# STELLAR-3D: A Structured Thermal- and Warpage-aware LLM-aided Floorplanning Framework for 3D ICs

Haotian Lu, Jincong Lu, Shanmukh Racha and Sheldon X.-D. Tan
Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521
*Email*: hlu123@ucr.edu, stan@ece.ucr.edu

***Abstract*— As technology nodes continue to scale, 3D ICs offer compelling advantages over conventional 2D ICs in power efficiency and performance. However, 3D integration introduces strong cross-tier coupling and complex multiphysics interactions, making reliability-aware floorplanning substantially more challenging than its 2D counterpart. In particular, reliability metrics such as hotspot temperature and thermal warpage often compete with conventional floorplanning objectives, necessitating careful trade-offs under expensive multiphysics evaluations. In this paper, we present *STELLAR-3D*, an LLM-aided closed-loop optimization framework for reliability-aware 3D IC floorplanning that considers coupled multiphysics effects, including hotspot temperature and thermal warpage. *STELLAR-3D* employs an LLM agent to intelligently generate candidate floorplan modifications and leverages efficient and accurate evaluators to provide quantitative feedback on both reliability and design objectives such as HPWL and legality, enabling iterative refinement toward improved reliability while maintaining competitive design quality. Numerical experiments on multiple benchmarks show that *STELLAR-3D* improves reliability metrics on small-to-moderate-scale circuits, reducing hotspot temperature and thermal warpage by 3.8% and 3.6%, respectively, while incurring an average HPWL overhead of 9%.**

## I. Introduction

2.5D/3D design paradigms have emerged as a promising alternative to traditional 2D IC by enabling vertical integration of multiple dies, thereby improving system-level performance. However, 3D stacking introduces tightly coupled reliability challenges that are difficult to address during physical design [1], [2]. In 3D floorplanning, designers must jointly determine tier assignment and in-tier placement of heterogeneous modules to optimize traditional objectives such as wirelength and legality while accounting for multiphysics reliability effects [3], [4]. Consequently, thermal- and warpage-aware reliability-constrained 3D floorplanning becomes a challenging multi-objective optimization problem.

Early research on 3D floorplanning mainly falls into heuristic search and learning-based paradigms. Simulated Annealing (SA) and related heuristic algorithms [5], [6] remain popular due to their simplicity and interpretability, but their effectiveness depends heavily on carefully designed local perturbations, handcrafted cost weights, and extensive trial-and-error tuning. More recently, deep reinforcement learning (RL) has been explored to learn placement policies that operate directly on grid-based representations [7]–[9], for example via actor–critic methods in hybrid action spaces [10]. While RL significantly reduces the cost of decision-making once trained, it typically incurs non-trivial training overhead and requires delicate reward shaping to balance multiple coupled objectives and constraints. These challenges motivate a decision-making framework that can incorporate multiphysics feedback and adapt to changing objectives with minimal re-engineering.

Recent advances in large language models (LLMs) have enabled a new class of agentic workflows that combine natural-language reasoning with tool use and external feedback, showing promising potential in EDA [11]–[14]. Unlike gradient-based training or handcrafted heuristics, an LLM agent can iteratively propose candidate actions and refine decisions based on quantitative evaluator feedback. This paradigm is appealing for reliability-constrained 3D floorplanning, where conflicting objectives such as HPWL, legality, hotspot temperature, and thermal warpage can be efficiently evaluated. By explicitly encoding these objectives and incorporating inference-time feedback, LLM agents can adapt to new circuits and constraints with minimal re-engineering.

We propose *STELLAR-3D*, an LLM-aided floorplan optimization framework with the LLM as an adaptive decision-making agent for reliability-aware 3D IC floorplanning. By incorporating thermal hotspot and thermal warpage into the optimization loop, the framework supports multiphysics reliability-aware design exploration while maintaining competitive wirelength. The major contributions of this work are summarized as follows:

- We propose a novel LLM-aided 3D IC floorplan optimization framework that integrates fast numerical thermal and warpage evaluators. The framework enables iterative floorplan refinement guided by feedback on design metrics including multiphysics reliability for 3D IC floorplan task.
- We develop an LLM-driven decision-making strategy based on a carefully designed system prompt and iterative feedback prompts. By incorporating information from previous trials and evaluator feedback, the LLM can adaptively propose and refine floorplan actions during the optimization.
- Experimental results show that *STELLAR-3D* improves reliability-oriented metrics on small-to-moderate-scale circuits, reducing hotspot temperature and thermal warpage by 3.8% and 3.6%, respectively, while maintaining competitive wirelength with an average HPWL overhead of 9%.

## II. Related Work

### A. Existing 3D IC Floorplanning Methods

Early studies on floorplanning mainly relied on representation-based optimization and analytical modeling. Classical representation schemes such as sequence-pair [15], corner-block-list [16] and B$^*$-tree [6] provided compact encodings of floorplan topology while analytical methods optimized block coordinates through gradient-based procedures [17], [18]. Although these approaches established the foundation of 2D/3D floorplanning, they primarily targeted wirelength and area rather than coupled reliability constraints in multi-tier 3D ICs.

