2024 International Symposium on Physical Design

Timing-Driven Analytical Placement According to Expected Cell Distribution Range

Jai-Ming Lin , You-Yu Chang, and Wei-Lun Huang

Speaker: Wei-Lun Huang
March 2024
Department of Electrical Engineering,
National Cheng Kung University, Tainan, Taiwan

Outline

- Introduction
- Our TDP Methodology
- Experimental Results
- Conclusion

Outline

- Introduction
- Our TDP Methodology
- Experimental Results
- Conclusion

Introduction

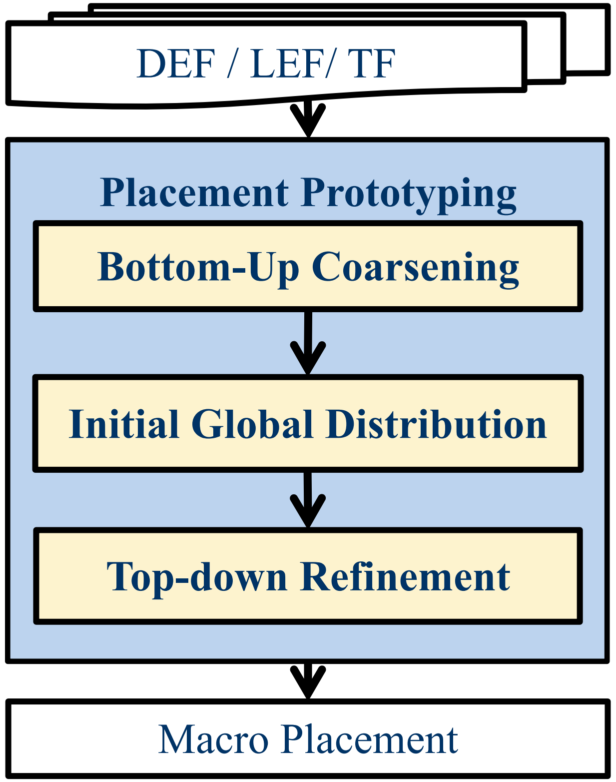
- Achieving timing closure is the largest concern to a modern design.
  - Tape-out time will be delayed if the timing requirement is not met.

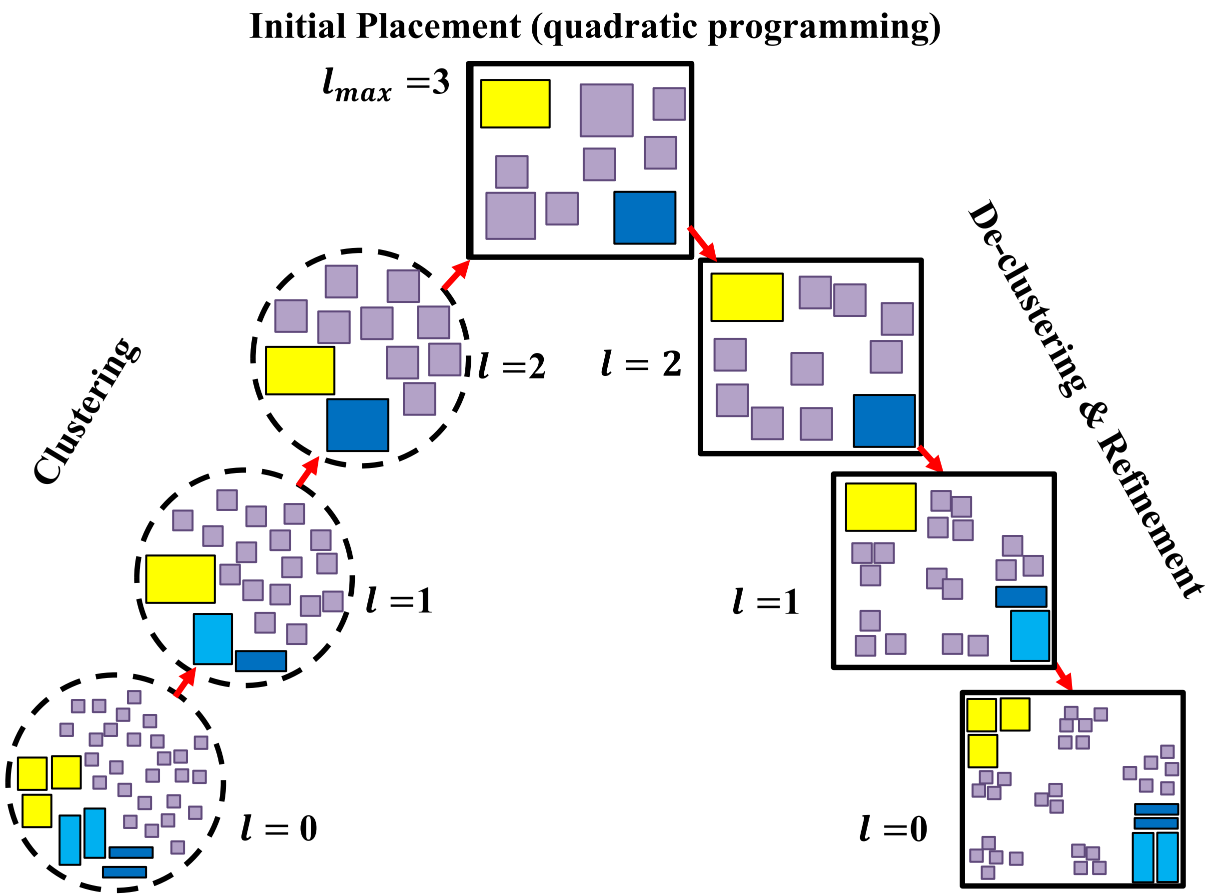
- Most of placement algorithms target on minimization of wirelength and routing congestion without considering timing.

