WHERE **INNOVATION** BEGINS

DESIGN AUTOMATION CONFERENCE

JULY 9–13, 2023

MOSCONE WEST CENTER
SAN FRANCISCO, CA, USA

DESIGN AUTOMATION CONFERENCE
WHERE INNOVATION BEGINS

On a Moreau Envelope Wirelength Model for Analytical Global Placement

Peiyu Liao[1,2], Hongduo Liu[2], Yibo Lin[1,3], Bei Yu[2], Martin Wong[2]

[1]Peking University
[2]Chinese University of Hong Kong
[3]Institute of Electronic Design Automation, Peking University

Global Placement

- Place standard cells (rectangles) onto valid region.
- **Legality**: no overlap; row alignment; site alignment.
- **Objective**: minimize total wirelength (routed? HPWL? Steiner-tree).

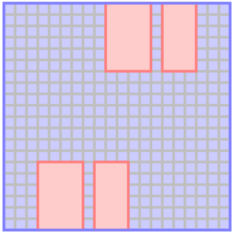


Figure: Fill standard cells into blue region.

Wirelength Models

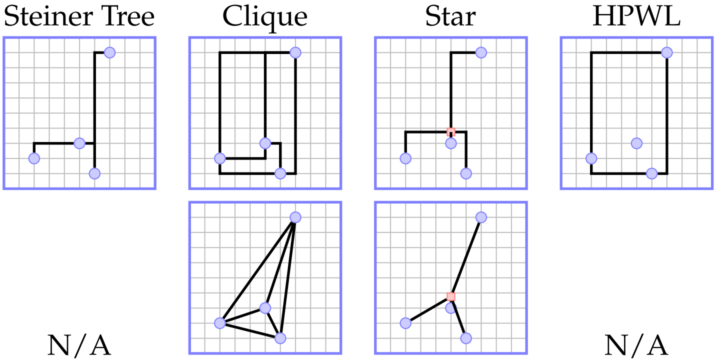


Table: Examples of different wirelength models. The first row shows models using ℓ_1 distance, while the second shows those using ℓ_2 distance.

Wirelength Models

- Steiner tree is the most accurate model. *HARD to optimize*.
- All ℓ_2 models are smooth and easy to optimize.
- HPWL is the most widely adopted one. *Why?*

Model	Steiner Tree	Clique		Star		HPWL
Distance	ℓ_1	ℓ_1	ℓ_2	ℓ_1	ℓ_2	ℓ_1
Accuracy	■■■■■	■	■	■■■	■■	■■■■
Smoothness	■	■■	■■■■■	■■	■■■■■	■■

Table: The approximation accuracy and smoothness of different models.

Differentiable Approximations

Two widely-used models: the *log-sum-exp* model[1] and the *weighted-average* model [DAC'11],

$$\begin{aligned} W_e(\boldsymbol{x}) &\approx W^{\gamma}_{e,\mathrm{LSE}}(\boldsymbol{x}) = \gamma \ln \sum_{i=1}^{n} \mathrm{e}^{\frac{x_i}{\gamma}} + \gamma \ln \sum_{i=1}^{n} \mathrm{e}^{-\frac{x_i}{\gamma}}, \\ W_e(\boldsymbol{x}) &\approx W^{\gamma}_{e,\mathrm{WA}}(\boldsymbol{x}) = \frac{\sum_{i=1}^{n} x_i \mathrm{e}^{\frac{x_i}{\gamma}}}{\sum_{i=1}^{n} \mathrm{e}^{\frac{x_i}{\gamma}}} - \frac{\sum_{i=1}^{n} x_i \mathrm{e}^{-\frac{x_i}{\gamma}}}{\sum_{i=1}^{n} \mathrm{e}^{-\frac{x_i}{\gamma}}}, \end{aligned} \tag{1}$$

where γ is the precision parameter. Uniform convergence?

[1] W. C. Naylor *et al.*, *Non-linear optimization system and method for wire length and delay optimization for an automatic electric circuit placer*, US Patent 6,301,693, Oct 9, 2001.

