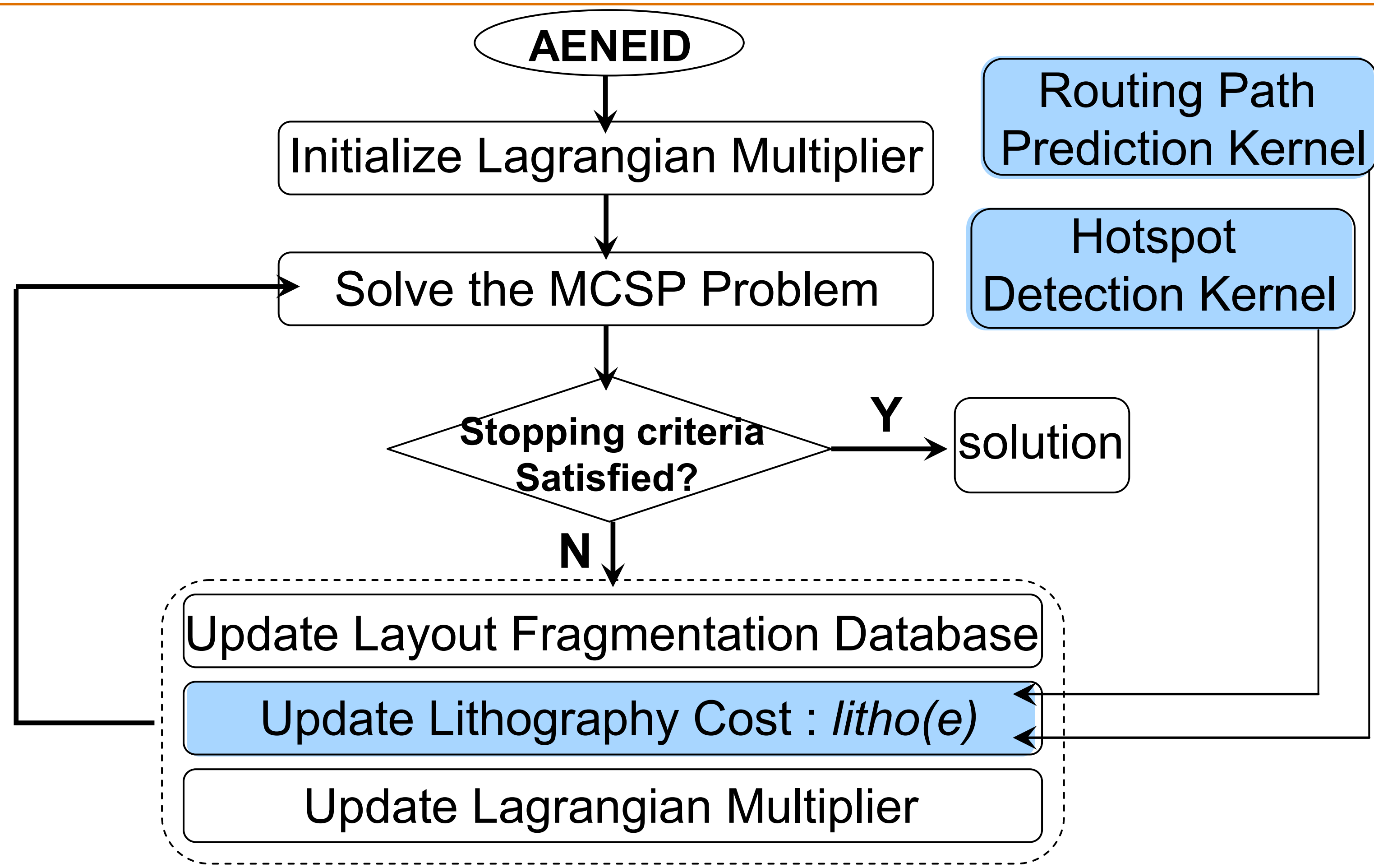


- Base classifier results are first collected
- Weighting functions to make the overall meta decision (e.g., quadratic programming)
- Threshold with accuracy and false-alarm trade-off

False-alarm Rate and Accuracy



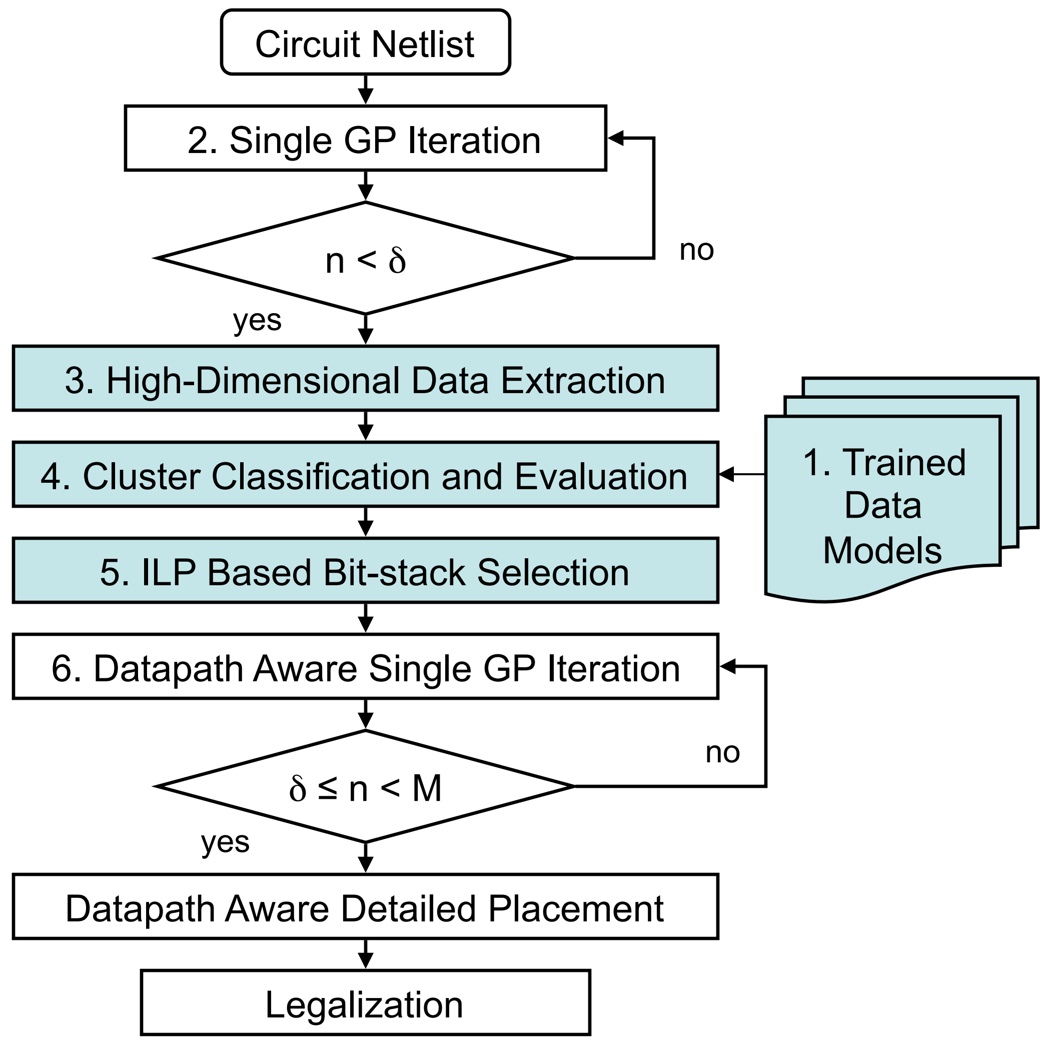
AENEID Router [Ding+, DAC'11]



- Using the machine learning models, we built a new detailed router AENEID to avoid hotspot patterns

Machine Learning for Placement

- Data mining and extraction based on not just graph but also physical information
- We can extract data-path like structures even for "random" logics
- Use them to explicitly guide placement
- Very good results obtained cf. other leading placers like simPL, NTUPlace, mPL, CAPO



[Ward+, DAC' 12]

Abstraction to Logic Synthesis & Above?