Heuristic methods remain dominant in floorplanning, typically combining topological representations with simulated annealing or rule-based search. These methods have been extended to jointly optimize tier assignment and in-tier placement in 3D ICs [5], while more recent work has explored black-box search methods such as wire-mask-guided macro placement [19]. However, their performance still depends

This work is supported in part by NSF grant under No. CCF-2007135 and in part by NSF grant No. CCF-2305437.

979-8-3195-1246-8-0/26/$31.00 ©2026 IEEE

heavily on handcrafted search operators and objectives, limiting adaptation to coupled multiphysics reliability constraints.

Learning-based methods, particularly reinforcement learning (RL), learn placement policies from graph or visual representations and reduce dependence on handcrafted search rules [7]–[9]. FlexPlanner extends this paradigm to 3D floorplanning using a hybrid action space and multi-modal representation [10], while end-to-end learning has also been applied to thermal-aware fixed-outline 3D floorplanning [20]. However, these methods typically require substantial training data and may generalize poorly to new circuit configurations without fine-tuning.

### B. LLM-aided Design Automation

Recent studies have explored LLMs in EDA for autonomous tool interaction, design-space exploration (DSE), and feedback-driven optimization. ChatEDA [21] uses LLMs as tool-augmented EDA agents, while ChatDSE [12] applies inference-time reasoning to microarchitecture DSE without task-specific retraining. LEDRO [13] and LATENT [14] further demonstrate iterative closed-loop optimization. LEDRO refines the search region for analog sizing using optimizer and simulator feedback, whereas LATENT follows the ReAct framework [22] to modify analog netlists using syntax, simulation, and diagnosis feedback. These works demonstrate the potential of LLM-guided EDA optimization, but mainly target general automation, analog sizing, or security-oriented netlist modification. Reliability-aware 3D IC floorplanning with coupled thermal and warpage feedback remains largely unexplored, which motivates this work.

## III. Proposed STELLAR-3D Framework for Reliability-aware 3D IC Floorplanning

### A. Problem Formulation

Let $\mathbf{P}$ denote a $\mathcal{V}$-module candidate floorplan, where each module $M_i$, $i \in \mathcal{V}$ with size $(w_i, h_i)$ and power density $p_i$ is characterized by its central location $(x_i, y_i)$ and tier assignment $z_i$, while each terminal in $\mathbf{P}$ has a fixed location $(x_t, y_t)$. The optimization jointly considers electrical quality, physical legality, and reliability-related metrics. The resulting formulation is given in Eq. (1),

$$\begin{aligned} \min_{\mathbf{P}} \quad & \Big[ f_1(\mathbf{P}) = \mathrm{HPWL}(\mathbf{P}),\ f_2(\mathbf{P}) = \mathrm{Overlap}(\mathbf{P}) + \mathrm{OOB}(\mathbf{P}), \\ & \quad f_3(\mathbf{P}) = \max \mathcal{T}(\mathbf{P}; \{p_i\}),\ f_4(\mathbf{P}) = \mathcal{W}(\mathbf{P}; \mathcal{T}(\mathbf{P}; \{p_i\})) \Big] \\ \text{s.t.} \quad & w_i/2 \le x_i \le W - w_i/2,\ \ h_i/2 \le y_i \le H - h_i/2, \\ & z_i \in \{1, \dots, T\}, \quad \forall i \in \mathcal{V}. \end{aligned} \tag{1}$$

where $f_1(\mathbf{P})$ denotes the total half-perimeter wirelength (HPWL) and is defined in Eq. (2), where $x_u = x_i$, $y_u = y_i$ for modules and $x_u = x_t$, $y_u = y_t$ for terminals,

$$f_1(\mathbf{P}) = \sum_{n \in nets} \left( \max_{u \in n} x_u - \min_{u \in n} x_u + \max_{u \in n} y_u - \min_{u \in n} y_u \right) \tag{2}$$

$f_2(\mathbf{P})$ represents the overlap and out-of-boundary penalties, which characterize placement legality. $\mathcal{T}(\mathbf{P}; \{p_i\})$ and $\mathcal{W}(\mathbf{P}; \mathcal{T}(\mathbf{P}; \{p_i\}))$ represent the thermal map and warpage, respectively. We further consider a budget-constrained optimization setting, where at most $N$ candidate floorplans can be evaluated throughout the process.

### B. Overall Framework of the Proposed STELLAR-3D

Fig. 1 illustrates the overall workflow of *STELLAR-3D*. Given a structured 3D floorplan netlist containing module attributes, net connectivity, die outline, and tier configuration, the framework performs iterative reliability-aware closed-loop optimization through an LLM agent and external evaluators.

