A Graph-learning-driven Prediction Method for Combined Electromigration and Thermomigration Stress on Multi-Segment Interconnects

Yunfan Zuo[1], Yuyang Ye[1], Hongchao Zhang[1], Tinghuan Chen[2,3], Hao Yan[1], Longxing Shi[1]
[1]Southeast University [2]CUHK [3]CUHK-Shenzhen

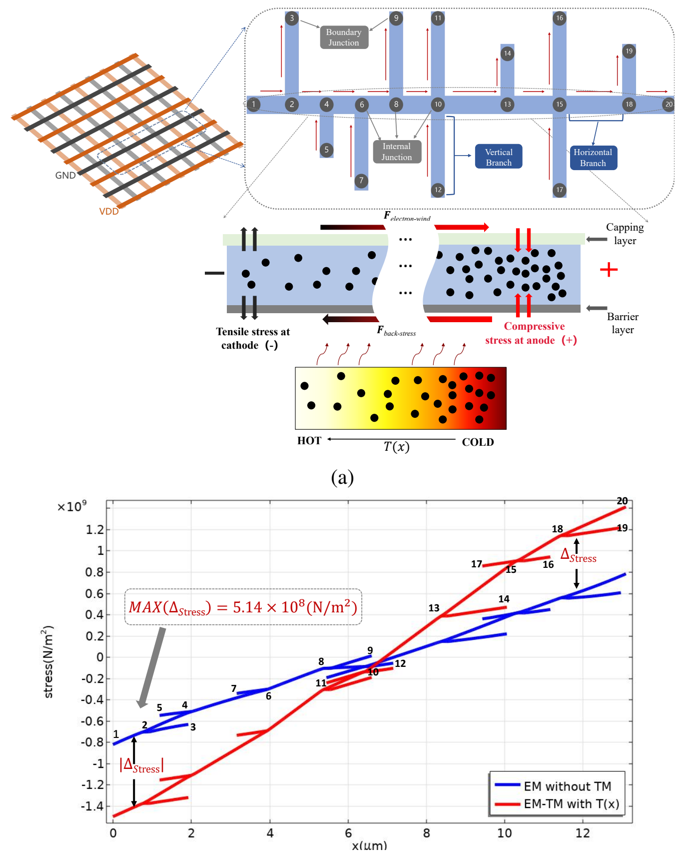
***Abstract*—As technology advances, the temperature gradient in the interconnects becomes more significant, which causes serious thermomigration. Simulating the coupling effects of thermomigration (TM) and electromigration (EM) on large-scale circuits is very time-consuming caused by a substantial increase in computational complexity. Recently, some researchers utilized graph learning-based methods to predict EM stress in medium-scale cases. Unfortunately, these works overlooked the effects of TM. To predict the EM-TM stress of large-scale interconnects accurately and efficiently, we propose a framework based on Graph Attention Networks (GATs) with a customized alternating aggregation method for collecting information in junctions and branches of interconnects jointly. The experimental results show that our work achieves an average relative error of less than 1% compared to the commercial software COMSOL for interconnects consisting of fewer than 200 segments. Furthermore, our method also achieves $9037\times$ speedup in predicting the OpenROAD test circuit with a maximum segment number reaching 10807.**

***Index Terms*—Electromigration, thermomigration, graph attention networks, multi-segment interconnect.**

I. Introduction

The rapid evolution of the integrated circuit industry has propelled the adoption of advanced technologies, resulting in the realization of high-density packaging solutions. It makes the electromigration (EM) and thermomigration(TM) effects of metal wires more serious. As predicted by the International Technology Roadmap for Semiconductors, EM-TM-induced reliability problems become a significant design constraint in future technology nodes [1]. Therefore, the development of more accurate and efficient EM-TM sign-off and validation processes is important.

In traditional design flow, Black and Blech's model is usually used to quickly predict EM stress and can only deal with a single wire segment [2]. However, this conservative worst-case design model results in severe over-design and 2X-3X enlarged guardbands, which significantly increases buffer size, powers, and chip costs [3]. To address the traditional limitations of EM models, the Korhonen equation, a partial differential equation approach, has gained widespread use in simulating the evolution of hydrostatic stress in multi-segment interconnects. Various numerical and analytical approaches have been proposed to efficiently and accurately solve partial differential equations in complex multi-segment interconnects [4]–[10]. Some numerical methods, such as integral transformation [4], accelerated separation of variables [5], and Krylov subspace method [6] exhibit strong performance in simulating EM stress of complex interconnect structures. Recently, some



(a)

(b)

Fig. 1 (a) shows the principle of stress on power grid multi-segment interconnects; (b) shows the comparison of the effects on stress considering only the EM effect and considering the EM-TM coupling effect.

analytical [7] and semi-analytical methods [8], [9] can help alleviate the computational costs of large-scale EM simulations. A semi-analytic approach [10] introduced thermomigration (TM) into EM simulations. However, this method has not yet demonstrated satisfactory simulation performance on large-scale interconnect structures. Since semi-analytical and numerical methods have significantly extended solution times when applied to large-scale interconnect structures, it is imperative to explore accurate and efficient EM-TM prediction models.

