Physical Design Techniques for Optimizing RTA-induced Variations

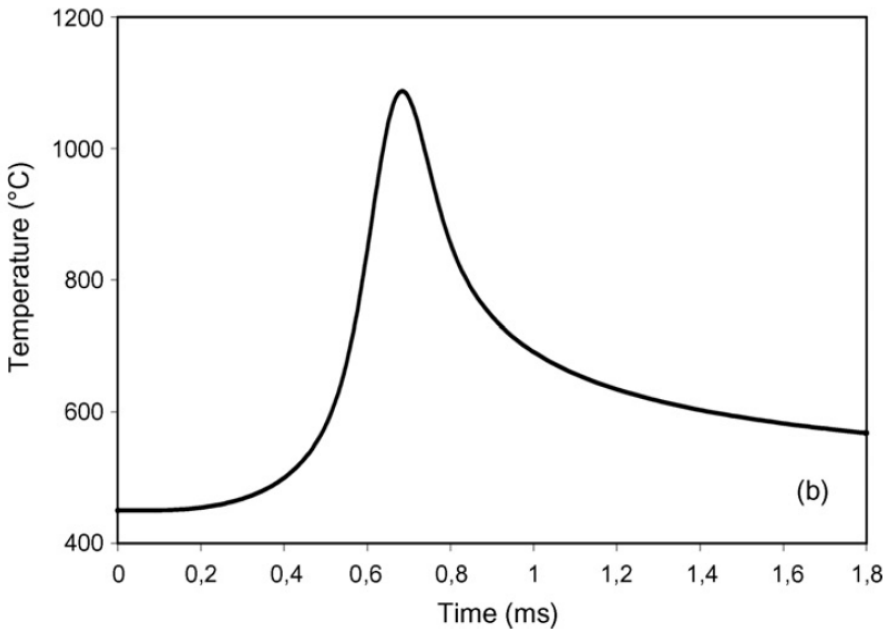
Yaoguang Wei	Jiang Hu	Frank Liu	Sachin S. Sapatnekar
University of Minnesota	Texas A&M University	IBM Austin Research Lab	University of Minnesota

Outline

- Background and motivations
- RTA-driven floorplanning
- Dummy polysilicon filling
- Experimental results
- Summary

Rapid thermal annealing (RTA)

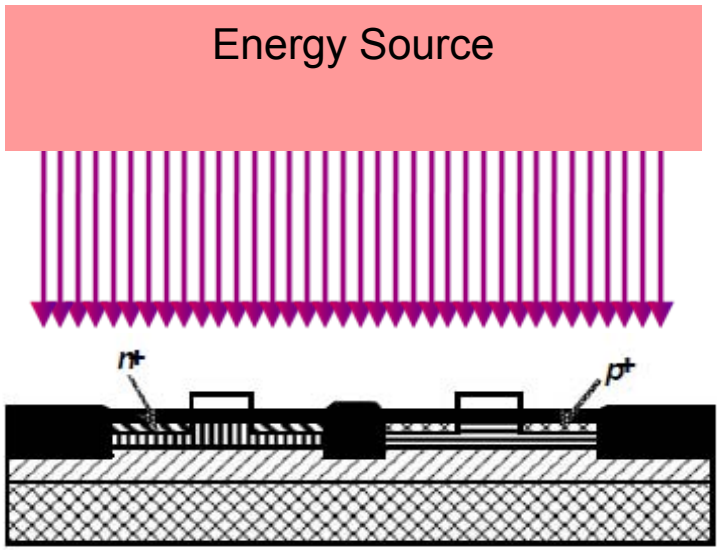
- High temperature (e.g., 1200°C)
  - Provides a high level of dopant activation after implantation
- A short period (e.g., 1ms)
  - Restricts dopant diffusion
- Used in contemporary ultra-shallow junction technologies



Temperature-time profile in a laser RTA [Colin, 2008]

Surface temperature variation

- Pattern effect
  - Temperature variation induced by different reflectivity coefficients
  - The reflectivity of exposed shallow trench isolation (STI) substantially lower
  - Density of the exposed STI primarily determines the annealing temperature



Calculated reflectivity coefficients in RTA

Region	Reflectivity Coefficient
N+ source/drain	0.57
P+ source/drain	0.57
Gate over isolation	0.54
Gate over transistor	0.45
Trench isolation	0.20

[http://bwrc.eecs.berkeley.edu/Classes/IcDesign/EE141_s02/Lectures/Lecture5-Manufacturing.pdf]; [Ye, 2009]

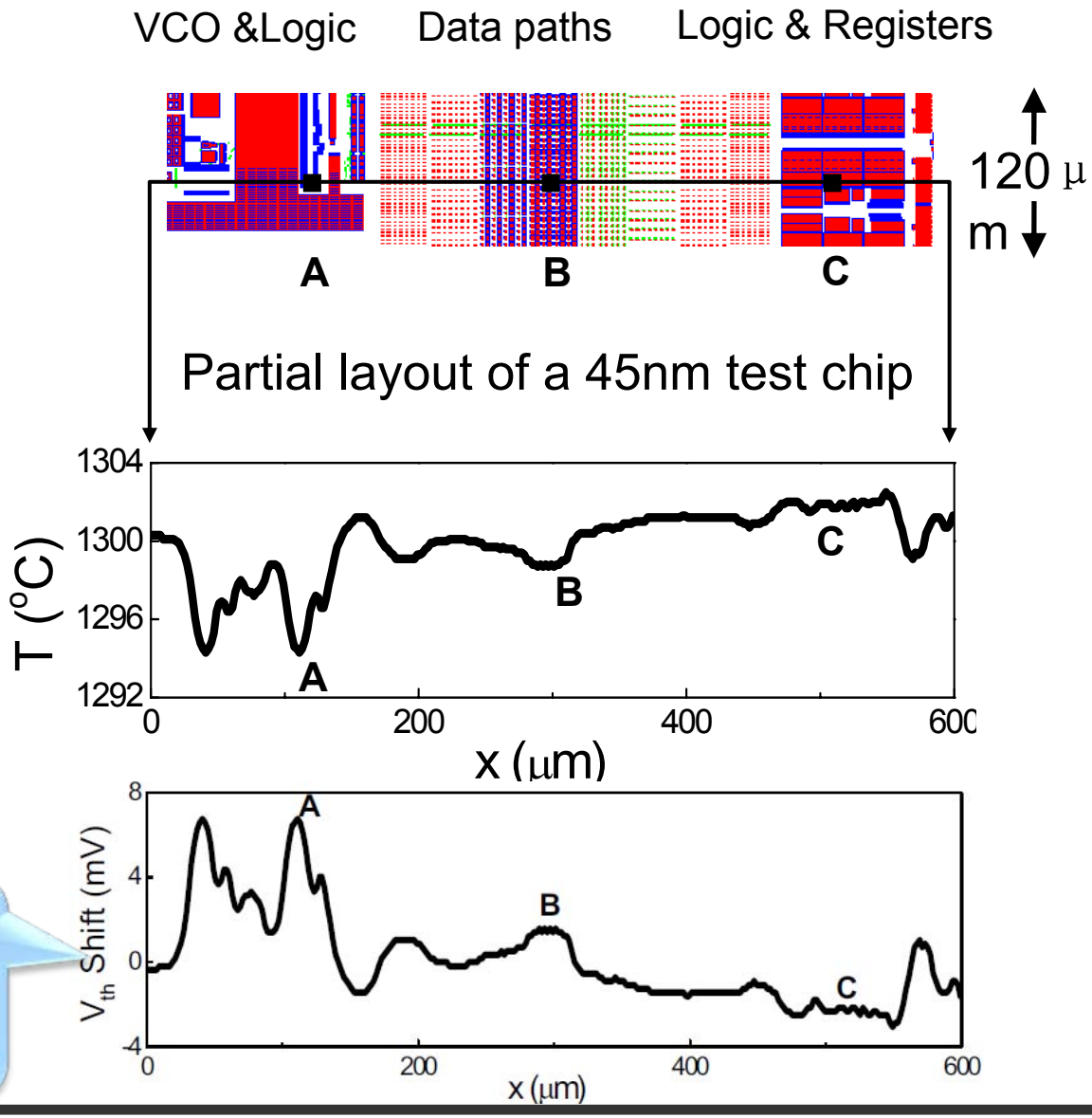
[Ahsan, 2006]