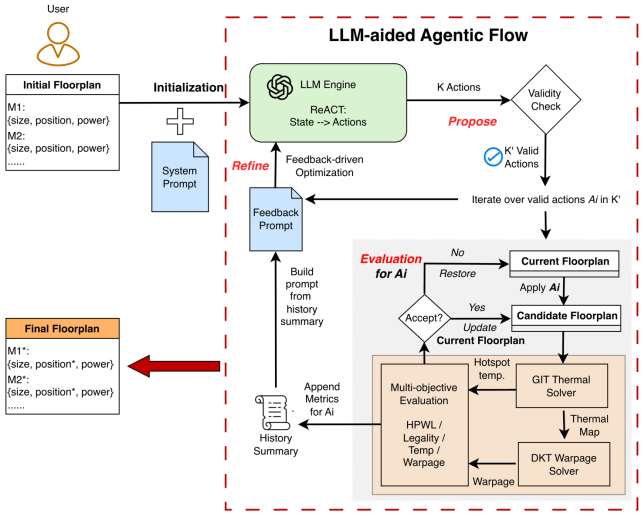


Fig. 1: Overall framework of *STELLAR-3D* for LLM-aided reliability-aware 3D IC floorplanning.

At each iteration, the current floorplan state and evaluator feedback are summarized into a structured prompt and passed to the LLM, which generates $K$ candidate actions for progressively improving the current solution. Generating multiple actions per query amortizes the LLM inference overhead and avoids invoking the model for every individual modification. The candidate actions are then sequentially checked, applied, and evaluated in terms of HPWL, legality, hotspot temperature, and thermal warpage. Minor geometric violations are legalized when possible, whereas malformed or semantically invalid actions are discarded before evaluation. Valid candidates are accepted or rejected using an SA-inspired policy, which balances the exploitation of promising solutions with the exploration of alternative floorplan configurations. Under a fixed total action-evaluation budget of $N$, the process runs for $\lfloor N/K \rfloor$ iterations.

The evaluation results are converted into structured feedback and returned to the LLM in the next iteration, forming a propose–evaluate–refine loop [22]–[24]. This closed-loop interaction enables the LLM to refine subsequent decisions based on the observed effects of previous actions, rather than generating modifications in an open-loop manner. Through this iterative process, *STELLAR-3D* progressively balances wirelength reduction, legality recovery, hotspot mitigation, and thermal warpage optimization. The process terminates when the evaluation budget is exhausted, and the best feasible floorplan found so far is returned as the final output.

### C. Prompt Implementation

Fig. 2 and Fig. 3 illustrate our two-stage prompt design, consisting of a fixed *system prompt* and an iteration-dependent *feedback prompt*. The former defines the task and behavioral constraints, while the latter provides evaluator-guided context at each iteration.

*a)* ***System prompt****:* The system prompt defines the LLM as a reliability-aware 3D floorplanning agent and specifies the optimization objectives, design representation, action space, and structured output format. The action space includes 1) `INTRA_MOVE`, which moves a module within its current tier; 2) `RELOCATE`, which assigns a module to a random position; 3) `INTER_MOVE`, which moves a module to another tier; and 4) `SWAP`, which exchanges the positions of two modules. Lightweight domain heuristics further guide the model toward physically feasible decisions, such as prioritizing legality recovery under severe overlap and focusing on highly connected or power-dense modules when wirelength or thermal violations dominate.

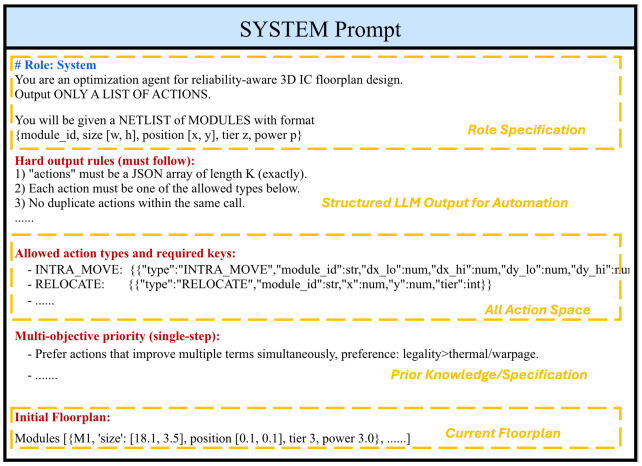


Fig. 2: Structured system prompt for floorplan information and problem instructions.

