HierPINN-EM: Fast Learning-Based Electromigration Analysis for Multi-Segment Interconnects Using Hierarchical Physics-informed Neural Network

Wentian Jin[1], Liang Chen[1], Subed Lamichhane[1], Mohammadamir Kavousi[1] and Sheldon X.-D. Tan[1]
[1]Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521
wjin018@ucr.edu,liangch@ucr.edu,slami002@ucr.edu,mkavo003@ucr.edu,stan@ece.ucr.edu

ABSTRACT

Electromigration (EM) becomes a major concern for VLSI circuits as the technology advances in the nanometer regime. The crux of problem is to solve the partial differential Korhonen equations, which remains challenging due to the increasing integrated density. Recently, scientific m achine l earning h as b een e xplored to solve partial differential e quations (PDE) d ue t o breakthrough success in deep neural networks and existing approach such as physics-informed neural networks (PINN) shows promising results for some small PDE problems. However, for large engineering problems like EM analysis for large interconnect trees, it was shown that the plain PINN does not work well due the to large number of variables. In this work, we propose a novel hierarchical PINN approach, *HierPINN-EM* for fast EM induced stress analysis for multi-segment interconnects. Instead of solving the interconnect tree as a whole, we first solve EM problem for one wire segment under different boundary and geometrical parameters using supervised learning. Then we apply unsupervised PINN concept to solve the whole interconnects by enforcing the physics laws in the boundaries for all wire segments. In this way, *HierPINN-EM* can significantly reduce the number of variables at plain PINN solver. Numerical results on a number of synthetic interconnect trees show that *HierPINN-EM* can lead to orders of magnitude speedup in training and more than 79× better accuracy over the plain PINN method. Furthermore, *HierPINN-EM* yields 19% better accuracy with 99% reduction in training cost over recently proposed Graph Neural Network-based EM solver, EMGraph.

KEYWORDS

Electromigration (EM), physics informed neural network (PINN), multisegment interconnect, hydrostatic stress assessment

ACM Reference Format:
Wentian Jin[1], Liang Chen[1], Subed Lamichhane[1], Mohammadamir Kavousi[1] and Sheldon X.-D. Tan[1]. 2022. HierPINN-EM: Fast Learning-Based Electromigration Analysis for Multi-Segment Interconnects Using Hierarchical Physics-informed Neural Network. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD '22), October 30–November 3, 2022, San Diego, CA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3508352.3549371

ICCAD '22, October 30-November 3, 2022, San Diego, CA, USA

ACM ISBN 978-1-4503-9217-4/22/10.
https://doi.org/10.1145/3508352.3549371

1 INTRODUCTION

Electromigration (EM) still remains the primary reliability killer for copper based connects in the current and foreseeable nanometer technology nodes. EM-related aging and other VLSI long-term reliability problems will become worse for current 3nm and below technologies. Therefore, it is crucial to have an accurate assessment of aging and reliability for both interconnects and devices during the design process.

For EM analysis, It is well known that existing Black and Blech-based EM models [1, 2] are overly conservative and can only work for single wire segment [3, 4]. To mitigate those problems, recently many physics-based EM models and simulation techniques have been proposed [5–19]. The crux of the problem is to solve Korhonen equations [20], which is the partial differential equations (PDEs) describing the hydrostatic stress evolution in the confined multi-segment interconnect trees subject to blocking materials boundary conditions. In general, solving the Korhonen equation in particular and PDEs by traditional numerical methods still remains a big challenge due to the inherent limitation of those methods.

Recently, scientific machine learning (SciML) has emerged as a promising and alternative solution to traditional numerical analysis techniques to solve partial differential equations (PDE) due to the breakthrough successes of deep neural networks on cognitive tasks [21, 22]. The main concept is to replace the traditional numerical discretization with a deep neural network (DNN) that approximates the solution of the PDE.

One important framework is so-called physics-informed neural networks (PINN), which use differentiable DNN to regularize the loss functions via back-propagation based training to obtain so-called physics-informed/constrained surrogate models [23, 24]. The resulting models can quickly infer the solutions of the PDE to all input coordinates and parameters. Recently PINN-based EM analysis approach has been proposed in [25]. The authors show that PINN can be applied to solving the PDE for stress evolution in the confined metal. However, it only demonstrated the solution on the simple interconnect straight wires. Also, our study show that the plain PINN method does not work well for large interconnect trees.

In this work, we propose a new learning-based solution, called *Hierarchical PINN* or *HierPINN-EM* framework, to solve the Korhonen equations for multi-segment interconnects for fast EM failure analysis. Our new contributions are as follows:

- First, we show that the plain PINN based unsupervised learning does not work very well for interconnects with a large number of segments in terms of accuracy and training

speed. To mitigate this problem, we propose a new *hierarchical PINN* solving strategy to reduce the number of variables, which can lead to faster training and more accurate models. The resulting PINN training framework is called *HierPINN-EM* for fast EM-induced stress analysis for multi-segment interconnects.

- In the *HierPINN-EM* framework, the solving process consists of two steps (levels). The first step is to find a parameterized solution for single-segment wires under different boundary conditions, geometrical parameters and stressing current densities. This step can be solved by different approaches. In this work, we apply supervised learning method to build the DNN model with parameterized input layer as a universal solution to different single-segment wires under various boundary conditions.
- In the second step of *HierPINN-EM*, we apply the existing unsupervised PINN concept to solve the stress and atom flux continuity problem in the interconnects by enforcing the physics laws at the boundaries of all wire segments. In this way, *HierPINN-EM* can significantly reduce the number of variables at the PINN solver, which leads to faster training speed and better accuracy than the plain PINN method.
- Our evaluation results show that *HierPINN-EM* outperforms plain PINN method in both accuracy and performance with more than 79× error reduction and orders of magnitude speedup, which suggests much better scalability. *HierPINN-EM* also yields 19% better accuracy with 99% reduction in training cost over the recently proposed state-of-the-art Graph Neural Network (GNN)-based EM solver, EMGraph.

The paper is organized as follows:

Section 2 reviews the physics-based EM model and its analysis techniques. Section 3 proposes the *HierPINN-EM* with detailed description of two levels inside the framework. Experimental results are presented in Section 4. Finally, section 5 concludes this paper.