- Timing closure can be achieved more quickly if a global placement algorithm can reduce WNS and TSN in addition to wirelength and routablity. Moreover, chip cost can be reduced as well since it does not have to insert a lot of buffers to timing violations nets.

Multilevel Framework in Our Work

- Multilevel Framework is composed of three stages.
  - Bottom-up coarsening.
  - Initial placement.
  - Top-down refinement.





Motivation

- **Lin *et al.*** [1] propose the design hierarchy-guided clustering which cluster objects layer by layer based on a design hierarchy tree.
- The score function to cluster two objects o_i & o_j is defined as:

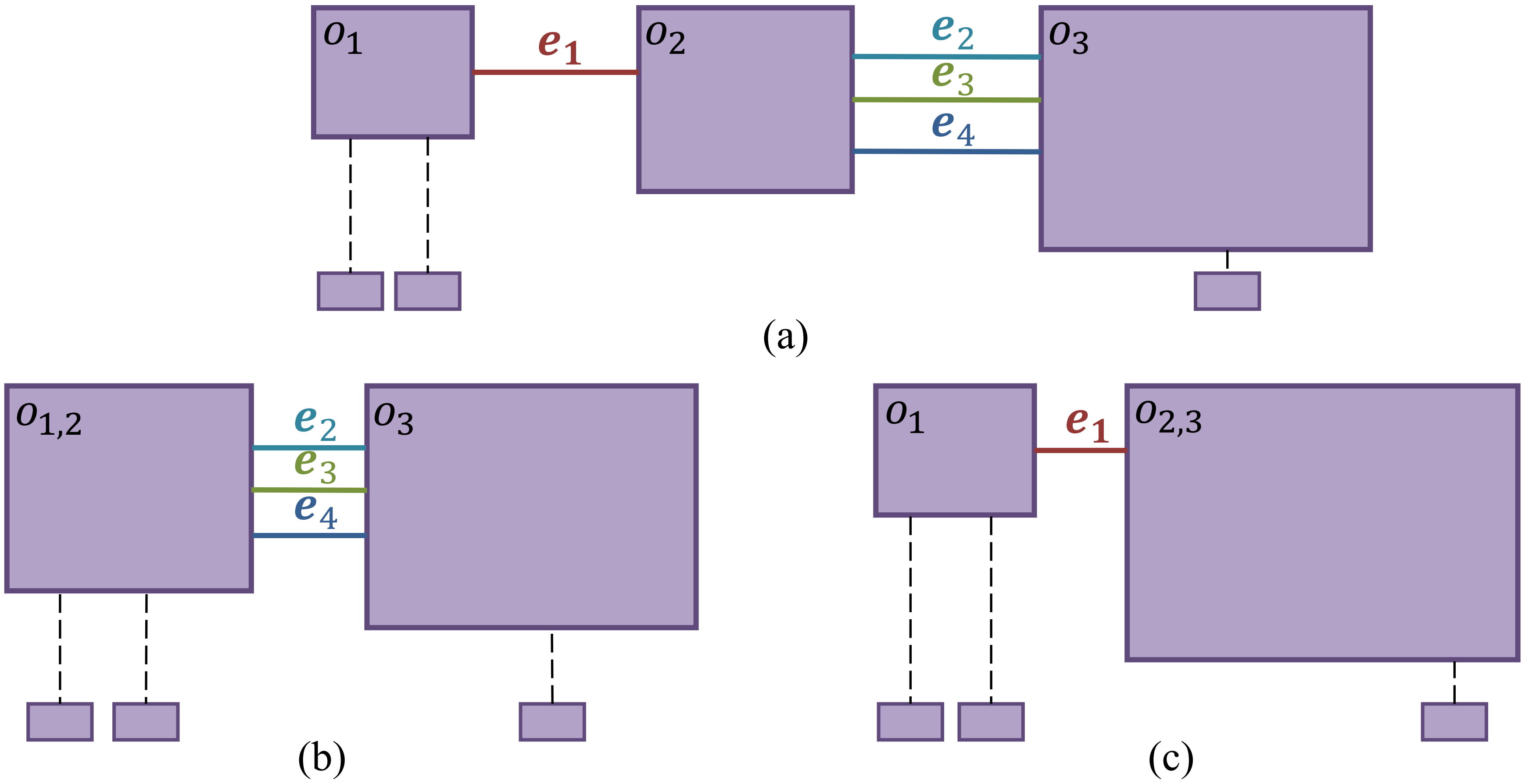
$$\boldsymbol{S(o_i, o_j)} = \rho(o_i, o_j) \times \Gamma(o_i, o_j) \times \left(\Phi(o_i, o_j) + \alpha \times \Psi(o_i, o_j)\right)$$

  - $\rho(o_i, o_j)$: Area function.
  - $\Gamma(o_i, o_j)$: Hierarchy function.
  - $\Phi(o_i, o_j)$: Connectivity function.
  - $\Psi(o_i, o_j)$: Indirect connectivity function.
  - α: The user-specified value.
- The value of score function is greatly determined by the combined area of o_i & o_j.