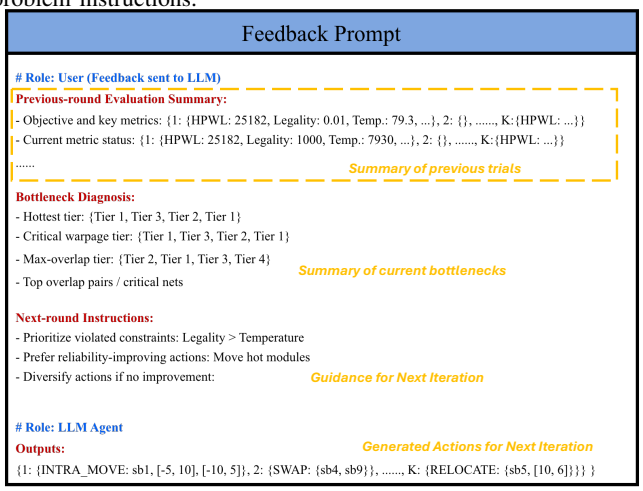


Fig. 3: Structured iterative feedback prompt for action proposal.

*b) **Iterative feedback prompt**:* The feedback prompt is updated at each iteration using the current floorplan and evaluator results, including HPWL, legality penalty, thermal and warpage costs, and diagnostics such as critical modules, congested regions, and dominant nets. When available, trends from previous iterations are included to reveal tradeoffs among objectives. Based on this feedback, the LLM generates the next $K$ candidate actions.

This two-level design separates persistent task semantics from dynamic search feedback, reducing prompt redundancy and improving action validity.

### D. Acceptance Policy and Search Strategy

To address the multi-objective problem, we employ a local search strategy to refine LLM-proposed modifications. To balance exploration and exploitation, we adopt the SA-inspired acceptance rule in Algo. 1, which permits occasional uphill moves early in the search and becomes progressively stricter as the temperature decreases.

At each iteration t, the LLM proposes a $K$-size action sequence $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,K})$. Unlike one-shot generation, we evaluate and gate each action sequentially. Each action $a_{t,j}$ is applied only if it passes the sanity check and SA acceptance criterion. Otherwise, the current floorplan is retained. Thus, the LLM serves as a proposal generator, while the evaluators and acceptance policy control the search trajectory and suppress invalid or overly aggressive modifications.

Let $\ell = \mathcal{L}(\mathbf{P})$ denote the scalar loss of the current floorplan $\mathbf{P}$, and $\ell' = \mathcal{L}(\mathbf{P}')$ be the loss after applying a candidate action. We compute $\Delta = \ell' - \ell$, and accept the candidate action with probability as shown in Eq. (3),

$$p_{\text{acc}}(\Delta;\tau) = \begin{cases} 1, & \Delta \le 0, \\ \exp(-\Delta/\tau), & \Delta > 0, \end{cases} \tag{3}$$

where $\tau$ is the temperature schedule.

**Algorithm 1** LLM-guided iterative floorplan optimization with SA-inspired acceptance policy

**Input:** initial floorplan $\mathbf{P}_0$, objective function $\mathcal{L}(\mathbf{P})$, floorplan update $\mathcal{U}(\mathbf{P}, a)$, validity checker $\textsc{Valid}(\mathbf{P}, a)$, LLM engine for action generation $\texttt{LLM}(\cdot)$, evaluation budget $N$, number of LLM-proposed actions per iteration $K$, temperature schedule $\tau(\cdot)$
**Output:** Best floorplan $\mathbf{P}^\star$ and loss $\ell^\star$

$\mathbf{P} \leftarrow \mathbf{P}_0$, $\ell \leftarrow \mathcal{L}(\mathbf{P})$, $x_0 \leftarrow \emptyset$
$\mathbf{P}^\star \leftarrow \mathbf{P}$, $\ell^\star \leftarrow \ell$
**for** $t = 1$ to $\lfloor N/K \rfloor$ **do**
  $x_t \leftarrow \textsc{Summarize}(\mathbf{P}, \ell, \ell^\star, x_{t-1})$
  $\mathbf{a}_t = (a_{t,1}, \ldots, a_{t,K}) \leftarrow \texttt{LLM}(x_t)$
  **for** $j = 1$ to $K$ **do**
    **if** $\textsc{Valid}(\mathbf{P}, a_{t,j})$ **then**
      $\mathbf{P}' \leftarrow \mathcal{U}(\mathbf{P}, a_{t,j})$, $\ell' \leftarrow \mathcal{L}(\mathbf{P}')$, update summary $x_t$
      Accept $\mathbf{P}'$ with probability $\min\left(1, \exp\left(-\frac{\ell' - \ell}{\tau(t)}\right)\right)$
      If accepted, update $(\mathbf{P}, \ell) \leftarrow (\mathbf{P}', \ell')$
      If $\ell < \ell^\star$, update $(\mathbf{P}^\star, \ell^\star) \leftarrow (\mathbf{P}, \ell)$
    **end if**
  **end for**
**end for**
**return** $\mathbf{P}^\star, \ell^\star$

### E. Temperature and Thermal Warpage Evaluation

Machine Learning (ML) has been applied to various EDA learning tasks [27]–[29]. In particular, ML-based thermal and warpage predictors [30]–[33] can serve as fast surrogates for evaluating the thermal and mechanical feasibility of candidate floorplans. However, their effectiveness depends on sufficient task-specific training data and reliable generalization to unseen designs. Therefore, in this work, we instead adopt efficient numerical estimators to assess temperature and thermal warpage behavior within the framework. For each LLM-generated candidate, the estimator evaluates reliability-related temperature and warpage metrics, such as hotspot temperature and overall warpage. In particular, to preserve both efficiency and fidelity in warpage evaluation, we adopt a CLT [34]-based modeling framework and further revise the efficient formulation in DKT warpage estimator [35]. For fast 3D IC thermal analysis, we build the thermal model based on GIT thermal solver [36]. Both the solvers take floorplan information as input, and output the corresponding hotspot temperature and warpage metrics.