2 REVIEW OF RELEVANT WORK

2.1 Review of the EM and EM modeling

EM is a diffusion phenomenon of metal atoms migrating from cathode to anode of confined metal interconnect wires due to the momentum exchange between the conducting electrons and metal atoms [1]. With the EM driving force, the hydrostatic stress increases over time. When the stress reaches a critical value, void is nucleated at the cathode and hillock is created at the anode of the interconnects. This eventually leads to an open or short circuit, which is an EM-induced reliability problem in modern VLSI circuits.

Black's equation predicts EM-induced time-to-failure (TTF) based on an empirical or statistical data fitting, which only works for one specific single wire [1]. Blech's limit, which is an immortality check method, can not estimate transient hydrostatic stress and is subject to growing criticism due to unnecessary overdesign [2]. To mitigate this problem, the physics-based EM model, Korhonen equations [20], is employed to describe the hydrostatic stress evolution for general multi-segment interconnects.

The general multi-segment interconnect consists of n nodes, including p interior junction nodes $x_r \in \{x_{r1}, x_{r2}, \cdots, x_{rp}\}$ and q block terminals $x_b \in \{x_{b1}, x_{b2}, \cdots, x_{bq}\}$, and several branches. The physics-based Korhonen's PDE for this general structure in nucleation phase can be formulated as follows [16, 26].

$$
\begin{aligned}
&\frac{\partial \sigma_{ij}(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\kappa_{ij}\left(\frac{\partial \sigma_{ij}(x,t)}{\partial x} + G_{ij}\right)\right], t > 0 \\
&BC: \sigma_{ij_1}(x_i,t) = \sigma_{ij_2}(x_i,t), t > 0 \\
&BC: \sum_{ij} \kappa_{ij}\left(\frac{\partial \sigma_{ij}(x,t)}{\partial x}\bigg|_{x=x_r} + G_{ij}\right)\cdot n_r = 0, t > 0 \\
&BC: \kappa_{ij}\left(\frac{\partial \sigma_{ij}(x,t)}{\partial x}\bigg|_{x=x_b} + G_{ij}\right) = 0, t > 0 \\
&IC: \sigma_{ij}(x,0) = \sigma_{ij,T}
\end{aligned}
\tag{1}
$$

where BC and IC are boundary and initial conditions respectively, ij denotes a branch connected to nodes i and j, n_r represents the unit inward normal direction of the interior junction node r on branch ij. $\sigma(x,t)$ is the hydrostatic stress, $G = \frac{Eq*}{\Omega}$ is the EM driving force, and $\kappa = D_a B\Omega/k_B T$ is the diffusivity of stress. E is the electric field, $q*$ is the effective charge. $D_a = D_0 \exp(\frac{-E_a}{k_B T})$ is the effective atomic diffusion coefficient. D_0 is the pre-exponential factor, B is the effective bulk elasticity modulus, Ω is the atomic lattice volume, k_B is the Boltzmann's constant, T is the absolute temperature, E_a is the EM activation energy. σ_T is the initial thermal-induced residual stress.

2.2 Existing numerical approaches for solving PDEs

In order to solve the PDE (1) , many conventional numerical and analytical methods are proposed to attempt to solve the PDEs efficiently and accurately [14, 16–18, 26–28]. Although the numerical methods, such as finite difference method [17, 27, 28] and finite element method (FEM) [14], can work for the complex interconnect structures and obtain EM stress accurately, they impose high computational costs due to discretization of space and time. Recently, semi-analytical method based on separation of variables method has been proposed [16, 26], which show promising performance in both accuracy and efficiency on general multi-segment interconnect. Furthermore, an analytic approach [18] was proposed recently, which is very fast. But it can not be applied to the interconnect line wires.

2.3 Learning based approaches for solving PDEs

Recently machine learning, especially deep learning based on deep neural networks have made breakthrough success in many cognitive applications such as image, text, speech and graph recognition [29, 30]. Inspired by these observations, neural networks are modified to solve the PDEs [31–33].

Recently a generative adversarial networks (GAN) based method, called *EM-GAN*, is proposed to perform a fast transient hydrostatic stress analysis by solving Korhonen equations [34]. It achieved an order of magnitude speedup over the efficient analytic based EM solver with good accuracy. However, this method only works for a fixed region because its output is an image with fixed size, which restricts its application in real chips. What is more,

the image is not a natural tool to represent the multisegment interconnects as the region with large areas are filled with nothing. Then Jin *et al.* further proposed an improved GNN-based EM solver, called *EMGraph* [35]. Since GNN represents more general and natural relationship among different design objectives, knowledges learned by GNN models tend to be more transferable for different designs, which is highly desirable. However, all those methods are still supervised learning approaches, which requires extensive training from numerical solvers or measured data.

To mitigate this drawback, recently unsupervised learning framework, called *physics-informed neural networks* [23, 36], PINN or *physics-constrained neural networks* [31, 37] have been proposed. The key concept is to frame the PDE solving process into a nonlinear optimization process coded by DNN with the loss functions to enforce the physics laws represented by the PDE and boundary conditions. However, only very simple PDE problems were demonstrated in [23, 24, 31, 36, 38] although some progresses were made for more complicated aerodynamics simulation recently [39]. Recently, a PINN-based approach for EM analysis has been proposed [25]. The method tries to improve the PINN method to better handle the temperature-dependent diffusivities for metal atom migrations. It tries to add more neurons representing some predetermined allocation points and time instances into the neural networks. This method slightly improves the plain PINN method by achieving better training accuracy at the cost of longer training time under the same number of neurons.

3 HIERPINN: HIERARCHICAL PHYSICS INFORMED NEURAL NETWORK

In this section, we present the new hierarchical PINN solving strategy, which takes multi-segment interconnect tree as input and predicts the EM-induced stress for arbitrary locations in the interconnects at any given aging time. The proposed model solves the stress evolution equations in a hierarchical way which consists of two levels (stages). The first stage, or the lower level, takes only single-segment straight wire as input and predicts the stress inside and at both ends of the given wire. The first stage can be viewed as implicitly enforcing the physics laws related to the stress evolution inside a segment wire and leaving the boundaries as the input and output variables to be used in the second stage. The second stage, or the upper level, takes all internal junctions or boundaries as inputs and matches the stress predicted by the lower level at each junction to meet the boundary conditions among adjacent wires (i.e. stress continuity and atom flux conservation) in the original PDE using the PINN optimization framework. Each level employs a multilayer perceptron (MLP) network with different configurations as the backbone. The rest of this section will introduce both levels in detail together with data preparation procedures.

