2025 62nd ACM/IEEE Design Automation Conference (DAC) | 979-8-3315-0304-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/DAC63849.2025.11132593

Late Breaking Results: A Geometric Diffusion Model for Macro Placement Generation

Jongho Yoon*[1], Jinsung Jeon*[2], and Seokhyeong Kang[1]
[1]*Pohang University of Science and Technology*, Pohang, Korea
[2]*University of California, San Diego*, San Diego, California, U.S
[1]{jh.yoon, shkang}@postech.ac.kr [2]jjsjjs0902@gmail.com

***Abstract*—Macro placement is crucial in VLSI design, directly impacting circuit performance. We introduce MacroDiff, a diffusion-based macro placement generative model that captures wirelength relationships instead of directly predicting macro coordinates. By leveraging wirelength as an intermediate representation, MacroDiff naturally preserves circuit connectivity, reduces placement constraints, and enhances solution flexibility while inherently handling rotational and translational invariances. Experiments on *ISPD2005* benchmarks show that MacroDiff reduces macro overlap by 91.6%, lowers macro legalization displacement by 74.4%, and improves half-perimeter wirelength (HPWL) by 7.0%. While maintaining the efficiency of generative approaches, MacroDiff generates high-quality placements more reliably, narrowing the gap with state-of-the-art methods. The source code for this work is available at https://github.com/jhy00n/MacroDiff.**

I. Introduction

Macro placement is a foundational step in VLSI design, determining optimal locations for large circuit blocks (macros) on a chip while balancing Power, Performance, and Area (PPA). In modern chips, comprising numerous macros such as memory and analog blocks, an effective placement strategy is critical for design productivity and product viability. Traditional approaches rely on human expertise, rendering the process labor-intensive and prone to variability. This motivates the development of machine learning–based techniques to automate the placement process and achieve more consistent results.

As a solution for automating macro placement, reinforcement learning (RL) has been explored [1]. However, RL-based methods face significant limitations. By formulating macro placement as a Markov Decision Process (MDP), these approaches enforce a sequential placement strategy where early decisions cannot be revised, often leading to suboptimal outcomes. This sequential nature complicates long-term dependency management, as errors made early in the process propagate throughout. Moreover, RL-based methods suffer from severe sample inefficiency; each new netlist demands extensive retraining, even with offline techniques, limiting its generalizability and preventing convergence to globally optimal solutions.

Recently, deep generative models presented a promising alternative to address these limitations of RL-based methods by generating solutions simultaneously with better sample efficiency. In particular, a diffusion-based coordinate model [2] for chip placement has been proposed, offering improved efficiency compared to RL approaches. However, while this model demonstrates higher efficiency, it has not yet achieved performance levels comparable to state-of-the-art RL methods. This gap motivates further exploration of generative models that can maintain efficiency while improving placement quality.

To address the challenges in macro placement, we propose **MacroDiff**, a wirelength-based generation model that captures circuit connectivity by modeling the wirelengths between connected macros. Our approach leverages a Denoising Diffusion Probabilistic Model (DDPM) [3], which utilizes graph neural networks to generate macro placement solutions. Instead of directly generating macro positions, which can be challenging due to rotational and translational invariance requirements, our method first generates the wirelengths associated with each net and then recovers the macro positions from these wirelengths. This indirect generation strategy naturally handles the invariance properties of placements while enabling more stable optimization compared to direct coordinate generation.

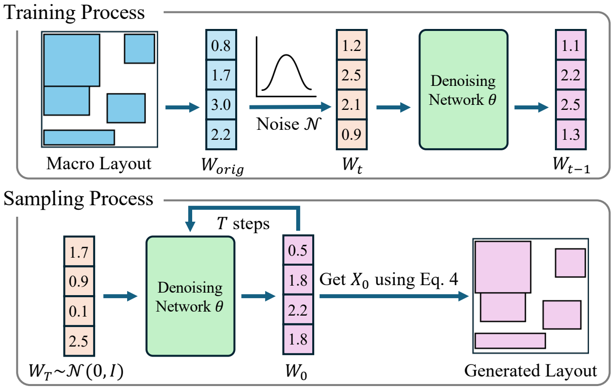


Fig. 1. Overview of the proposed **MacroDiff** method.

II. Proposed Method

A. Problem Definition

We formulate macro placement as a conditional generative modeling problem. The input netlist is represented as a heterogeneous graph $G = (V, E)$, where $V = V_{macro} \cup V_{net}$ consists of macro nodes V_{macro} and net nodes V_{net}. Each macro node $m_i \in V_{macro}$ contains its size attributes, while each net node $n_j \in V_{net}$ is characterized by its degree. For a netlist G with N macros, a placement solution is represented as a matrix $X \in \mathbb{R}^{N\times 2}$ specifying the 2D coordinates. The corresponding wirelength of each net can be represented as a vector $W \in \mathbb{R}^{|V_{net}|}$, where each element is computed based on the macro positions in X. Our objective is to learn a generative model $p_\theta(X|G)$ that can generate feasible placements.

