Modern Circuit Placement

Yao-Wen Chang

張耀文

ywchang@ntu.edu.tw

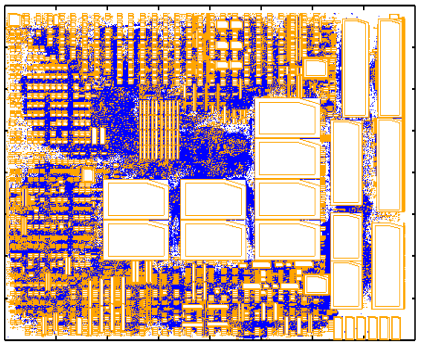
http://cc.ee.ntu.edu.tw/~ywchang

National Taiwan University

August 6, 2015

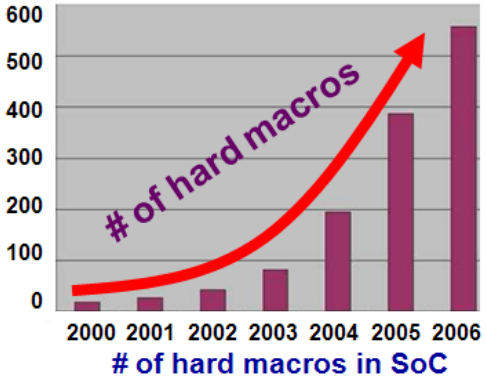
Modern Placement Challenges

- High complexity
  - Millions of objects to be placed
- Placement constraints
  - Preplaced blocks
  - Routability/density
  - Regions
- Mixed-size placement
  - Hundreds/thousands of large macros with millions of small standard cells
- Many more
  - 3D IC, datapath, FPGA, etc.

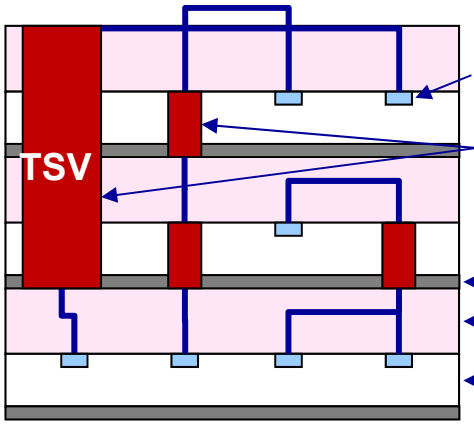


2.5M placeable objects

mixed-size design



Macros have revolutionized SoC design



device

TSV

dielectric

routing region

substrate

Hot Topic Ranking from EDApedia (IEEE CEDA)

More than 20 new academic placers developed since 2005!!

Topic	Importance	Difficulty	Well-defined	Average
Wirelength-driven Standard-Cell Placement	4.60	4.40	3.80	4.27
Global Routing	4.00	4.00	4.50	4.17
Clock Network Synthesis	3.75	4.00	4.50	4.08
Through Silicon Via and Cell Co-Placement	3.50	4.00	3.60	3.70
Power Gating	4.00	3.00	4.00	3.67
Partitioning	3.33	3.50	3.33	3.39
3DIC Thermal Management	3.33	3.67	3.00	3.33
Stochastic Behavioral Modeling	4.00	3.00	3.00	3.33
Static Timing Analysis	3.00	2.00	4.00	3.00
Multi-Voltage Mode Clock Skew Minimization	2.50	3.50	2.50	2.83

Recent contests

49 DESIGN AUTOMATION CONFERENCE

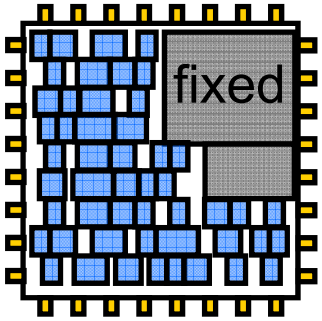
IEEE/ACM ICCAD 2014 INTERNATIONAL CONFERENCE ON COMPUTER-AIDED DESIGN

IC/CAD Contest

ISPD Contests

SIGda

Outline

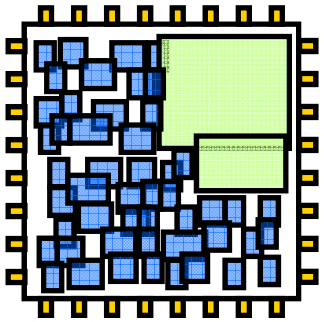


- Placement Basics
- Standard Cell Placement
- Mixed-Size Placement
- Future Research Directions

NTUplace3/4 Placement Flow

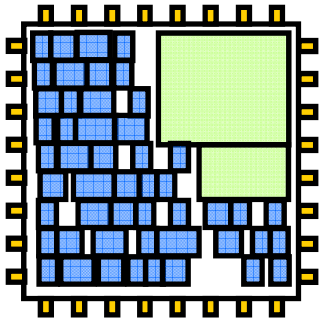
- Chen et al., “A high quality **analytical placer** considering preplaced blocks and density constraint,“ ICCAD-06 (TCAD-08)

Global Placement (GP)



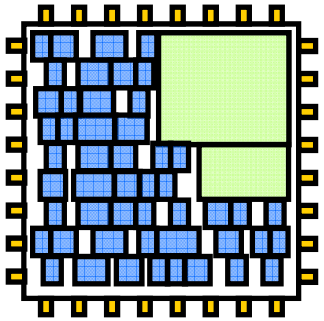
Computes the best position for each cell to minimize the cost (e.g., wirelength), ignoring cell overlaps

Legalization (LG)



Places cells into rows & removes all overlaps among cells

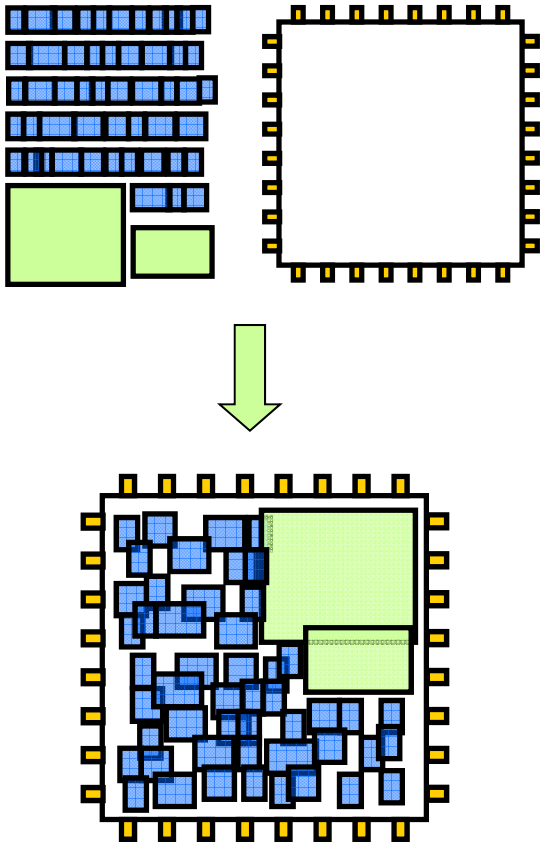
Detailed Placement (DP)



Refines the solution

Global Placement

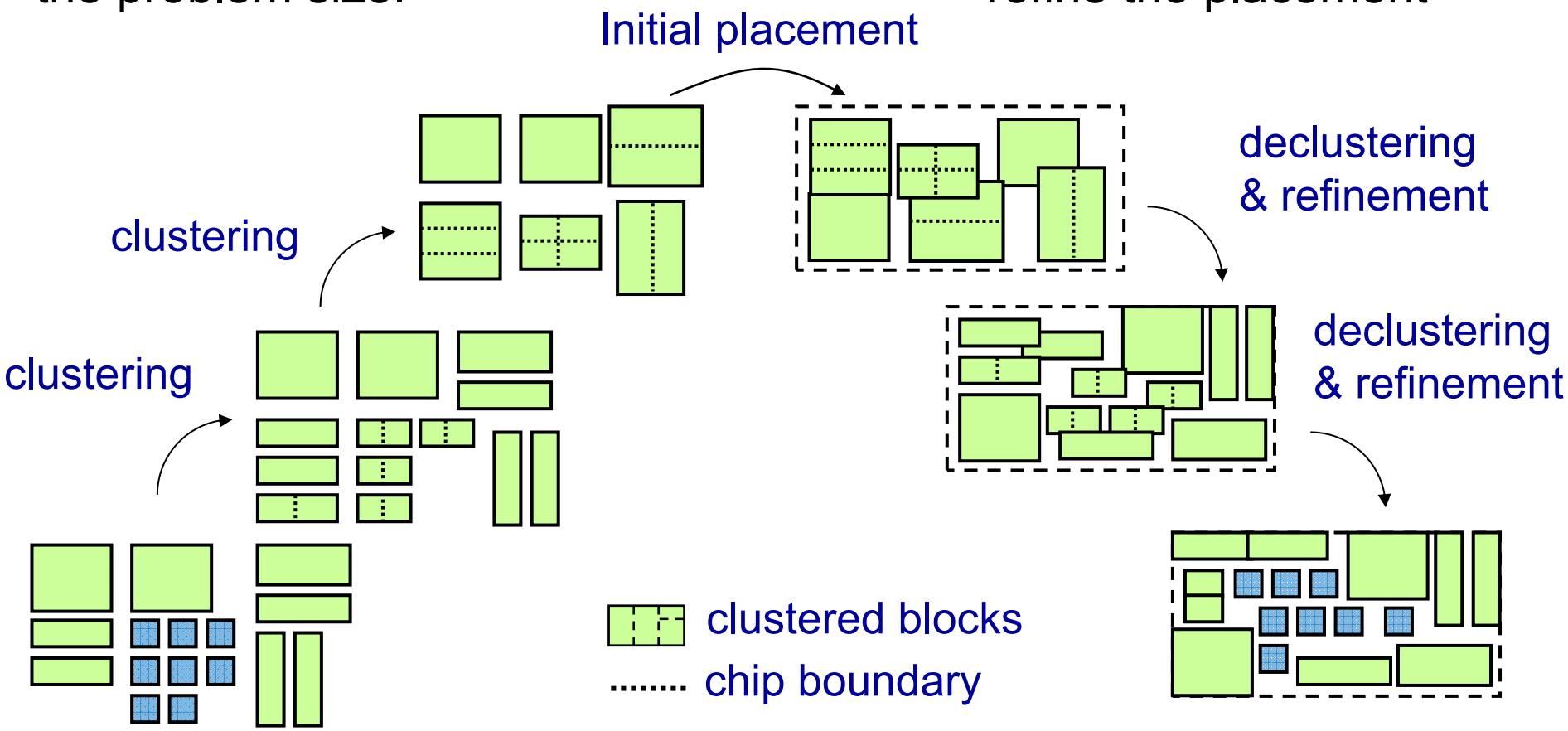
- Placement flow
  - ***Global placement***
    - ***Multilevel framework***
    - ***Analytical formulation with a nonlinear objective function***
    - ***Smoothing techniques for preplaced blocks***
    - ***Overlap removal for density control***
    - *Many more….*
  - Legalization
  - Detailed placement



Multilevel Global Placement

Cluster cells based on connectivity/area to reduce the problem size.

Iteratively decluster the clusters and further refine the placement



Analytical Placement Model

- Analytical placement during declustering
- Global placement problem (allow overlaps)

$$\textbf{min} \quad W(x, y)$$
$$\textbf{s.t.} \quad D_b(x, y) \le M_b$$

Minimize wirelength
D_b: density for bin *b*
M_b: max density for bin *b*

- Relax the constraints into the objective function

$$\textbf{min} \quad W(x, y) + \boldsymbol{\lambda}\Sigma(D_b(x, y) - M_b)^2$$

  - Apply gradient search to solve it, which needs a smooth & differentiable function
  - Increase **λ** gradually to reduce the density penalty to find cell positions (x, y) with desired wirelength

Gradient Solver

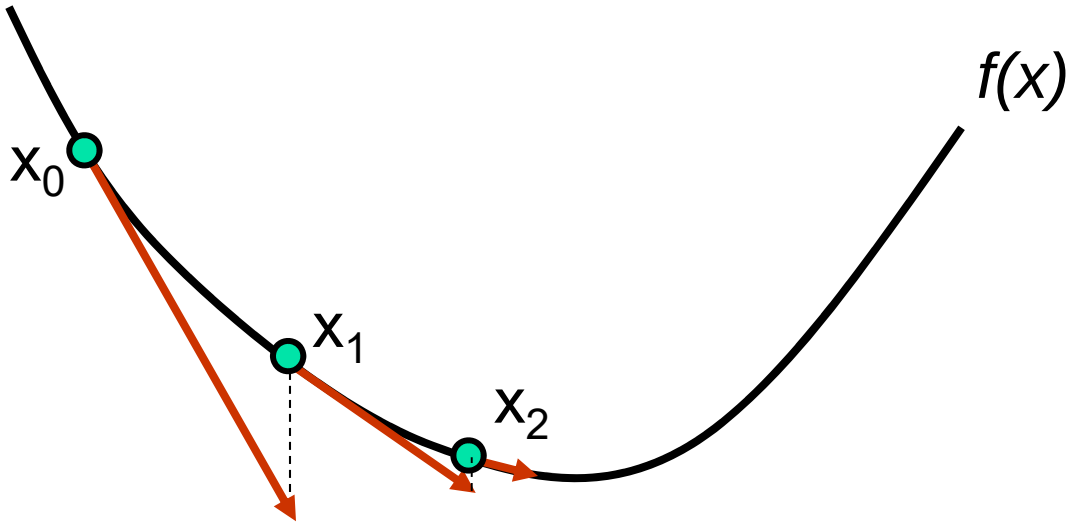
min $f(x)$

[Gradient Solver]

$x_0 \leftarrow$ initial value
Repeat until convergence

$$x_{i+1} = x_i - f'(x)|_{x=xi} * \text{stepsize}$$

Function f must be **smooth and differentiable** to apply the gradient solver.