3.1 Lower level: single-segment straight wire stress predictor

The lower level of the proposed hierarchical PINN is a stress predictor/solver which takes single-segment straight wire as input and predicts the EM-induced stress for any location on the wire at a given aging time instant. We note that multi-segment interconnects always consist of many wire segments with different widths and lengths, stress currents and atomic fluxes at the two terminals. For one wire segment, once the geometrical parameters, current density and boundary conditions are given, EM-induced stresses are determined at all locations including terminals for a given time instant. For one wire segment under those parameterized conditions, one way to obtain the fast and compact model is by means of DNN networks via supervised learning. The whole process is illustrated in Fig. 1: The backbone of the stress predictor is a multilayer

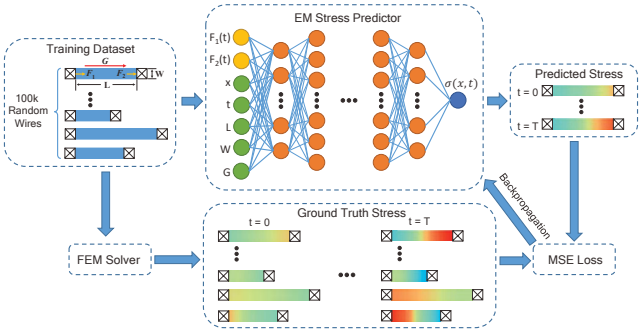

Figure 1: Framework of proposed stress predictor in the first stage.

perceptron network with 7 layers. The input layer has 7 neurons corresponding to location, aging time, stress current and geometrical parameters of a straight wire. The input vector is called wire feature vector denoted as u. As is shown in Fig. 1, we use L, W and G to denote the length, width, and driving force of the input wire. F_1 and F_2 are the atom fluxes at left and right end of the wire segment. x and t are the location and aging time at which the stress is to be predicted. For any input wire, the positive direction of location, driving force and atom flux is always pointing from left to right, such that negative values are allowed to represent the opposite direction. The output layer of stress predictor has only a single neuron with no non-linear activation function attached. The scalar value of this output neuron indicates the predicted stress at the input location x and aging time t.

For any given combination of wire geometries, driving force and boundary conditions at both ends of a wire, the stress evolution is uniquely determined according to the EM stress PDE. With that being said, the task of the proposed stress predictor is to solve the PDE in a single-wire case. The boundary conditions at left and right ends of the wire are defined by F_1 and F_2 respectively, while the rest of the parameters in the stress evolution equation are set by the other 5 input neurons. The input neurons are then processed through layers of non-linear forward propagation operations and the final output neuron is a scalar value indicating the predicted stress. In this process, the stress predictor serves as an approximation to the single-wire EM stress solver. The hidden layers inside the MLP learn to capture the governing physics laws and convert the input features into stress results.

We want to stress that for practical use of this method, the first stage needs to be trained *ONLY ONCE*. Then it can be used for different multi-segment wires with different number of wires and topologies and stressing current density conditions. This is a

real benefit of our approach as the training cost of this stage can be ignored for sufficient applications of this method. Second, our method is open to other solving solutions at the first stage (such as analytic solutions, fast numerical solutions). Third, we can pursue more accurate DNN modeling even at high computing costs.

Regarding the accuracy, the accuracy of stress predictor as an approximated PDE solver is guaranteed in two senses. First, we limit the stress predictor to work only on simple single-wire cases instead of overcomplicated interconnect trees. Although a multi-layer perceptron, as a universal approximator, can theoretically approximate any complicated non-linear solver, it is not practical to implement it in real cases with limited resources and strict speed requirements. Thus, such limitation substantially reduces the complexity of the problem given to the model, which makes it possible to achieve both high accuracy and performance.

Secondly, the stress predictor is trained on a large dataset consisting of 80k random wires, as is shown on the lefthand side in Fig. 1. To obtain abundant high-quality training data, we generated 80k single straight wires and randomly assigned length, width and current density to each wire. These wires are then randomly connected together to create interconnect trees. Due to the randomness in both generation and connection procedures, the resulting interconnect trees are also completely random with various topologies. These interconnects are then passed into COMSOL, which is a commercial FEM solver, to do the EM stress simulation. The COMSOL-simulated stress evolutions in every single wire, as illustrated at the bottom of Fig. 1, are saved as ground truth results. During the FEM simulation process, the time-variant atom fluxes at both ends of each wire ($F_1(t)$ and $F_2(t)$) were recorded and then concatenated with wire geometries to serve as input features in the training dataset. Again, such one-time training cost does not add significant overall computing cost to the final *HierPINN-EM* solution as mentioned earlier.

To verify the accuracy of the proposed stress predictor, we generated another 20k single wires to serve as the test set. These wires were generated using the same methodology as we used in generating the training dataset. These test data were never seen by the model during the training process and the trained stress predictor showed impressive accuracy on this test set with details demonstrated in Section 4.1.

3.2 Upper level: atom flux predictor for all the wire segments

After the lower level stress predictor is trained, we can then build the upper level on top of it as illustrated in Fig. 2. Similar to the stress predictor, the backbone of the upper level is still an MLP but with completely different objective and configurations.

The goal of the upper level is to predict correct boundary conditions, i.e. atom fluxes, for every single wire in the given interconnect tree, so that the EM stress in internal junctions or boundaries are continuous and EM stress in each wire can be independently derived using the stress predictor already trained in the lower level. To achieve this, we implemented the atom flux predictor using an MLP model with 7 layers.