- Can we further extend the abstraction up to logic synthesis?
  - Not just lithography hotspot, but other hotspots such as reliability metrics including BTI, oxide breakdown
- Machine learning to raise the abstraction
- NSF/SRC FRS program (started April 1, 2003)
  - E.g., Deming Chen and I have a collaborative project across lower level PD to high level synthesis
- NSF/SRC/DFG Cross-Layer Resilience Workshop in Austin, July 11 and 12

Extreme Scaling and Beyond

- Continuing pushing the envelope, 14nm, 11nm, 7nm (ITRS)
  - Double/triple Patterning
  - Emerging Nanolithography
  - Novel design tools and methodologies
- Vertically – 3D IC integration

Thermal/Mechanical Stress

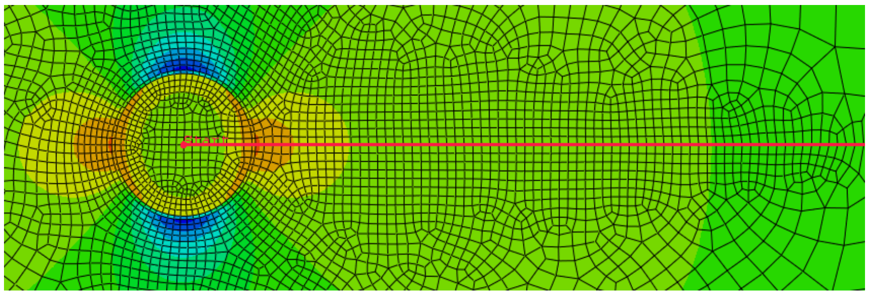
Material	CTE in 10^{-6}/K at 20°C
Si	3
W	4.5
Cu	17

CTE : Coefficient of thermal expansion

TSV: 250 °C ~400 °C process (Higher than operating temperature)
Since Cu has larger CTE than Si ➔ tensile stress in Si near TSV.

< Tensile stress >

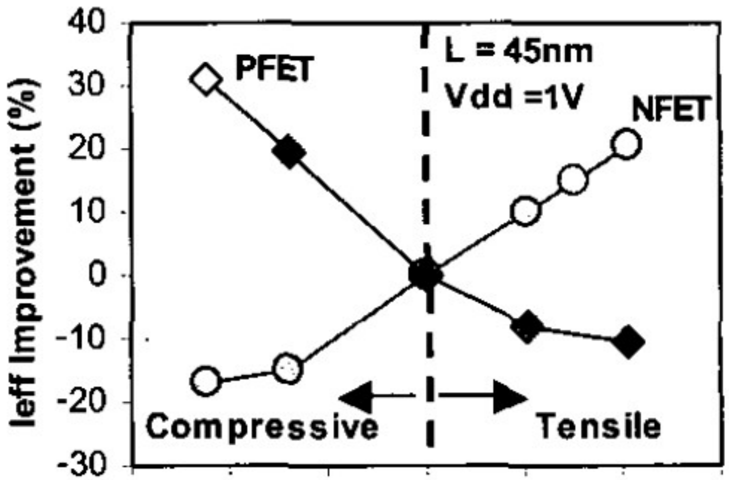
Cu TSV Silicon



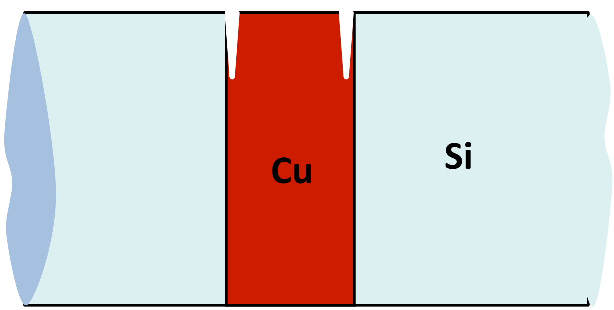
FEA simulation structure of a single TSV
variables: 400K, Memory: 2GB
runtime: 40min

Stress => Variability/Reliability

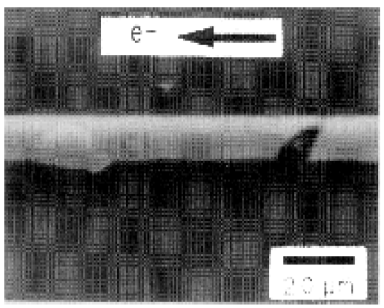
- Systematic Variations
  - Mobility
  - Timing



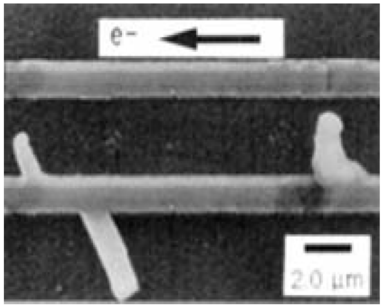
- Reliability (interfacial crack, EM, etc.)