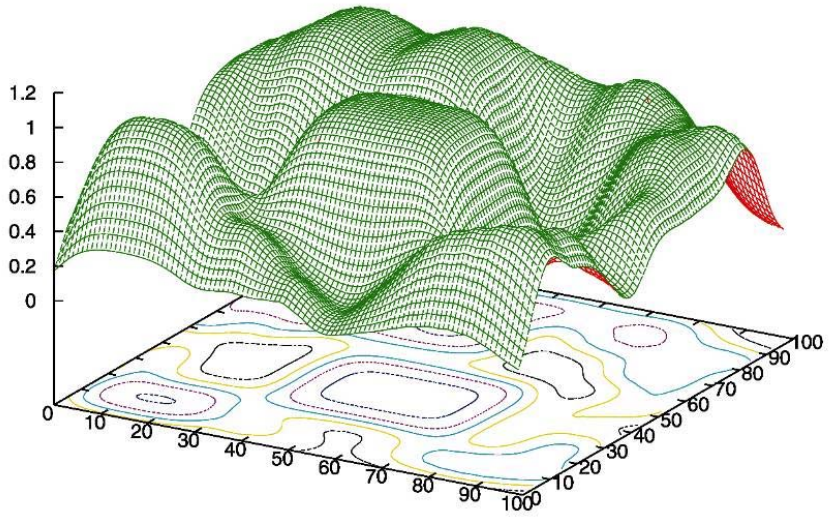
Two Ingredients in Analytical Formulation

$$\min \boxed{W(x, y)} + \lambda \boxed{\Sigma (D_b(x, y) - M_b)^2}$$

Wirelength (↑ $W(x, y)$)

Density (↑ $\Sigma (D_b(x, y) - M_b)^2$)

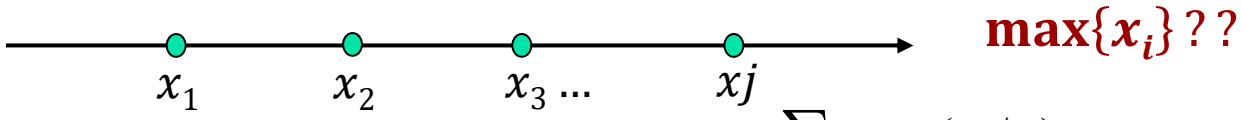
Both functions must be smooth & differentiable!!



Our Weighted-Average (WA) Model

$$W_{WA}(\mathbf{x},\mathbf{y}) = \sum_{e\in E}\left(\frac{\sum_{v_i\in e} x_i \exp(x_i/\gamma)}{\sum_{v_i\in e}\exp(x_i/\gamma)} - \frac{\sum_{v_i\in e} x_i \exp(-x_i/\gamma)}{\sum_{v_i\in e}\exp(-x_i/\gamma)} + \frac{\sum_{v_i\in e} y_i \exp(y_i/\gamma)}{\sum_{v_i\in e}\exp(y_i/\gamma)} - \frac{\sum_{v_i\in e} y_i \exp(-y_i/\gamma)}{\sum_{v_i\in e}\exp(-y_i/\gamma)}\right)$$

- **1st model that outperforms LSE theoretically & empirically**
  [Hsu, Chang, Balabanov, DAC-11, TCAD-13; US Patent, 2014]
- Weighted average of a set of x coordinates, $\mathbf{x}_e$, of a net e:
  - $X(\mathbf{x}_e)$ can approximate the maximum value of $\mathbf{x}_e$ by setting the weight function of x_i: $\boldsymbol{F(x_i) = \exp(x_i/\gamma)}$, a fast growing function



$$X(\mathbf{x}_e) = \frac{\sum_{v_i\in e} x_i F(x_i)}{\sum_{v_i\in e} F(x_i)} \xrightarrow{\max} X_{\max}(\mathbf{x}_e) = \frac{\sum_{v_i\in e} x_i \exp(x_i/\gamma)}{\sum_{v_i\in e}\exp(x_i/\gamma)}$$

- Is an effective smooth, differentiable function for HPWL approximation
- Approaches exact HPWL when $\gamma \rightarrow 0$

Popular Wirelength Models

HPWL

$$W(\mathbf{x},\mathbf{y}) = \sum_{\text{net } e} \left(\max_{v_i, v_j \in e} |x_i - x_j| + \max_{v_i, v_j \in e} |y_i - y_j| \right)$$

quadratic

$$\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} [(x_i - x_j)^2 + (y_i - y_j)^2]$$

Log-sum-exp

$$\gamma \sum_{e \in E} \left(\log \sum_{v_k \in e} \exp(x_k/\gamma) + \log \sum_{v_k \in e} \exp(-x_k/\gamma) + \right.$$
$$\left. \log \sum_{v_k \in e} \exp(y_k/\gamma) + \log \sum_{v_k \in e} \exp(-y_k/\gamma) \right).$$

L_p-norm

$$\sum_{e \in E} \left(\left(\sum_{v_k \in e} x_k^p \right)^{\frac{1}{p}} - \left(\sum_{v_k \in e} x_k^{-p} \right)^{-\frac{1}{p}} + \left(\sum_{v_k \in e} y_k^p \right)^{\frac{1}{p}} - \left(\sum_{v_k \in e} y_k^{-p} \right)^{-\frac{1}{p}} \right)$$

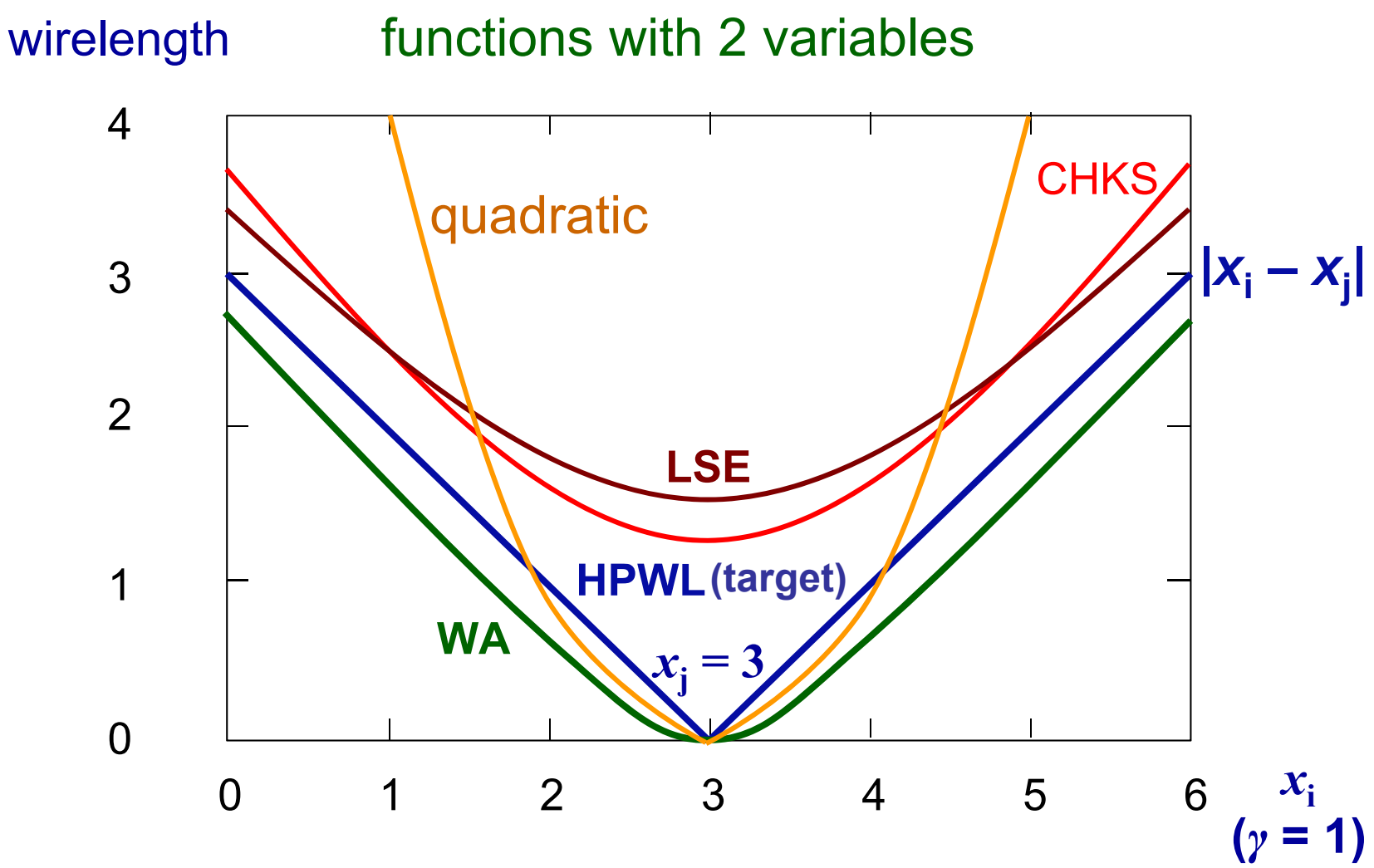
CHKS

$$CHKS(x_1, x_2) = \frac{\sqrt{(x_1 - x_2)^2 + t^2} + x_1 + x_2}{2},$$

Weighted-average

$$\sum_{e \in E} \left(\frac{\sum_{v_i \in e} x_i \exp(x_i/\gamma)}{\sum_{v_i \in e} \exp(x_i/\gamma)} - \frac{\sum_{v_i \in e} x_i \exp(-x_i/\gamma)}{\sum_{v_i \in e} \exp(-x_i/\gamma)} + \right.$$
$$\left. \frac{\sum_{v_i \in e} y_i \exp(y_i/\gamma)}{\sum_{v_i \in e} \exp(y_i/\gamma)} - \frac{\sum_{v_i \in e} y_i \exp(-y_i/\gamma)}{\sum_{v_i \in e} \exp(-y_i/\gamma)} \right).$$

Popular Wirelength Model Comparisons



Theoretical Comparisons

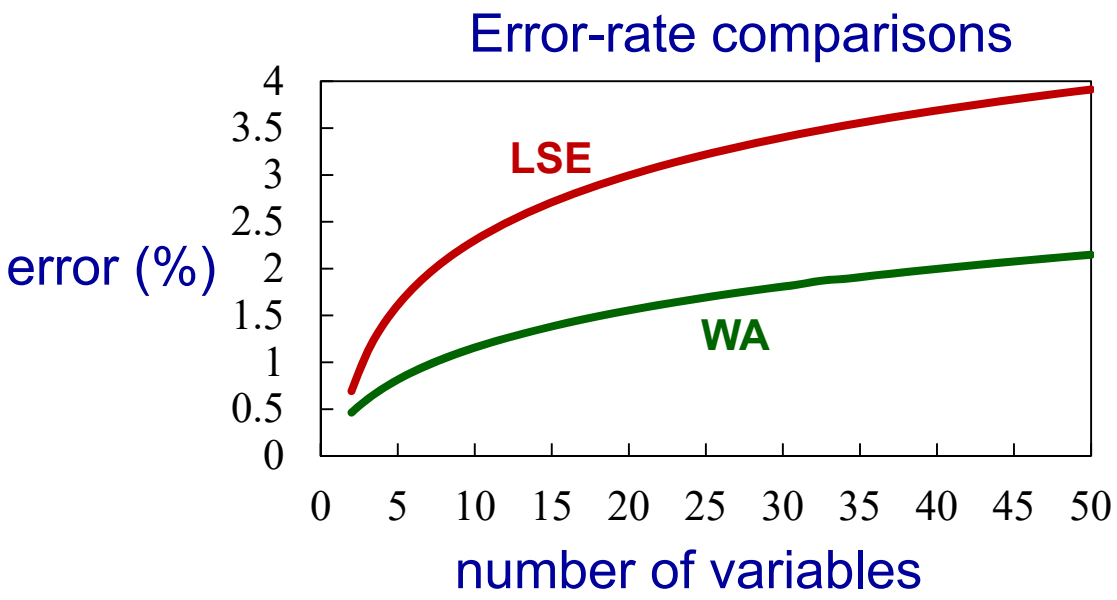
- Theorem: The estimation error bound of the WA model is

$$0 \le \varepsilon_{WA}(\mathbf{x}_e) \le \frac{\gamma(x_{\max} - x_{\min})}{1 + e^{(x_{\max} - x_{\min})}/n}.$$

$\mathbf{x}_e = \{x_i \mid v_i \in e\}$: a set **x** of coordinates associated with net e

- Theorem: The error upper bound of the WA model is smaller than that of the LSE model:

$$\varepsilon_{WA}(\mathbf{x}_e) \le \varepsilon_{LSE}(\mathbf{x}_e) = \gamma \ln n$$



Wirelength Model Comparison

- Integrated both the LSE and WA models into NTUplace3, a leading academic placer
- Used ISPD-06 placement benchmark circuits

  #Cells: 330K—2481K, #Nets: 338K—2636K
- The WA model can achieve shorter wirelength than LSE

Wirelength Model	Wirelength	CPU Time
LSE	1.000	1.000
WA	**0.980**	**1.066**

- The results show that WA outperforms LSE consistently

Recent works from other groups show the same effects!!