This model takes internal junctions instead of wire segments as input. Therefore, the first stage in the upper level, as illustrated in the lefthand side of Fig. 2, is to label all internal junctions in the input interconnect tree starting from 1. Each internal junction is then assigned a feature vector and the feature vector of j-th junction is denoted as x_j. The first two entries in the feature vector are the label number j and aging time t. Since each internal junction may have up to 4 wires connected to it from 4 directions, i.e. left-side, upside, rightside, downside, we appended the driving forces in these four directions (G_L, G_U, G_R, and G_D) to the feature vector and the driving force would be set to zero if there was no connected wire in the corresponding direction. The resulting feature vectors are then passed into the atom flux predictor.

The output of atom flux predictor for j-th internal junction is a vector denoted as z_j. There are 3 entries in z_j, which correspond to the predicted atom fluxes in left, up and right directions respectively. As suggested by the second and third boundary conditions in EM PDE (1), the atom fluxes at each internal junction must satisfy the flux conservation law. As a result, we apply the flux conservation law to z_j so that we can calculate the fourth atom flux in the downward direction using the other three predicted results. In this way, the flux conservation law is strictly enforced at every internal junction in the interconnects. The predicted atom fluxes are then filled into corresponding F_1 and F_2 entries of wire feature vectors u, in which the other entries are directly obtained from the original input interconnects.

A huge advantage of our proposed hierarchical method is that the upper level just requires to be trained at internal junctions instead of the whole interconnects. As a result, the x entries in the wire feature vectors are set to 0 or 1 which corresponds to the left/down or right/up end of each wire separately. This avoids random sampling inside each wire ($0 < x < 1$), which would otherwise lead to a large amount of training points and significantly increase the training cost.

The wire feature vectors are then passed into the trained stress predictor to obtain the stress results at all internal junctions. For each internal junction at any aging time, there would be 2 to 4 predicted stress results depending on how many wires are connected at this junction. According to the first boundary condition in EM PDE (1), stress values should be continuous at boundaries, which means that the predicted stress results should equal to each other at any internal junction. This leads to the following physics-informed loss function which we proposed to train the atom flux predictor,

$$L = \frac{1}{N_I \times K_i} \sum_{i=1}^{N_I} \sum_{k=2}^{K_i} (\sigma_k(t) - \sigma_{k-1}(t))^2 \tag{2}$$

where N_I denotes the number of internal junctions in the interconnects, K_i represents the number of connected wires at i-th junction, which ranges between 2 and 4. $\sigma_k(t)$ is the k-th predicted stress result for the current junction at aging time t. The loss function is the mean squared error (MSE) of all predicted internal junction stress, which serves as a measurement of stress discontinuity at boundaries. When training the upper level, the lower level stress predictor is fixed and the loss is only back-propagated back to the atom flux predictor to update the weights and biases in the model.

We note that the proposed hierarchical PINN approach bears some similarity to the domain decomposition method [40] in which hierarchical solving strategies are employed. However,

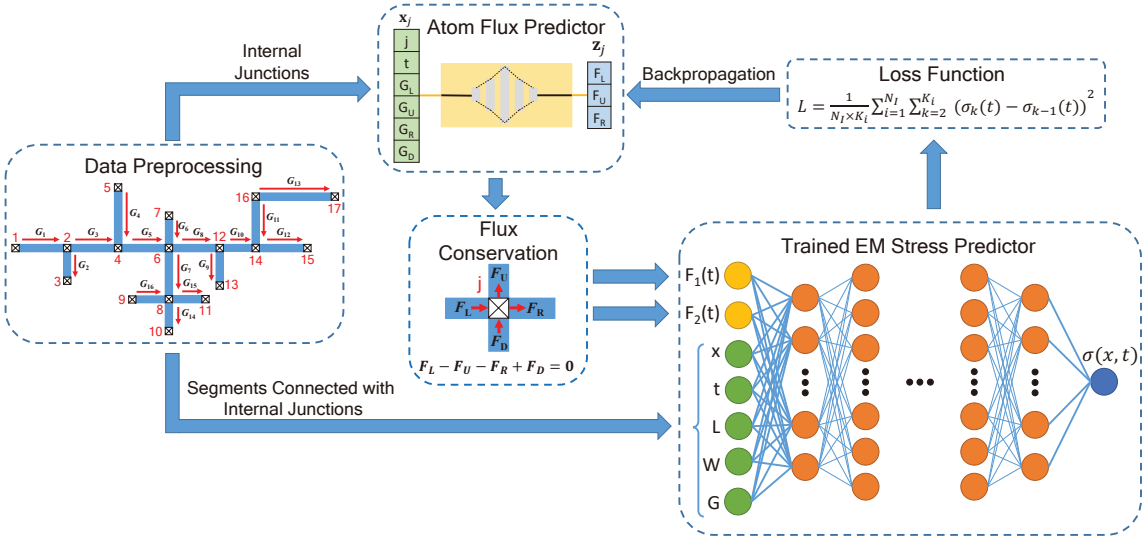

Figure 2: Framework of proposed hierarchical PINN based EM Predictor.

there are several distinguished differences between the two approaches. First, in domain decomposition, the subcircuits are typically obtained by partitioning and have to be solved for each subcircuit every time when the whole circuit is to be solved. While for *HierPINN-EM*, solutions of the single wire, whose boundary or junctions are naturally defined, can be obtained much more efficiently via inference on the DNN network, which only needs to be trained *ONCE*. At the top level, domain decomposition method tries to solve more dense matrices due to the subcircuit reduction via matrix solving processes like LU decomposition, while *HierPINN-EM* uses the unsupervised PINN framework to find the solution, which are meshless and easy for design space parameterizations. It uses the differential nature of the DNN model trained in the first stage to guide backpropagation process in the second stage.

4 EXPERIMENTAL RESULTS

In this section, we demonstrate the prediction accuracy, speed and scalability of the proposed *HierPINN-EM* by testing it on both straight wires and interconnect trees that were randomly generated. Both Atom Flux Predictor and Stress Predictor in *HierPINN-EM* are implemented in Python 3.8.12 with PyTorch 1.7.1. The training and testing of both models were run on a Linux server with 2 Xeon E5-2699v4 2.2 GHz processors and Nvidia TITAN RTX GPU. In the training phase, Adam optimizer was used to update the model and the learning rate was set to 10^{-4}. The cross-validation technique was employed in the training process.