Interfacial Crack



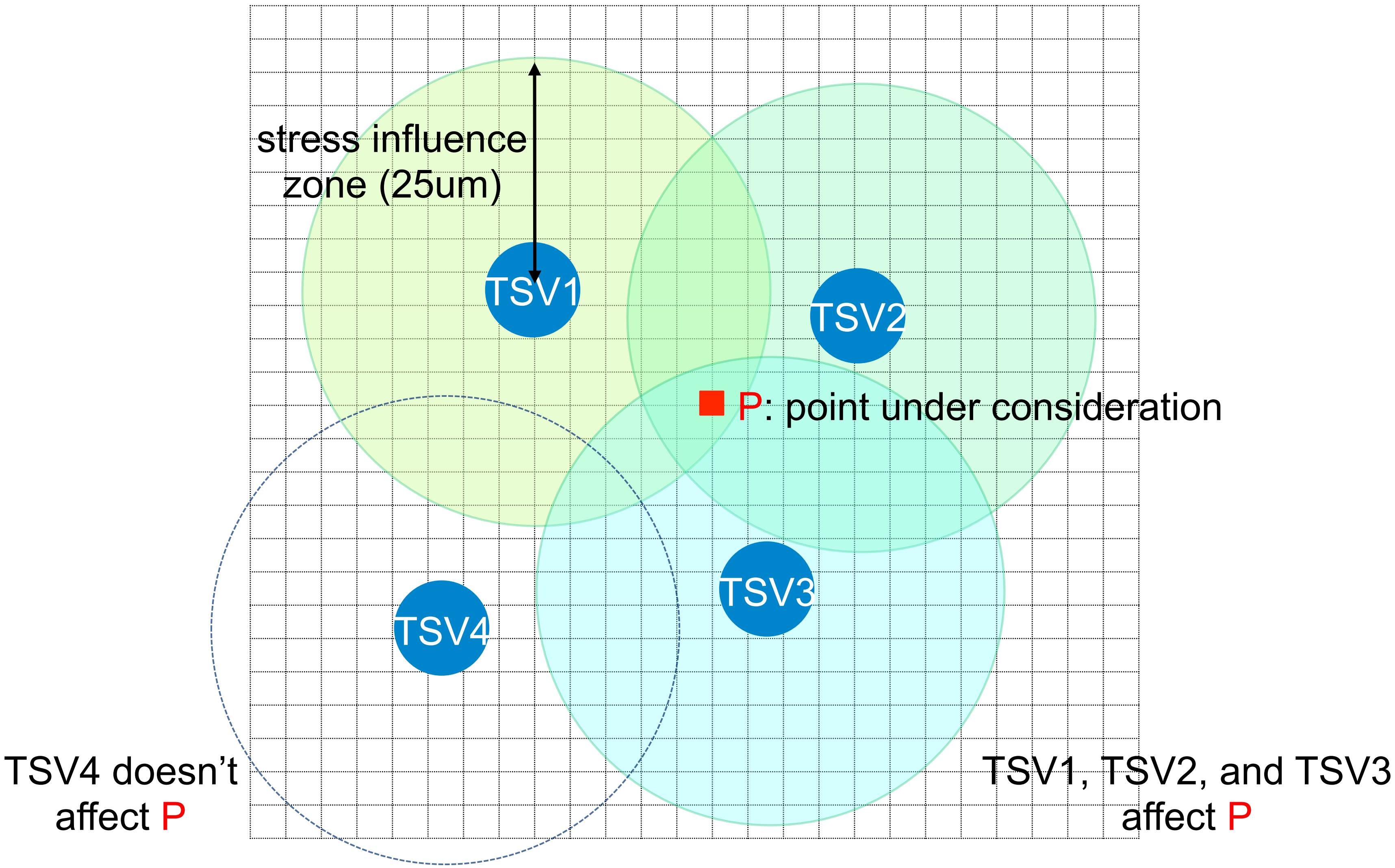
Electromigration Effect – Open



Electromigration Effect – Short

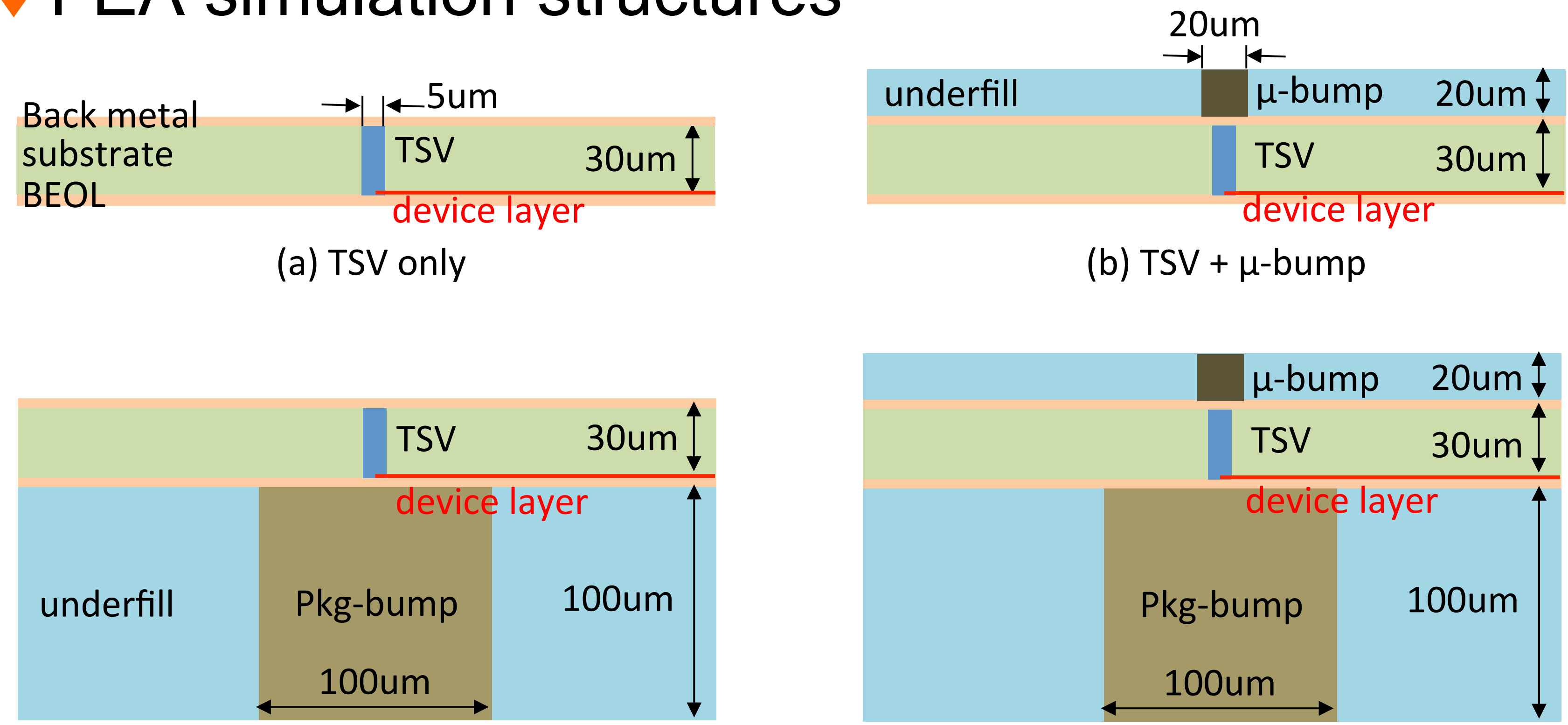
Lateral Linear Superposition [ECTC'11, DAC'11]

- Full-chip stress analysis considering multiple TSVs



Chip-Package Co-Analysis of Stress

- FEA simulation structures



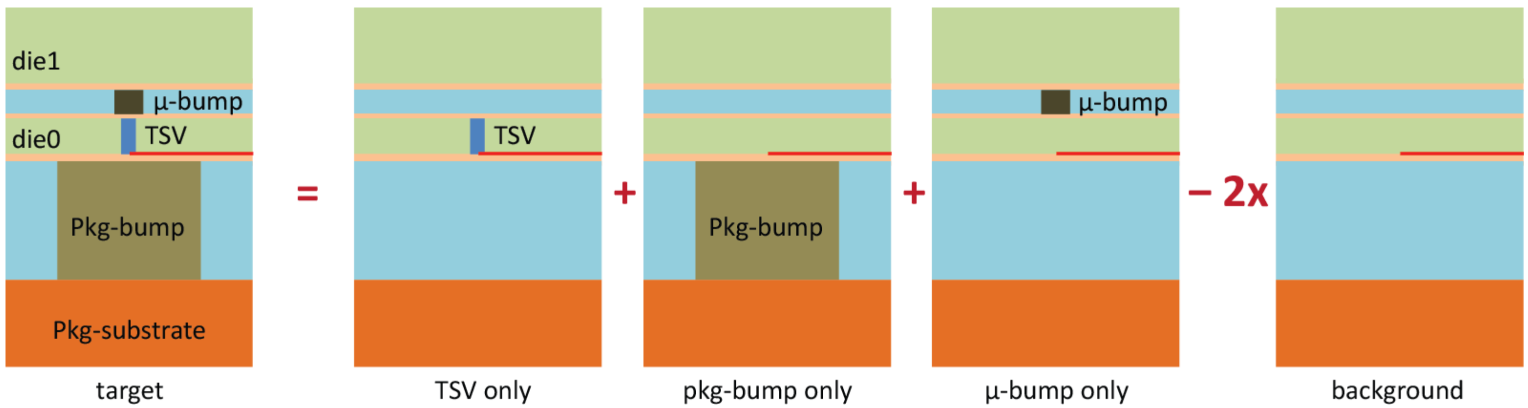
(a) TSV only (b) TSV + μ-bump (c) TSV + pkg-bump (d) TSV + μ-bump + pkg-bump

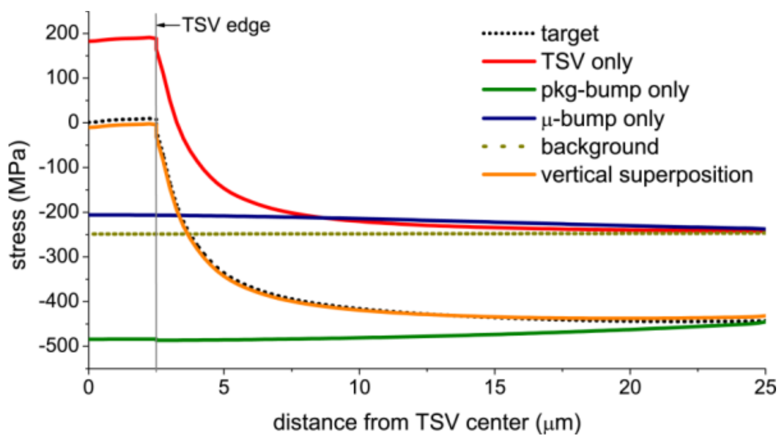
- All structures undergo ΔT = -250°C of thermal load (Annealing/reflow 275°C → room temperature 25°C)