[1] J.-M. Lin, S.-T. Li and Y.-T. Wang, "Routability-driven Mixed-size Placement Prototyping Approach Considering Design Hierarchy and Indirect Connectivity Between Macros," *in Proc. of DAC*, pp. 1-6, 2019

Motivation (cont'd)

- The score function cannot get balance the combined area and pin-connectivity strength.
  - Two objects with stronger connectivity cannot be clustered because they have a larger combined area.

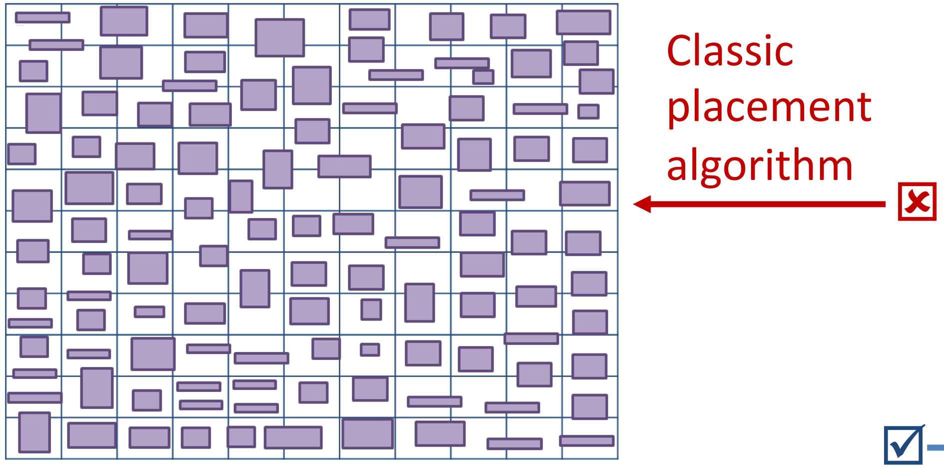


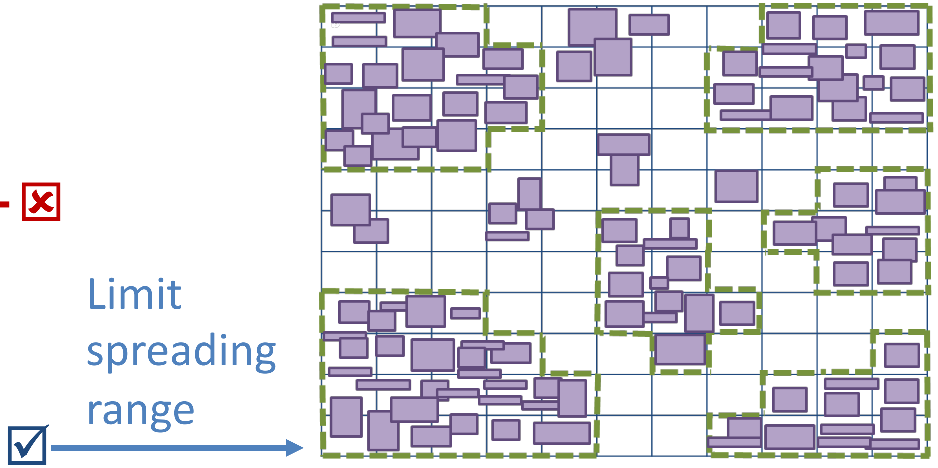
☒ S → Prohibition of large combined area. ☑ $\hat{S}$ → Balance of the connectivity and area term.

Motivation (cont'd)

- The classic multilevel placement placer [2] may generate longer wirelength since its formulation forces each bin to have an equal occupied area.

- A smaller wirelength can be obtained if the spreading ranges of objects are limited.





[2] T.-C. Chen, Z.-W. Jiang, T.-C. Hsu, H.-C. Chen and Y.-W. Chang, "NTUplace3: An Analytical Placer for Large-Scale Mixed-Size Designs with Preplaced Blocks and Density Constraints," *IEEE Trans. of TCAD*, Vol. 27, No. 7, pp. 1228-1240, 2008.