4.1 Accuracy of lower level on single-segment wires

The lower level stress predictor is the foundation of our proposed hierarchical method. A 7-layer multilayer perceptron with configurations of [7, 256, 512, 1024, 512, 256, 1] is employed as the backbone of the stress predictor. The stress predictor takes wire feature vector consisting of wire geometries, EM driving force and boundary conditions as input, and outputs the predicted EM-induced stress at given position and aging time. As mentioned in Section 3.1, we generated a large dataset composed of 100k single-segment straight wires, and 80k of them were used to train the model while the rest 20k wires were reserved for testing. The model was trained for 20 epochs which cost approximately 23 hours.

Fig. 3 shows the predicted stress vs ground truth at 2576392 locations for aging time of 10^6 seconds. These locations were randomly sampled in all 20k wires from the test set which were never seen by the stress predictor during the training process.

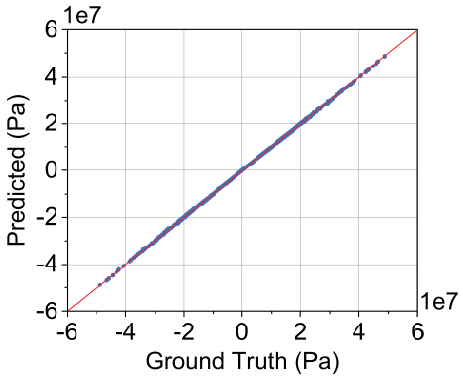

Figure 3: Stress predictor vs ground truth on 20k single-segment wires.

As is shown in Fig. 3, the predicted stress values all lie close to the red line (slope=1, intercept=0). For all 20k wires in the test set, the trained stress predictor yields root-mean-square error (RMSE) from 7.5×10^3 to 5.7×10^4 Pa and achieves an averaged RMSE of

2.7×10^4 Pa. Both the worst and best predicted wire segments are shown in Fig. 4.

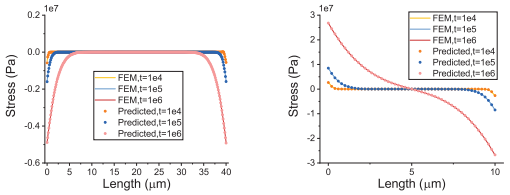

Figure 4: Comparison of predicted stress and ground truth in (a) best and (b) worst wire segments.

The predicted results agree very well with the ground truth even in the worst case, and such accuracy is even more impressive considering that the ground truth stress values vary in a large range between -5×10^7 and 5×10^7 Pa. By dividing the RMSE with the full range of the stress (i.e. 10^8 Pa), the worst and mean error rates of the stress predictor can be calculated as 0.008% and 0.057% separately. Its low error rate in predicting stress in each segment will serve as the foundation of accurate predictions in larger interconnect trees.

We want to note that this stress predictor only needs to be trained *ONCE* and can then be embedded into the upper level for EM analysis of different interconnect trees with different topologies. As a result, it can be viewed as a library, which needs to be built once and can actually be generated using different methods as long as the stress evolution physics laws in a single wire are learned and enforced.

4.2 Accuracy of EM stress prediction on straight wires

To verify the accuracy of *HierPINN-EM*, we first test it on 121 randomly generated multi-segment straight wires. All test cases have random numbers of segments ranging from 10 to 130. The stressing current density and geometrical parameters of each segment are also randomly assigned. Fig. 5 shows the comparison of EM stress predicted by *HierPINN-EM* and ground truth FEM results simulated by COMSOL. Fig. 5(a) shows the results of the smallest case in the test set which has 10 segments, and Fig. 5(b) shows the results of the largest case with 130 segments. To show the evolution of the EM stress as aging time increases, we plotted stress results at three aging time instants (10^4, 10^5 and 10^6 seconds) for each case. The RMSEs of the *HierPINN-EM* predicted stress results are $5.9{\times}10^4$ and 2.0×10^5 Pa separately for these two cases. For all 121 test cases, *HierPINN-EM* yields RMSEs ranging from 4.7×10^4 to 4.0×10^5 Pa and the mean error is 1.9×10^5 Pa which can be converted to 0.19% mean error rate when taking the full 10^8 Pa stress range into consideration.

To compare *HierPINN-EM* with existing PINN method, we also implemented a plain PINN model using a 7-layer MLP with exactly the same structure as *HierPINN-EM*. Note that we did not compare our method with [25], as this method can be essentially viewed as

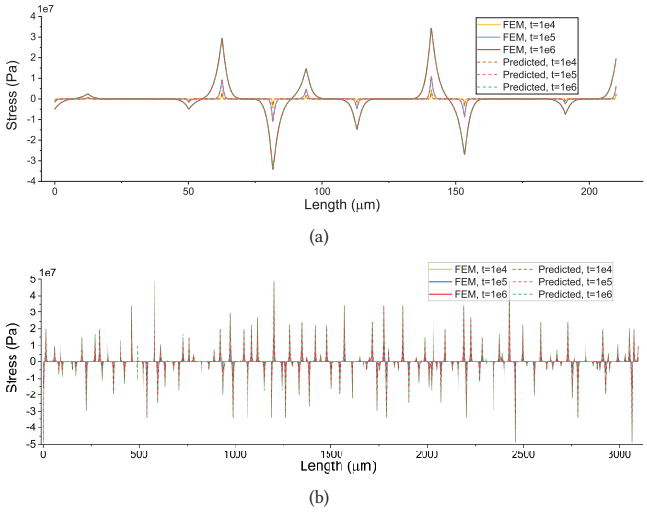

Figure 5: Stress comparisons of (a) a 10-segment and (b) a 130-segment straight wire between *HierPINN-EM* and COMSOL at 3 aging time instants: 1e4, 1e5 and 1e6 seconds.

a plain PINN method with some trade-offs between training accuracy and training time due to more complicated neuron representations.