Impact on performance

Pattern variation in layout

Temperature variation in RTA

Circuit performance variation

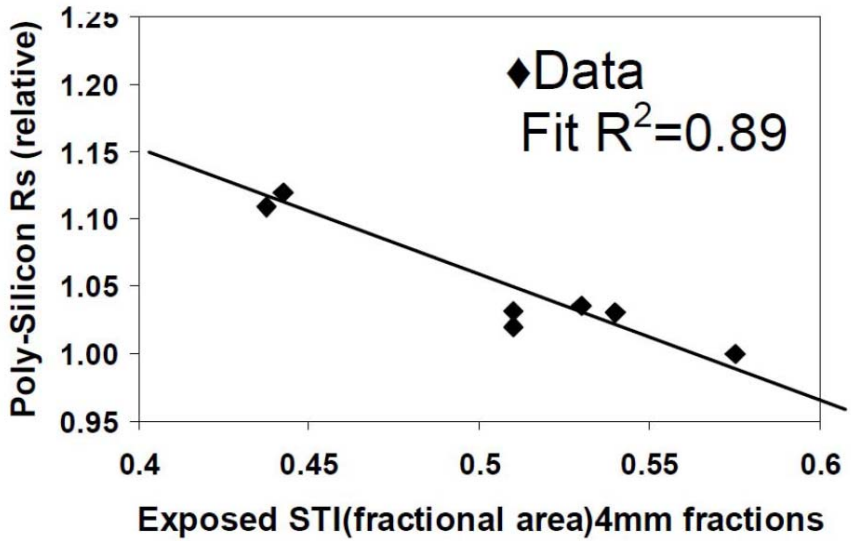


ΔV_{th} variation:
10mV across 600 μ m
20% leakage variation

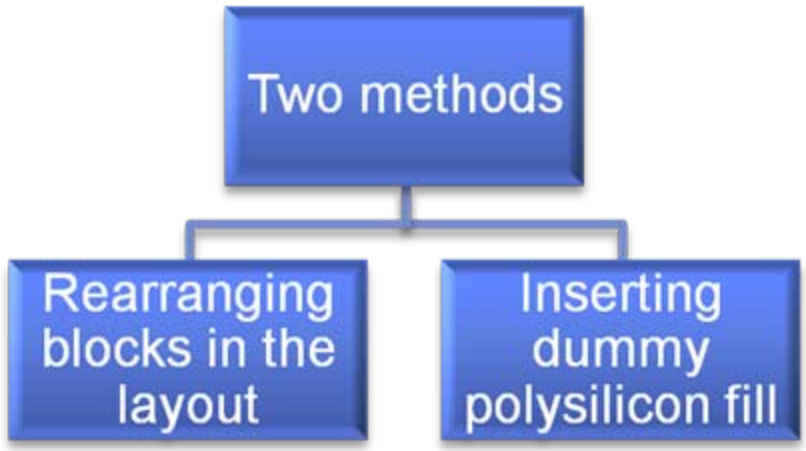
[Ye, 2009]

STI density vs. circuit parameters

- 20% variation in speed observed [Ahsan, 2006]
- Variations in circuit parameters linearly correlates with exposed STI density
- Reduce RTA variations
  - Achieve an even distribution of exposed STI throughout the layout



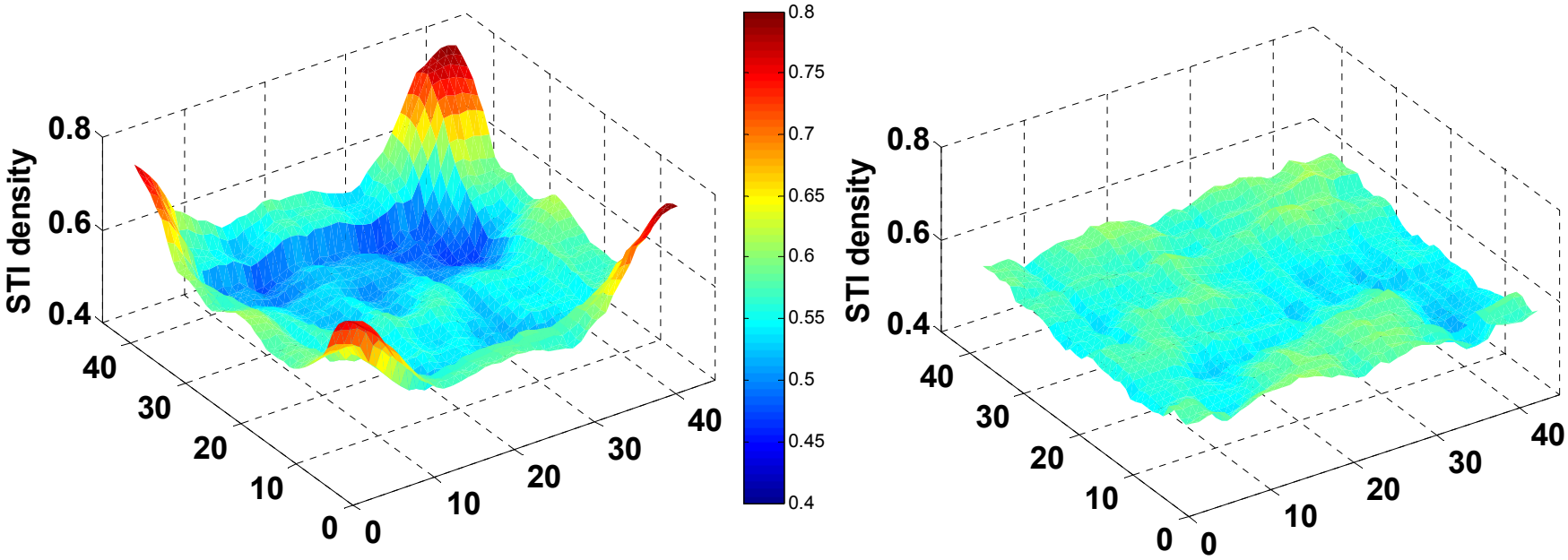
[Ahsan, 2006]