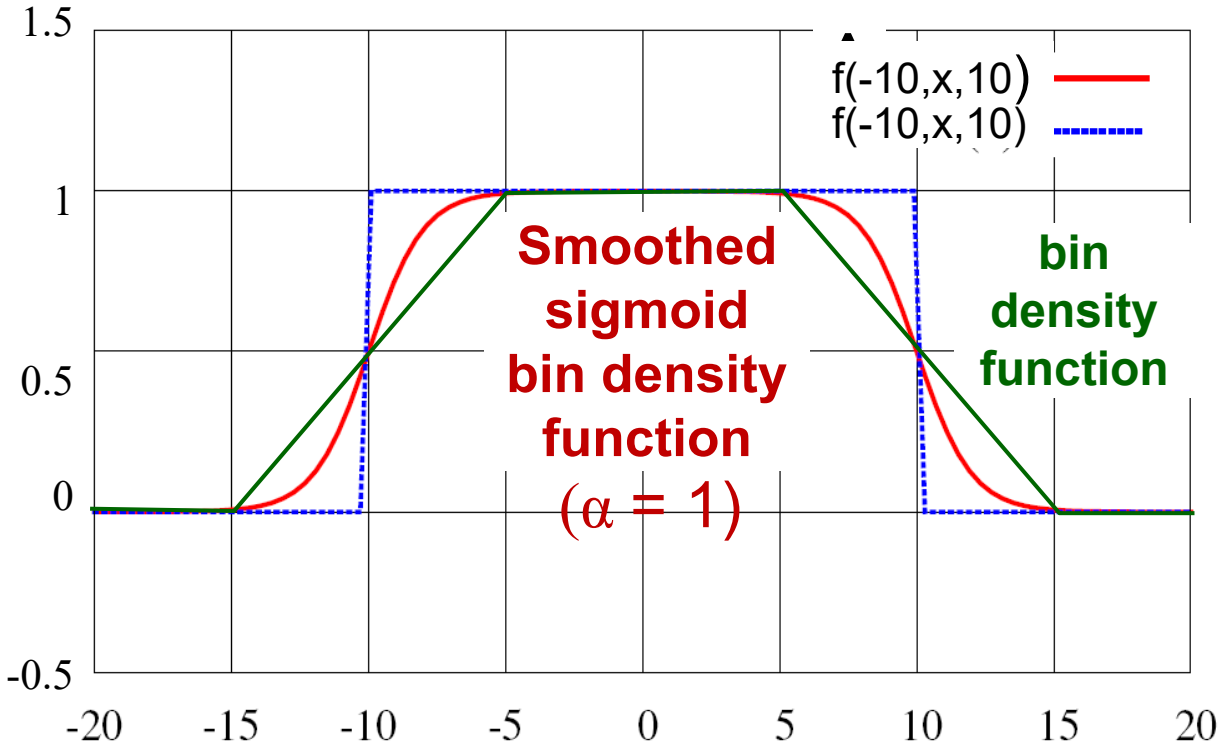
Our Density Smoothing: Sigmoid Function

$$f(l,x,u)=\begin{cases}1, & \text{if } l<x<u\\ 0, & \text{otherwise}\end{cases}$$

$$p(t)=\frac{1}{1+e^{-\alpha t}}$$

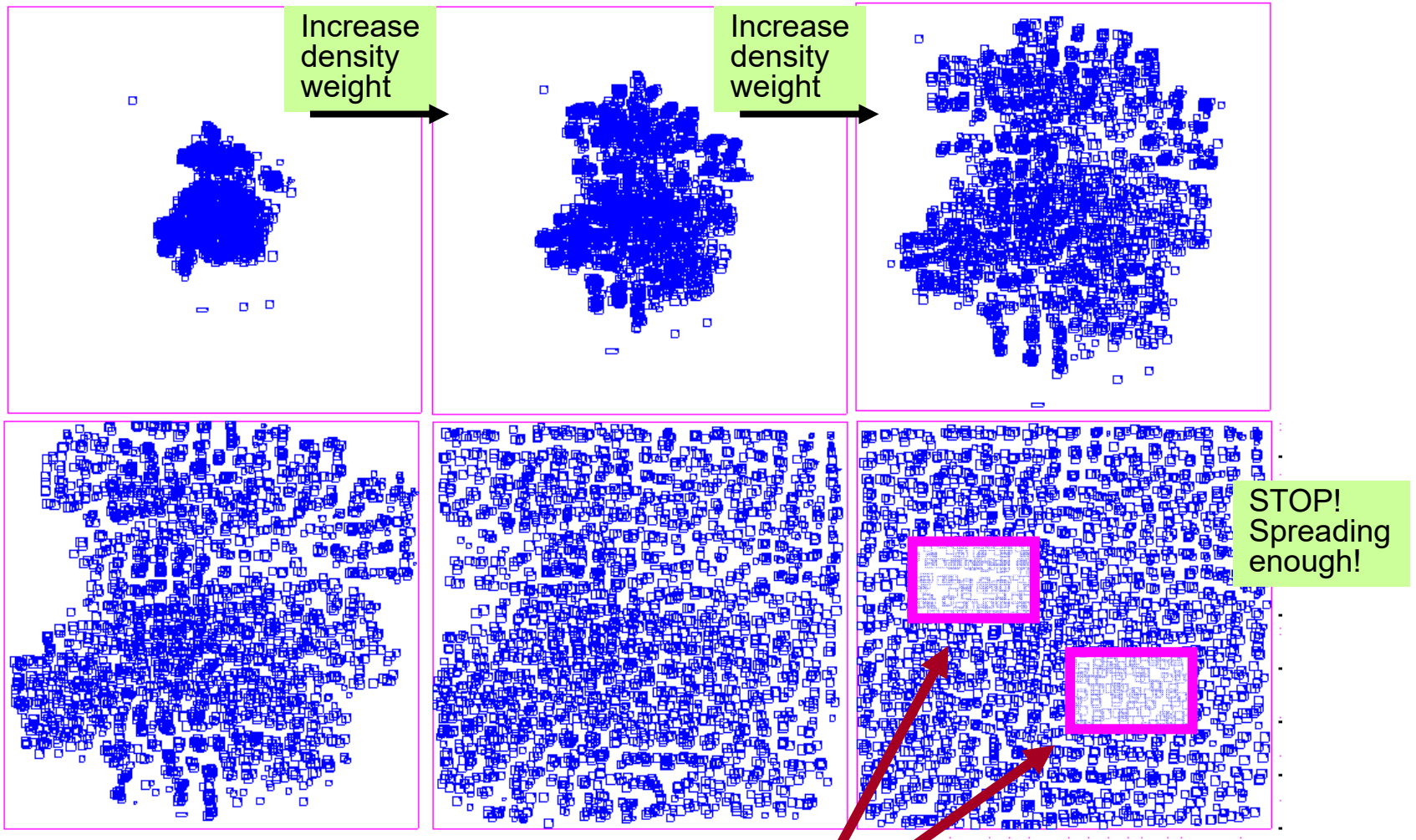
[Hsu & Chang, ICCAD'11]

$$f(l,x,u)\cong p(x-l)\,p(u-x)$$



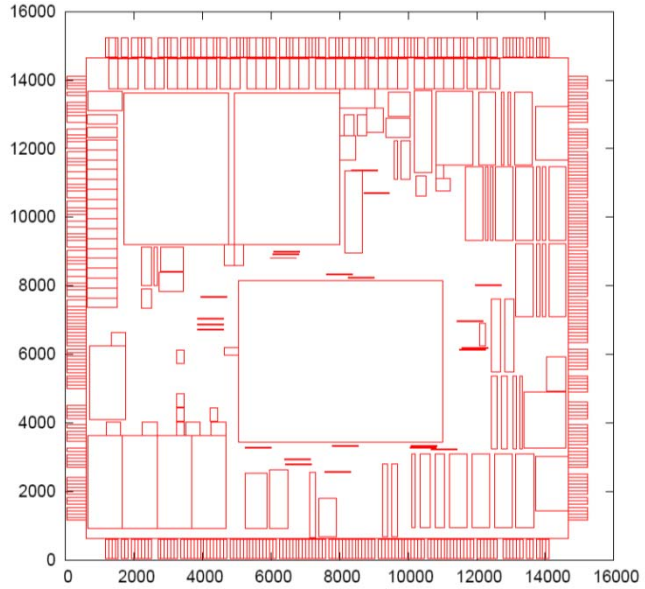
- Is an effective smooth & differentiable approximation for the bin density function
- Approximates exact 0-1 logistic function when $\alpha \rightarrow 0$

Cell Spreading During Global Placement

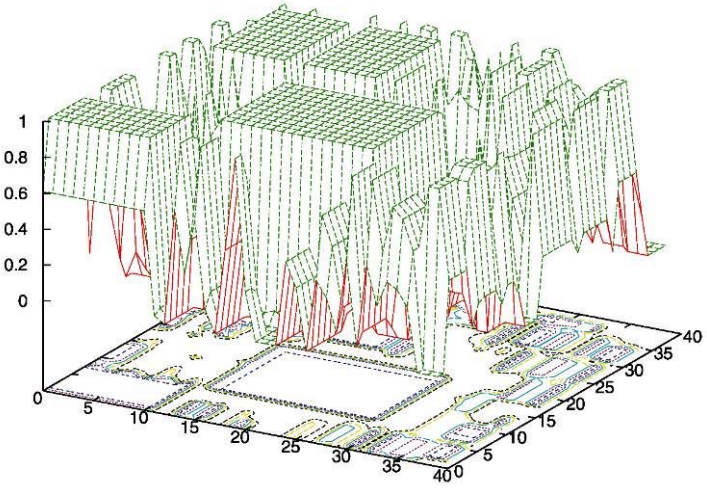


- How to handle preplaced blocks?
  - Pre-defined density makes cell spreading harder.

Density Map with Preplaced Blocks



Placement with preplaced blocks



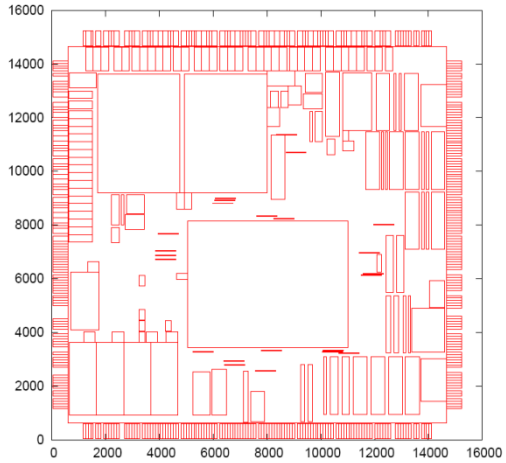
Corresponding density map

Two major problems

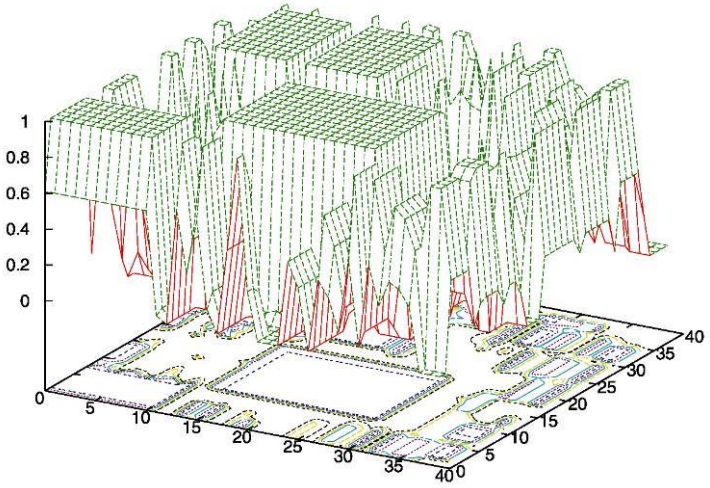
☹ the density map is not smooth

☹ some densities are too high to spread cells over the mountains/high plateaus

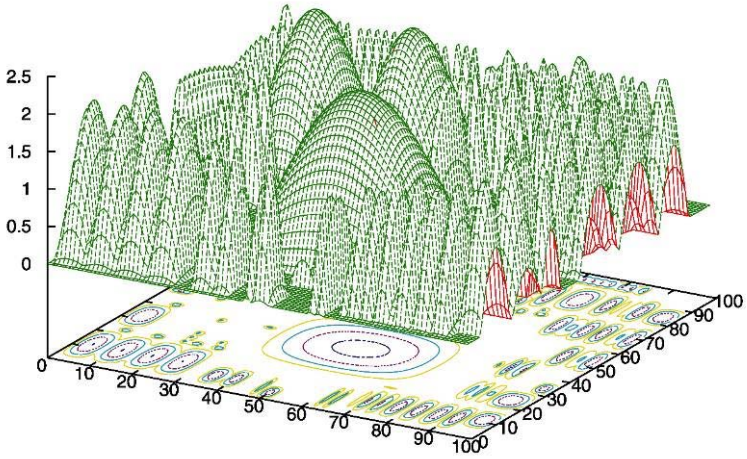
Bell-Shaped Block Smoothing??



☹ mountain heights are quite different
☹ there are valleys among mountains
➔ might mis-guide the cells to be placed in the valleys (overlaps!!)



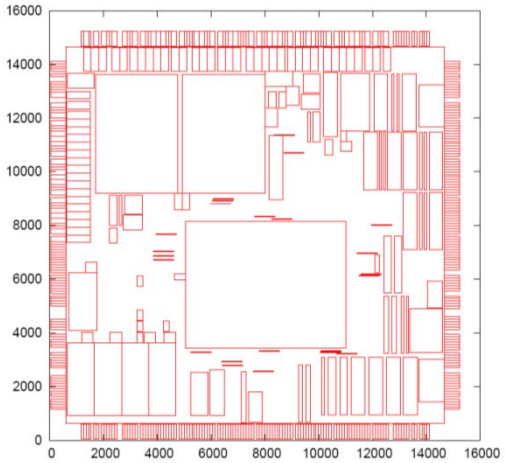
Apply the bell-shaped function??