All equations from (1) are formulated into a single physics-informed loss function to train the plain PINN. We tested the plain PINN on the same 10-segment and 130-segment test cases and the comparison of plain PINN and FEM stress results at aging time of 10^6s are shown in Fig. 6. The RMSEs of plain PINN results in these two cases are 6.4×10^6 and 1.6×10^7 Pa separately, which are 107× and 79× worse than that of *HierPINN-EM*. Due to the large number of collocation points required in each segment in the training process of the plain PINN model, the training cost is around 30 minutes on the 10-segment case and it leaps significantly to around 10 hours on the 130-segment case. As a comparison, the training costs of *HierPINN-EM* are only 5.5s and 39.1s separately in these two cases. The result verifies the great advantage in scalability of our proposed hierarchical method over the plain PINN model.

The physics-informed loss function used to train the plain PINN model contains all PDE equations for all domains which leads to a complex training process. This means that the PINN model has to be trained simultaneously in all segments and boundaries to minimize every single error in the loss function to satisfy different physics laws. This makes the training process of PINN highly unstable and a successful convergence is not always guaranteed. In contrast, our proposed *HierPINN-EM* overcomes this issue by splitting the equations into two levels so that the model at each level is separately trained to satisfy a simpler physics law. The lower level is focused on a single segment while the upper level is based on a trained lower level model, so that the upper level only has to be trained on a few internal junction points which significantly alleviates the training load. Once the upper level is trained, the physics laws inside each segment are automatically satisfied thanks to the

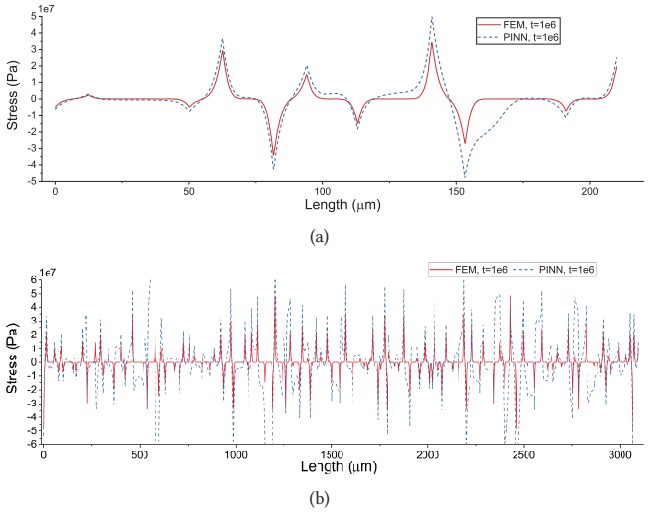

Figure 6: Stress comparisons of (a) a 10-segment and (b) a 130-segment straight wire between plain PINN and COMSOL at aging time instant 1e6 seconds.

highly accurate lower level model which is already shown in Section 4.2.

Moreover, different physics laws in a single loss function also require careful consideration in weight assignment to balance the influence of each equation [33]. Such weight balancing process adds extra overheads to the training process of plain PINN model and further limits its scalability in large interconnects.

4.3 Accuracy of EM stress prediction on interconnect trees

To demonstrate the generalizability of the proposed model, we further test *HierPINN-EM* on 2-D interconnect trees with more complicated topologies. The test set consists of 165 interconnect trees with random numbers of segments ranging from 10 to 105. Similar to the straight wire test set, the stressing current density and geometrical parameters of each segment are also randomly assigned.

To make an apple-to-apple comparison with GNN-based method EM-Graph, we convert all prediction results into stress maps, which are generated by projecting the predicted stress results for every single segment onto the interconnects topology and are shown in 3-D formats. We show the comparison between *HierPINN-EM*, EM-Graph and COMSOL results of the smallest (10 segments) and largest (105 segments) designs from the test set in Fig. 7. The evolution of the stress at 3 time instants, i.e. 10^4, 10^5 and 10^6 seconds, are illustrated from left to right in each row.

Comparing the predicted stress maps in Fig. 7 with COMSOL ground truth results, *HierPINN-EM* yields 1.2×10^5 and 3.4×10^5 Pa RMSE for these two cases while EM-Graph yields 3.1×10^5 and 3.9×10^5 Pa which are slightly worse than *HierPINN-EM*. For all 165 interconnect trees in the test set, *HierPINN-EM* achieves better accuracy with a mean RMSE of 2.8×10^5 Pa while EM-Graph achieves 3.6×10^5 Pa. Thus, *HierPINN-EM* outperforms EM-Graph

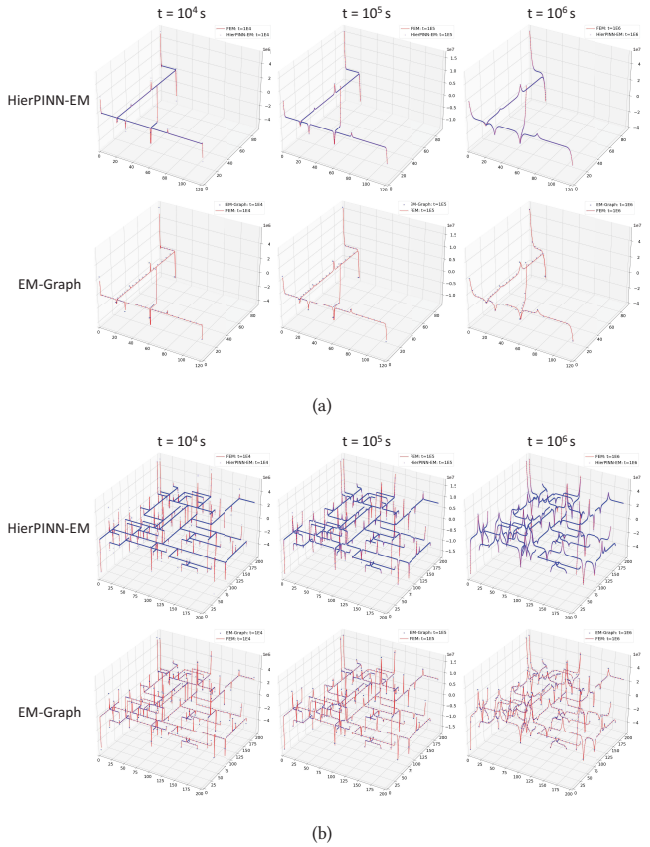

Figure 7: Stress comparisons of (a) a 10-segment and (b) a 105-segment interconnect tree between *HierPINN-EM* and COMSOL at 3 aging time instants: 1e4, 1e5 and 1e6 seconds.

in accuracy with 19% better RMSE when predicting EM-induced stress for 2-D interconnect trees.