[Jung et al, DAC'12]

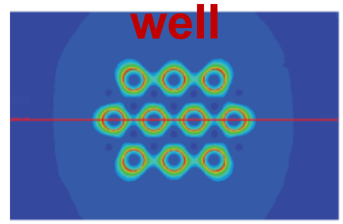
(Lateral &) Vertical Superposition

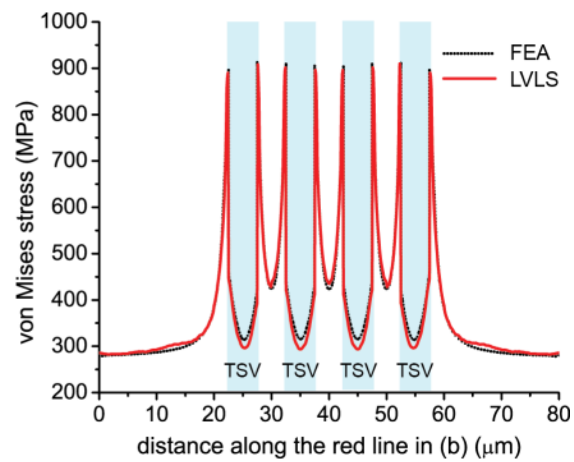
- Stress components are added up "vertically"





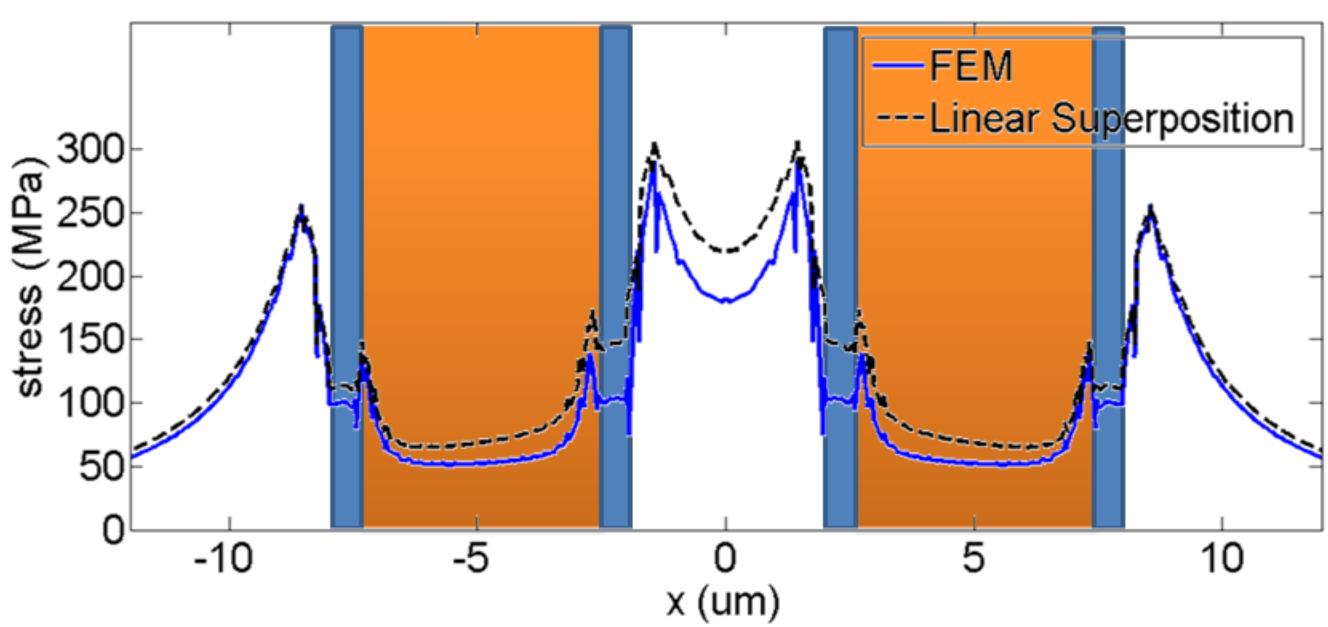
LVLS works well





Interactive Stress & Modeling

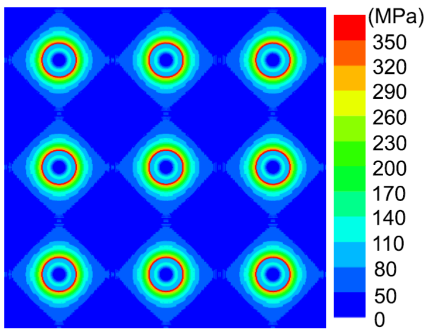
- Linear superposition:
  - Consider the stress contribution of TSV separately
  - May not be accurate enough for very dense TSVs with BCB liner
- Semi-analytical model developed [Li and Pan, DAC'13]
  - Still run fast
  - Can reduce the error by 50%



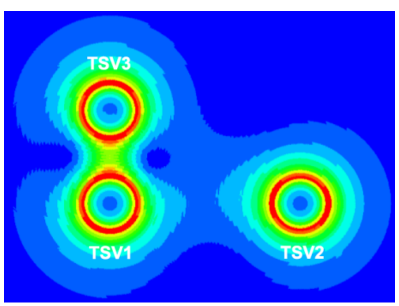
Reliability/Variability Impact of Stress

1. Von Mises Reliability

- Von Mises Yield is function of stress tensor



(a) Von Mises stress with TSV array

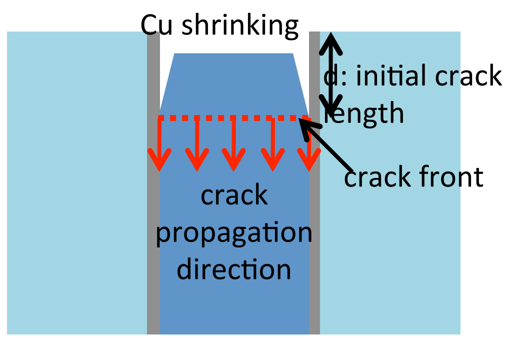


(b) Von Mises stress with three TSVs

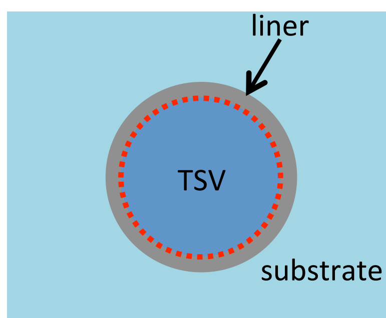
[J. Mitra et al., ECTC'11]

2. Crack: Energy release rate (ERR)

- TSV stress affects ERR of TSV structure → aggravate crack



(a) Side view

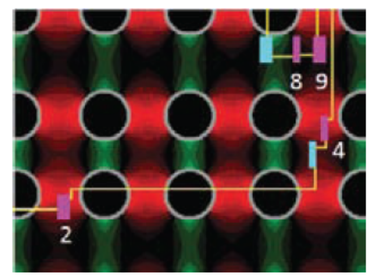


(b) Top view

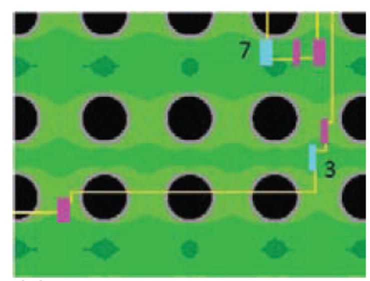
[M. Jung et al., ICCAD'11]