## IV. Experimental Results

### A. Experimental Setup

***LLM Agentic Flow***: API-based *GPT-4.1* is employed as the backbone model of *STELLAR-3D* [37], [38]. The temperature is set to 0.7, with $K = 10$ candidate actions per iteration and a total evaluation budget of $N = 1000$.

***Metrics and Benchmarks***: We use the following metrics to evaluate the overall performance of *STELLAR-3D* and existing works: 1) *HPWL*, 2) *Legality*, 3) *Temperature* and 4) *Warpage*. We utilize 1) MCNC [25] and GSRC [26] benchmarks from realistic designs to demonstrate the real-world applicability of our proposed framework, and 2) Synthesized small-to-moderate-scale 3D floorplans adopted from [25]. Power densities from 0 to 0.08 $\mu$W/$\mu$m$^2$ are randomly assigned to modules, creating thermally stressed but plausible designs. Number of tiers $T$ is fixed at 4 and the floorplan outline is determined so that the general area utilization is 60%.

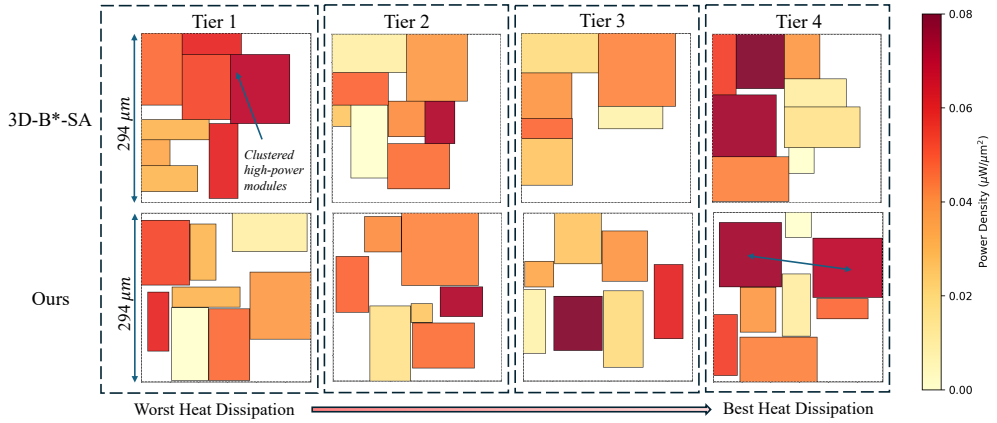


Fig. 4: Comparison of four-tier floorplans generated by our proposed framework and 3D-B*-SA [5] on the n30 benchmark.

TABLE I: Comparison between our proposed STELLAR-3D and 3D-B*-SA method [5].

| Circuits | | HPWL ($\mu m$) | | Legality | | Temperature ($^\circ C$) | | Warpage ($\mu$m) | |
|---|---|---|---|---|---|---|---|---|---|
| | | 3D-B*-SA | Ours | 3D-B*-SA | Ours | 3D-B*-SA | Ours | 3D-B*-SA | Ours |
| Realistic [25], [26] | ami33 | 42238.20±4087.78 | 44203.64±1179.96 | 0.000±0.000 | 0.002±0.001 | 76.96±3.13 | 73.40±5.02 | 35.21±1.67 | 34.01±1.09 |
| | ami49 | 769594.85±40198.87 | 933967.59±46108.67 | 0.006±0.012 | 0.007±0.006 | 79.01±2.54 | 77.31±3.46 | 543.97±25.67 | 533.28±12.66 |
| | n10 | 22701.46±1591.90 | 22570.77±2516.25 | 0.000±0.000 | 0.000±0.000 | 76.19±4.06 | 71.64±2.29 | 3.51±0.16 | 3.43±0.04 |
| | n30 | 67047.46±987.13 | 73675.88±1459.45 | 0.000±0.000 | 0.000±0.000 | 74.59±4.13 | 70.36±1.45 | 3.19±0.12 | 3.19±0.10 |
| | n50 | 86702.11±2280.50 | 104306.35±373.11 | 0.000±0.000 | 0.001±0.000 | 77.19±2.65 | 74.32±3.46 | 3.05±0.07 | 2.94±0.08 |
| Synthesized | syn15 | 41060.65±1010.55 | 43777.34±525.67 | 0.000±0.000 | 0.000±0.000 | 73.62±4.24 | 71.59±2.32 | 3.72±0.10 | 3.64±0.14 |
| | syn20 | 51849.59±1338.35 | 55147.36±966.49 | 0.003±0.005 | 0.000±0.000 | 73.81±3.14 | 71.05±3.31 | 3.14±0.17 | 2.85±0.07 |
| | syn25 | 66873.48±937.27 | 74698.21±1427.21 | 0.000±0.000 | 0.000±0.000 | 73.97±2.00 | 72.74±2.22 | 2.88±0.09 | 2.71±0.02 |
| **Average (*Norm.*)** | | **1**× | **1.09**× | **0.001** | **0.001** | **1**× | **0.962**× | **1**× | **0.964**× |