Moreover, EM-Graph sets the number of sampling points in each segment to 5 which leads to a coarse granularity in predicted stress results. This is not a big concern when the lengths of segments are relatively small. However, when the segments get much longer, it may introduce huge errors into the prediction since there are large spaces between 5 sampling points and the interpolations between them become much less reliable. This problem is solved in *HierPINN-EM* as the location input x is a scalar value in float format which can represent any point in the segment. The granularity of the results can be easily controlled by altering the sampling density of input x. The better inference flexibility gives *HierPINN-EM* more potential in generalizability to larger interconnect trees.

4.4 Speed of inference

The training process of the *HierPINN-EM* is conducted in stages, the lower level was trained for 23 hours, while the training cost of the upper level varies case by case between 4 to 57 seconds, which is mainly determined by the number of internal junctions in the interconnect tree. Although the training of lower level seems quite

time-consuming, it only has to be trained *ONCE* and provides a universal predictor that can be repeatedly used in the upper level with no further tuning effort required.

Once the *HierPINN-EM* is trained, both levels will be set to inference mode. All sampling points in the interconnects will be passed simultaneously into the model to leverage the parallel computation advantage of GPU. We tested the training and inference speeds of both *HierPINN-EM* and EM-Graph on the interconnect tree test set, and the results are summarized in Table 1.

Table 1: Accuracy and speed comparison

Metrics	HierPINN-EM	EM-Graph	COMSOL
Max RMSE	8.9×10^5 Pa	5.3×10^5 Pa	Ground Truth
Min RMSE	8.4×10^4 Pa	1.9×10^5 Pa	
Mean RMSE	2.8×10^5 Pa	3.6×10^5 Pa	
Mean Error Rate	0.28%	0.36%	
Training Speed	<1min	2hr	-
Inference Speed	0.8ms	0.27ms	30min

Both learning-based methods are able to achieve magnitudes of speedup against COMSOL. Specifically, the mean inference speed of *HierPINN-EM* for each interconnect tree is 0.8ms, which is 3× slower than the 0.27ms inference speed of EM-Graph. However, such difference in the inference speed is mainly caused by the difference in sampling density as is already shown in Section 4.3. During the inference test, the number of sampling points in each segment varies according to the wire length in *HierPINN-EM* but is fixed to only 5 in EM-Graph due to its fixed input layer structure. This leads to approximately 30× more sampling points in *HierPINN-EM*, and thus, more computational cost. However, with such adaptive sampling ability, *HierPINN-EM* is able to predict more accurate stress map with much better granularity. As a result, the loss in the inference speed is actually an acceptable tradeoff.

Another major advantage of *HierPINN-EM* over EM-Graph lies in its better flexibility in the inference phase. Limited by the message-passing structure of GNN, to predict EM stress in any single segment, the whole interconnect graph requires to be fed into EM-Graph so that the target segment can receive useful information from its multi-hop neighbors which can be leveraged to predict stress. This means that EM-Graph is only able to predict stress for interconnects as a whole but is not able to do predictions for small local regions. On the contrary, *HierPINN-EM* takes position x and time t as input parameters, which enables it to do stress predictions with much better flexibility. It can predict stress for any segment or even a single point in the interconnects at any aging time by simply passing the interested location and time into the model. This enables *HierPINN-EM* to achieve more significant speedups in local stress analysis. The better accuracy, inference flexibility and results granularity make *HierPINN-EM* a better learning-based approach for transient EM stress analysis.

5 CONCLUSION

In this paper, we propose a hierarchical learning-based method, called *HierPINN-EM*, to solve the Korhonen equations for multi-segment interconnects for fast EM failure analysis. *HierPINN-EM* split the physics laws into two levels and solve the PDE equations step by step. The lower level employs supervised learning to train a DNN model which takes parameterized neurons as inputs and serves as a universal parameterized EM stress solver for single-segment wires. The upper level employs physics-informed loss function to train a separate DNN model at the boundaries of all wire segments to enforce the stress and atom flux continuities at internal junctions in interconnects. Numerical results on a number of synthetic interconnect trees show that *HierPINN-EM* can lead to orders of magnitude speed up in training and more than 79× better accuracy over the plain PINN method. Furthermore, *HierPINN-EM* yields 19% better accuracy with 99% reduction in training cost over recently proposed Graph Neural Network-based EM solver, EMGraph.

ACKNOWLEDGMENTS

This work is supported in part by NSF grants under No. CCF-1816361, in part by NSF grants under No. OISE-1854276, No. CCF-2007135 and CCF-2113928.