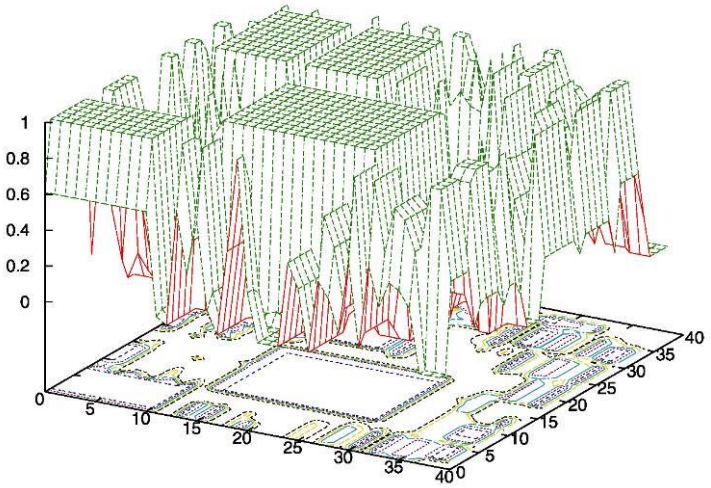


Bell-shaped smoothing

[Kahng & Wang, ICCAD-04]

Preplaced Block Smoothing



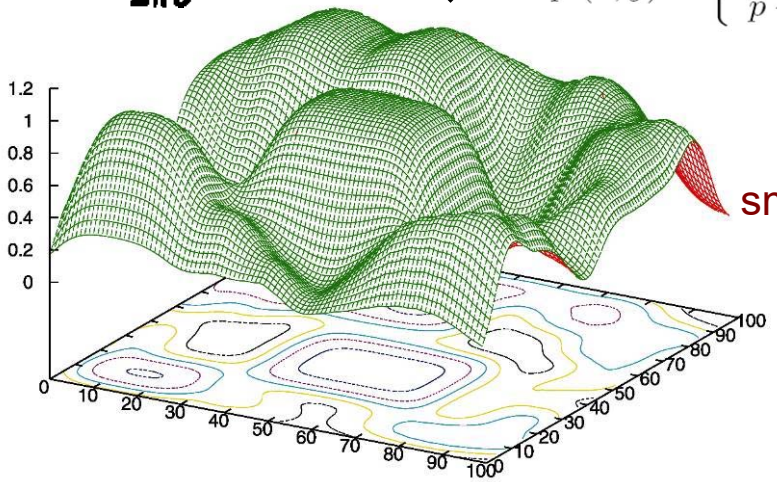


or Sigmoid-function smoothing + level smoothing

Convolute with Gaussian

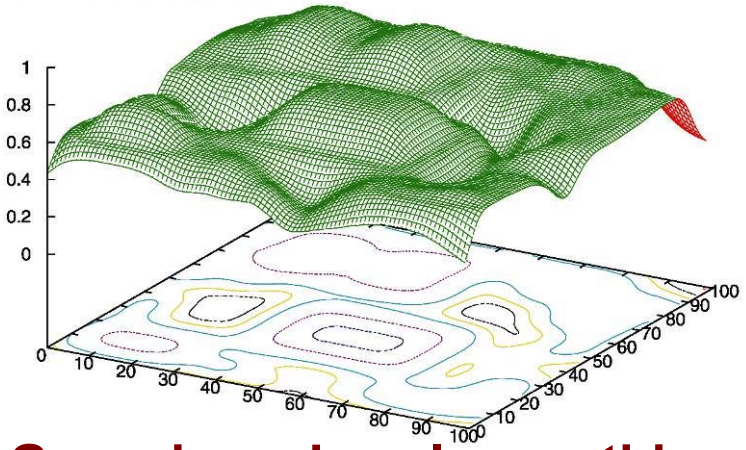
$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$p'(x,y) = \begin{cases} \overline{p} + (p(x,y) - \overline{p})^{\delta} & if\ p(x,y) \geq \overline{p} \\ \overline{p} - (\overline{p} - p(x,y))^{\delta} & if\ p(x,y) \leq \overline{p} \end{cases}$$



Gaussian smoothing

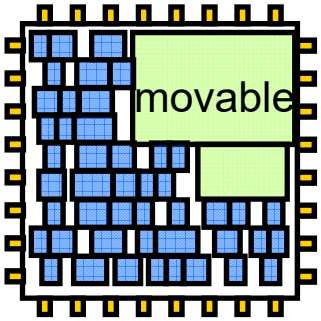
Level smoothing



Gaussian + Level smoothing

Outline

- Placement Basics
- Standard Cell Placement
- Mixed-Size Placement
- Future Research Directions



Methods on Mixed-Size Placement

Type 1: Constructive approach

- Combine floorplanning and placement
- Capo, PATOMA, FLOP

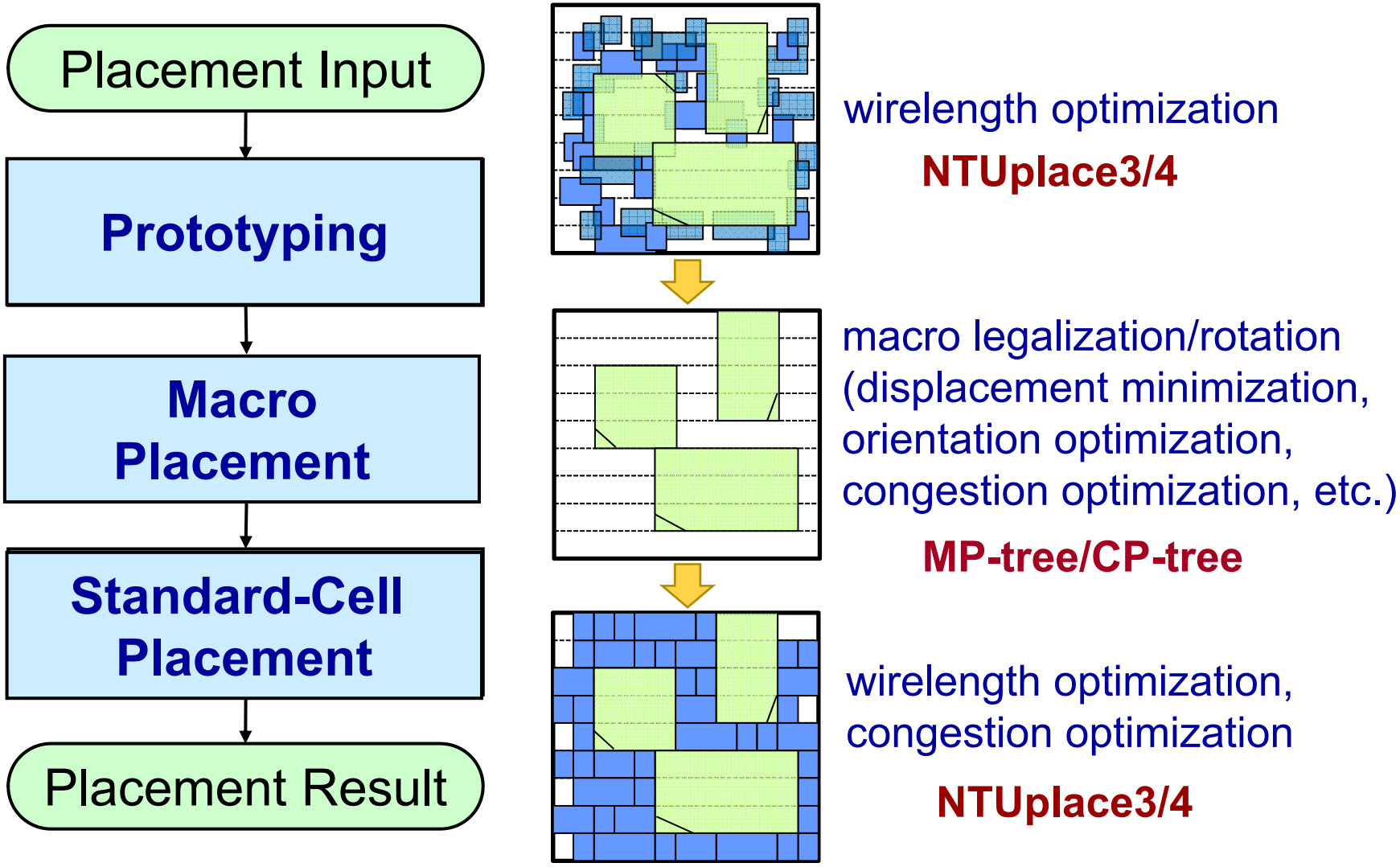
Type 2: Three-stage approach

- Perform (1) prototyping, (2) macro placement, and (3) cell placement
- MP-tree [DAC'07], CG, CP-tree [DAC'14], ePlace-MS [TCAD, 2015]

Type 3: Unified approach

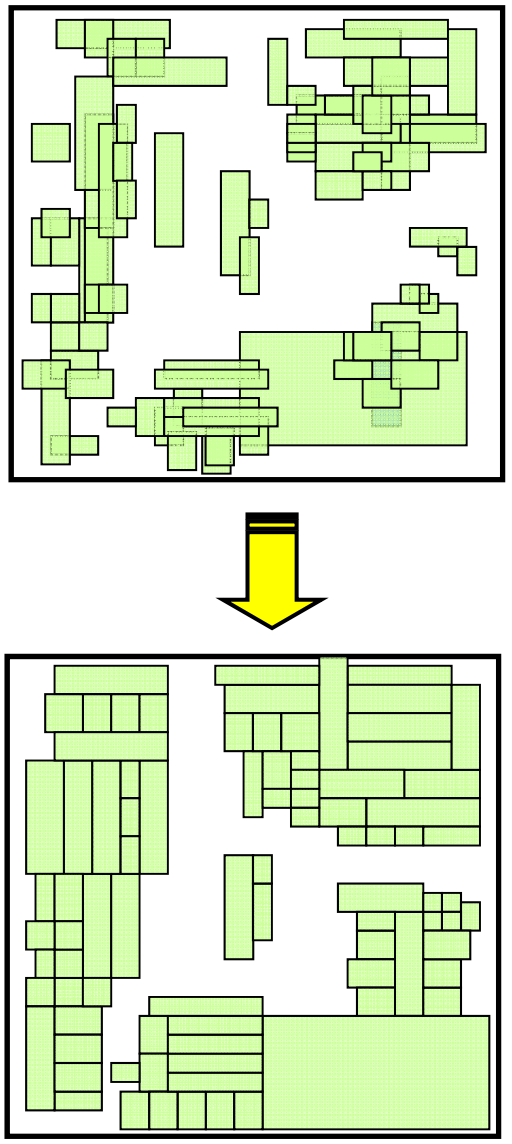
- Place macro and cell simultaneously
- mPG-MS, APlace, mPL, UPlace, NTUplace3 [ICCAD'10], etc.

Three-Stage Approach

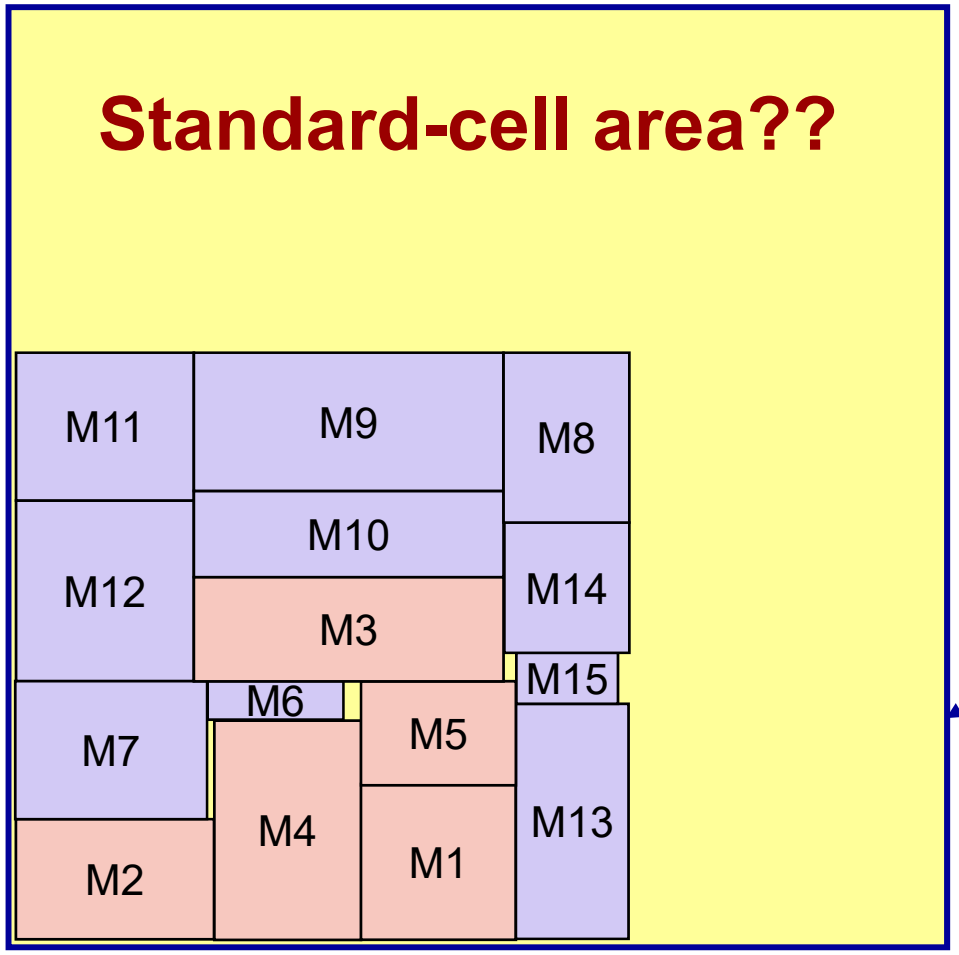


Macro Placement

- Input
  - An initial placement that considers both macros and standard cells and optimizes a simplified cost metric (e.g., wirelength)
- Objectives
  - Remove all overlaps between macros
  - Find macro position and orientation
  - Minimize macro displacement
- Popular approaches
  - Packing-based method: MP-tree [DAC'07, TCAD'08], CP-tree [DAC'14]
  - Constraint graph-based method: CG [ICCAD'08]