B. MacroDiff

We introduce a novel wirelength-based generation model that leverages circuit connectivity information to address the inherent geometric constraints of placement. We employ a DDPM [3] framework composed of heterogeneous graph neural networks to capture the relation between macros and nets. Fig. 1 illustrated the overall workflow of the **MacroDiff**.

Training Process Based on the wirelength-based formulation, our model learns to generate a vector $W_0 \in \mathbb{R}^{|V_{net}|}$ representing the wirelength of each net, measured using the weighted-average metric [4]. The forward process gradually adds Gaussian noise to the wirelength vector through t timesteps according to:

$$q(W_t|W_{t-1}) = \mathcal{N}(W_t; \sqrt{1-\beta_t}W_{t-1}, \beta_t I), \quad (1)$$

where β_t represents the noise schedule at timestep t and I is the identity matrix. The reverse process, parameterized by neural networks θ, learns to iteratively denoise corrupted samples:

$$p_\theta(W_{t-1}|W_t) = \mathcal{N}(W_{t-1}; \mu_\theta(W_t, t), \sigma_t^2 I), \quad (2)$$

where, μ_θ represents the predicted mean, and σ_t is the fixed variance. The model is trained to minimize the L2 loss between the predicted and target noise at each timestep.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2023-00222609, Development of thermal modelling and heat-spreading architecture for heterogeneous integrating package of AI chiplets) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00405991).
*Both authors contributed equally to this paper.

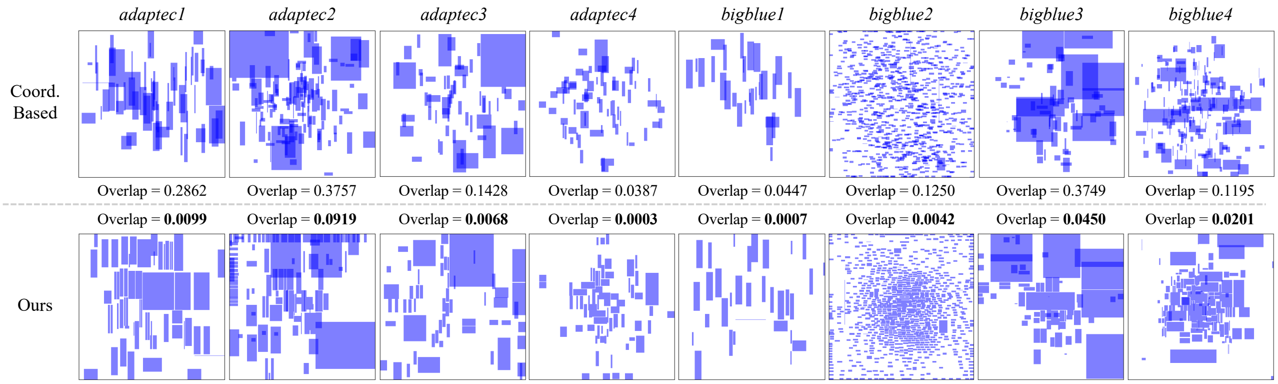
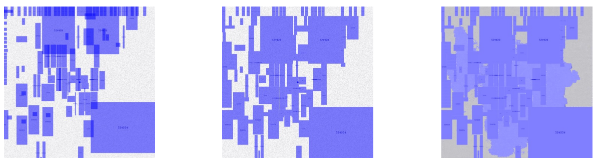


Fig. 2. Generated macro layouts from the trained diffusion model on *ISPD2005* benchmarks. Overlap values, representing the ratio of overlapping area to total area, are averaged across 10 samples for each benchmark.



(a) Generated Layout (b) After Macro Legalization (c) After Placement

Fig. 3. Overview of the placement process after macro legalization.

Sampling Process Given a trained model, we first generate the wirelength vector W_0 by iteratively applying the reverse process denoising steps. Starting from an random noise W_T sampled from $\mathcal{N}(0, I)$, we apply the learned reverse diffusion steps:

$$W_{t-1} = \mu_\theta(W_t, t) + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \tag{3}$$

After T steps of denoising, we obtain the final wirelength vector W_0. We then determine the macro positions $X_0 \in \mathbb{R}^{N\times 2}$ by solving the following objective function:

$$J(X) = \|f(X) - W_0\|^2 + \alpha \mathrm{WL}(X) + \beta D(X), \tag{4}$$

where, $f(\cdot)$ computes the wirelength vector based on the macro positions, $\mathrm{WL}(\cdot)$ represents the total wirelength between macros, $D(\cdot)$ measures the overlap between macros, and α, β are weighting factors. This position recovery process yields the final macro placement that satisfies the density constraint to ensure legality, while maintaining reasonable macro wirelength.