3. Mobility/ V_{th} variation of MOS

- TSV stress changes mobility of hole/electron → timing, V_{th} variation



(a) Hole mobility variation



(b) Electron mobility variation

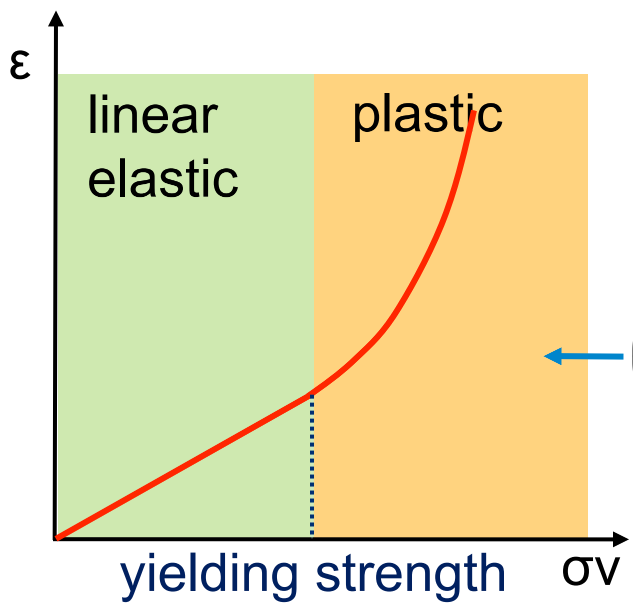
[J. Yang et al., DAC'10]

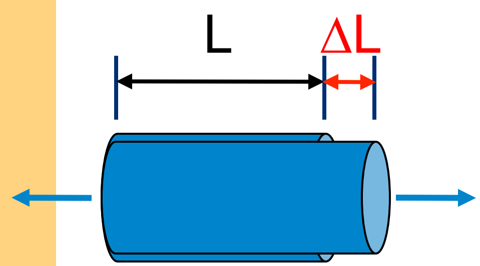
From Stress to Reliability

- Von Mises Reliability Metric

$$\sigma_v = \sqrt{\frac{(\sigma_{xx}-\sigma_{yy})^2+(\sigma_{yy}-\sigma_{zz})^2+(\sigma_{zz}-\sigma_{xx})^2+6(\sigma_{xy}^2+\sigma_{yz}^2+\sigma_{zx}^2)}{2}}$$

- Physical meaning



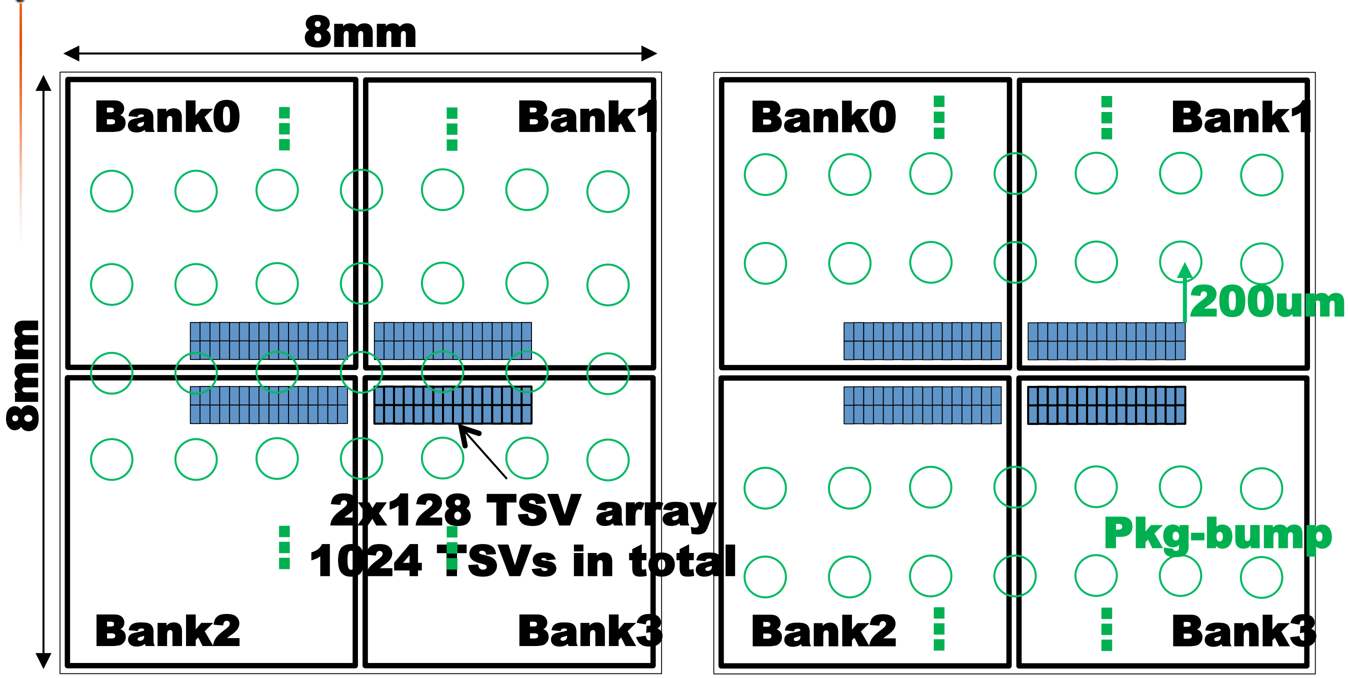


strain $\varepsilon = \frac{\Delta L}{L}$

If σ_v > yielding strength, deformation will be permanent and non-reversible

Yielding strength
- Cu: 225 ~ 600 MPa
- Si: 7,000 MPa

Wide I/O 3D DRAM



(a) Pkg-bumps are placed underneath TSV arrays

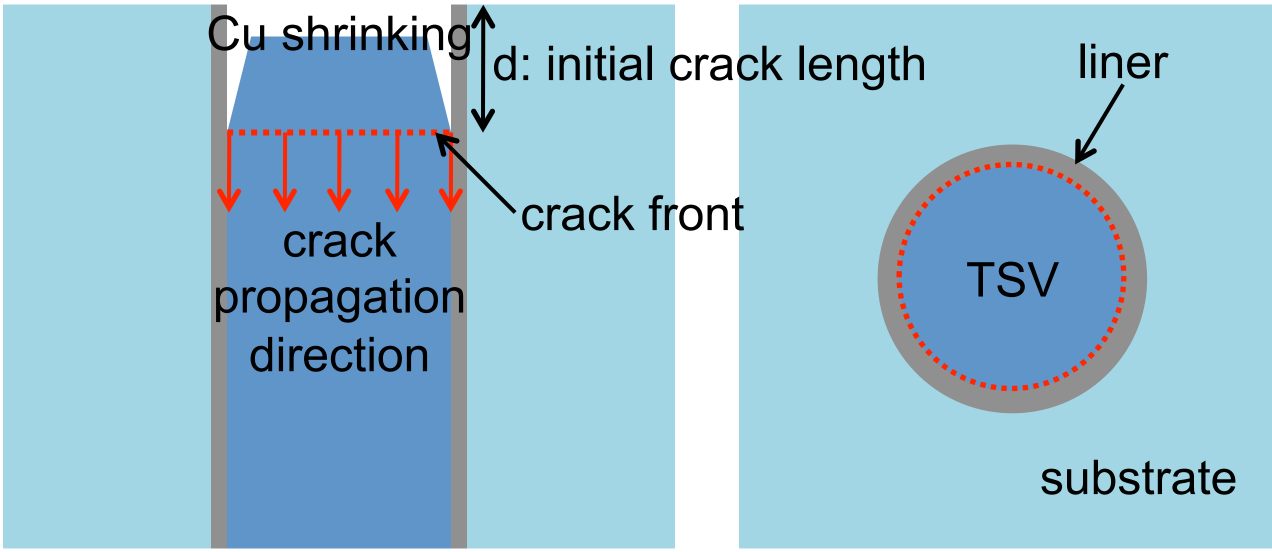
(b) Pkg-bumps are placed 200um apart from TSV arrays

case (b) shows that chip/package co-design can greatly reduce mechanical reliability problem in TSV-based 3D ICs

case	von Mises stress distribution (MPa)				
	780-810	810-840	840-870	870-900	900-930
(a)	30	114	52	220	608
(b)	182	842	0	0	0

[Jung et al, DAC’12]