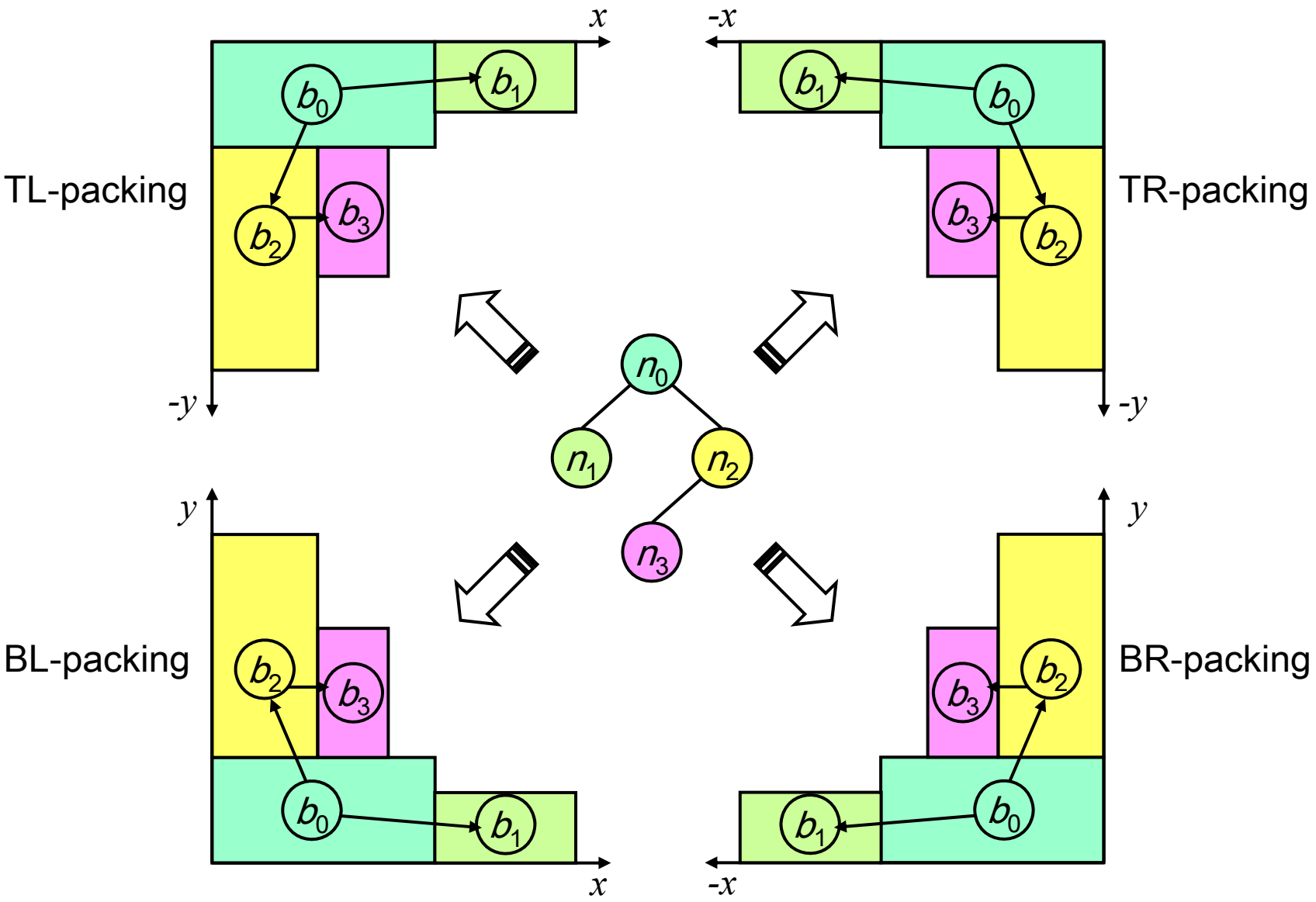
But What If %Macro Area Is Not High?



All macros will be packed together!!

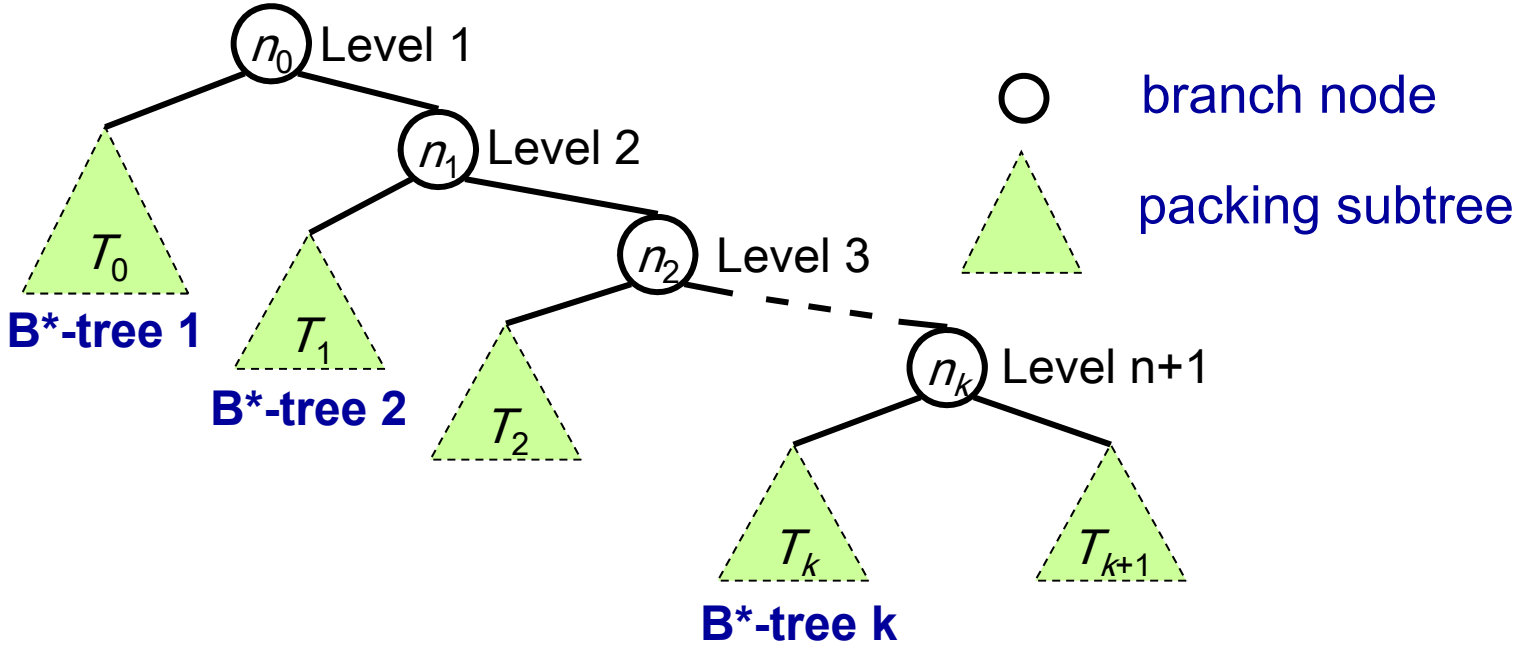
Chip outline

Multi-Packing (MP) Tree Representation



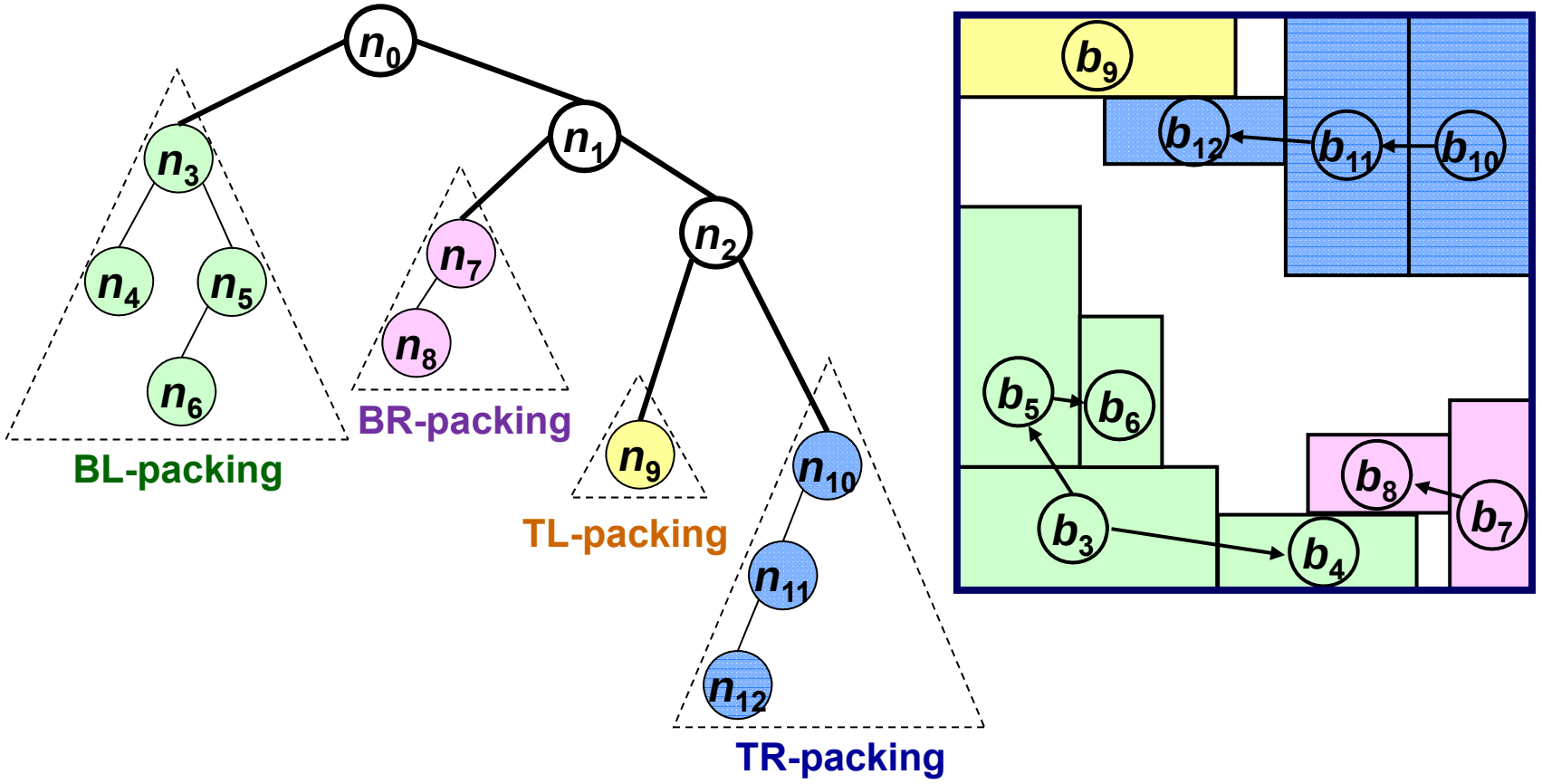
Generalized MP-Trees

- Working on four independent packing trees may not get a desired solution
  - Lack global interactions among different subproblems
- Key: **Combine packing trees** packing to different corners
  - Chen *et al.*, "MP-trees: A packing-based macro placement algorithm for modern mixed-size designs," DAC'07 & TCAD'08
- Use the right skewed branch for easier implementation



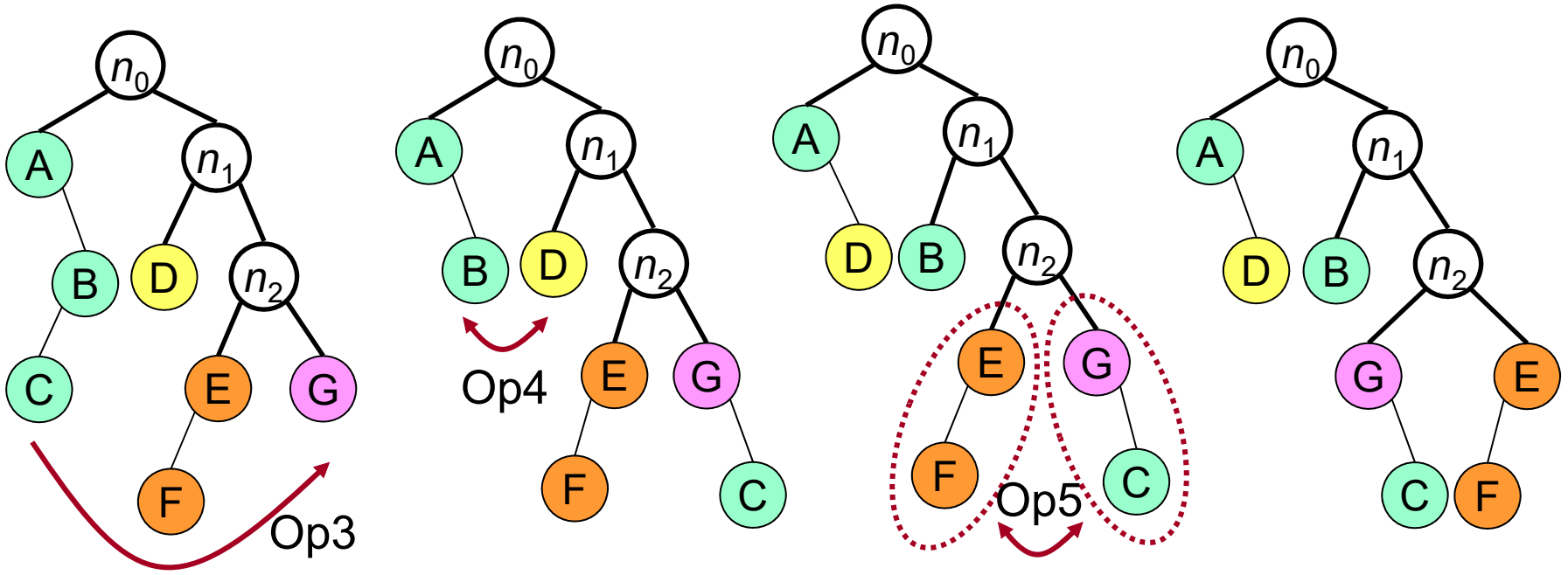
MP-tree Macro Placement Example

- Use four packing subtrees to handle a rectangular chip
- Applies to a placement region with any number of corners