In recent years, with the rapid development of machine learning, deep learning is increasingly being applied to solving problems of Electronic Design Automation [11]. Some deep

979-8-3503-4859-0/DATE24/© 2024 EDAA

learning-based methods have also begun to be applied to accelerate EM prediction, focusing primarily on discovering efficient alternatives for solving partial differential equations [12]–[15]. Almost none considered the influence of the temperature gradient-induced TM problem. Traditional methods require significantly more time to solve EM-TM coupling problems compared to addressing EM problems alone. However, current machine-learning methods cannot model temperature gradients accurately in interconnects to help consider TM problem. Therefore, it is necessary to explore a learning-based method to predict EM-TM coupling stress in a fast and accurate way.

In this paper, we propose an accurate and efficient method to predict the evolution of transient hydrostatic stress. Our method considers both EM and TM for multi-segment interconnects using graph attention networks (GAT). GAT naturally captures the relationships between junctions and branches in multi-segment interconnects, which can be represented as graphs with edges. The key contributions of our proposed method are as follows:

- To our knowledge, we are the first to use a graph learning-based approach to predict stress caused by the combination of electromigration and thermomigration in large-scale interconnects.
- We implement a prediction model using an enhanced GAT. This model utilizes alternating aggregation of junction and branch information, leading to a significant improvement in the speed of EM-TM stress prediction for the interconnection segment as a whole.
- Our method shows impressive efficiency in predicting power network stresses from OpenROAD test circuits, with the maximum test size reaching 10,807 segments. It achieves a speed advantage of three orders of magnitude over commercial software, while still maintaining a high level of prediction accuracy.

II. Preliminaries

A. Review of Transient EM and TM Analysis

EM is a diffusion phenomenon of metal atoms that migrate from the cathode to the anode of metal interconnect wires due to momentum exchange between conductor electrons and meta atoms [2]. This momentum exchange is mainly determined by the two opposing forces [1], as is shown in Fig. 1(a).

However, new technology leads to higher power densities, increased Joule heating, and greater thermal resistance. Consequently, chip temperatures and temperature gradients are also on the rise [16]. The introduction of a temperature gradient alters the state of atomic motion. It brings a notably distinct stress distribution within the multi-segment interconnects when compared to only considering electromagnetic effects, as shown in Fig. 1(b). Commercial EM-TM simulation software COMSOL predicts the hydrostatic stress in each segment by solving Korhonen's PDE [17]. This classic diffusion PDE expresses the evolution of stress $\sigma(x,t)$, where x is the location along the wire and t is the time. The EM-TM equation is shown in Equation (1):

$$\frac{\partial \sigma}{\partial t} = \frac{\partial}{\partial x}\left[k(x)(\frac{\partial \sigma}{\partial x} - \frac{eZ\rho j}{\Omega} - \frac{Q}{\Omega T}\frac{\partial T}{\partial x})\right] \tag{1}$$

where $k(x) = \frac{D_a(T(x))B\Omega}{K_B T(x)}$ is the diffusivity of stress with $D_a = D_0 e^{\frac{-E_a}{K_B T(x)}}$ being the effective diffusion coefficient of metal atoms. It is evident from the equation that compared to the PDE equation solely addressing EM problems, the equations for EM and TM coupling problems necessitate the inclusion of an additional term $\frac{Q}{\Omega T}\frac{\partial T}{\partial x}$. When employing traditional methods, the temperature term significantly escalates the overall solution's complexity, especially when simulating stress in large-scale multi-segment interconnects. This process is indeed much more time-consuming than a simulation that only considers the effects of electromigration (EM).

B. EM and TM Prediction Using Machine Learning

While some numerical and analytical methods demonstrate effectiveness in EM analysis, they face limitations in memory and speed when performing large-scale circuit simulations [6], [8], [9], [18]. Recently, due to the rapid advancements in machine learning technology, ML-based techniques for solving PDE equations have found extensive application in fast EM stress analysis. A generative adversarial network (GAN)-based method, called EM-GAN, was initially employed to predict EM stress by treating multi-segment interconnect structures as images [13].

