Extraction of IC Interconnect Parasitic Capacitance Based on Adaptive Incremental Learning

Ziwei Yu[1, 2], Shuai Yan[1], Xiaoyu Xu[1], Zhuoxiang Ren[1, 3]

1: *Institute of Electrical Engineering, Chinese Academy of Sciences*. Beijing, China
2: *University of Chinese Academy of Sciences*. Beijing, China
3: *GeePs, Sorbonne Université, Université Paris-Saclay, CentraleSupélec, CNRS*. Paris, France
yanshuai@mail.iee.ac.cn

***Abstract*—The extraction of interconnect parasitic capacitance is a critical issue for designing VLSI circuits. In this paper, we propose an effective method to build the deep neural network (DNN) model based on adaptive incremental learning (AIL) for parasitic capacitance extraction. This method is based on AIL algorithm, which makes the DNN continuously and incrementally learn and update the network. It can improve the efficiency and accuracy of DNN training based on a smaller number of samples, and the trained DNN model can be used for the rapid extraction of parasitic capacitance. The results show that the size of training dataset required by the DNN based on adaptive incremental learning is about 40% less than the traditional DNN with similar accuracy.**

I. INTRODUCTION

With the down scaling of process technologies, the interconnect wires in integrated circuits (IC) become smaller, closer to each other, and the integration density becomes higher. Therefore, the extraction of interconnect parasitic parameters is very important to guarantee the performance of IC [1]. Parasitic parameters consist of resistance, capacitance, and inductance. The extraction of parasitic capacitance is mainly concerned in this work. The proposed method can be readily applied to extraction of resistance and inductance as well.

Field solvers, based on numerical methods including the finite element method (FEM), the boundary element method (BEM) etc., are considered as "gold standard" in parasitic extraction (PEX). However, the computational cost of field solvers is large and thus can hardly meet the requirement of on-the-fly extraction in VLSI circuits. Therefore, the industry needs to establish pattern libraries based on field solvers for the fast parasitic extraction. Recently, the deep learning (DL) has become an important tool for establishing pattern libraries for fast extraction [2]. Kasai et al. extracted parasitic capacitance of three-dimensional interconnect wires based on multilayer perception (MLP) [3]. Yang et al. built a capacitance model based on convolutional neural network (CNN) to extract parasitic capacitance from two-dimensional structures of the whole chip [1]. Ma et al. compared the performance of a deep neural network (DNN) and a deep residual network (ResNet) on the establishment of pattern models for parasitic capacitance extraction [2]. These works are all based on a fixed training dataset. However, the size of the dataset required to train a model with an expected accuracy is unknown *a-priori*. The proper size of the dataset needs to be obtained by numerical experiments, which leads to the cost of extra computation resources.

In this work, we propose an interactive method for the data generation and training process of the parasitic capacitance prediction model, based on the adaptive incremental learning. With the proposed approach, a small batch of samples is generated by the field solver in each training round and the trained model is updated iteratively to meet the accuracy requirement, so that the computational cost can be reduced both in the generation of training data and in the process of training.

This work was partially supported by the National Natural Science Foundation of China under Grant 52277019.

II. PARASITIC CAPACITANCE EXTRACTION BASED ON ADAPTIVE INCREMENTAL LEARNING

A. Data Sampling and Acquisition

First, the geometric models of the patterned interconnects in IC are parameterized. The range of the variables are set in accordance with the IC fabrication process requirement. The geometric variables (such as the positions and widths of conductors) are sampled within the value range by Latin hypercube sampling (LHS). Then, simulation models are built for each sample. For capacitance extraction, the electrostatic equation:

$$\nabla \cdot \epsilon \nabla \phi = 0 \tag{1}$$

is solved by FEM, and the parasitic capacitances are then calculated. In (1), ϕ is the electric potential and ϵ is the electric permittivity. The samples of the geometric variables of the patterned interconnects and their corresponding parasitic capacitances work together as the dataset to train the neural network. The samples of the geometric variables correspond to the feature data, while the calculated parasitic capacitances correspond to the label data.

B. Deep Neural Network Based on Adaptive Incremental Learning

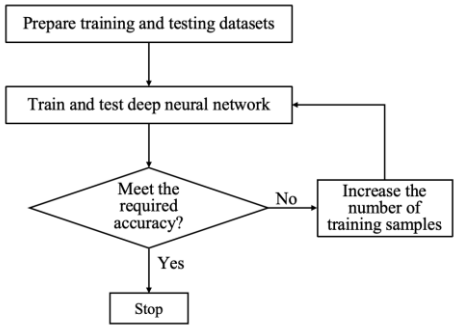

Fig. 1 Workflow of a traditional deep neural network

The traditional DNN based prediction model is trained with a given dataset. The procedure is shown in Fig. 1. The proper size of the dataset is unknown. Therefore, when the trained network fails to achieve the expected accuracy, a larger dataset is prepared and the network is re-trained. It

leads to extra computational cost in both data preparation and training. Incorporating with *a-posterior* error estimation, re-sampling and incremental learning, we propose an adaptive training process of the DNN based prediction model as shown in Fig. 2. The key implementation steps are as follows.