***Baselines***: For numerical comparison, we adopt 3D-B*-SA [5] as the baseline. For fairness, it uses the same allowable actions and the same thermal and warpage evaluators as *STELLAR-3D*, and is further augmented with hotspot temperature and thermal warpage terms in the weighted objective. The penalty factors are set to $\{\lambda_{\text{HPWL}}, \lambda_{\text{Leg}}, \lambda_{\text{T}}, \lambda_{\text{W}}\} = \{1, 10^3, 2, 2\}$ and are kept the same in the acceptance policy of *STELLAR-3D*. The temperature- and warpage-related parameters in the evaluators are adopted from [35], [36], [39]. The thermal evaluation further considers top-side cooling for heat dissipation in the 3D stack, with tier 4 corresponding to the best cooling condition and tier 1 to the worst [40], [41].

### *B. Ablation Study: Configurations of the Numerical Solvers*

We evaluate the impact of GIT thermal-solver and DKT warpage-estimator resolutions on reliability assessment and runtime. As shown in Tab. II, higher resolutions improve the estimated hotspot temperature and warpage but substantially increase runtime, particularly for the DKT warpage estimator. Since the results change only marginally beyond moderate resolutions, we use a resolution of 16 as a practical accuracy–efficiency trade-off in the following experiments.

TABLE II: Effects of FEM Resolution on Reliability Assessment.

| Solver Resolution | Circuit | Average Runtime (s) | | Metrics in Initial Floorplan | |
|---|---|---|---|---|---|
| | | GIT [36] | DKT [35] | Temperature ($^\circ$C) | Warpage ($\mu m$) |
| 4 | n30 | 0.015 | 0.004 | 67.58 | 7.33 |
| | ami49 | 0.015 | 0.005 | 66.41 | 1093.42 |
| 8 | n30 | 0.030 | 0.014 | 87.96 | 6.76 |
| | ami49 | 0.034 | 0.016 | 70.94 | 896.62 |
| 16 | n30 | 0.089 | 0.078 | 95.02 | 6.30 |
| | ami49 | 0.091 | 0.079 | 72.81 | 889.36 |
| 32 | n30 | 0.333 | 1.081 | 97.36 | 6.13 |
| | ami49 | 0.344 | 1.128 | 73.44 | 836.45 |
| 64 | n30 | 1.338 | 65.240 | 97.36 | 6.16 |
| | ami49 | 1.328 | 65.845 | 73.44 | 828.19 |

### *C. Overall Performance of STELLAR-3D*

Tab. I compares *STELLAR-3D* with the SA-based 3D-B*-SA baseline [5] on HPWL, legality, hotspot temperature, and thermal warpage over 10 random seeds per circuit to eliminate the impact of randomness. Overall, *STELLAR-3D* consistently improves the reliability-related metrics, particularly temperature and warpage, across the evaluated benchmarks. On average, *STELLAR-3D* reduces hotspot temperature and thermal warpage by 3.8% and 3.6%, respectively. Fig. 4 shows an example of the final 4-tier floorplans for the n30 benchmark generated by *STELLAR-3D* and 3D-B*-SA. The comparison shows that *STELLAR-3D* tends to place more high-power modules on upper tiers with better heat dissipation and to separate hot modules while maintaining a relatively compact floorplan, thereby producing a more reliability-aware solution.

These reliability improvements come at the cost of a 9% average increase in HPWL. The degradation is more pronounced on larger benchmarks such as ami49 and n50. This suggests that the current framework tends to favor a reliability-oriented solution, but still faces difficulty in maintaining an ideal HPWL–reliability trade-off as circuit size increases, which remains an important direction for future work.

## V. Conclusion

In this paper, we introduced *STELLAR-3D*, an LLM-aided framework for reliability-aware 3D IC floorplanning. In the proposed framework, the LLM serves as an adaptive decision-making agent that iteratively proposes floorplan actions, while fast numerical thermal and warpage evaluators provide quantitative feedback for floorplan refinement. Through this closed-loop optimization process, *STELLAR-3D* enables reliability-aware design exploration under multiphysics considerations. Numerical experiments on multiple 3D IC benchmarks show that *STELLAR-3D* reduces hotspot temperature and thermal warpage by 3.8% and 3.6%, respectively, on small-to-moderate-scale circuits, while maintaining competitive wirelength performance with an average HPWL overhead of 9%. These results suggest that LLM-aided closed-loop optimization is a promising direction for reliability-aware 3D IC physical design.