However, due to the fixed-size input and output constraints, this method struggles to accurate EM predictions in multi-segment interconnections of varying sizes. Subsequently, an EM prediction method based on a graph convolutional network (GCN), called EMGraph [14], is introduced to address the issues present in EM-GAN. This ML-based model has less memory consumption and faster prediction speed than GAN-based methods, and prediction accuracy is also greatly improved. Although this method effectively accelerates the prediction of EM, it does not factor in the influence of thermomigration on hydrostatic stress.

Hydrostatic stress in interconnects undergoes significant changes when thermomigration is considered. Thermomigration is mainly caused by temperature gradients, where longer branches and junctions connected to more branches generate larger temperature gradients. As shown in Fig. 1(a), the type of junction is one of the crucial influencing factors. Internal junctions are influenced by branches in multiple directions, while boundary junctions connected to standard cells have limitations in temperature changes and atomic diffusion due to differences in media. For branches, it is important to pay attention to the discrete temperature distribution values and the type of branches. Horizontal branches are influenced by internal junction nodes at both ends, leading to more complex changes. Vertical branches are typically connected to boundary junctions at one end, resulting in more limited variations in their states. The stress result at the junction is influenced by the combined effects of its connected branches, whereas the stress result of a branch is influenced by the junctions at both ends. Thus, it becomes essential to choose

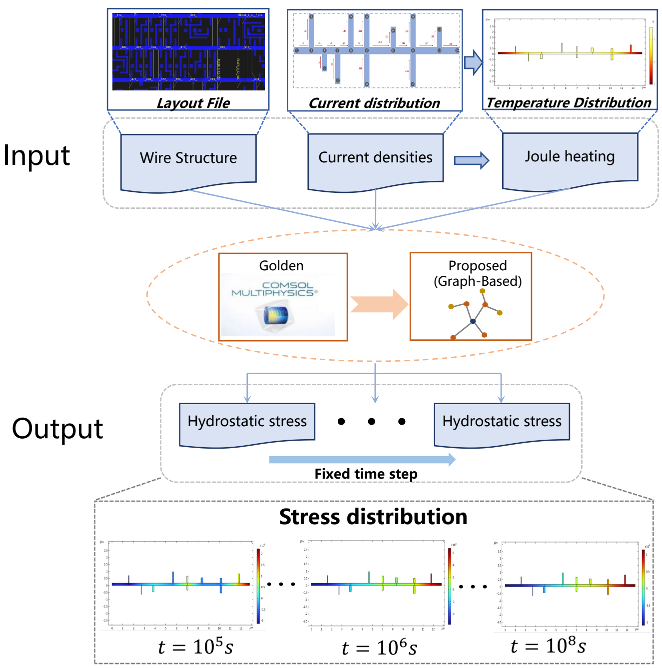


Fig. 2 Overall flow of a fast and accurate EM prediction method

the appropriate features of different junctions and branches, and to accurately model them in machine learning methods.

C. Problem Formulation

Taking into account the impact of both EM and TM factors, simulating hydrostatic stress in large-scale multi-segment interconnects is limited by runtime cost and memory usage.

To replace the traditional process of solving PDE equations, we can employ alternative models based on graph attention networks. In the context of multi-segment interconnects the general structure can be depicted as a combination of junctions and branches shown in Fig. 1(a). This representation enables us to transform it into a graph, denoted $G = (V, E)$, consisting of a set of nodes V (junctions) and a set of directed edges E (branches). The primary challenge in this problem is to devise suitable attention aggregation methods that simultaneously account for both node and edge predictions. This facilitates accurate and efficient EM-TM stress prediction for large-scale multi-segment interconnects like power grids.

III. Proposed Models

In this section, we focus on the development of an improved Graph Attention Network (GAT) model, incorporating node and edge embedding, to predict the EM-TM stress across the entire multi-segment interconnects. We illustrate the pivotal role of our model in EM-TM stress simulations and offer an in-depth analysis of the overall process. A more detailed illustration of our work is shown below.

A. Overall Flow

This paper introduces a fast and accurate framework to predict EM-TM stress in multi-segment interconnects based on customized GATs. The general flow is shown in Fig. 2. First, we obtain the message about the multi-segment interconnect structure (W, L) and use the SPICE simulator to compute the current density J for each wire segment. Meanwhile, we simulate the temperature distribution $T(x)$ resulting from Joule heating as a part of the original input.