Our Contributions

- Improve the classic analytical placement approach to obtain a smaller wirelength.
  - Propose a pin-connectivity-aware score function to consider the pin relationship between two objects to balance the connectivity and area during clustering.
  - Propose an approach to identify expected distribution ranges of objects before the analytical formulation is solved.
- Propose the net-based timing-driven placement (TDP for short) algorithm under the multilevel framework.
  - Re-define timing-critical nets and only adjust weights for TCNs to reduce its impact on wirelength.
  - Estimate the weight of a net according to its degree in addition to timing slack.
  - Propose a new historical net weight adjustment equation to get a stable weight each time a new placement is obtained.
- Experimental results demonstrate that our TDP can get better timing result than the previous timing-driven placers [3], [4].

[3] Zizheng Guo and Yibo Lin. 2022. ”Differentiable-timing-driven global placement,” in *Proc. of DAC*, pp. 1315-1320, 2022.
[4] P. Liao, S. Liu, Z. Chen, W. Lv, Y. Lin and B. Yu, “DREAMPlace 4.0: Timing-driven global placement with momentum-based net weighting,” in *Proc. of DATE*, pp.939-944, 2022.

Outline

- Introduction
- Our Methodology
  - Pin-connectivity-aware Score Function
  - Expected Distribution Range in The Analytical Placement
  - Timing-driven Global Placement Algorithm
- Experimental Results
- Conclusion

Pin-connectivity-aware Score Function

- To balance pin-connectivity and object area, we propose pin-connectivity-aware score function $\hat{S}(o_i, o_j)$ to improve the score function [1] as follows:

$$\hat{\boldsymbol{S}}(\boldsymbol{o_i}, \boldsymbol{o_j}) = \mathcal{G}\left(\rho(o_i, o_j), \zeta(o_i, o_j)\right) \times \Gamma(o_i, o_j) \times \left(\Phi(o_i, o_j) + \alpha \times \Psi(o_i, o_j)\right)$$

  - $\mathcal{G}(\rho, \zeta)$: To trade-off between area and pin connective according to the Gompertz curve function.
    - $\rho(o_i, o_j)$: Area function.
  - Pin−connectivity function $\zeta(o_i, o_j)$: To estimate the true interconnection strength between o_i and o_j by considering the pin count.

Pin-connectivity-aware Score Function (cont'd)

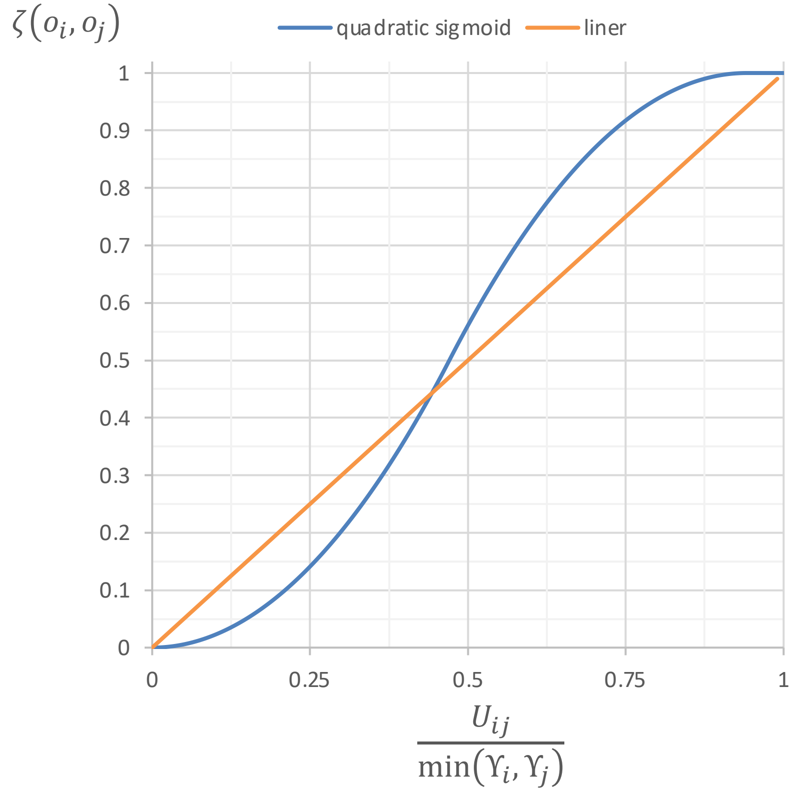
- The classic interconnection function $\Phi(o_i, o_j)$ simply counts the number of nets between two objects o_i and o_j.
- We propose a new pin connectivity function $\zeta(o_i, o_j)$ to measure the interconnection strength of o_i and o_j by considering their pin count numbers as follows:

$$\zeta(o_i, o_j) = \eta\left(\frac{U_{ij}}{\min(\Upsilon_i, \Upsilon_j)}\right)$$

  - U_{ij}: The number of connection between o_i and o_j.
  - Υ_i (Υ_j): The number of nets which are connected to o_i (o_j).
  - η denotes the quadratic sigmoid function, where

$$\eta(x) = \begin{cases} 1, & 1 \le \theta \times x \\ 1 - 2(\theta \times x - 1)^2, & 0.5 \le \theta \times x \le 1 \\ 2\theta \times x^2, & 0 \le \theta \times x \le 0.5 \\ 0, & \theta \times x < 0 \end{cases}$$

    - θ is a user-specified value (θ is set as 1.0625 in our experiment).