REFERENCES

[1] J. R. Black, "Electromigration-A Brief Survey and Some Recent Results," *IEEE Trans. on Electron Devices*, vol. 16, no. 4, pp. 338–347, Apr. 1969.
[2] I. A. Blech, "Electromigration in thin aluminum films on titanium nitride," *Journal of Applied Physics*, vol. 47, no. 4, pp. 1203–1208, 1976.
[3] M. Hauschildt, C. Hennesthal, G. Talut, O. Aubel, M. Gall, K. B. Yeap, and E. Zschech, "Electromigration Early Failure Void Nucleation and Growth Phenomena in Cu And Cu(Mn) Interconnects," in *IEEE Int. Reliability Physics Symposium (IRPS)*, 2013, pp. 2C.1.1–2C.1.6.
[4] V. Sukharev, "Beyond Black's Equation: Full-Chip EM/SM Assessment in 3D IC Stack," *Microelectronic Engineering*, vol. 120, pp. 99–105, 2014.
[5] R. De Orio, H. Ceric, and S. Selberherr, "Physically based models of electromigration: From black's equation to modern tcad models," *Microelectronics Reliability*, vol. 50, no. 6, pp. 775–789, 2010.
[6] X. Huang, A. Kteyan, S. X.-D. Tan, and V. Sukharev, "Physics-Based Electromigration Models and Full-Chip Assessment for Power Grid Networks," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 11, pp. 1848–1861, Nov. 2016.
[7] V. Sukharev, A. Kteyan, and X. Huang, "Postvoiding stress evolution in confined metal lines," *IEEE Transactions on Device and Materials Reliability*, vol. 16, no. 1, pp. 50–60, 2016.
[8] H. Chen, S. X.-D. Tan, X. Huang, T. Kim, and V. Sukharev, "Analytical modeling and characterization of electromigration effects for multibranch interconnect trees," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 11, pp. 1811–1824, 2016.
[9] V. Mishra and S. S. Sapatnekar, "Predicting Electromigration Mortality Under Temperature and Product Lifetime Specifications," in *Proc. Design Automation Conf. (DAC)*, Jun. 2016, pp. 1–6.
[10] H.-B. Chen, S. X.-D. Tan, J. Peng, T. Kim, and J. Chen, "Analytical modeling of electromigration failure for vlsi interconnect tree considering temperature and segment length effects," *IEEE Transaction on Device and Materials Reliability (T-DMR)*, vol. 17, no. 4, pp. 653–666, 2017.
[11] S. Chatterjee, V. Sukharev, and F. N. Najm, "Power Grid Electromigration Checking Using Physics-Based Models," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 7, pp. 1317–1330, Jul. 2018.
[12] 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 Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 5, pp. 969–980, May 2018.
[13] S. Wang, Z. Sun, Y. Cheng, S. X.-D. Tan, and M. Tahoori, "Leveraging recovery effect to reduce electromigration degradation in power/ground TSV," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*. IEEE, Nov. 2017, pp. 811–818.
[14] H. Zhao and S. X.-D. Tan, "Postvoiding fem analysis for electromigration failure characterization," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*,

vol. 26, no. 11, pp. 2483–2493, Nov. 2018.
[15] A. Abbasinasab and M. Marek-Sadowska, "RAIN: A tool for reliability assessment of interconnect networks—physics to software," in *Proc. Design Automation Conf. (DAC)*. New York, NY, USA: ACM, 2018, pp. 133:1–133:6.
[16] 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*, pp. 1–12, 2019.
[17] Z. Sun, S. Yu, H. Zhou, Y. Liu, and S. X.-D. Tan, "EMSpice: Physics-Based Electromigration Check Using Coupled Electronic and Stress Simulation," *IEEE Transactions on Device and Materials Reliability*, vol. 20, no. 2, pp. 376–389, Jun. 2020.
[18] 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)*, 2021, pp. 1–8.
[19] S. X.-D. Tan, M. Tahoori, T. Kim, S. Wang, Z. Sun, and S. Kiamehr, *VLSI Systems Long-Term Reliability – Modeling, Simulation and Optimization*. Springer Publishing, 2019.
[20] 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.
[21] L. Lu, "Theory, algorithms, and software for physics-informed deep learning," Ph.D. dissertation, Brown University, May 2020.
[22] N. Baker, F. Alexander, T. Bremer, A. Hagberg, Y. Kevrekidis, H. Najm, M. Parashar, A. Patra, J. Sethian, S. Wild, K. Willcox, and S. Lee, "Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence," *Technical Report, US DOE Office of Science*. [Online]. Available: https://www.osti.gov/biblio/1478744
[23] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations," *arXiv e-prints*, p. arXiv:1711.10561, Nov. 2017.
[24] Y. Yang and P. Perdikaris, "Physics-informed deep generative models," *arXiv e-prints*, p. arXiv:1812.03511, Dec. 2018.
[25] T. Hou, N. Wong, Q. Chen, Z. Ji, and H.-B. Chen, "A space-time neural network for analysis of stress evolution under dc current stressing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp. 1–1, 2022.
[26] X. Wang, Y. Yan, J. He, S. X.-D. Tan, C. Cook, and S. Yang, "Fast physics-based electromigration analysis for multi-branch interconnect trees," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*. IEEE, Nov. 2017, pp. 169–176.
[27] 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 Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 5, pp. 969–980, May 2018.
[28] V. Sukharev and F. N. Najm, "Electromigration Check: Where the Design and Reliability Methodologies Meet," *IEEE Transactions on Device and Materials Reliability*, vol. 18, no. 4, pp. 498–507, Dec. 2018.
[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[30] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016, http://www.deeplearningbook.org.
[31] J. Sirignano and K. Spiliopoulos, "DGM: A deep learning algorithm for solving partial differential equations," *Journal of Computational Physics*, vol. 375, pp. 1339 – 1364, 2018.
[32] J. Tompson, K. Schlachter, P. Sprechmann, and K. Perlin, "Accelerating Eulerian fluid simulation with convolutional networks," ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. International Convention Centre, Sydney, Australia: PMLR, 06–11 Aug 2017, pp. 3424–3433.
[33] W. Jin, S. Peng, and S. X.-D. Tan, "Data-driven electrostatics analysis based on physics-constrained deep learning," in *Proc. Design, Automation and Test In Europe Conf. (DATE)*, Feb. 2021, pp. 1–6.
[34] W. Jin, S. Sadiqbatcha, Z. Sun, H. Zhou, and S. X.-D. Tan, "Em-gan: Data-driven fast stress analysis for multi-segment interconnects," in *Proc. IEEE Int. Conf. on Computer Design (ICCD)*, Oct. 2020, pp. 296–303.
[35] 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)*, 2021, pp. 919–924.
[36] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
[37] X. Meng and G. E. Karniadakis, "A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse pde problems," *Journal of Computational Physics*, vol. 401, p. 109020, 2020.
[38] J. Berg and K. Nyström, "A unified deep artificial neural network approach to partial differential equations in complex geometries," *Neurocomputing*, vol. 317, pp. 28 – 41, 2018.
[39] S. Choudhary, "Physics informed neural networks," Electronic Design Process Symposium Lecture, NVIDIA, 10 2019. [Online]. Available: http://edpsieee.ieeesiliconvalley.org/Papers/4-3_Sanjay_Choudhry_AI_Computational_Science%2
[40] Q. Zhou, K. Sun, K. Mohanram, and D. C. Sorensen, "Large power grid analysis using domain decomposition," in *Proc. Design, Automation and Test In Europe. (DATE)*. 3001 Leuven, Belgium, Belgium: European Design and Automation Association, 2006, pp. 27–32.