TABLE I Node and edge features used in the prediction model.

Type	Features	Definition
Node	T_N	Temperature(K)
	t	Simulation time (s)
	C_N	Node types including boundary nodes and internal nodes
Edge	J	Current density (A/m^2)
	T_E	Discrete temperature distribution(K)
	L	Length (μm)
	W	Width (μm)
	t	Simulation time (s)
	C_E	Edge types including horizontal edges and vertical edges

The conventional approach uses the commercial software COMSOL to simulate the aforementioned parameters and structure, resulting in a series of time-varying EM-TM stress distribution outcomes. This procedure requires the solution of numerous partial differential equations (PDEs). In practice, employing this traditional method for simulating long-term stresses in large-scale multi-segment interconnects is both time-consuming and resource-intensive.

Our work focuses on substituting the PDE-solving phase of COMSOL with the initial input features. We encapsulate the simulation's structure and parameters as a graph, taking into account feature aggregation for both nodes and edges based on GATs Subsequently, we achieve fast EM-TM stress prediction based on these aggregated features. In this model, stress values simulated using COMSOL serve as the gold standard for network labels, ensuring accurate and reliable predictions.

B. Graph Feature Definition

Before engaging in specific data processing and graph neural network embedding, we define significant node features and edge features for each branch as shown in TABLE I. In general, we consider three primary node features: temperature T_N, simulation time t, and node type C_N. Additionally, we include six types of edge features, encompassing current density J, discrete temperature distribution T_E, length L, width W, simulation time t, and edge type C_E. The selection of these features is based on circuit knowledge and parameter comparison experiments, including critical parameters designated for utilization in COMSOL simulation. Labels are the stress values $\sigma_{(comsol)}$ of nodes and edges computed using COMSOL software, which is known as the gold standard.

We utilize a directed graph, denoted $G = (V, E)$, to represent the interconnect structure while taking into account the direction of the current flow. In this representation, V represents a set of nodes N_ν and E represents a set of edges N_e. The node embedding feature of input can be represented as $V : \{v_i = (T_N^{(i)}, t^{(i)}, C_N^{(i)}) | i \in N_\nu\}$. The edge embedding feature of input can be represented as $E : \{e_{(i,j)} = (J^{(i,j)}, T_E^{(i,j)}, L^{(i,j)}, W^{(i,j)}, t^{(i,j)}, C_E^{(i,j)}) | (i,j) \in N_e\}$. The output of the node prediction and edge prediction can be denoted as σ_{pre}^V and σ_{pre}^E, respectively. The output feature for nodes consists solely of the stress value at that point, whereas for edges, it encompasses the stress values of four equidistant