RTA-driven floorplanning

- Floorplanning can improve the variations in the STI density significantly

Topography of STI density for circuit n300

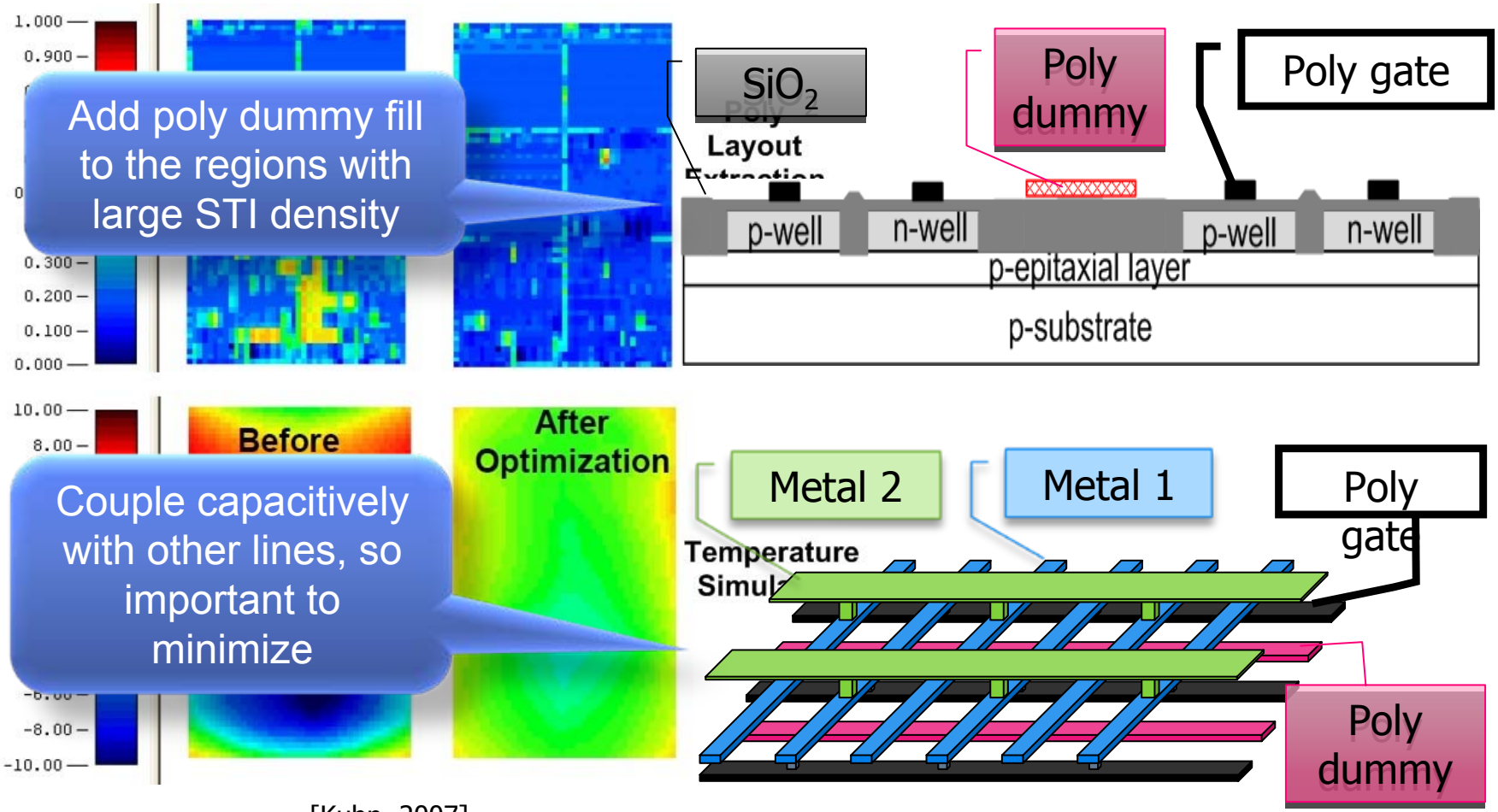


Floorplan without RTA optimization

Floorplan with RTA optimization

Dummy polysilicon filling for RTA

- Insert poly dummies to improve RTA temperature uniformity and reduce variations in circuit parameters

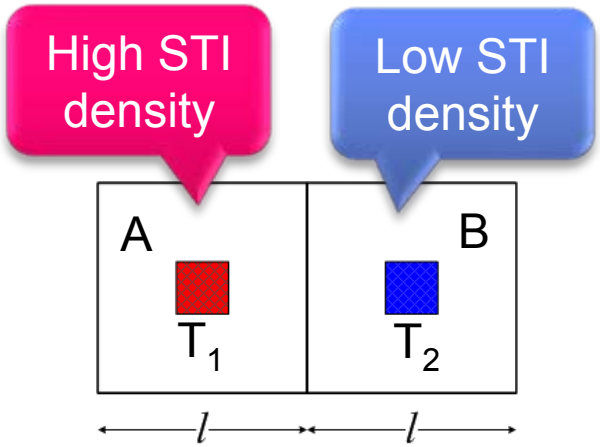


[Kuhn, 2007]

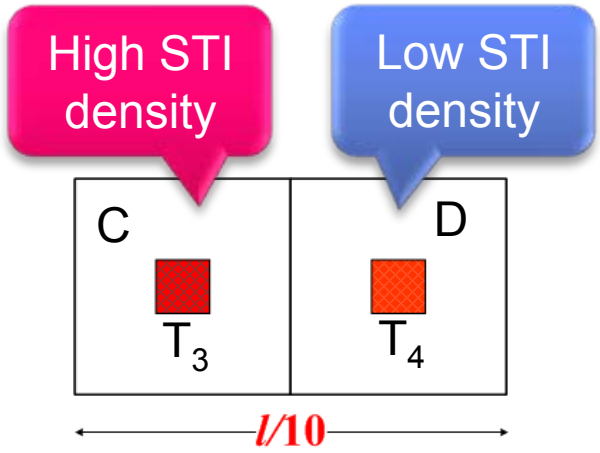
Heat diffusion length

$$l = \sqrt{\frac{\sigma}{c_v \rho} t}$$

σ is thermal conductivity; c_v is specific heat; ρ is density; t is heat diffusion interaction time [Colin, 2008]