## REFERENCES


[1] S. Naffziger, K. Lepak, M. Paraschou, and M. Subramony, "2.2 amd chiplet architecture for high-performance server and desktop products," pp. 44–45, 2020.

[2] F. Herrault, J. C. Wong, Y. Tang, H. Y. Tai, and I. Ramos, "Heterogeneously integrated rf circuits using highly scaled off-the-shelf gan hemt chiplets," *IEEE Microwave and Wireless Components Letters*, vol. 30, no. 11, pp. 1061–1064, 2020.

[3] J. Cong and Y. Zhang, "Thermal-aware physical design flow for 3-d ics," in *Proc. 23rd International VLSI Multilevel Interconnection Conference*, pp. 73–80, 2006.

[4] P. Emma, A. Buyuktosunoglu, M. Healy, K. Kailas, V. Puente, R. Yu, A. Hartstein, P. Bose, and J. Moreno, "3d stacking of high-performance processors," in *2014 IEEE 20th International Symposium on High Performance Computer Architecture (HPCA)*, pp. 500–511, 2014.

[5] P. Falkenstern, Y. Xie, Y.-W. Chang, and Y. Wang, "Three-dimensional integrated circuits (3d ic) floorplan and power/ground network co-synthesis," in *Proc. Asia South Pacific Design Automation Conf. (ASP-DAC)*, pp. 169–174, 2010.

[6] Y.-C. Chang, Y.-W. Chang, G.-M. Wu, and S.-W. Wu, "B*-trees: a new representation for non-slicing floorplans," in *Proc. Design Automation Conf. (DAC)*, pp. 458–463, 2000.

[7] A. Mirhoseini, A. Goldie, M. Yazgan, J. W. Jiang, E. Songhori, S. Wang, Y.-J. Lee, E. Johnson, O. Pathak, A. Nova, *et al.*, "A graph placement methodology for fast chip design," *Nature*, vol. 594, no. 7862, pp. 207–212, 2021.

[8] R. Cheng and J. Yan, "On joint learning for solving placement and routing in chip design," *Advances in Neural Information Processing Systems*, vol. 34, pp. 16508–16519, 2021.

[9] Y. Lai, Y. Mu, and P. Luo, "Maskplace: fast chip placement via reinforced visual representation learning," in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2022.

[10] R. Zhong, X. Du, S. Kai, Z. Tang, S. Xu, J. Hao, M. Yuan, and J. Yan, "Flexplanner: Flexible 3d floorplanning via deep reinforcement learning in hybrid action space with multi-modality representation," in *Advances in Neural Information Processing Systems*, vol. 37, pp. 49252–49278, 2024.

[11] Y. Yin, Y. Wang, B. Xu, and P. Li, "Ado-llm: Analog design bayesian optimization with in-context learning of large language models," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 1–9, 2024.

[12] M. Tang, W. Chen, L. Wu, L. Huang, and K. Zeng, "Chatdse: A zero-shot microarchitecture design space explorer powered by gpt4.0," *ACM Trans. on Design Automation of Electronics Systems*, vol. 30, no. 4, 2025.

[13] D. V. Kochar, H. Wang, A. P. Chandrakasan, and X. Zhang, "Ledro: Llm-enhanced design space reduction and optimization for analog circuits," in *IEEE International Conference on LLM-Aided Design (ICLAD)*, pp. 141–148, 2025.

[14] J. Chaudhuri, A. Chaudhuri, and K. Chakrabarty, "Latent: Llm-augmented trojan insertion and evaluation framework for analog netlist topologies," in *2025 IEEE International Conference on LLM-Aided Design (ICLAD)*, pp. 133–140, 2025.

[15] H. Murata, K. Fujiyoshi, S. Nakatake, and Y. Kajitani, "Vlsi module placement based on rectangle-packing by the sequence-pair," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 12, pp. 1518–1524, 1996.

[16] X. Hong, G. Huang, Y. Cai, J. Gu, S. Dong, C.-K. Cheng, and J. Gu, "Corner block list: an effective and efficient topological representation of non-slicing floorplan," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 8–12, 2000.

[17] X. Li, K. Peng, F. Huang, and W. Zhu, "Pef: Poisson's equation-based large-scale fixed-outline floorplanning," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 42, no. 6, pp. 2002–2015, 2023.

[18] F. Huang, D. Liu, X. Li, B. Yu, and W. Zhu, "Handling orientation and aspect ratio of modules in electrostatics-based large scale fixed-outline floorplanning," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 1–9, 2023.