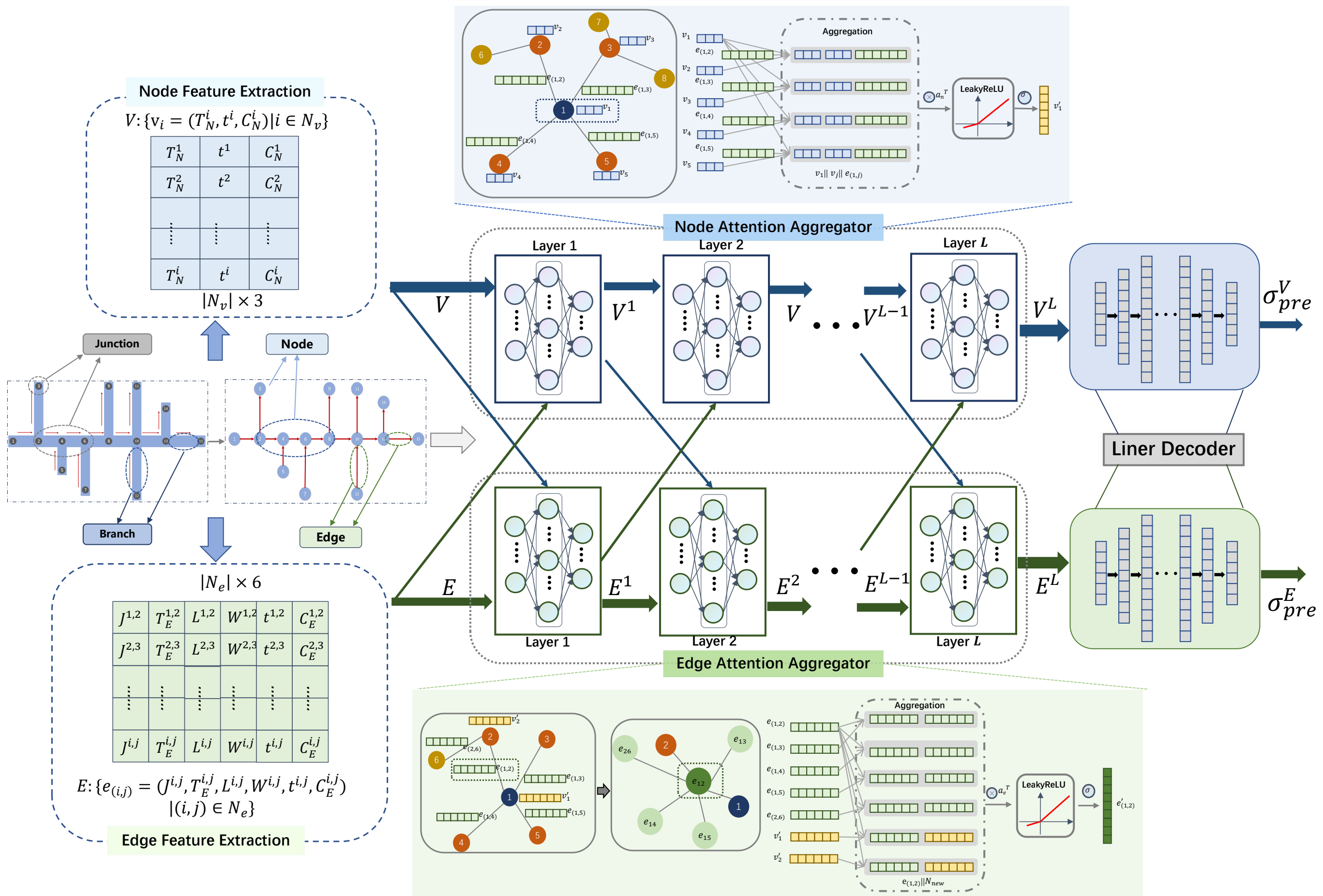


Fig. 3 The training architecture of the EM-TM stress prediction model, including feature extraction, node and edge aggregation, and the liner decoding step.

sampling points.

C. Data Preparation and Preprocessing

To construct a robust dataset of circuit structure samples, we leveraged OpenRoad to simulate some example circuits, and then extract genuine interconnection structures and feature ranges. Subsequently, we employed a random generation method to create 3,000 distinct structures within the specified feature ranges, with the number of branches fewer than 200.

The labels correspond to transient stress values computed using the commercial software COMSOL, spanning a time range of 10^6 to 10^8 seconds. We divided the entire dataset, allocating 80% as the training set and the remaining 20% as the validation set. Furthermore, we gathered comprehensive power network information from several example circuits provided by OpenRoad to predict large-scale network stress.

Due to the considerable variation in magnitude and range among the features of both nodes and edges, the normalization operation is the most important. We employ the Min-Max normalization method shown as Equation (2) to scale both the input features and the output labels in the $(0, 1)$.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

Since the output label affects error calculation, if it becomes zero after normalization, it can cause infinite mean squared error (MSE). To fix this, we convert zero values in our labels to a minimum of 10^{-5}.

D. Model Training Based on Improved GAT.

We have introduced a rapid EM-TM stress prediction model based on the graph attention network (GAT), which incorporates the alternating aggregation of edges and nodes. This approach is advantageous since traditional GAT methods are designed primarily for node classification tasks [19]. Fig. 3 illustrates the training flow of the EM-TM stress prediction model, encompassing feature extraction, node and edge aggregation, and the liner decoding step.

Firstly, we construct feature matrices for nodes and edges, as well as a connectivity matrix, using the pre-processed data mentioned earlier. Then the node attention aggregator in l-th layer accepts the transformed node and edge representations generated through the liner transformer as inputs, $\{\boldsymbol{v}_i^l, \forall i \in N_v\}$ and $\{\boldsymbol{e}_{(i,j)}^l, \forall (i,j) \in N_e\}$, and produces aggregated node representations $\{\boldsymbol{v}_i^l, \forall i \in N_v\}$. For node updates, we execute a convolution operation to aggregate information from a node's neighboring nodes and connected edge data. The node attention coefficients α_{ij}^l indicate the importance of neighborhood information to the target node.