Different annealing temperature observed



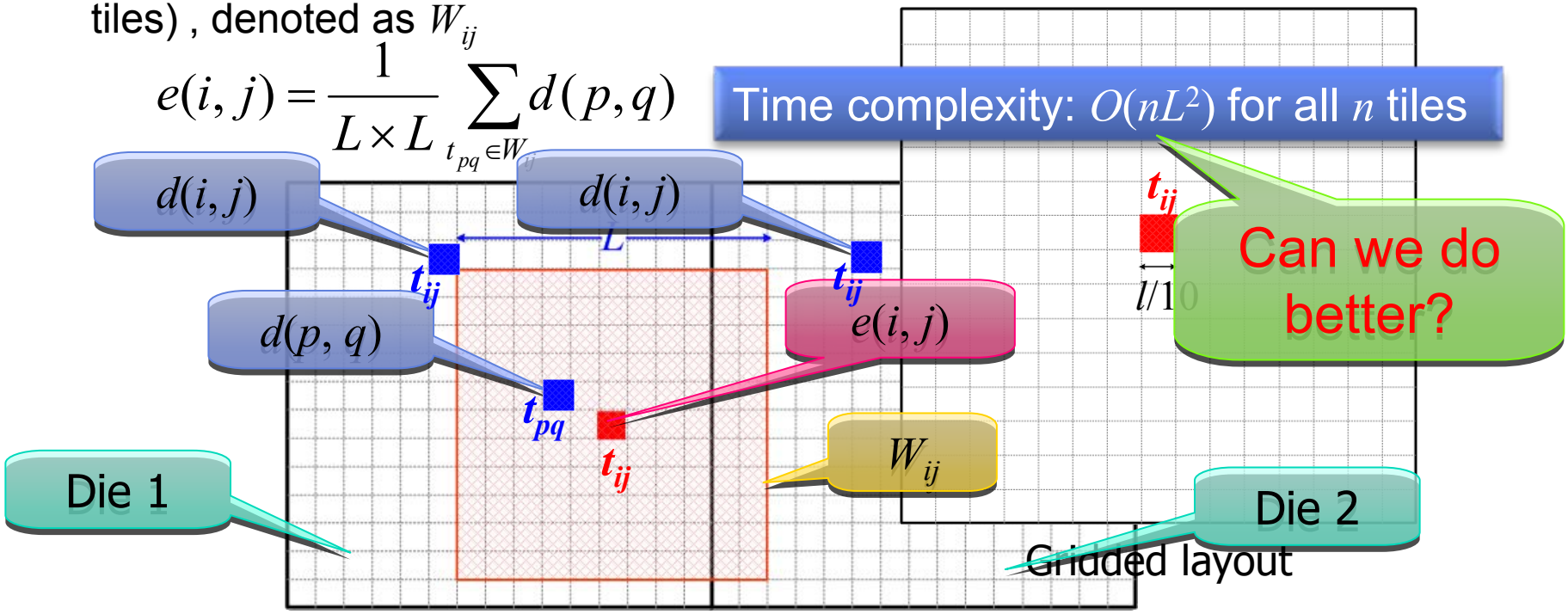
Temperature variation smoothed by heat diffusion

- Empirical model [Ahsan, 2006]
  - Spike RTA in 65nm technology
  - l~ 4mm

Effective STI density

- Tessellate the layout to $M \times N$ tiles
  - Side equal to l/10
  - Each tile t_{ij} has local STI density $d(i,j)$ and effective STI density $e(i,j)$
    - Assume the die is periodically repeated, and then $d(i,j)$ is periodic.
- Local STI density: averaged over the tile
- Effective STI density: averaged over the gating window $L \times L$ (in unit of tiles) , denoted as W_{ij}

$$e(i,j) = \frac{1}{L \times L} \sum_{t_{pq} \in W_{ij}} d(p,q)$$



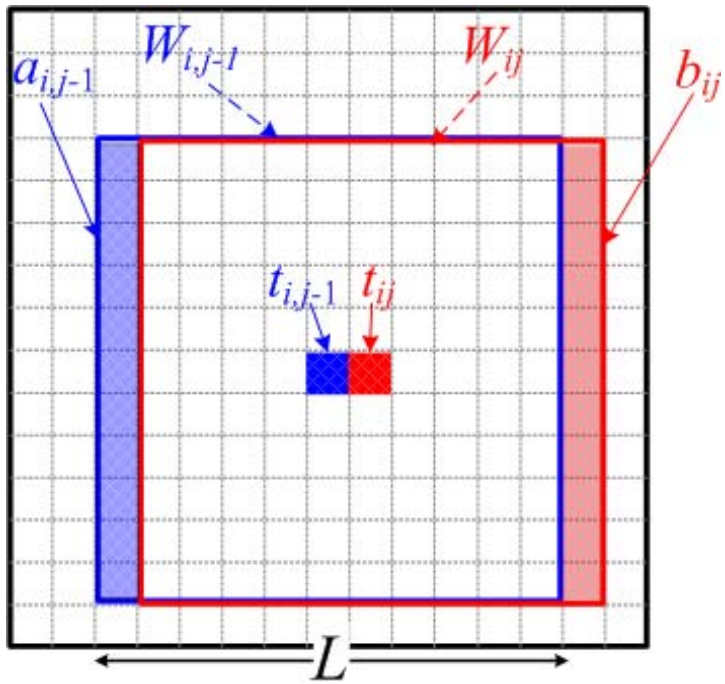
Incrementally compute effective STI density

- Only two rows/columns different between the gating windows of adjacent tiles
- Reuse the previous computation

$$e(i,j) = e(i,j-1) + \frac{D(b_{ij}) - D(a_{i,j-1})}{L^2}$$

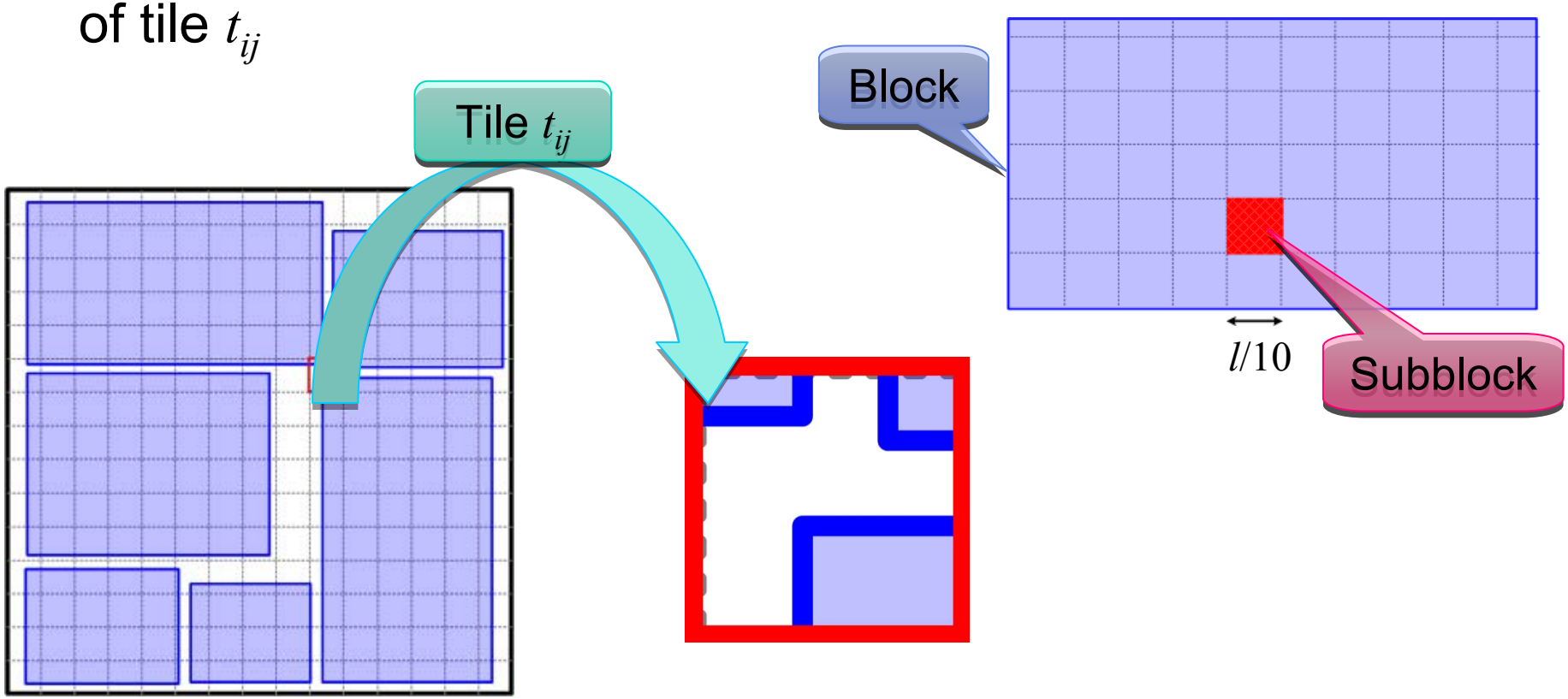
where $D(r) = \sum_{t_{pq} \in r} d(p,q)$ provides the sum of local STI density for the tiles in a region r.