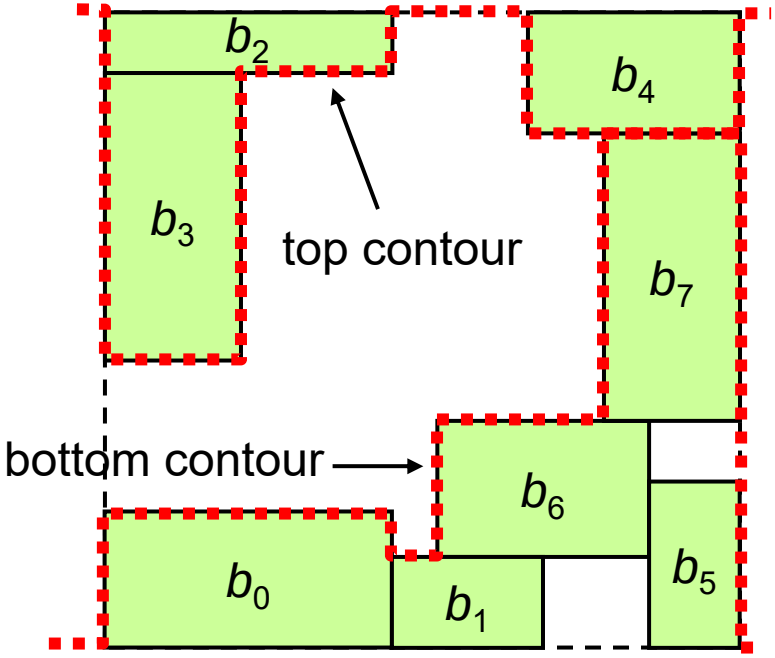
Operations on MP-tree

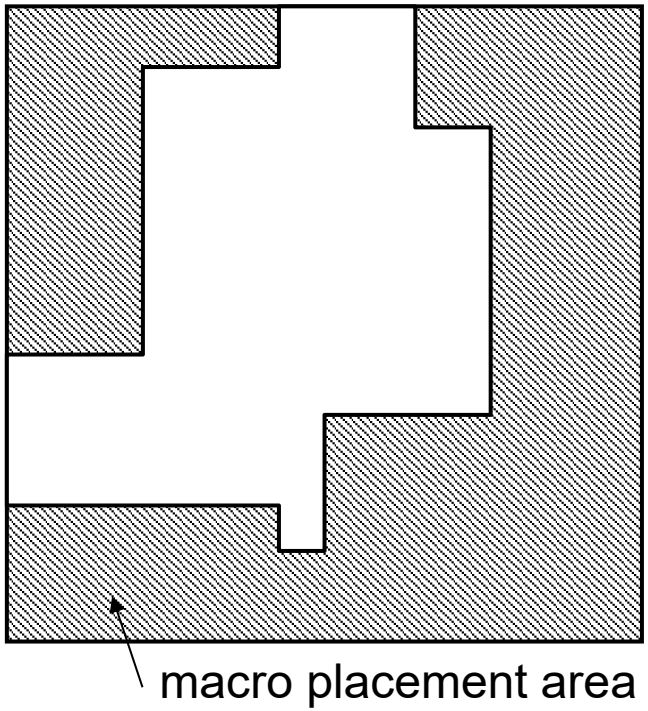
- Operations for iterative improvement or simulated annealing to perturb one MP-tree to another
  - Op1: Rotate a macro/cluster
  - Op2: Resize a cluster
  - Op3: Move a node to another place
  - Op4: Swap two nodes
  - Op5: Swap two packing subtrees



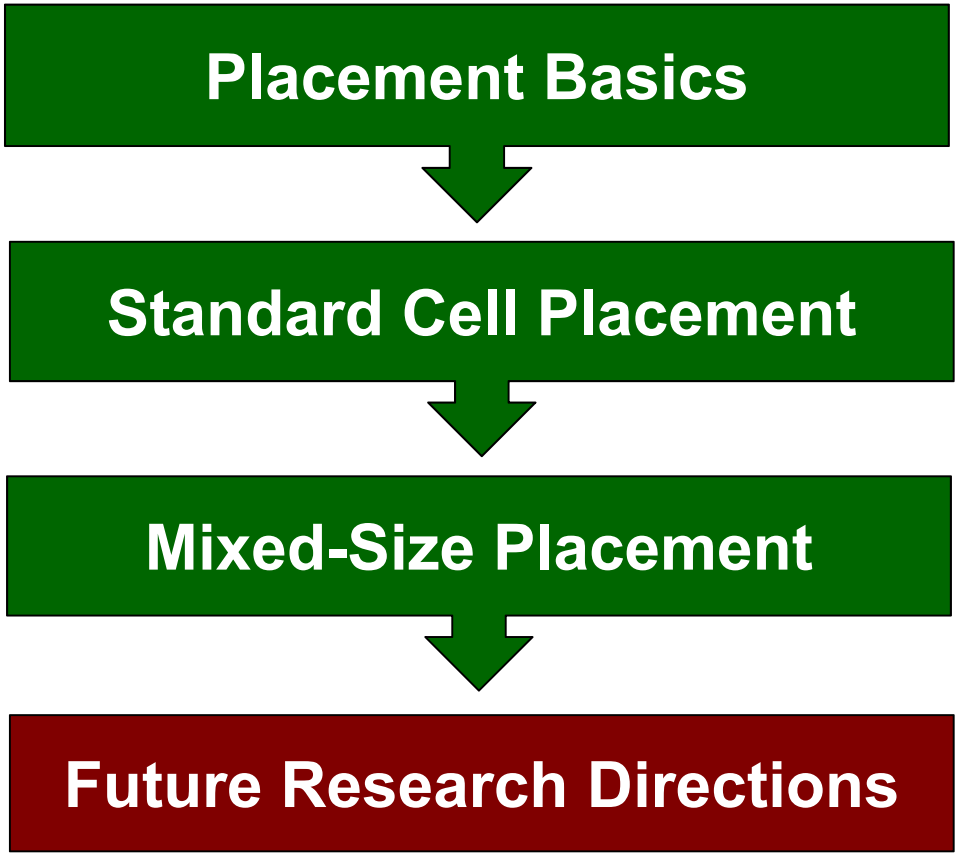
Evaluation of a Macro Placement

- Macro placement area
- Wirelength
- Macro displacement





Outline


Placement Basics
Standard Cell Placement
Mixed-Size Placement
Future Research Directions

Potential Research Directions

- **Scalability**
  - Faster differentiable wirelength model?
  - Multi-level mixed-size placement
- **Multi-dimension**
  - Routability-driven placement
  - Timing-driven placement
  - Cosynthesis with power/ground & clock networks
- **Heterogeneity**
  - Datapath-intensive mixed-size design
  - Analytical FPGA placement
- **Technology**
  - 3D IC placement
  - Manufacturability-aware placement
  - Reliability-aware placement: FinFET self-heating, stress

Many more!!

1. Faster Differentiable Wirelength Model

- **Golden:** Half-perimeter wirelength (HPWL) model

$$W(\mathbf{x},\mathbf{y}) = \sum_{e \in E} \left(\max_{v_i, v_j \in e} \left| x_i - x_j \right| + \max_{v_i, v_j \in e} \left| y_i - y_j \right| \right)$$

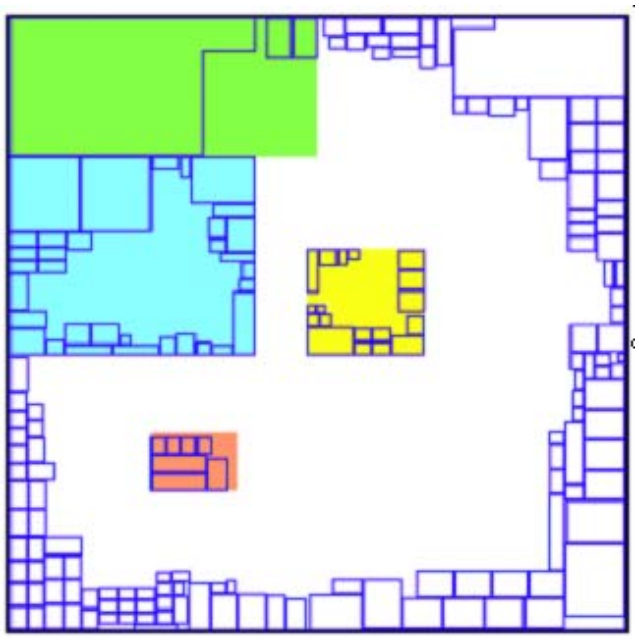
- **Key:** Faster smooth & differentiable approximation for the **max** function??

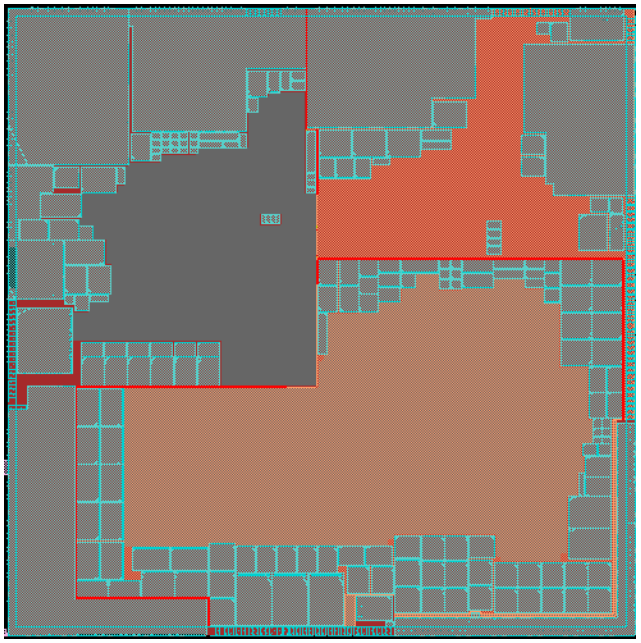
$$W_{LSE}(\mathbf{x},\mathbf{y}) = \gamma \sum_{e \in E} \left(\ln \sum_{v_k \in e} \exp(x_k / \gamma) + \ln \sum_{v_k \in e} \exp(-x_k / \gamma) + \right.$$
$$\left. \ln \sum_{v_k \in e} \exp(y_k / \gamma) + \ln \sum_{v_k \in e} \exp(-y_k / \gamma) \right)$$

$$W_{WA}(\mathbf{x},\mathbf{y}) = \sum_{e \in E} \left(\frac{\sum_{v_i \in e} x_i \exp(x_i / \gamma)}{\sum_{v_i \in e} \exp(x_i / \gamma)} - \frac{\sum_{v_i \in e} x_i \exp(-x_i / \gamma)}{\sum_{v_i \in e} \exp(-x_i / \gamma)} + \frac{\sum_{v_i \in e} y_i \exp(y_i / \gamma)}{\sum_{v_i \in e} \exp(y_i / \gamma)} - \frac{\sum_{v_i \in e} y_i \exp(-y_i / \gamma)}{\sum_{v_i \in e} \exp(-y_i / \gamma)} \right)$$

2. Large-Scale Mixed-Size Placement

- Still have a long way to go for mixed-size placement!!
  - Find best trade-offs among existing approaches?
  - Need to consider many other placement constraints
  - Could be multiple mixed-size domains:
    recursive MP-trees? Multi-domain unified formulation?