$$\alpha_{ij}^l = \frac{\exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{a}^l)^\top \left[\boldsymbol{v}_i^l \| \boldsymbol{v}_j^l \| \boldsymbol{e}_{ij}^l\right]\right)\right)}{\sum_{k \in N_i} \exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{a}^l)^\top \left[\boldsymbol{v}_i^l \| \boldsymbol{v}_k^l \| \boldsymbol{e}_{ik}^l\right]\right)\right)} \tag{3}$$

where i is the target node and the j is its neighbor belongs to neighborhood set N_i. $\|$ is the concatenation operation, and $.T$ represents transposition. e_{ij} and e_{ik} represent the edges that connect node i, j or i, k . The node attention coefficient is parametrized by a weight vector $\boldsymbol{a}^l \in \mathbb{R}^{2K_V^l + K_E^l}$ and applies LeakyReLU nonlinearity (with a negative input slope of 0.2). All node attention coefficients of node i are normalized with the function $softmax$. Once obtained, the normalized attention coefficients α_{ij}^l are used to compute a linear combination of the features corresponding to them as the final output features for each node (after potentially applying a nonlinearity, σ):

$$\boldsymbol{v}_i^{l+1} = \sigma \left(\sum_{j \in N_i} \alpha_{ij}^l \boldsymbol{v}_j^l \right), \quad \forall i \in N_v \tag{4}$$

Then for edge updates, we introduce an attention aggregator to collect information from neighboring nodes and edges and form new edge features. The two-end nodes and the adjacent edges of each edge in the original graph will become their new neighboring nodes. We integrate the node features of the input $\{\boldsymbol{v}_i^l, \forall i \in N_v\}$ and the edge features $\{\boldsymbol{e}_{(i,j)}^l, \forall (i,j) \in N_e\}$ into a new set of features $\{\boldsymbol{f}_p^l, \forall p \in N'\}$. Unlike the node attention module, the edge-attention module produces aggregated features based on edge attention coefficients β.

$$\beta_{pq}^l = \frac{\exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{b}^l)^\top \left[\boldsymbol{f}_p^l \| \boldsymbol{f}_q^l\right]\right)\right)}{\sum_{k \in N_i'} \exp\left(\mathrm{LeakyReLU}\left((\boldsymbol{b}^l)^\top \left[\boldsymbol{f}_p^l \| \boldsymbol{f}_k^l\right]\right)\right)} \tag{5}$$

where $\boldsymbol{b}^l \in \mathbb{R}^{2(K_V^l + K_E^l)}$. We can get the aggregated representation $\boldsymbol{f}_p^{l+1}$:

$$\boldsymbol{f}_p^{l+1} = \sigma \left(\sum_{q \in N_p} \beta_{pq}^l \boldsymbol{f}_q^l \right), \quad \forall p \in N'. \tag{6}$$

Then the updated edge features $\boldsymbol{e}_{(i,j)}^{l+1}$ can be selected from $\boldsymbol{f}_p^{l+1}$ where $p \longrightarrow (i,j), \forall (i,j) \in N_e$.

After multi-layer alternating attention aggregation of nodes and edges, we obtain the final node embeddings denoted as $V^L : \{v_i, i \in N_\nu\}$ and the final edge embeddings denoted as $E^L : \{e_{(i,j)}, (i,j) \in N_e\}$. These embeddings are then passed to a decoder to generate stress prediction results. The decoder consists of multiple linear transformation layers, and each layer can be expressed as:

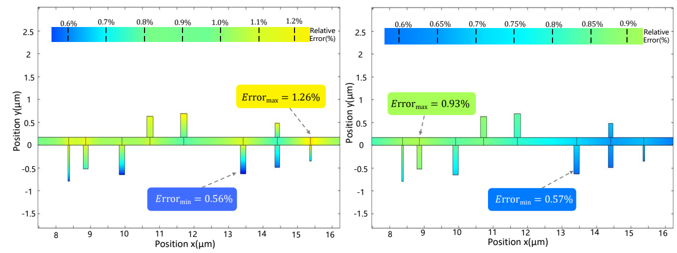
$$h^{(l+1)} = \sigma(W^{(l+1)} h^{(l)} + b^{(l+1)}) \tag{7}$$

where $h^{(l)}$ and $h^{(l+1)}$ represent the input and output respectively. $W^{(l+1)}$ represents the weight matrix and $b^{(l+1)}$ represents the bias.