Time complexity: $O(n)$ for all n tiles



Determine the local STI density

- Allow different local STI densities to different subblocks with size $l/10 \times l/10$
- $d(i,j)$ calculated by the STI density averaged over the area of tile t_{ij}



Formulation for RTA-driven floorplanning

- Inputs
  - Circuit blocks, local STI densities of all the subblocks, netlists
- Outputs
  - A legal floorplan
- Objectives
  - Minimize area, wirelength and effective STI density uniformity
- Metrics of the effective STI density uniformity
  - Global variation $R = \max_{\text{all } t_{ij}}\left(e(i,j)\right) - \min_{\text{all } t_{ij}}\left(e(i,j)\right)$
  - Local variation $G = \sum_{\text{all } t_{ij}}\left(\left|h_{ij}\right| + \left|v_{ij}\right|\right)$

    where h_{ij} (v_{ij}) is the horizontal (vertical) gradient of effective STI density of t_{ij}

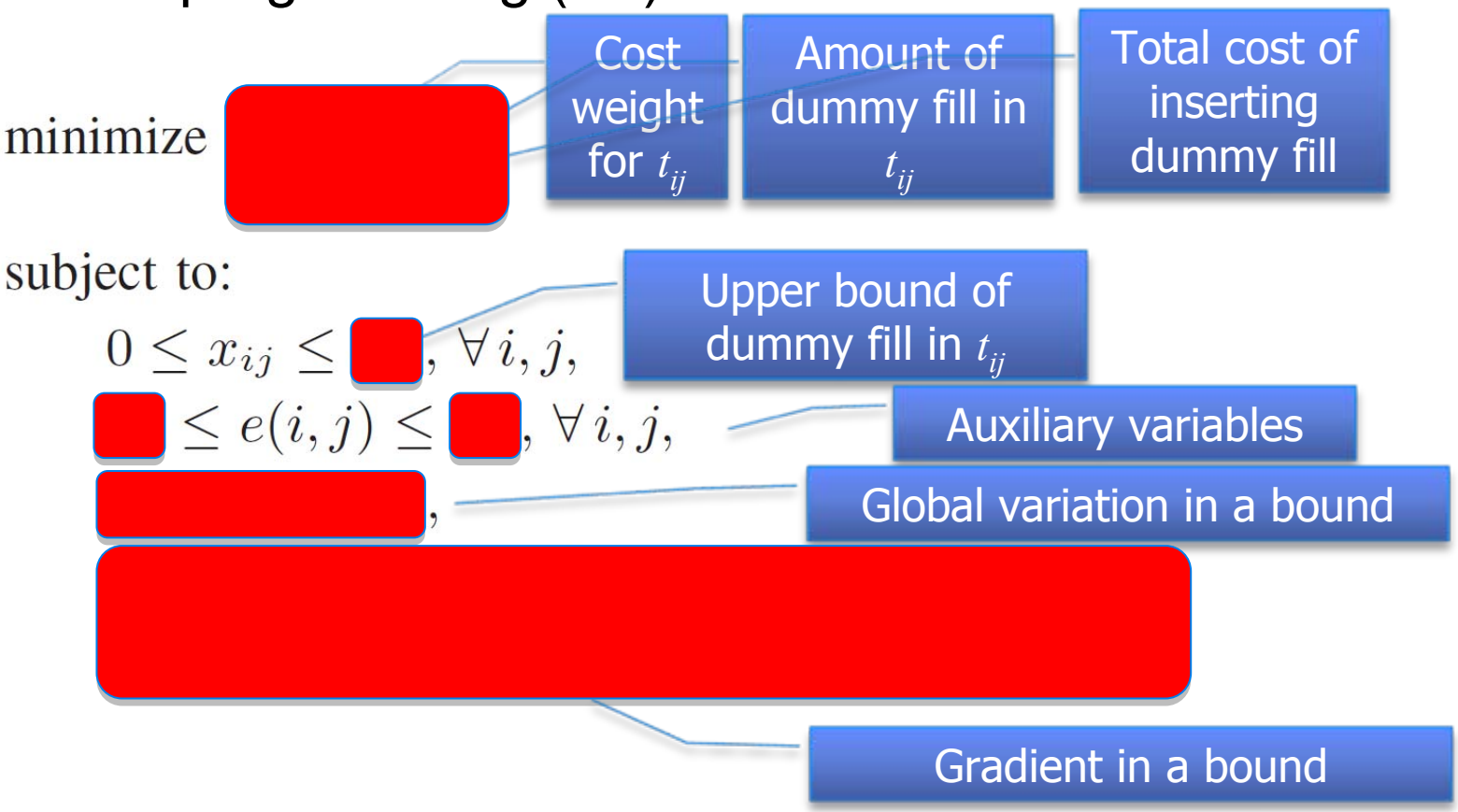
RTA-driven floorplanning

- Cost function: $C = \overline{A} + \alpha\overline{W} + \beta(\overline{R} + \gamma\overline{G})$

  $\overline{A}, \overline{W}, \overline{R}, \overline{G}$: normalized version of area, wirelength, global/local variations in effective STI density
- Adaption of Parquet package [Adya, 2001]
- Two-stage simulated annealing (SA)
- Heuristic
  - Observation
    - A floorplan with larger area contains more dead space
    - Mitigate the variations in STI density
  - Add barriers for accepting a floorplan with better STI uniformity but much worse area

Dummy polysilicon filling

- Based on an adaptation of the oxide CMP fill solution [Tian, 2001]
- Linear programming (LP) formulation