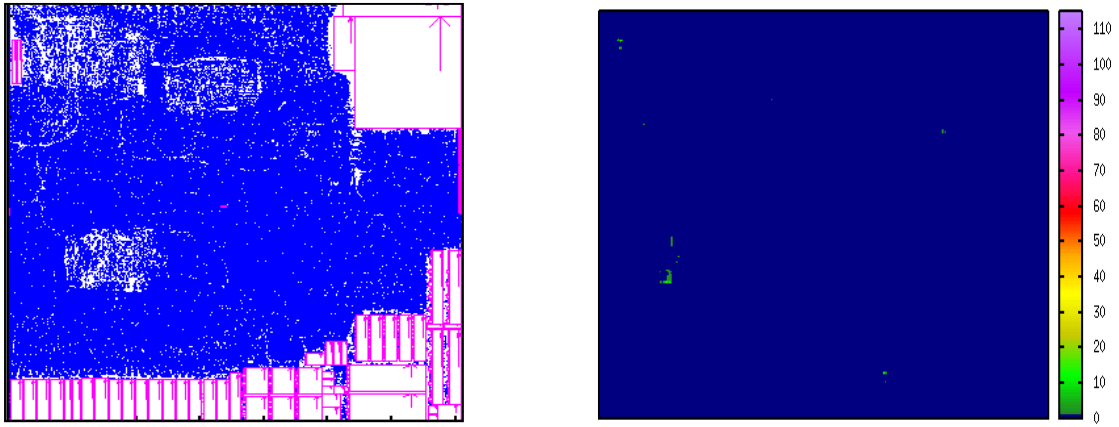




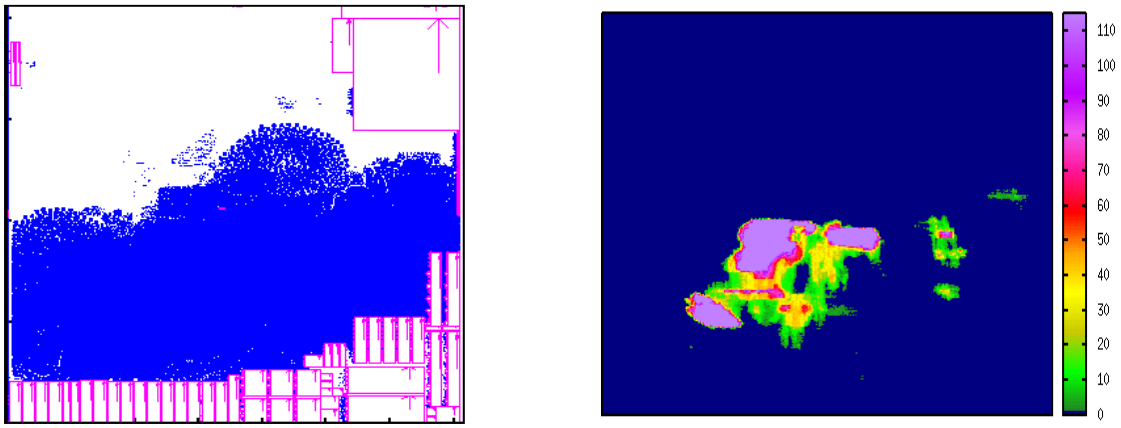
3. Routability-Driven Mixed-Size Placement

- Hsu *et. al*, "NTUplace4h: A novel routability-driven placement algorithm for hierarchical mixed-size circuit designs," IEEE TCAD, 2014

Routability-Driven



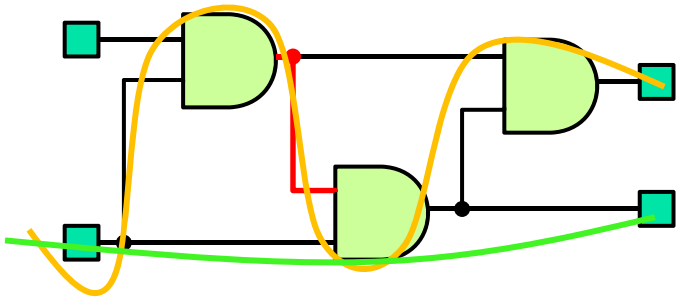
Wirelength-Driven



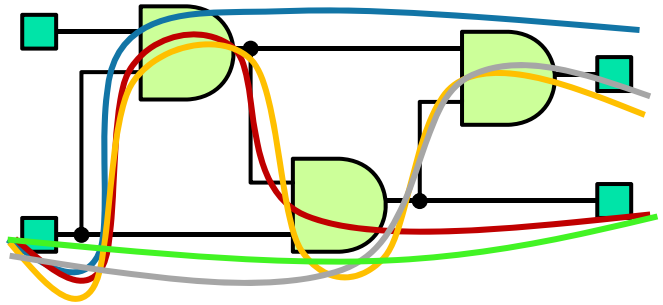
DAC'12 & ICCAD'13 & ISPD'15 Contests!!

4. Timing-Driven Placement

- Timing does not always co-relate with wirelength well
  - Path-based vs. net-based methods

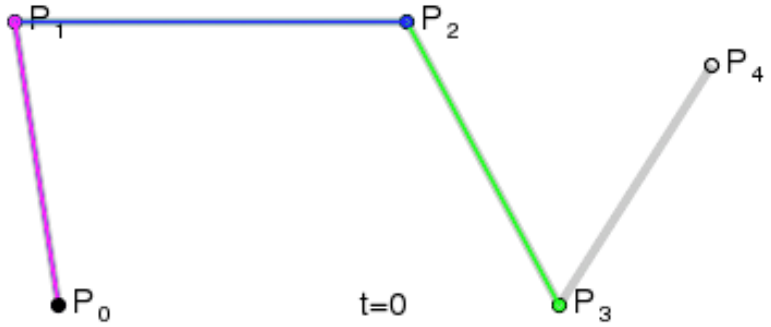


Net based method may not optimize timing effectively



timing paths increases exponentially

- Bézier curve based timing optimization

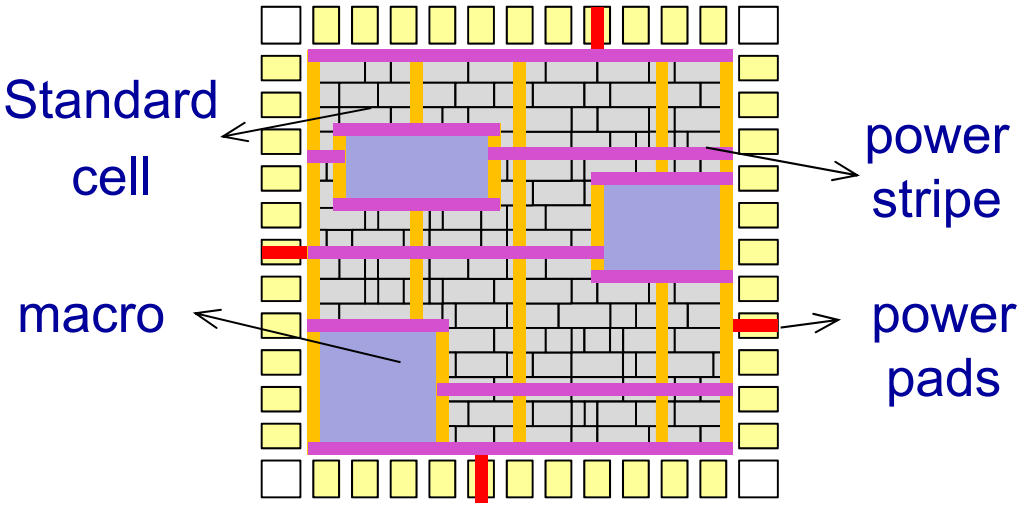


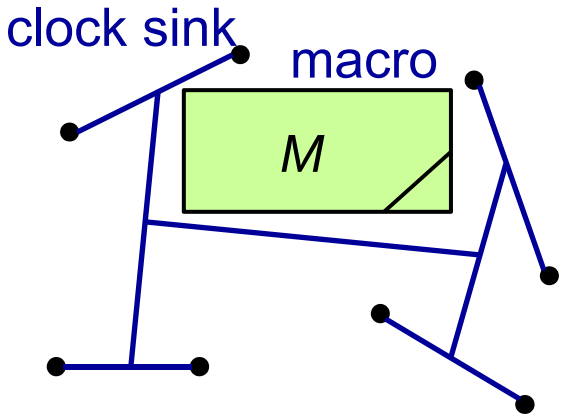
Source: http://en.wikipedia.org/wiki/Bézier_curve

ICCAD'15 Contests!!

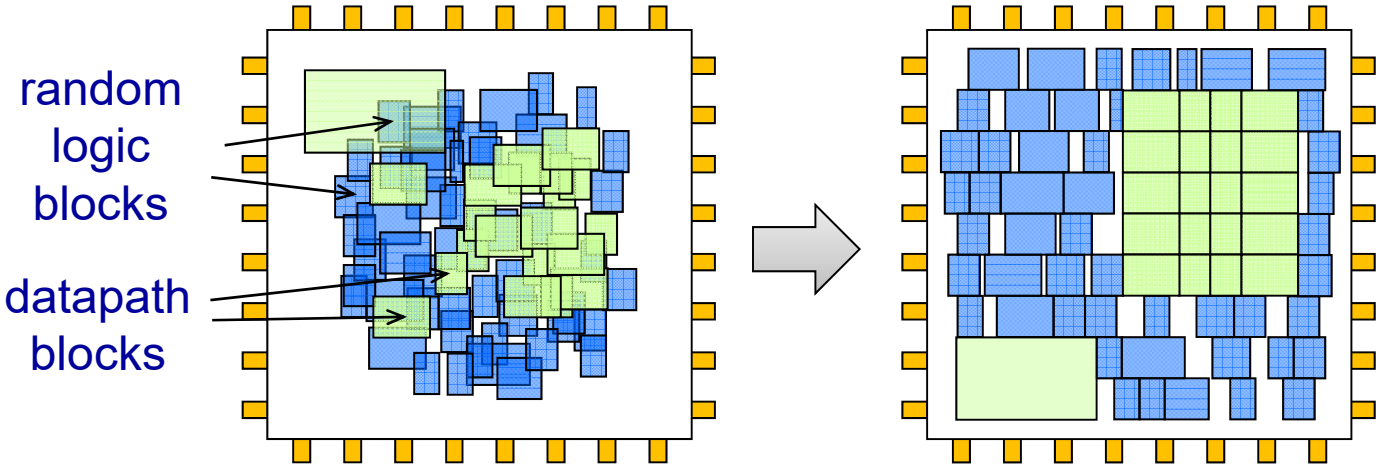
5. Co-synthesis w. Power/Ground/Clock Networks

- Chuang, Lee, and Chang, "Voltage-drop aware analytical placement by global power spreading for mixed-size circuit designs," ICCAD'09, TCAD'11
- Chuang, Lin, Ho, Chang, and Marculescu, "PRICE: Power reduction by placement and clock-network co-synthesis for pulsed-latch designs," ICCAD-2011.





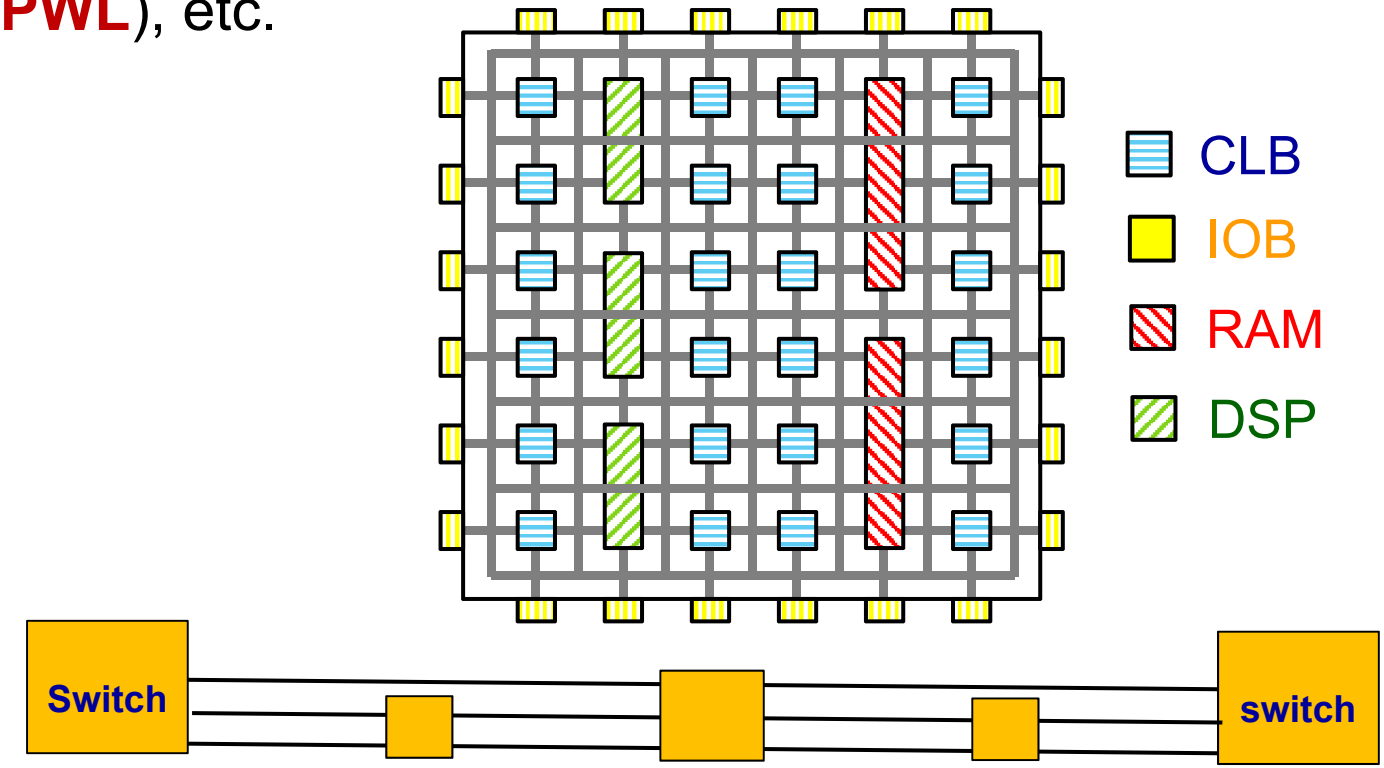
6. Datepath-Intensive Mixed-Size Placement



- Extract datapaths, and then place cells, macros, and datapath with regular structure
- Keys: regular/discrete structure, functional stage alignment, high-degree nets (**inaccurate HPWL**), etc.