Differentiable Approximations

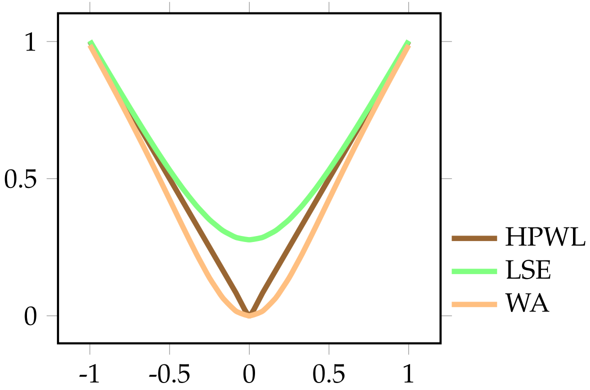


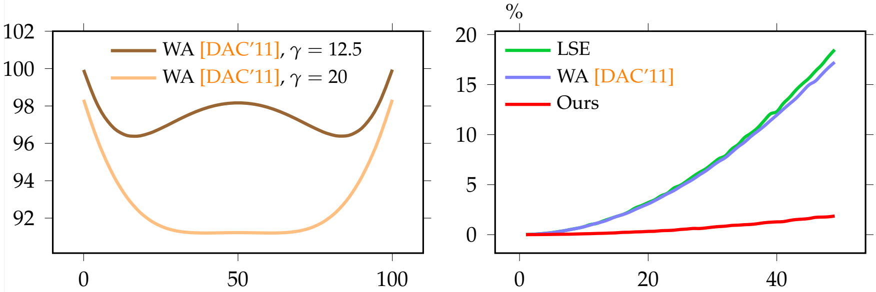
Figure: A simple example of approximating $\max\{x, 0\} - \min\{x, 0\} = |x|$ (2-pin net) with differentiable models, under $\gamma = 0.2$.

Differentiable Approximations

The WA model is good, but are there any drawbacks?

- **Numerical Stability**. It occurs due to the exponential function.
- **Non-Convexity**. It can be easily verified in the above figure. The non-convexity may get more complicated in high-dimensional cases of real designs.
- **Approximation Error**. The exponential terms provide high precision when γ is very small, but mostly γ is not an ϵ-like value.

Differentiable Approximations



(a) (b)

Figure: (a) The non-convexity of WA [DAC'11] on a simple 3-pin net to approximate $\Delta x = \max\{x_{\min}, x, x_{\max}\} - \min\{x_{\min}, x, x_{\max}\}$. (b) The average approximation against γ or t for 4-pin nets, under fixed $\Delta x = 200$.

Moreau Envelope

In our placement applications, we only consider *closed convex* functions $h(\boldsymbol{x})$ defined in $\mathbb{R}^n$ to simplify the notations.

Definition 1 (Moreau Envelope)

For any $t > 0$, the **Moreau envelope** function h^t is defined by

$$h^t(\boldsymbol{x}) = \min_{\boldsymbol{u} \in \mathbb{R}^n} \left\{ h(\boldsymbol{u}) + \frac{1}{2t} \|\boldsymbol{u} - \boldsymbol{x}\|_2^2 \right\}. \tag{2}$$

Usually, it does not have an explicit closed-form representation.

Moreau Envelope

We have the following facts.

Fact 2 (Point-wise Convergence)

We always have the point-wise convergence $\lim_{t \to 0^+} h^t(\boldsymbol{x}) = h(\boldsymbol{x})$.

Fact 3 (Differentiability)

The envelope-theorem states that $\nabla_{\boldsymbol{x}} h^t(\boldsymbol{x}) = \frac{1}{t}(\boldsymbol{x} - \mathrm{prox}_{th}(\boldsymbol{x}))$.

Moreau Envelope

Replace h^t with the HPWL function approximation W_e^t, defined as

$$W_e^t = \min_{\boldsymbol{u}\in\mathbb{R}^n} \left\{ W_e(\boldsymbol{u}) + \frac{1}{2t}\|\boldsymbol{u}-\boldsymbol{x}\|_2^2 \right\}, \tag{3}$$

where W_e is the horizontal (vertical) half-perimeter wirelength function of net $e \in E$,

$$W_e(\boldsymbol{x}) = \max_{1\le i\le n} x_i - \min_{1\le i\le n} x_i. \tag{4}$$

Moreau Envelope

Replace h^t with the HPWL function approximation W_e^t, defined as

$$W_e^t = \min_{\boldsymbol{u} \in \mathbb{R}^n} \left\{ W_e(\boldsymbol{u}) + \frac{1}{2t} \|\boldsymbol{u} - \boldsymbol{x}\|_2^2 \right\}, \tag{3}$$

where W_e is the horizontal (vertical) half-perimeter wirelength function of net $e \in E$,

$$W_e(\boldsymbol{x}) = \max_{1 \le i \le n} x_i - \min_{1 \le i \le n} x_i. \tag{4}$$

The forward computation of $W_e^t(\boldsymbol{x})$ and the backward computation $\nabla W_e^t(\boldsymbol{x})$ is **possible**.

Moreau Envelope

The key problem is to find how to compute the proximal point,

$$\operatorname{prox}_{tW_e}(\boldsymbol{x}) = \operatorname*{argmin}_{\boldsymbol{u}\in\mathbb{R}^n} \left\{ W_e(\boldsymbol{u}) + \frac{1}{2t}\|\boldsymbol{u}-\boldsymbol{x}\|_2^2 \right\}. \tag{5}$$

We have to solve the above optimization problem.