[19] Y. Shi, K. Xue, L. Song, and C. Qian, "Macro placement by wire-mask-guided black-box optimization," 2023.

[20] W. Guan, X. Tang, H. Lu, Y. Zhang, and Y. Zhang, "Thermal-aware fixed-outline 3-d ic floorplanning: An end-to-end learning-based approach," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 31, no. 12, pp. 1882–1895, 2023.

[21] H. Wu, Z. He, X. Zhang, X. Yao, S. Zheng, H. Zheng, and B. Yu, "Chateda: A large language model powered autonomous agent for eda," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 43, no. 10, pp. 3184–3197, 2024.

[22] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao, "React: Synergizing reasoning and acting in language models," 2023.

[23] A. Madaan, N. Tandon, P. Gupta, S. Hallinan, L. Gao, S. Wiegreffe, U. Alon, N. Dziri, S. Prabhumoye, Y. Yang, S. Gupta, B. P. Majumder, K. Hermann, S. Welleck, A. Yazdanbakhsh, and P. Clark, "Self-refine: Iterative refinement with self-feedback," 2023.

[24] X. Wei, Y. Dong, X. Wang, X. Zhang, Z. Zhao, D. Shen, L. Xia, and D. Yin, "Beyond react: A planner-centric framework for complex tool-augmented llm reasoning," 2025.

[25] "MCNC Benchmark." http://vlsicad.eecs.umich.edu/BK/MCNCbench. Accessed: May 26, 2007.

[26] "GSRC Benchmark." http://vlsicad.eecs.umich.edu/BK/GSRCbench. Accessed: May 26, 2007.

[27] J. Lu, J. Zhang, W. Jin, S. Sachdeva, and S. X.-D. Tan, "Learning based spatial power characterization and full-chip power estimation for commercial tpus," in *Proceedings of the 28th Asia and South Pacific Design Automation Conference*, ASPDAC '23, (New York, NY, USA), pp. 98–103, Association for Computing Machinery, 2023.

[28] J. Lu, J. Zhang, and S. X.-D. Tan, "Real-time thermal map estimation for amd multi-core cpus using transformer," in *IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, pp. 1–7, 2023.

[29] S. Lamichhane, H. Lu, and S. X.-D. Tan, "Bpinn-em-post: Bayesian physics-informed neural network based stochastic electromigration damage analysis in the post-void phase," in *Proc. Int. Symposium. on Quality Electronic Design (ISQED)*, pp. 1–8, 2026.

[30] L. Chen, J. Lu, W. Jin, and S. X.-D. Tan, "Fast full-chip parametric thermal analysis based on enhanced physics enforced neural networks," in *IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, pp. 1–8, 2023.

[31] J. Lu, S. Sachdeva, H. Lu, and S. X.-D. Tan, "Power map characterization and modeling for commercial cpu/gpus considering temperature dependence," in *Proc. Int. Symp. on Low Power Electronics and Design (ISLPED)*, pp. 1–7, 2025.

[32] S.-Y. Lo, C.-M. Chang, and Y.-W. Chang, "Advanced packaging warpage modeling with deeponet-based operator learning," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 1–9, 2025.

[33] H. Lu, J. Lu, S. Sachdeva, and S. X. D. Tan, "Warpgnn: A parametric thermal warpage analysis framework with physics-aware graph neural network," in *Proc. Int. Symp. on Low Power Electronics and Design (ISLPED)*, pp. 1–7, 2026.

[34] J. N. Reddy, *Mechanics of laminated composite plates and shells: theory and analysis*. CRC press, 2003.

[35] S.-Y. Lo, M. Liu, and Y.-W. Chang, "Efficient high-fidelity two-dimensional warpage modeling for advanced packaging analysis," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, pp. 1–9, 2024.

[36] P.-Y. Huang and Y.-M. Lee, "Full-chip thermal analysis for the early design stage via generalized integral transforms," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 5, pp. 613–626, 2009.

[37] OpenAI, "Api pricing," *https://openai.com/api/pricing/*.

[38] OpenAI, "All models," *https://developers.openai.com/api/docs/models/all*.

[39] M.-Y. Tsai and Y.-W. Wang, "A theoretical solution for thermal warpage of flip-chip packages," *IEEE Transactions on Components, Packaging and Manufacturing Technology (TCPMT)*, vol. 10, no. 1, pp. 72–78, 2020.

[40] S. K. Samal, S. Panth, K. Samadi, M. Saedi, Y. Du, and S. K. Lim, "Fast and accurate thermal modeling and optimization for monolithic 3d ics," in *Proc. Design Automation Conf. (DAC)*, pp. 1–6, 2014.

[41] K. Cao, J. Zhou, T. Wei, M. Chen, S. Hu, and K. Li, "A survey of optimization techniques for thermal-aware 3d processors," *Journal of Systems Architecture*, vol. 97, pp. 397–415, 2019.