TSV Interfacial Crack



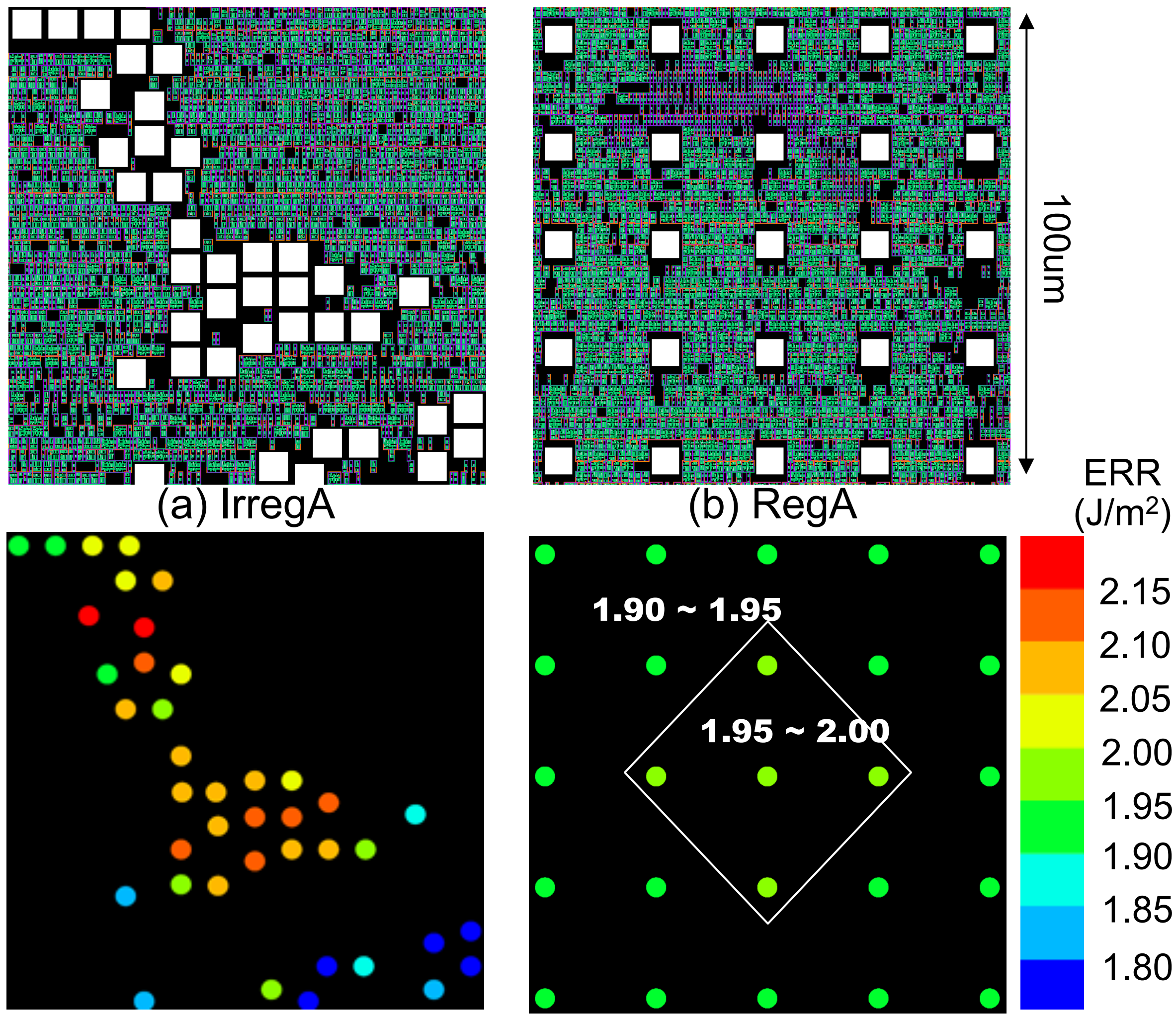
(a) Side view (b) Top view

- **Cu shrinks faster than Si under negative thermal load (ΔT = -250°C)**
- **Model through Energy Release Rate (ERR)**
- **Full chip model with design-of-experiments of different layout styles and multiple TSV structures**

[Jung et al, ICCAD'11]

Full-Chip Crack Analysis and Study

- Regular vs. irregular TSV arrays

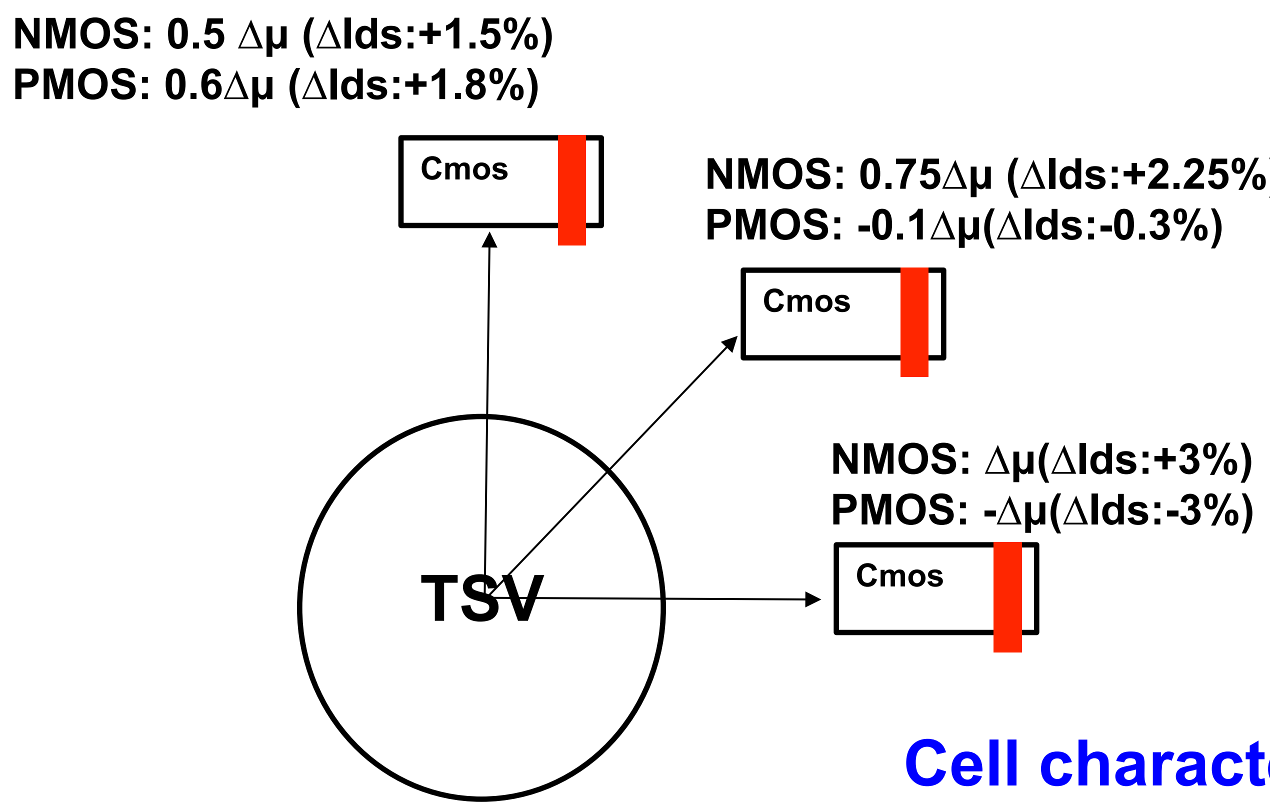


(a) IrregA (b) RegA
(c) ERR map of IrregA (d) ERR map of RegA

Stress Effect on Mobility & Current

CMOS (Stress: 200MPa, R=r)