E. Prediction of Node and Edge EM-TM Stress

Following feature extraction, attention aggregation, and decoding through linear transformation layers, we obtain the predicted stress values for node σ_{pre}^V and edge σ_{pre}^E. To iteratively optimize network parameters, we utilize the mean square error(MSE) of node and edge stress prediction results as the loss function:

$$\begin{aligned}
\mathcal{L}_{\S} &= \frac{1}{n} \sum_{i \in N_v} (\sigma_{i_pre}^V - \sigma_i^V)^2 \\
\mathcal{L}_{\mathcal{E}} &= \frac{1}{m} \sum_{(i,j) \in N_e} (\sigma_{(i,j)_pre}^E - \sigma_{(i,j)}^E)^2 \\
\mathcal{L} &= \mathcal{L}_{\S} + \mathcal{L}_{\mathcal{E}}
\end{aligned} \tag{8}$$



(a) (b)

Fig. 4 (a) shows the stress prediction error rates at various locations within an interconnect structure comprising 130 segments at $t = 10^6 s$;(b) shows the error rates at $t = 10^8 s$.

TABLE II Performance comparison between our method and others.

Number of the segments	Mean Error Rate(%)			
	GCN [20]	EGNN [21]	EMgraph [14]	**Proposed**
$n < 20$	12.7%	4.2%	1.5%	**0.8%**
$20 < n < 100$	16.2%	6.3%		**0.9%**
$100 < n < 200$	19.3%	7.8%	—	**0.9%**

IV. Experimental Results

We have implemented the model using PyTorch and constructed our stress prediction flow based on the proposed graph attention network framework. Our models are trained on a Linux machine with an Intel Xeon CPU with 2.20GHz and 4 NVIDIA Tesla V100 GPUs. Each GPU has 32G of memory. The golden EM-TM stress results were generated using the commercial tool COMSOL on a Windows machine equipped with 16GB of memory and a 2.3GHz Intel Core i7-11800H processor.

A. Prediction Accuracy and Speed of Proposed Model

We employ the trained GAT-based model to make EM-TM stress predictions on the multi-segment interconnect structures present in the test set. Subsequently, we compare the prediction accuracy and speed of the model with some state-of-the-art graph learning methods such as GCN [20], EGNN [21]. These methods introduce edge indices and consider full graph aggregation and local aggregation of nodes, respectively. However, these aggregation updates do not fully consider the multi-dimensional features of real edges. We also compare with an existing graph-based EM prediction method called EMgraph [14].

Fig. 4(a) and Fig. 4(b) show the stress prediction error rates for a portion of a 130-segment interconnects at $t = 10^6 s$ and $t = 10^8 s$, respectively. As depicted in TABLE II, our method achieves an impressive mean error rate of less than 1% to

TABLE III Comparison between our method and COMSOL on several OpenROAD benchmarks.

Tech.	Design	Segments	Runtime(s)			Accuracy
			COMSOL	**Proposed**	**Speed-up**	**Error**(%)
$Nang45nm$	gcd	960	219	0.25	876×	0.96%
	ibex	4322	1753	0.38	4613×	1.04%
	aes	10807	7682	0.85	9037×	1.09%

predict the EM-TM stress for the interconnection segments with fewer than 200 segments and over time spans from 10^6s to 10^8s. This result significantly outperforms GCN and EGNN, which do not consider the alternating aggregation of edge features. The improvement in the model's prediction results can be observed when taking into account the influence of the multi-dimensional features of edges on node features. Furthermore, our model exhibits higher average accuracy compared to EMgraph, which focuses solely on EM stress prediction. Compared to COMSOL, our proposed GAT-based method achieves a substantial speedup of $812\times$ in stress prediction of the test set.

B. Stress prediction on OpenROAD power grid circuits.

To analyze large-scale multi-segment interconnect circuits, we employed several power grids found in part of OpenROAD circuits. These circuits were implemented using the $Nangate45$ nm technology and have a maximum scale of 10,807 segments. We set the simulation step size to 10^6s, and the maximum simulation time is 10^8s. The results of the comparison of EM-TM stress prediction for several circuits are shown in TABLE III. As can be seen, our method maintains a mean error rate lower than 1.1%, while achieving a maximum speedup of $9037\times$ compared to COMSOL.

V. Conclusion

In this paper, we propose a framework to predict EM-TM stress of multi-segment interconnects accurately and efficiently based on customized Graph Attention Networks (GATs) which can collect information of junctions and branches jointly. For randomly selected interconnected samples comprising fewer than 200 segments, the average prediction error is less than 1%, surpassing the performance of other graph-based methods. Furthermore, when applied to the large-scale OpenROAD power grid, our approach shows a substantial speed-up of $9037\times$ compared to COMSOL on structures containing up to $10,807$ segments. In future research, we aim to address challenges related to memory constraints and slow training resulting from the increasing number of samples and scale.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No. 62274034).