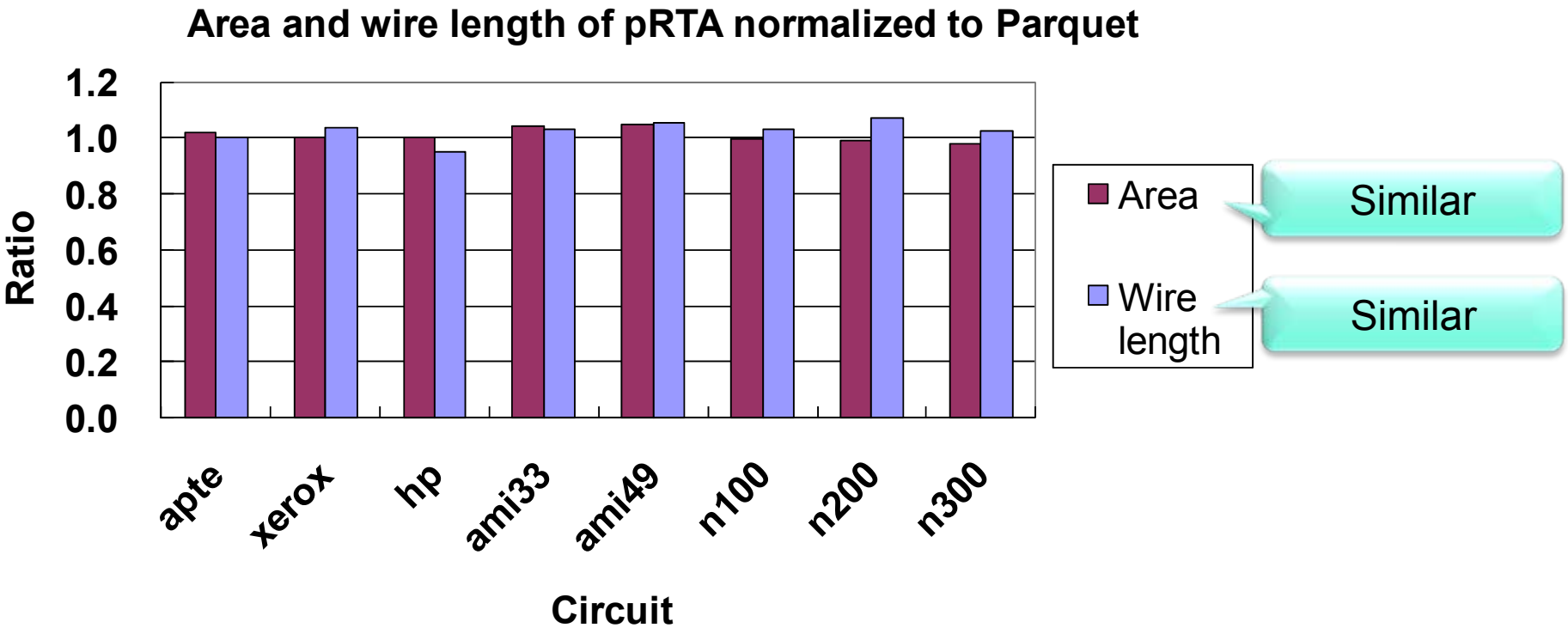


Experiment setup for RTA-aware floorplanning

- Platform
  - A 64-bit Linux machine with a 2.6 GHz AMD Opteron 2218 processor and 2GB memory
- Assumptions
  - Assume RTA profile for 65nm technology with l=4mm [Ahsan, 2006]
  - STI density of dead space to be 1
  - The space between the dies on the wafer is negligible
- Scaling of the benchmarks
  - Scaled areas from about 7mm×7mm to 15mm×15mm
- Generate the local STI density for each subblock randomly
- Parameters
  - Default parameters for Parquet
  - For our algorithm, parameters tuned for ami49
  - Run 10 times

Results for RTA-aware floorplanning

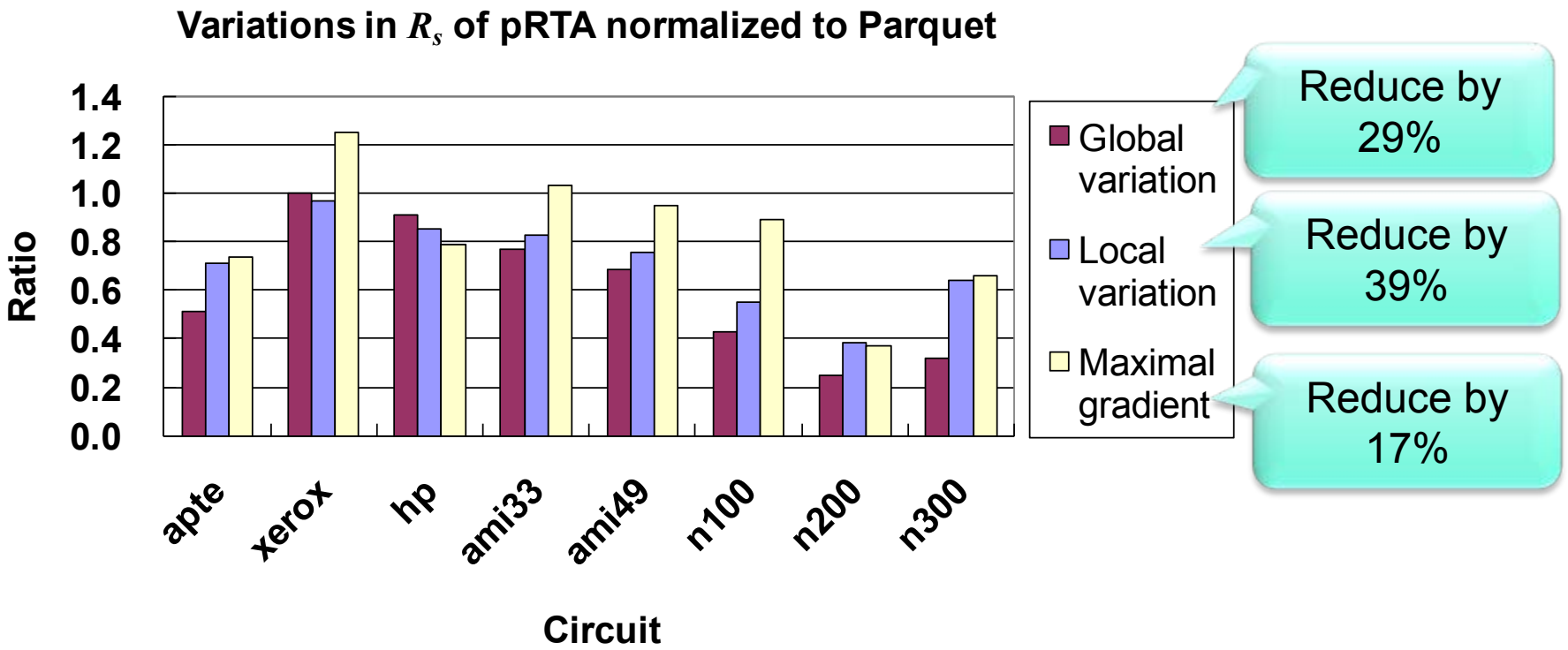
- Name our algorithm as pRTA
- Show the results of pRTA normalized to Parquet