- **Cheap**. The approximation will work.
- **Expensive**. The approximation is only symbolic, and it is hard to make it practical.

$W_e^t + t$ will be considered to be the approximation.

Moreau Envelope

Gradient Property: the gradient of Moreau envelope function W_e^t is $\boldsymbol{g} = \nabla W_e^t(\boldsymbol{x})$ where

$$g_i = \begin{cases} \frac{1}{t}(x_i - \tau_2), & \text{if } x_i > \tau_2; \\ 0, & \text{if } \tau_1 \leq x_i \leq \tau_2; \\ \frac{1}{t}(x_i - \tau_1), & \text{otherwise} \end{cases} \tag{6}$$

is defined for any $i = 1, \cdots, n$, such that

$$\sum_{i=1}^{n}(x_i - \tau_2)^+ = \sum_{i=1}^{n}(\tau_1 - x_i)^+ = t, \tag{7}$$

if the solution τ_1, τ_2 to (7) satisfy $\tau_1 \leq \tau_2$, otherwise $\boldsymbol{g} = \nabla W_e^t(\boldsymbol{x})$ is determined by the average coordinate: $g_i = \frac{1}{t}x_i - \frac{1}{tn}\sum_{i=1}^{n} x_i$ for any index $i = 1, \cdots, n$.

Moreau Envelope

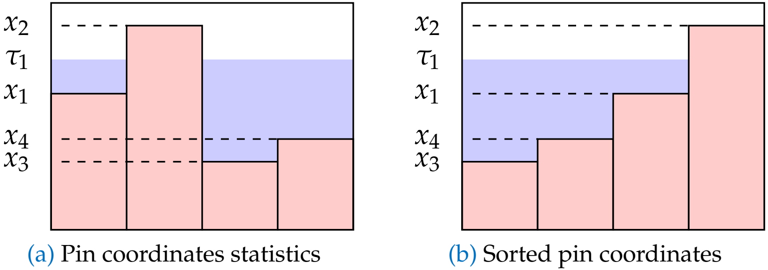


Figure: The illustration of water-filling to solve τ_1 in Equation (7).

Moreau Envelope

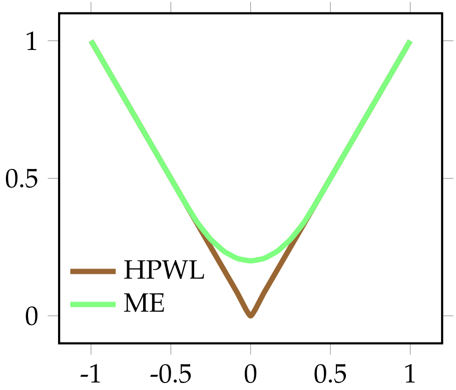


Figure: A simple example of approximating $\max\{x, 0\} - \min\{x, 0\} = |x|$ (2-pin net) with ME model, under $t = 0.2$.

Moreau Envelope

The closed-form representation is

$$W_e(x) \approx W_e^t(x) + t = \begin{cases} \frac{x^2}{4t} + t, & \text{if } |x| \leq 2t, \\ |x|, & \text{otherwise} \end{cases} \tag{8}$$

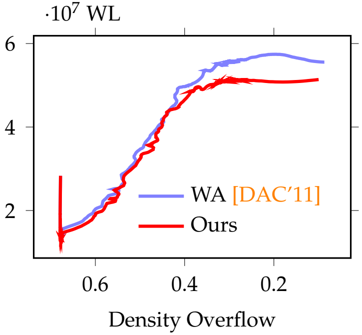
You may know about the **Huber loss**,

$$L_\delta(a) = \begin{cases} \frac{1}{2}a^2, & \text{if } |a| \leq \delta, \\ \delta(|a| - \frac{1}{2}\delta), & \text{otherwise} \end{cases} \tag{9}$$

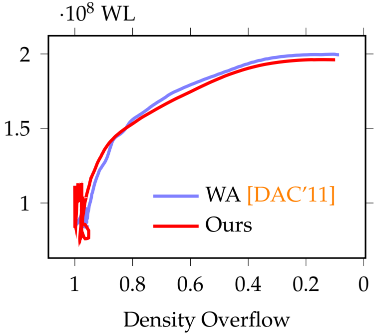
Then we have equality $W_e^t(x) = \frac{1}{2t}L_{2t}(x)$. The model is the multidimensional generalization of Huber loss.

A Test on ISPD'06 and ISPD'19

The wirelength curve against density overflow.



(a) ISPD2006 `newblue1`



(b) ISPD2019 `ispd19_test10`

THANK YOU!