References

[1] J. Lienig and M. Thiele, *Fundamentals of electromigration-aware integrated circuit design*. Springer, 2018.

[2] J. R. Black, "Electromigration failure modes in aluminum metallization for semiconductor devices," *Proceedings of the IEEE*, vol. 57, no. 9, pp. 1587–1594, 1969.

[3] S. X.-D. Tan, M. Tahoori, T. Kim, S. Wang, Z. Sun, and S. Kiamehr, "Vlsi systems long-term reliability–modeling, simulation and optimization," 2019.

[4] H.-B. Chen, S. X.-D. Tan, X. Huang, T. Kim, and V. Sukharev, "Analytical modeling and characterization of electromigration effects for multibranch interconnect trees," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 11, pp. 1811–1824, 2016.

[5] L. Chen, S. X.-D. Tan, Z. Sun, S. Peng, M. Tang, and J. Mao, "Fast analytic electromigration analysis for general multisegment interconnect wires," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 2, pp. 421–432, 2019.

[6] O. Axelou, G. Floros, N. Evmorfopoulos, and G. Stamoulis, "Accelerating electromigration stress analysis using low-rank balanced truncation," in *2022 18th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*. IEEE, 2022, pp. 1–4.

[7] M. A. Al Shohel, V. A. Chhabria, N. Evmorfopoulos, and S. S. Sapatnekar, "Analytical modeling of transient electromigration stress based on boundary reflections," in *2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*. IEEE, 2021, pp. 1–8.

[8] P. Stoikos, G. Floros, D. Garyfallou, N. Evmorfopoulos, and G. Stamoulis, "A fast semi-analytical approach for transient electromigration analysis of interconnect trees using matrix exponential," in *Proceedings of the 28th Asia and South Pacific Design Automation Conference*, 2023, pp. 1–6.

[9] O. Axelou, N. Evmorfopoulos, G. Floros, G. Stamoulis, and S. S. Sapatnekar, "A novel semi-analytical approach for fast electromigration stress analysis in multi-segment interconnects," in *Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design*, 2022, pp. 1–7.

[10] L. Chen, S. X.-D. Tan, Z. Sun, S. Peng, M. Tang, and J. Mao, "A fast semi-analytic approach for combined electromigration and thermomigration analysis for general multisegment interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 40, no. 2, pp. 350–363, 2020.

[11] M. Rapp, H. Amrouch, Y. Lin, B. Yu, D. Z. Pan, M. Wolf, and J. Henkel, "Mlcad: A survey of research in machine learning for cad keynote paper," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 10, pp. 3162–3181, 2021.

[12] S. Dey, S. Nandi, and G. Trivedi, "Machine learning approach for fast electromigration aware aging prediction in incremental design of large scale on-chip power grid network," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 25, no. 5, pp. 1–29, 2020.

[13] W. Jin, S. Sadiqbatcha, Z. Sun, H. Zhou, and S. X.-D. Tan, "Emgan: Data-driven fast stress analysis for multi-segment interconnects," in *2020 IEEE 38th International Conference on Computer Design (ICCD)*. IEEE, 2020, pp. 296–303.

[14] W. Jin, L. Chen, S. Sadiqbatcha, S. Peng, and S. X.-D. Tan, "Emgraph: Fast learning-based electromigration analysis for multi-segment interconnect using graph convolution networks," in *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2021, pp. 919–924.

[15] W. Jin, L. Chen, S. Lamichhane, M. Kavousi, and S. X.-D. Tan, "Hierpinn-em: Fast learning-based electromigration analysis for multi-segment interconnects using hierarchical physics-informed neural network," in *Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design*, 2022, pp. 1–9.

[16] A. Todri, S. Kundu, P. Girard, A. Bosio, L. Dilillo, and A. Virazel, "A study of tapered 3-d tsvs for power and thermal integrity," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 2, pp. 306–319, 2012.

[17] M. A. Korhonen, P. Bo/Rgesen, K.-N. Tu, and C.-Y. Li, "Stress evolution due to electromigration in confined metal lines," *Journal of Applied Physics*, vol. 73, no. 8, pp. 3790–3799, 1993.

[18] C. Cook, Z. Sun, E. Demircan, M. D. Shroff, and S. X.-D. Tan, "Fast electromigration stress evolution analysis for interconnect trees using krylov subspace method," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 5, pp. 969–980, 2018.

[19] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[20] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[21] L. Gong and Q. Cheng, "Exploiting edge features for graph neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 9211–9219.