7. Analytical FPGA Placement

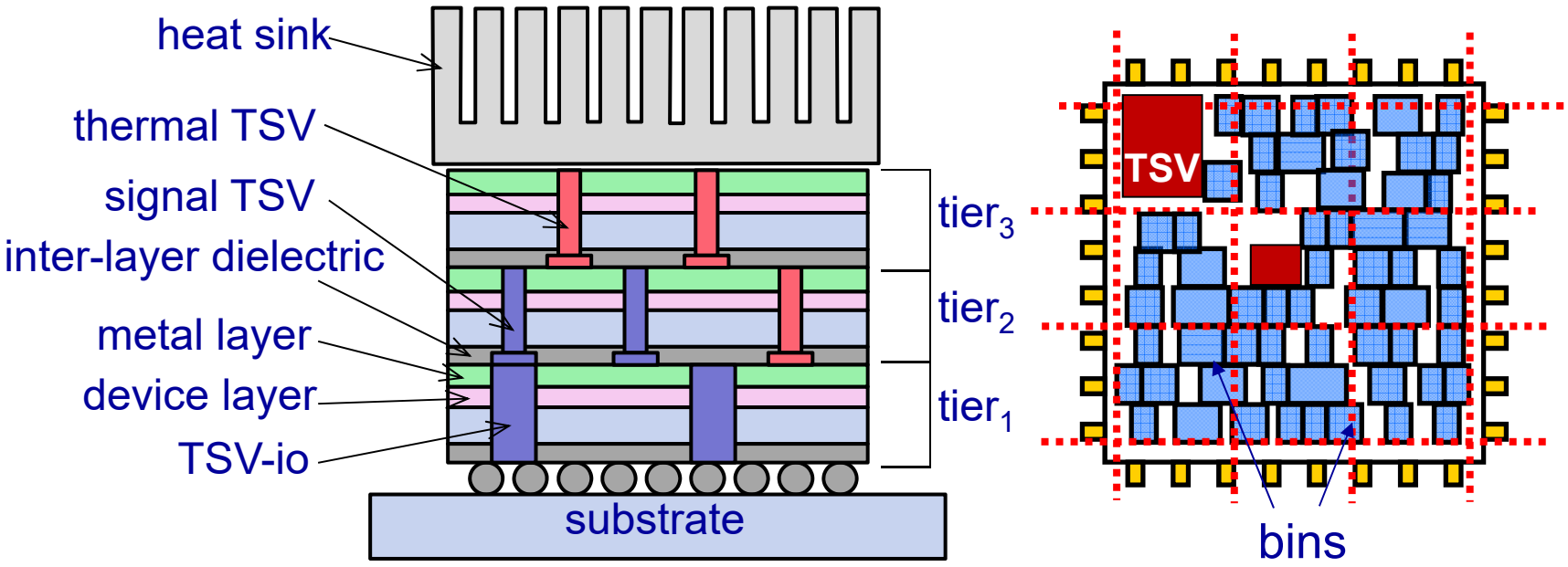
- Shift from combinatorial methods and simulated annealing to analytical placement [DAC'13, ICCAD'14, DAC'15]
- Keys: **heterogeneous logic structure, segmented wiring,** regular (discrete) structure, high-degree nets (**inaccurate HPWL**), etc.



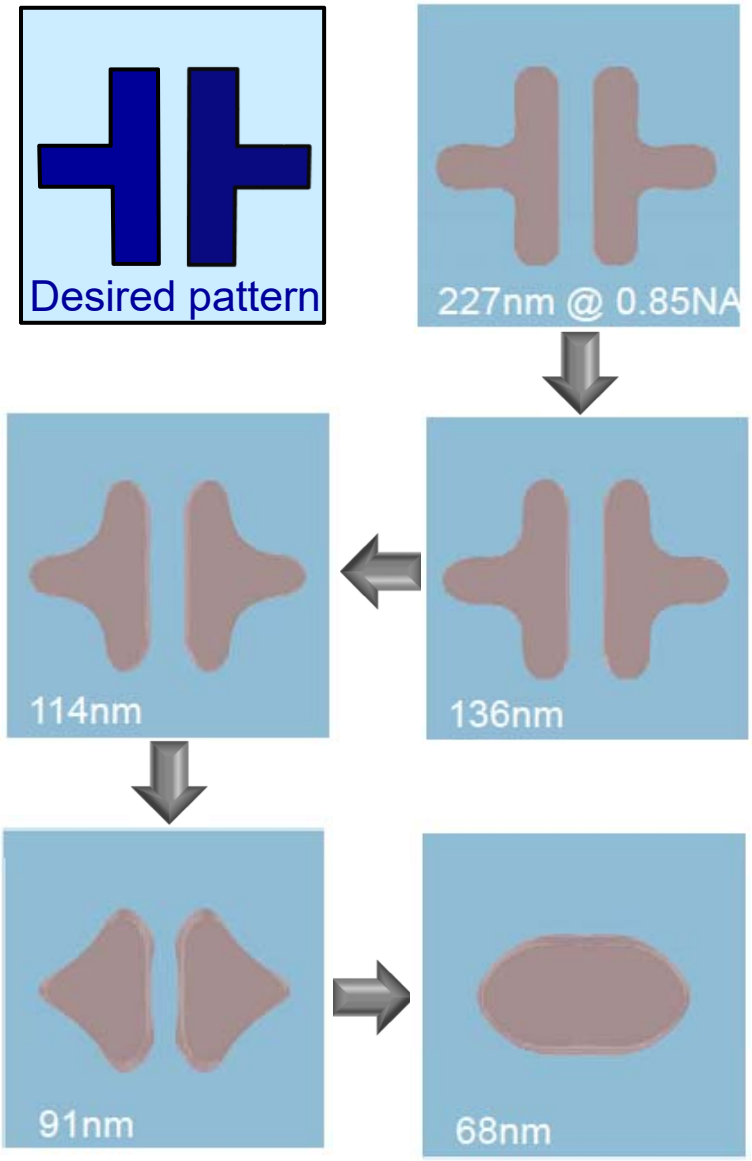
Segmented wiring **(HPWL is not accurate!!)**

8. Thermal-aware 3D IC Placement

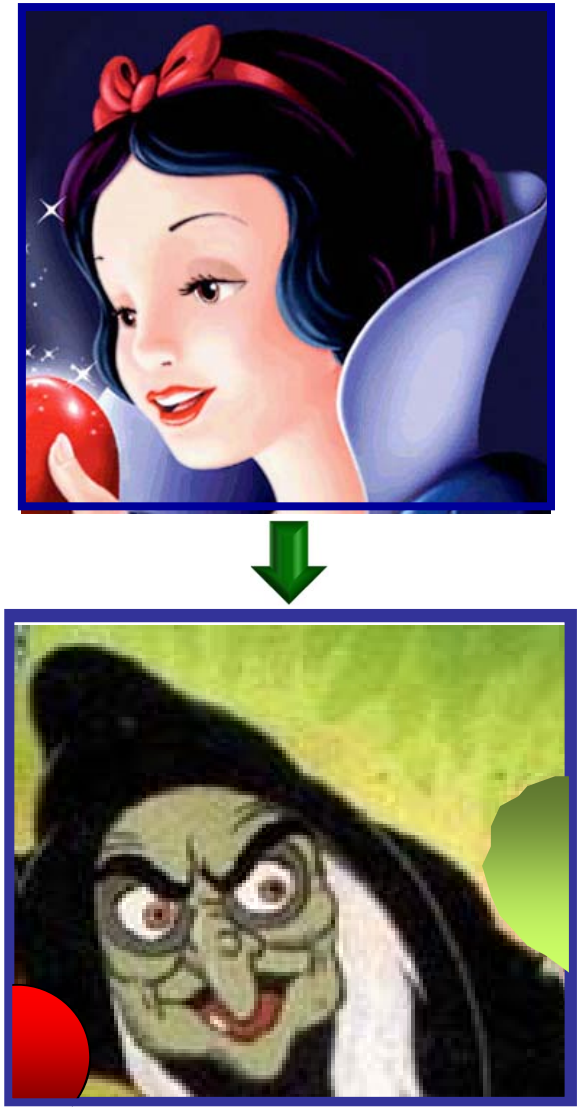
- Problem: Place cells into multiple tiers (dies) to optimize wirelength
- Important issues: reliability, thermal, routability, mixed-size design, etc.



9: Manufacturability: Sub-wavelength Lithography

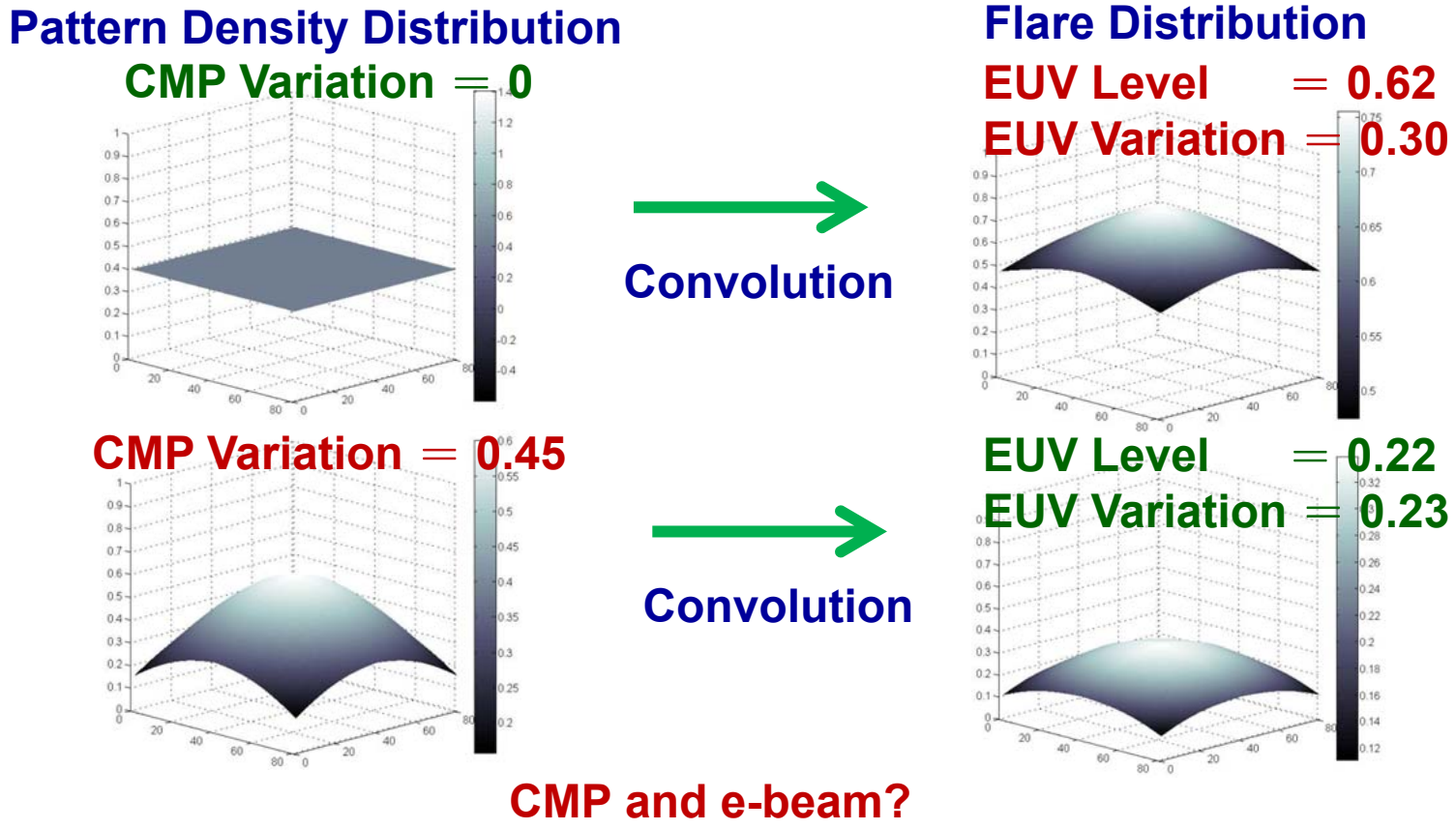


[Numerical Technologies]



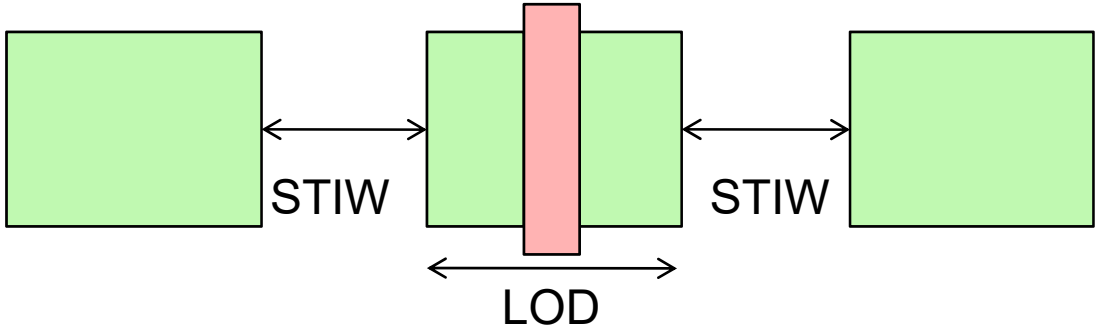
Manufacturability-Aware Placement

- Layout density for better CMP & EUV flare control
  - Control wire density for uniformity [Chen *et al.*, ISPD-08; Liu *et al.*, DAC'14]
- Multiple-pattering-aware placement & routing



10. STI Stress / FinFET Self-heating Aware Placement

- Shallow trench isolation (STI) is the mainstream CMOS isolation technique for advanced circuit design
  - By exploiting STI wells between device active regions, STI stress can effectively improve transistor performance
    - STI width (STIW) and length of diffusion (LOD)



  - $MOB_{L,R} = \gamma[(LOD/2)^{\alpha} + \beta/STIW_{L,R}]$ [Kahng et al., ICCAD'07]
    - If STIW↑ or LOD↓, then pMOS mobility ↑, but nMOS mobility ↓
- Problem: place cells to optimize STIW between neighboring cells for better performance