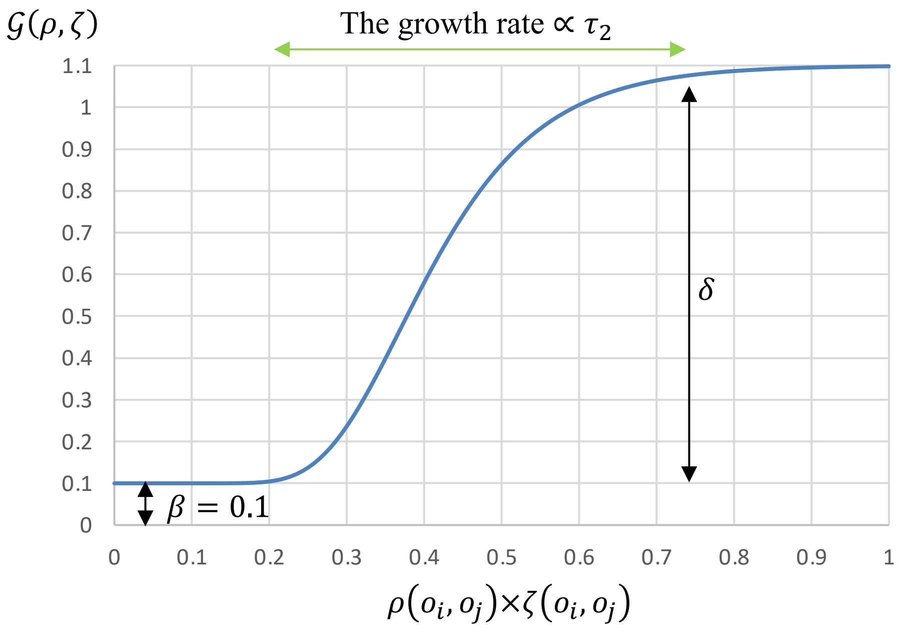
Pin-connectivity-aware Score Function (cont'd)

- Apply the Gompertz curve function $\mathcal{G}(\rho,\zeta)$ to balance the two values $\rho(o_i,o_j)$ and $\zeta(o_i,o_j)$ to avoid the result being overly determined by $\rho(o_i,o_j)$ as follows:

$$\mathcal{G}(\rho,\zeta) = \beta + \delta \times e^{-\tau_1 e^{-\left(\tau_2 \times \rho(o_i,o_j) \times \zeta(o_i,o_j)\right)}}$$

- β: The Gompertz curve function bias, default as 0.1.
- δ: The magnification of the function,

  where $\delta = \max\left(\frac{Number\ of\ nets}{Number\ of\ objects}, 1\right)$.
- τ_1: The displacement along the x-axis, default as 40.
- τ_2: The growth rate, default as 10.



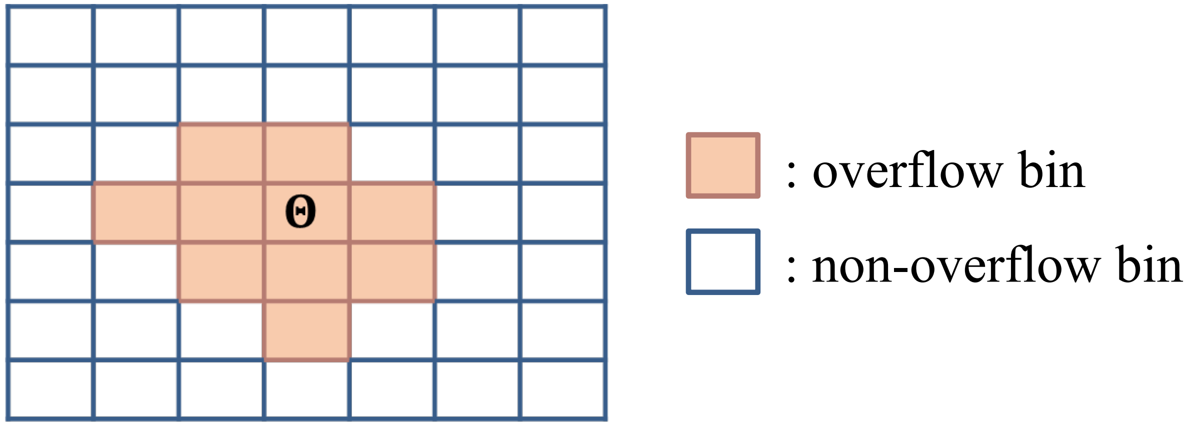
Outline

- Introduction
- Our Methodology
  - Pin-connectivity-aware Score Function
  - Expected Distribution Range in The Analytical Placement
  - Timing-driven Global Placement Algorithm
- Experimental Results
- Conclusion

Preliminaries

- Identify overflow zones Θ's after an initial placement is obtained.
  - A bin b_i is considered as overflowed if $D_{b_i} > M_{b_i}$.
  - Θ is composed of overflow bins at contiguous locations.