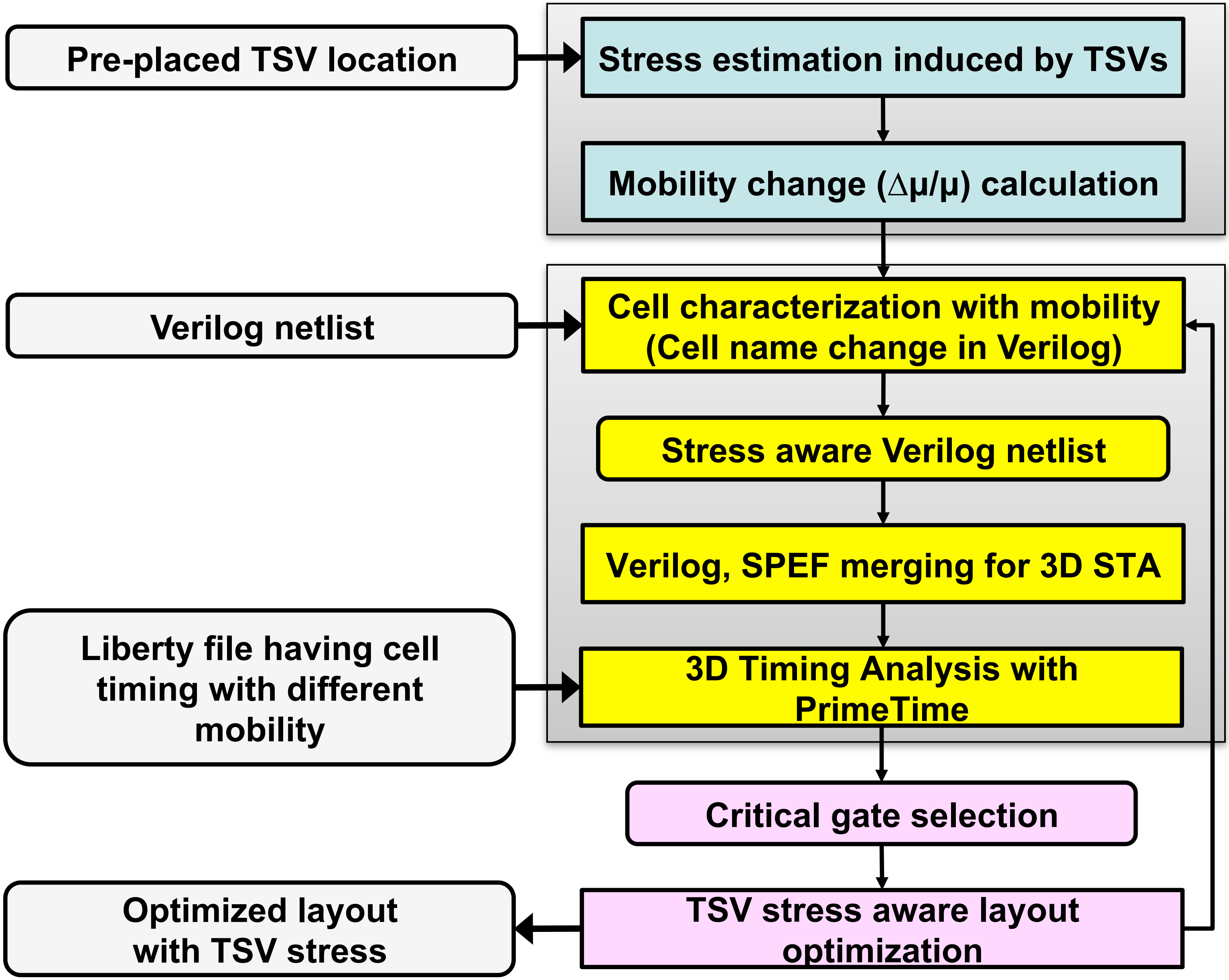
[Yang+, DAC' 10]



FS corner

Cell characterizations based on distance and orientation are needed

Stress Aware Design Flow [Yang+, DAC' 10]