Results for RTA-aware floorplanning

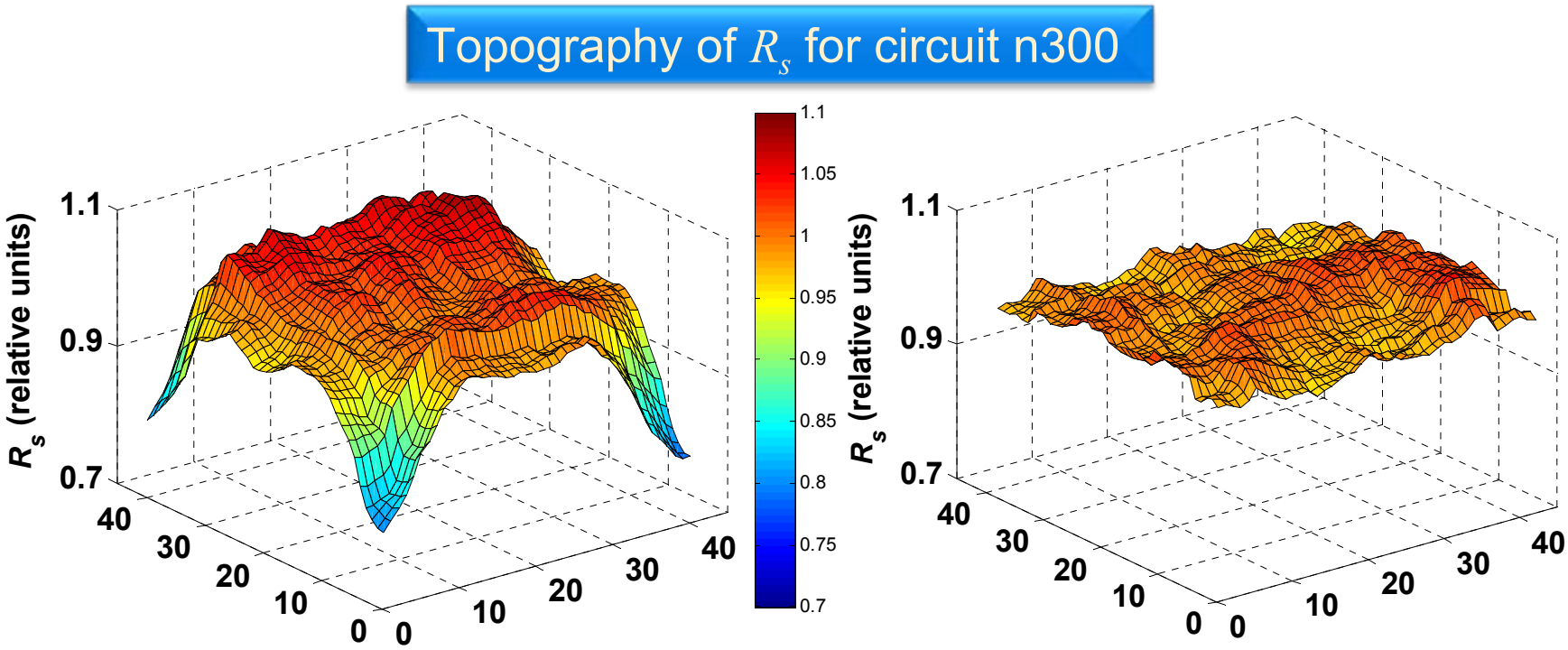
- Report polysilicon sheet resistance instead of STI density

$$R_s(i,j) = -0.9267 \times e(i,j) + 1.5223$$



- For xerox and ami33, the maximal gradient in R_s is a little larger, since it is not optimized directly in objective.

Topography of R_s

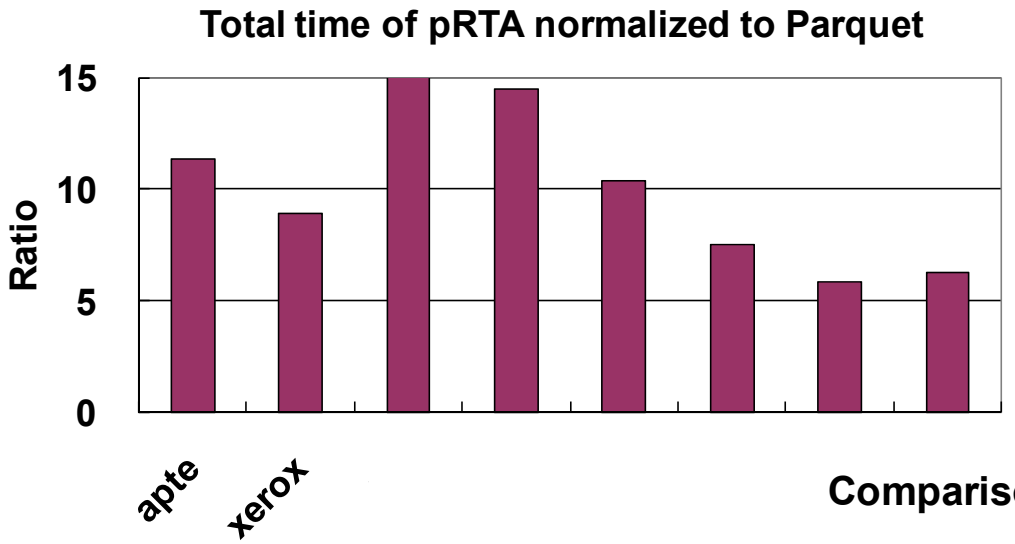


Floorplan without RTA optimization

Floorplan with RTA optimization

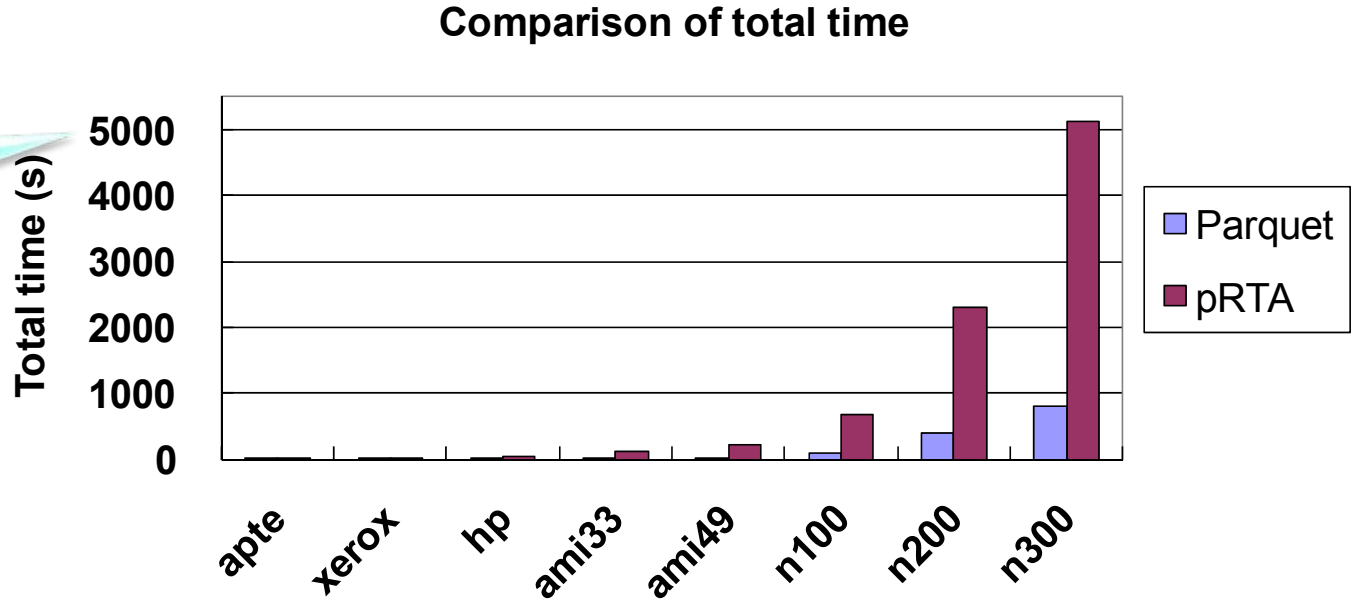
- Our algorithm reduces the variations in R_s significantly

Results for RTA-aware floorplanning



- Runtime is about 10x on average
  - Evaluation of the effective STI density costs about 8x longer than that of area and wire length