- Distribute objects from each overflow zone Θ to the corresponding diffusion region, denoted by $\widehat{\Theta}$.
  - A diffusion region is the largest area where the objects in Θ can be spread.



Procedure to Determine Expected Distribution Ranges of Objects

- $\widehat{\Theta}$ is set as Θ in the beginning.

- A diffusion region $\widehat{\Theta}$ is generated by gradually expanding the area of Θ ring by ring until $M_{\widehat{\Theta}} \geq D_{\widehat{\Theta}}$.
  - $M_{\widehat{\Theta}}$ denotes the target placeable area in $\widehat{\Theta}$.
  - $D_{\widehat{\Theta}}$ denotes the total cell potential in $\widehat{\Theta}$.

- Divide an overflow zone Θ into κ sub-regions (κ is a user-specified value and is set as 4 or 8 in our experiment) and spread objects to each direction according to the ratio of the area of original objects in the direction to the total area in Θ.
  - The spreading speed in each direction should be different since objects distribution is unbalanced.

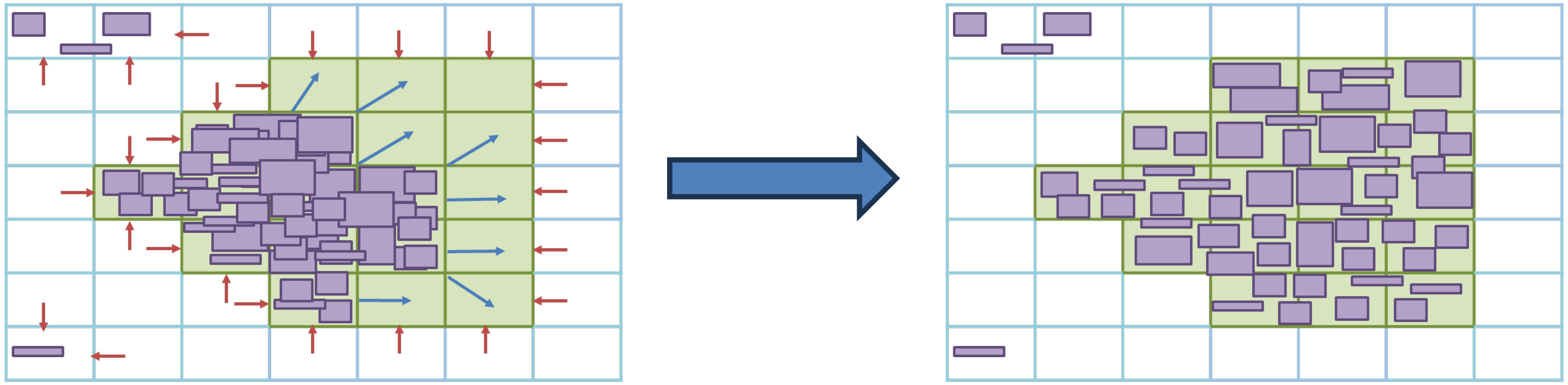
Maximum Placeable Area of Each Bin in a Diffusion Region

- Objects are spread according to the following analytical placement formulation after all diffusion regions $\widehat{\Theta}$'s are determined:

$$\min \sum_{e_j \in E} W_j(x,y) + \lambda \sum_{b} \left(D_b(x,y) - \widehat{M}_b\right)^2$$

  - $\widehat{M}_b$ is determined by the following equation:

$$\widehat{M}_b = \begin{cases} M_b, & if\ b \in \widehat{\Theta} \\ D_b, & otherwise \end{cases}$$



: bin $\notin \widehat{\Theta}$: bin $\in \widehat{\Theta}$: object : repulsion force : attraction force

Outline

- Introduction
- Our Methodology
  - Pin-connectivity-aware Score Function
  - Expected Distribution Range in The Analytical Placement
  - Timing-driven Global Placement Algorithm
- Experimental Results
- Conclusion