III. Experimental Setup and Result

A. Experimental Setup

We evaluate our method on eight designs from the *ISPD2005* [5] modern mixed-size placement (MMS) [6] benchmarks. For each design, we generate 1,000 variants using the configuration model [7] which preserves node degrees while randomizing connections, yielding 8,000 training samples. We use DREAMPlace [4] to obtain reference macro placements. For baseline comparison, we implemented the coordinate-based diffusion model [2].

B. Experimental Result

We evaluate our proposed method against a coordinate-based model using eight *ISPD2005* benchmark designs. Fig. 2 illustrates the comparison between the two approaches. Our method consistently achieves lower overlap ratios across all generated benchmarks when compared to the coordinate-based model, showing an average overlap reduction of 91.6%. By decoupling wirelength generation and position recovery while incorporating density constraints (Eq. 4), our approach provides more flexibility in meeting constraints, leading to substantially less overlap compared to directly generating positions.

To validate the practical applicability of our solutions, we integrate them into the placement flow by performing macro legalization followed by standard cell placement using DREAMPlace [4], as shown in Fig. 3. We evaluate performance on three metrics: Macro Displacement on macro legalization and both the number of Iterations and Half-Perimeter Wirelength (HPWL) on placement. As shown in Table I, our approach demonstrates substantial improvements in all metrics. The significant reduction in displacement, which averages 74.4% compared to the coordinate-based model, directly results from our minimal overlap solutions, requiring less legalization adjustment. In terms of efficiency, our approach requires 4.6% fewer iterations on average compared to the coordinate-based model. Moreover, our method consistently converges without any divergence failures across all designs, further demonstrating its efficiency and robustness. Furthermore, our approach achieves better HPWL with an average 7.0% reduction across all benchmarks, as generated wirelength vectors help maintain better connectivity relationships between macros serving as a superior starting point for further placement process.

TABLE I
Comparison of Macro Legalization Results on Displacement ($\times 10^3$) and Placement Results on Iteration, HPWL ($\times 10^6$).

Design	Displacement ↓		Iteration ↓		HPWL ↓	
	[2]	Ours	[2]	Ours	[2]	Ours
adaptec1	67.5	**11.5**	**566**	574	95.9	**90.1**
adaptec2	245.8	**133.8**	318*	**600**	936.2*	**118.9**
adaptec3	65.7	**20.3**	752	**715**	174.2	**138.9**
adaptec4	24.6	**0.4**	748	**734**	204.9	**183.3**
bigblue1	9.8	**2.2**	**646**	650	**101.2**	103.0
bigblue2	345.3	**6.8**	674	**672**	**159.1**	164.5
bigblue3	235.4	**113.6**	687*	**787**	3130.8*	**412.7**
bigblue4	338.7	**94.9**	1095	**847**	1072.0	**958.4**
Avg. Reduction	-	74.4%	-	4.6%	-	7.0%

* denotes divergence failure. Not included in *Avg. Reduction*.

IV. Conclusion

In this paper, we present **MacroDiff**, a novel approach for macro placement using a wirelength-based diffusion model. While a previous coordinate-based diffusion model improved efficiency over RL-based methods, it fell short in performance. Our approach addresses the geometric constraints of placement by leveraging wirelength information, resulting in a model that demonstrates both higher efficiency and performance compared to the coordinate-based model. For future work, we plan to incorporate RL fine-tuning [8] to optimize for actual downstream objectives like PPA metrics, enabling our method to generate placement solutions tailored to different design scenarios such as low power or high performance requirements.

References

[1] C.-K. Cheng, *et al.*, "Assessment of Reinforcement Learning for Macro Placement", *Proc. ISPD*, 2023.
[2] V. Lee, *et al.*, "Chip Placement with Diffusion", *arXiv:2407.12282*, 2024.
[3] J. Ho, *et al.*, "Denoising Diffusion Probabilistic Models", *Proc. NeurIPS*, 2020.
[4] Y. Lin, *et al.*, "DREAMPlace: Deep Learning Toolkit-Enabled GPU Acceleration for Modern VLSI Placement", *Proc. DAC*, 2019.
[5] G.-J. Nam, *et al.*, "The ISPD2005 Placement Contest and Benchmark Suite", *Proc. ISPD*, 2005.
[6] J. Z. Yan, *et al.*, "Handling Complexities in Modern Large-Scale Mixed-Size Placement", *Proc. DAC*, 2009.
[7] M. E. J. Newman, "The structure and function of complex networks", *SIAM Rev.*, 45, 167-256, 2002. *ICCAD*, 2023.
[8] K. Black, *et al.*, "Training Diffusion Models with Reinforcement Learning", *arXiv:2307.02514*, 2023.