Reasonable in absolute terms



Experiment setup for dummy polysilicon filling

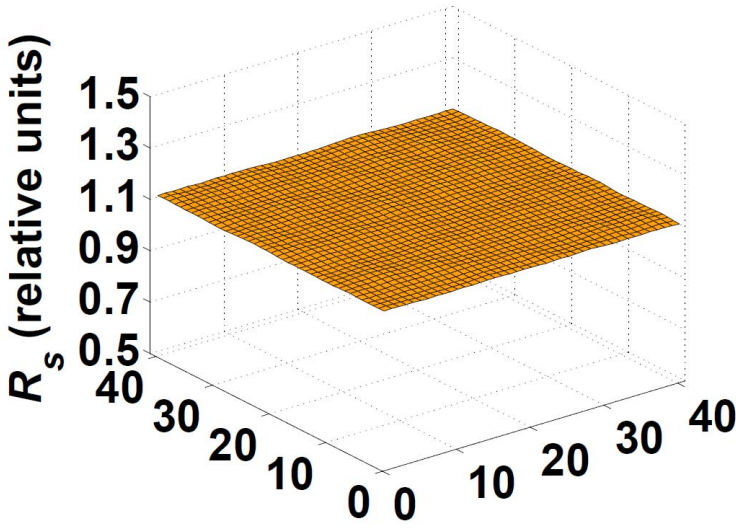
- Calculate upper bound for dummy filling
  - Assume minimal STI density after dummy filling is 0.15
  - Then

$$x_{ij}^{a} = \max\left(d^{0}(i,j) - 0.15, 0\right)$$

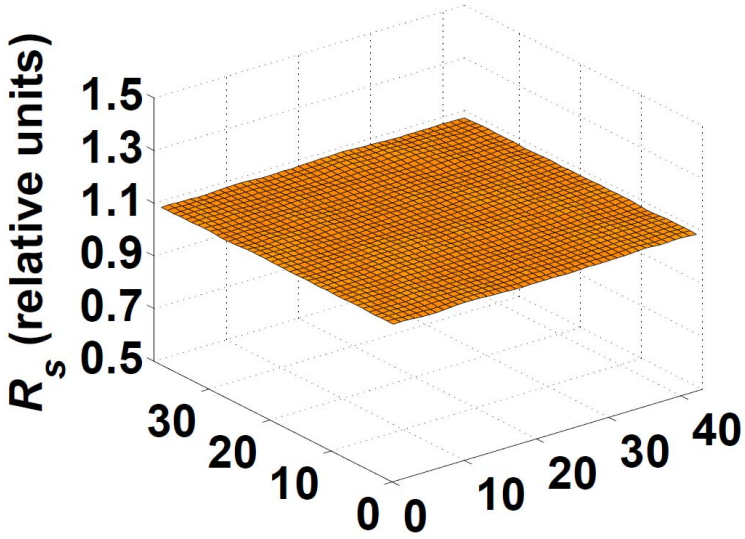
    where $d^{0}(i,j)$ is the initial local STI density before dummy filling
- Choose parameters such as after dummy filling:
  - Global variation in R_s no larger than 1%
  - Maximal gradient for R_s no larger than 0.2%
- Use commercial LP solver ILOG CPLEX
- Use same hardware as before

Topography of R_s after dummy filling

Topography of R_s for circuit n300



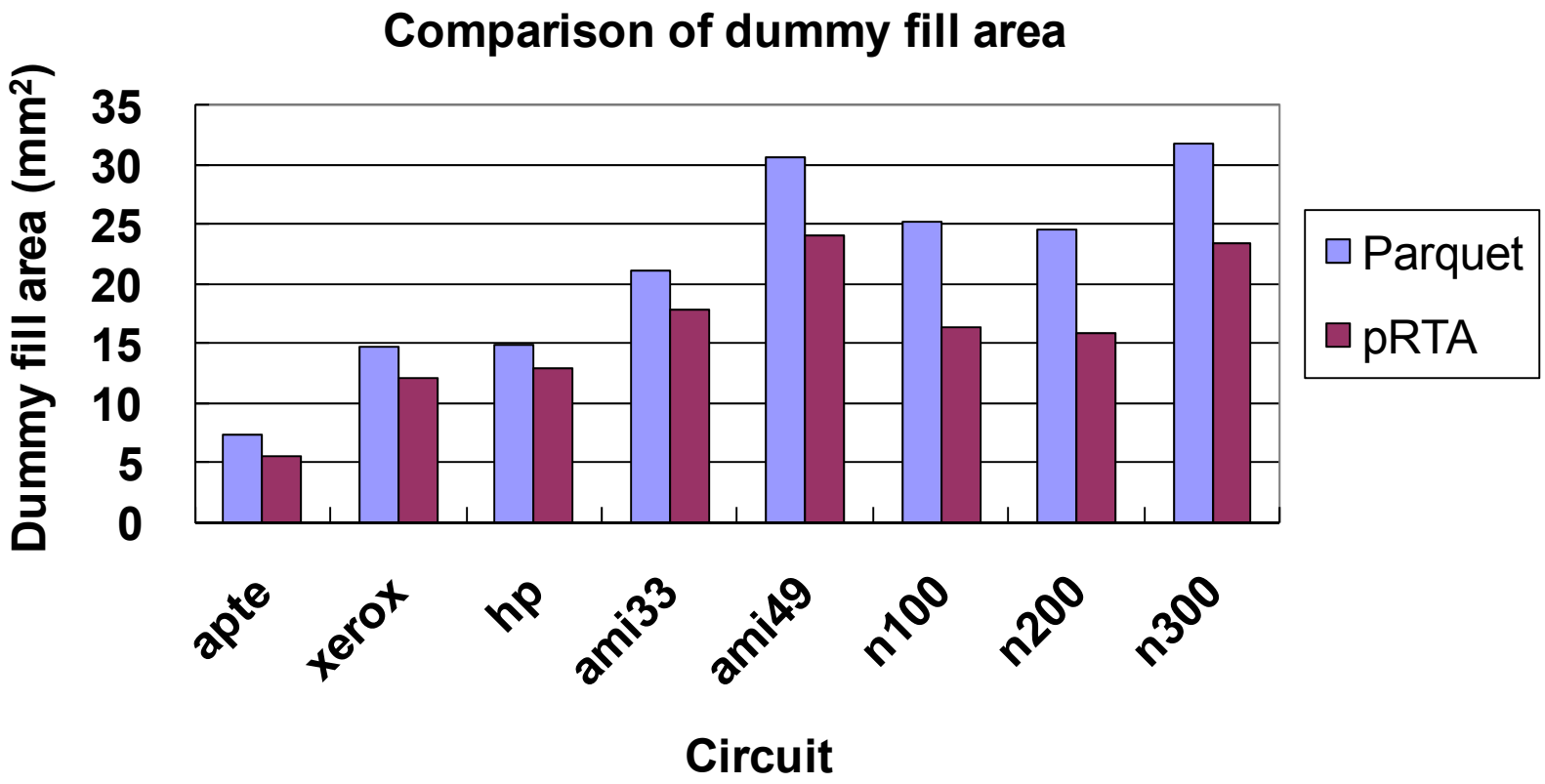
Floorplan without RTA optimization



Floorplan with RTA optimization

- The variations in R_s are negligible after dummy filling

Dummy polysilicon filling



- The amount of dummy fill reduced by 24% on average
  - To achieve the same requirements of STI uniformity

Summary

- RTA-induced variations strongly depend on circuit layout patterns
- First solve a floorplanning problem that aims to reduce the RTA variations by evening out the STI density distribution
- Next insert dummy polysilicon fill to further improve the uniformity of the STI density

Thank You!