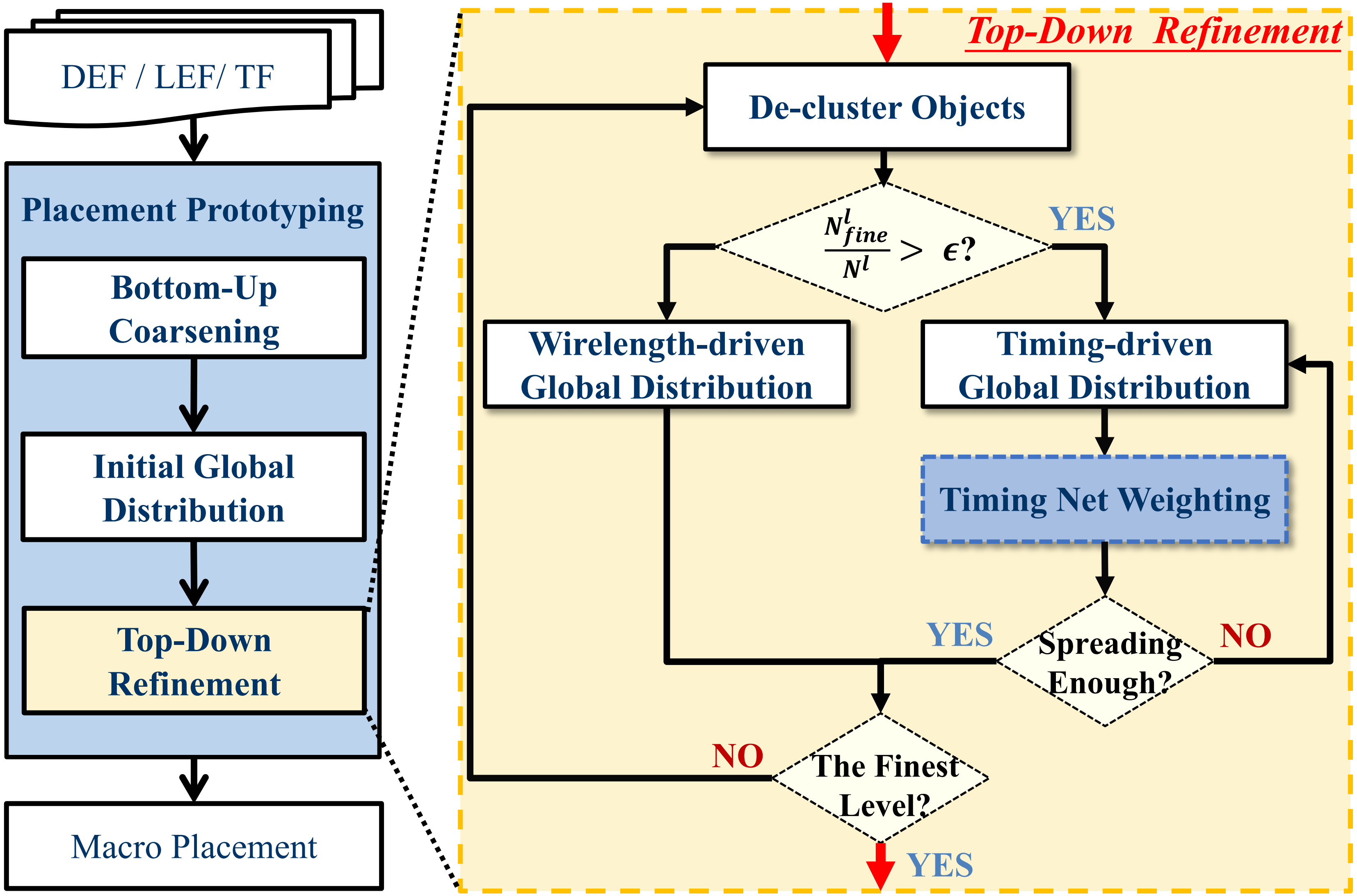
Overview for Our TDP Algorithm

- Apply the wirelength-driven global distribution in the beginning.

- Apply the timing-driven global placement when the netlist of the current level contains enough number of fine objects.

$$\min \sum_{e_j \in E} \omega_j W_j(x,y) + \lambda \sum_b (D_b(x,y) - M_b)^2$$

- ω_j denotes the weight of $W_j(x,y)$ for e_j.



Calculation of Net Weights

- A net e_j is considered as a timing-critical net (TCN) if one segment s_i of e_j belongs to a timing violation path p_u and its delay π_i meets the following condition:
  - ς is a user-specified parameter (ς is set as 0.05 in our experiment).

$$\pi_i > \varsigma \times \text{path delay of } p_u$$

- The weight ω_j of a TCN e_j is proportional to the slack of the path p_u ($p_u \in P_j$) which has the smallest value as follows, where P_j denotes a set of paths which contain the segment s_i of e_j.

$$\omega_j \propto \max_{p_u \in P_j}\left(\frac{\sigma_u}{\sigma_{min}}, \Omega\right)$$

  - σ_u denotes the timing slack of p_u and σ_{min} represents the minimum slack among all timing violation paths.
  - Ω is a user-specified parameter (Ω is set as 0.2 in our experiment).

Calculation of Net Weights (cont'd)

- The effect of weight adjustment will be reduced as the degree of a net increases.
- The function $L(d_j)$ adjust the weight of e_j according to its degree by a logistic function as follows:

$$L(d_j) = \frac{1}{1 + e^{(\frac{d_j}{\mu_1} - \mu_2)}}$$

  - d_j: the degree of e_j.
  - μ_1, μ_2: user-specified values (μ_1 and μ_2 are set as 3, 5 in our experiment, respectively).
  - The value of $L(d_j)$ is in the range of [0,1] and the value reduces as the value d_j increases.
- The weight ω_j of a net e_j is adjusted by the following equation:

$$\omega_j = \begin{cases} 1 + L(d_j) \max\limits_{p_u \in P_j} \left(\frac{\sigma_u}{\sigma_{min}}, \Omega \right), & if\ e_j\ is\ a\ TCN \\ 1, & otherwise \end{cases}$$

Outline

- Introduction
- Our Methodology
- Experimental Results
- Conclusion

Experimental Results

- Environments:

Programming Language		C++
Linux Workstation	CPU	Intel® Xeon® Silver 4110 2.1GHz
	Memory	400GB
	System	Cent OS 6.9

- This experiment is divided into two parts.
  - Effect of improved analytical placement approach with respect to classic analytical placement approach.
  - Comparison with other timing-driven placers.