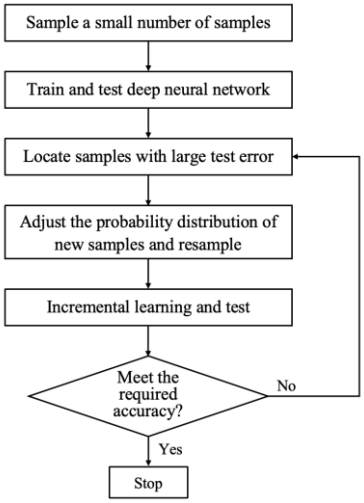

Fig. 2 Implementation process of the adaptive incremental learning

- Initial training: A small batch of samples is used to train an initial DNN. The hyper parameters are selected by hyperparameter optimization algorithms, such as the tree-structured parzen estimator (TPE). The best model is obtained after trained with a given number of epochs.
- Error estimation: Test the accuracy of the current trained model with a test set. If the expected accuracy is not achieved, locate the samples with relatively large error.
- Resampling: According to the located samples with relatively large error, adjust the probability distribution of LHS, so that the sampling is denser in the surrounding of the located samples. Then, resample a new batch of training data with a certain size. We can use the same size of the initial batch empirically.
- Model update with incrementally learning: The best model saved in the previous round of training is invoked to incrementally learn the new training data. Until the accuracy of the model meeting the expectation, the process of adaptive incremental learning is stopped.

III. EXAMPLE

The construction of a DNN-based parasitic capacitance prediction model of a two-dimensional interconnect pattern is studied as an example. The geometry of the pattern is shown in Figure 3. It consists of three metal layers and a substrate ground. The geometric parameters include the positions and widths of the conductors, denoted as $x_1, \dots, x_5$ and $w_1, \dots, w_5$. In this preliminary example, $x_1, x_2, x_3, x_4, x_5, w_2$ and w_4 are fixed, w_1, w_3 and w_5 are considered as parametric variables, and their ranges are set as [0.09, 0.9], [0.081, 2] and [0.09, 2] respectively. Thus, the inputs of the DNN-based parasitic capacitance prediction model are w_1, w_3 and w_5, and the outputs are the self-capacitance C_{11}, C_{22}, C_{33}, C_{44}, and C_{55} of the 5 conductors.

Firstly, the traditional DNN model is trained with different sizes of dataset. With each dataset, the hyperparameter is selected by the TPE method. The accuracy drops below 1% when the size of the dataset is increased to 1800. The average relative errors of prediction on the test dataset are plotted in the black line in Fig. 4.

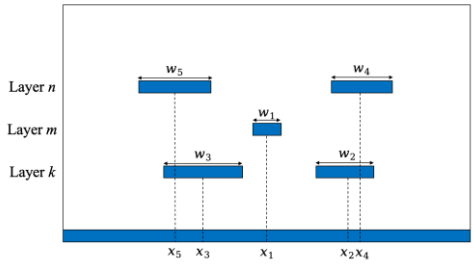

Fig. 3 Geometric structure of the example

Then, the proposed incremental learning procedure is used to train the model. The size of the initial batch is set as 200. Using the TPE method, the hyperparameters of the network with the best performance is chosen with the following parameters: the number of hidden layers is 1; each hidden layer has 300 neurons; the activation function is swish; the batch size is 16, and the learning rate is 0.1.

According to the workflow introduced in II. B, it takes 5 rounds of incremental learning to reach the expected accuracy. The average relative error of prediction on the test dataset are plotted in the red line in Fig. 4.

It can be observed that the proposed method achieve an average relative prediction error less than 1% with only about 40% of the dataset size required by the traditional process.

A comprehensive evaluation of the time cost in both training and data acquisition of the proposed method as well as its application on patterns with more geometric parameters will be presented in the full work.

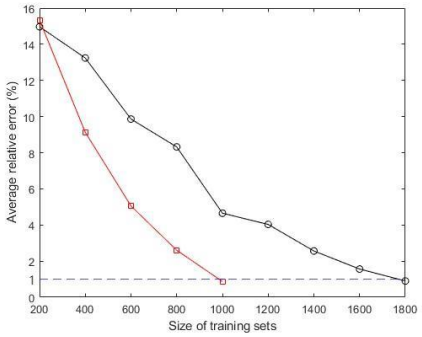

Fig. 4 Average relative errors of prediction with the traditional DNN model and the model trained by the incremental learning method

REFERENCES

[1] Yang D, Yu W, Guo Y, et al. CNN-Cap: Effective Convolutional Neural Network Based Capacitance Models for Full-Chip Parasitic Extraction[C]//2021 IEEE/ACM International Conference on Computer Aided Design (ICCAD). Munich, Germany, 2021: 1-9.

[2] Ma Y, Xu X, Yan S, Zhou Y, Zheng T, Ren Z, Chen L. Extraction of Interconnect Parasitic Capacitance Matrix Based on Deep Neural Network. *Electronics*. 2023; 12(6):1440.

[3] Kasai R, Kanamoto T, Imai M, et al. Neural network-based 3D IC interconnect capacitance extraction[C]//2019 2nd International Conference on Communication Engineering and Technology (ICCET). Nagoya, Japan, 2019: 168-172