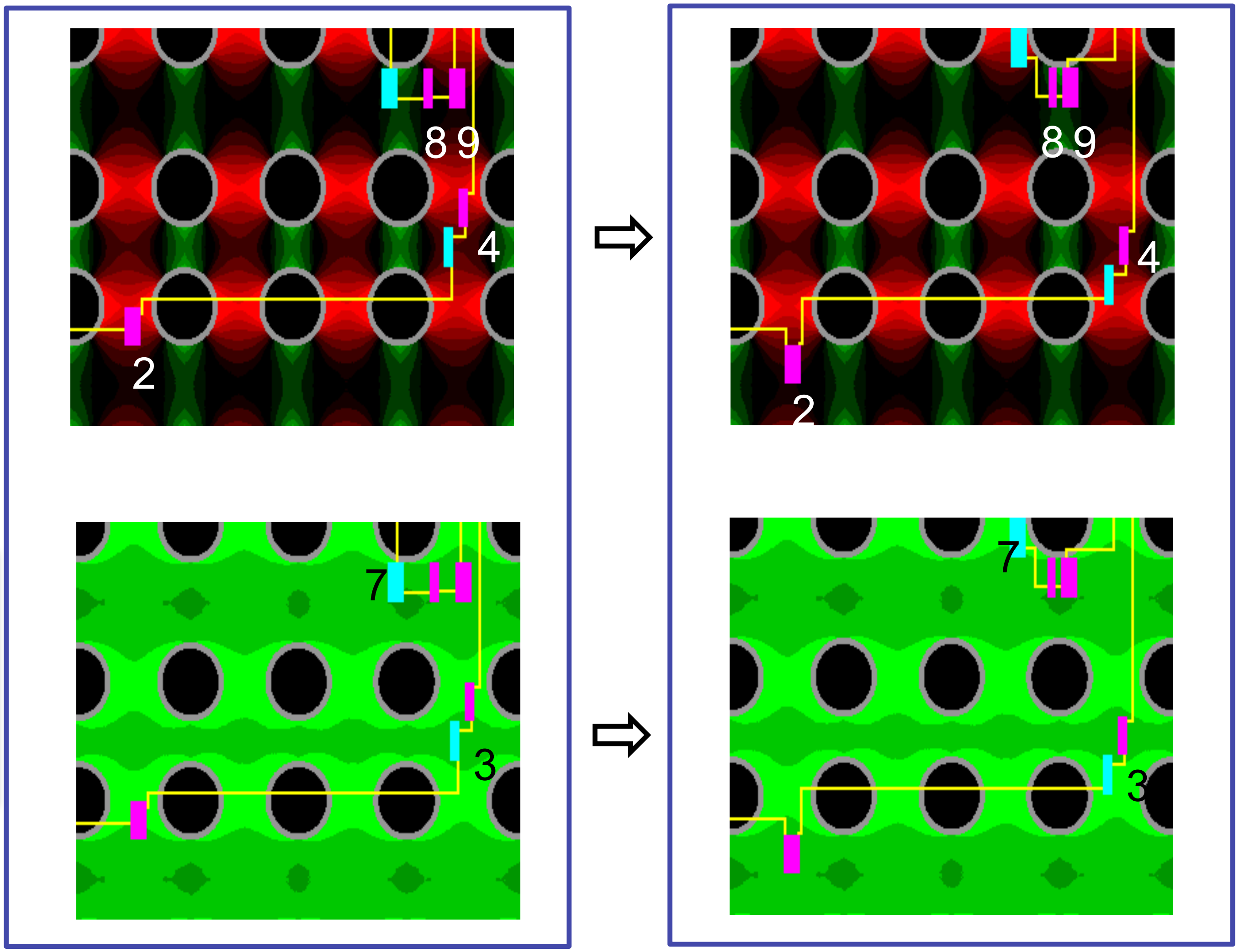


Stress-Aware ECO

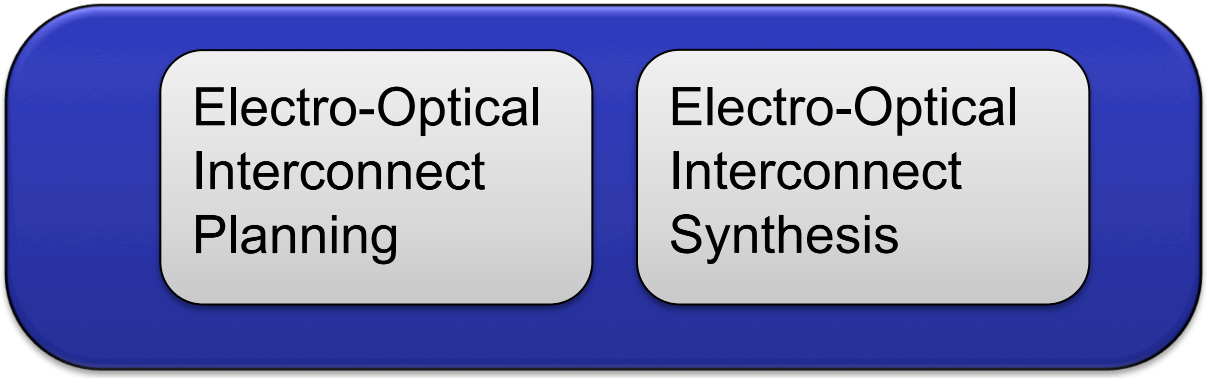
Original cell placement | **After cell perturbation**

Rising critical optimization with hole contour

Falling critical optimization with electron contour

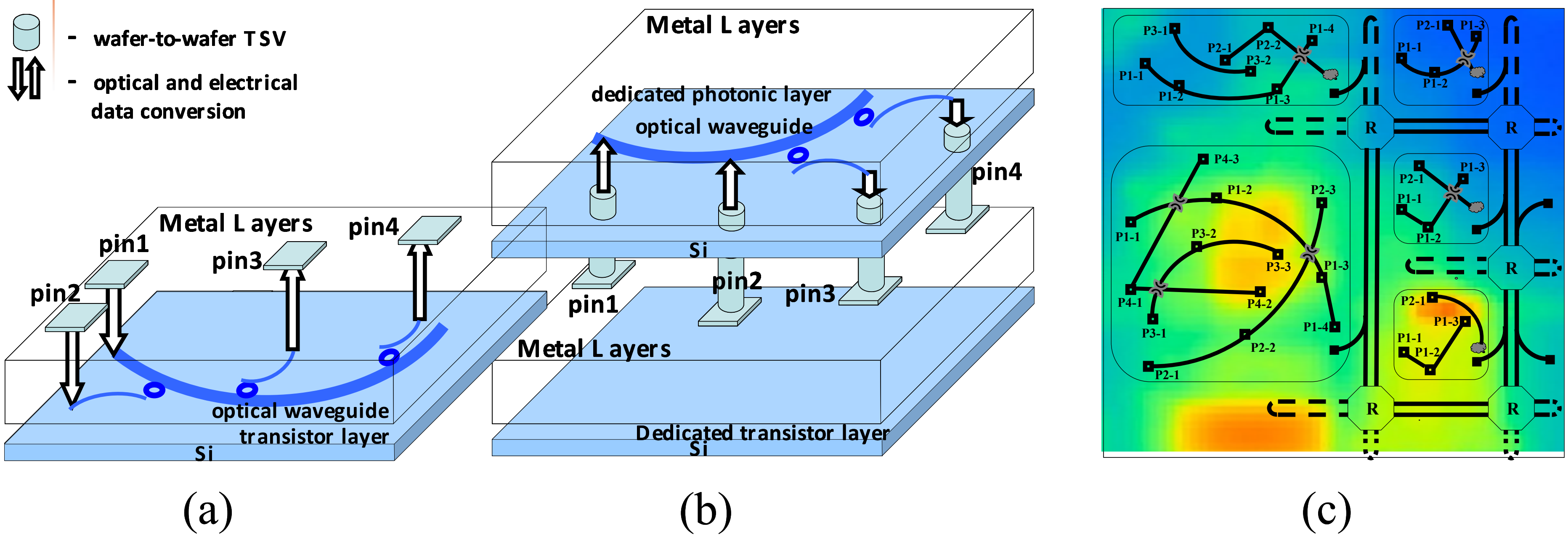


Nanophotonics On-chip Integration



- Holistic Optical Interconnect Planning and Synthesis
  - Co-design and optimization with electrical interconnect
  - Optical interconnect library (OIL)) [Ding+, DAC'09, SLIP'09, and available http://www.cerc.utexas.edu/~ding/oil.htm]
  - WDM, partitioning, routing, …
- Nanophotonics is a very active field
- Many new research problems for CAD community!

Case Study 1: O-Router [Ding et al, DAC'09]

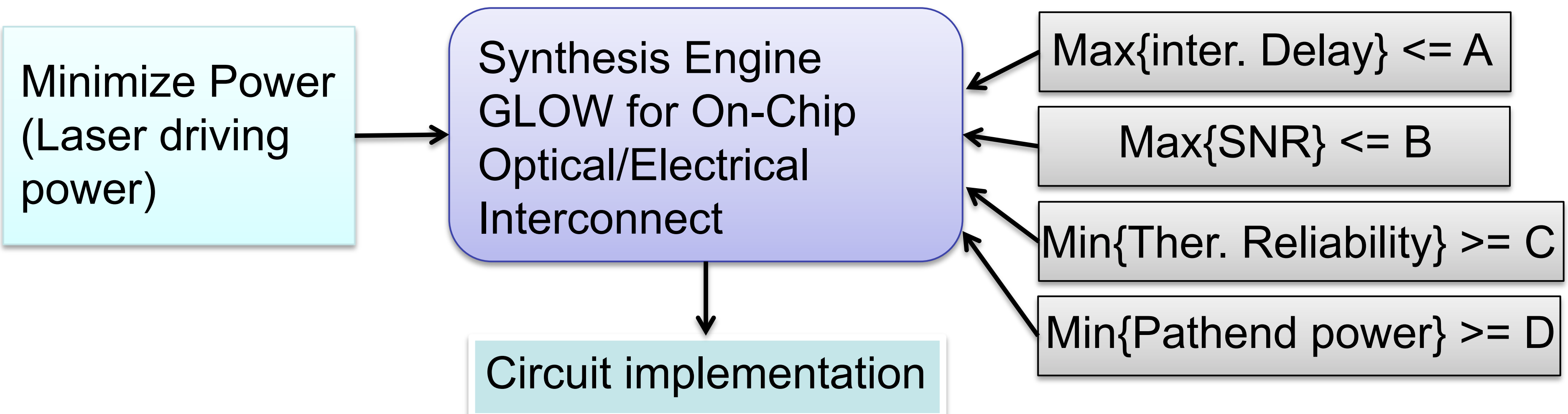


- Objectives: performance (throughput, latency, power), cost ($$, economics)
- Constraints (SNR, signal integrity, reliability, system-level reqs.)

Case Study 2: GLOW

[Ding et al, ASPDAC'12]

- Global router for low-power thermal-reliable optical interconnect synthesis using Wavelength Division Multiplexing (WDM)



Conclusion

- Optical lithography still pushing ahead for 14nm, 11nm, 7nm ➔ extreme scaling
  - Multiple patterning, EUV, DSA, and hybrid lithography
- Design enablement with lithography capability co-optimization from mask to physical synthesis (and logic/high-level synthesis?)
  - Cross-layer resilience
- Horizontal scaling ➔ Vertical scaling: 3D-IC
  - Reliability/Variability issues
- New material/devices ➔ new CAD paradigms and tools

Acknowledgment

- Support/collaboration from NSF, SRC, NSFC, Sematch, IBM, Intel, Oracle, Fujitsu, Qualcomm, Synopsys, Mentor Graphics, …
- The materials in this talk include results from many former/current PhD students at UTDA
  - They are the ones who did the real work!
- Collaborators
  - DFM: IBM, Intel, Globalfoundries, Mentor, Synopsys, etc.
  - 3D-IC: Prof. Sung Kyu Lim's group at Georgia Tech
  - Optical: Prof. Ray Chen's group at UT Austin