Effect of Improved Analytical Placement Approach

- Benchmarks: Industrial designs.
  - Placement legalization and signal routing are completed by IC Compiler 2.

Benchmark	# Movable Macros	# Preplaced Macros	# IO Pads	# Standard Cells	# Nets	# Pins
Cir. 1	29	2	684	111339	116928	477911
Cir. 2	81	5	1219	340130	365605	1328667
Cir. 3	48	101	19323	497704	509005	2182323
Cir. 4	146	36	190	1425892	1471214	6037846
Cir. 5	855	69	263	5555610	5797026	22340831

Effect of Improved Analytical Placement Approach (cont'd)

- Demonstrate the effectiveness of two techniques associated with wirelength.

	Classic Analytical Placement		with Pin-connectivity-aware Score Function		with Expected Distribution Range of Objects		Our Analytical Placement	
Benchmark	**WL**	**RT**	**WL**	**RT**	**WL**	**RT**	**WL**	**RT**
Cir. 1	0.622	205.73	0.634	149.91	0.646	204.57	0.655	169.51
Cir. 2	1.453	1158.76	1.417	1085.73	1.251	1307.15	1.103	1142.53
Cir. 3	1.799	333.17	1.764	257.85	1.728	281.93	1.705	277.46
Cir. 4	8.877	3051.78	8.713	2985.94	8.517	3158.65	8.428	2925.18
Cir. 5	42.447	24569.8	40.432	20320.6	34.072	19944.4	30.072	21410.49
Nor.	1.16	1.12	1.13	0.95	1.06	1.08	1.00	1.00

- WL: wirelength ($\times 10^7$ μm).
- RT: runtime (sec).

Comparison with Other Timing-driven Placers

- Benchmarks: ICCAD 2015 contest benchmark suite.
  - STA engine: an open-source timer, OpenTimer.
  - Placement legalization is performed using Innovus.
  - The timing slack was estimated by the evaluation script provided by the ICCAD 2015 contest.

Benchmark	#Cells	#Nets	#Pins	#Rows
superblue1	1209716	1215710	3767494	1829
superblue3	1213253	1224979	3905321	1840
superblue4	795645	802513	2497940	1840
superblue5	1086888	1100825	3246878	2528
superblue7	1931639	1933945	6372094	3163
superblue10	1876103	1898119	5560506	3437
superblue16	981559	999902	3013268	1788
superblue18	768068	771542	2559143	1788

Comparison with Other Timing-driven Placers (cont'd)

- We compare our TDP with state-of-the-art timing-driven placers including DREAMPlace4.0 [4] and Differentiable TDP [3].

	DREAMPlace4.0 [3]		Differentiable TDP [1]		Our Work	
Benchmark	**WNS**	**TNS**	**WNS**	**TNS**	**WNS**	**TNS**
superblue1	-14.10	-85.03	-10.77	-74.85	-9.26	-42.10
superblue3	-16.43	-54.74	-12.37	-39.43	-12.19	-26.59
superblue4	-12.78	-144.38	-8.49	-82.92	-8.86	-123.28
superblue5	-26.76	-95.78	-25.21	-108.08	-31.64	-70.35
superblue7	-15.21	-63.86	-15.22	-46.43	-17.24	-95.89
superblue10	-31.88	-768.75	-21.97	-558.05	-25.86	-691.10
superblue16	-12.11	-124.18	-10.85	-87.02	-12.21	-55.99
superblue18	-11.87	-47.25	-7.99	-19.31	-5.25	-19.23
Nor.	1.32	1.63	1.01	1.17	1.00	1.00

- WNS: worst negative slack ($\times 10^3\ ps$).
- TNS: total negative slack ($\times 10^5\ ps$).

[3] Zizheng Guo and Yibo Lin. 2022. "Differentiable-timing-driven global placement," in *Proc. of DAC*, pp. 1315-1320, 2022.
[4] P. Liao, S. Liu, Z. Chen, W. Lv, Y. Lin and B. Yu, "DREAMPlace 4.0: Timing-driven global placement with momentum-based net weighting," in *Proc. of DATE*, pp.939-944, 2022.

Outline

- Introduction
- Our Methodology
- Experimental Results
- Conclusion

Conclusion

- We have proposed an improved analytical placement algorithm.
  - Pin-connectivity-aware score function.
  - Expected distribution range of objects.

- We have proposed a TDP algorithm based on an improved analytical placement algorithm.
  - Determine timing net weights only for TCNs.
  - Estimate net weights based on the degree of a net in addition to consideration of timing slacks.
  - Propose a new equation to update net weights based on their historical values.

- The experimental results have demonstrated that our approach can obtain better results than the previous approaches.

End

Thank You For Your Attention