Deep Neural Network for Generation of the Initial Electrostatic Potential Profile

Seung-Cheol Han and Sung-Min Hong
School of Electrical Engineering and Computer Science
Gwangju Institute of Science and Technology
Gwangju, Republic of Korea
Email: smhong@gist.ac.kr

***Abstract*—A deep neural network is trained to learn the electrostatic potential of the semiconductor device. In order to demonstrate its feasibility, pn diodes are considered. Various pn diodes with different doping densities are generated and the numerical solutions are calculated. The resultant electrostatic potential profiles are used in the training phase. Our numerical results clearly demonstrate that the trained neural network can provide the initial electrostatic potential reasonably well. Since the initial electrostatic potential is improved, the Newton-Raphson loop for the nonlinear Poisson equation can be converged within a smaller number of iterations.**

***Keywords*—Deep neural network; Initial electrostatic potential; Newton-Raphson method**

I. Introduction

Recently, a deep neural network has been widely applied to many application areas such as image analysis, natural language processing, and expert systems. For example, in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2017, 29 of 38 competing teams had accuracy greater than 95 %. Inspired by its superior performance demonstrated in those applications, great research efforts have been made to apply the deep neural network to other unexplored fields.

As far as the semiconductor industry is involved, a deep neural network has been mainly used as a tool for the design-technology co-optimization (DTCO) [1], [2]. In [2], the neural network is applied to the yield estimation. Basically, the neural network is considered as just a component of an optimizer for predicting a better technology option, based upon the pre-existing results from the semiconductor device simulator. Therefore, in those works, the semiconductor device simulator itself is not modified. It is treated as a given building block of the technology development.

Except for the DTCO application discussed so far, it is difficult to find a report on application of the deep neural network to the semiconductor device simulation. In [3], the machine learning technique has been used in predicting the Hamiltonian operators for the density-functional theory. However, the density-function theory is quite different from the conventional device simulation based on the drift-diffusion equation. Also, the adopted machine learning technique is not the deep neural network.

In the present authors' opinion, the deep neural network can be employed in the semiconductor device simulation field in an alternative manner. To be specific, it can be used to solve a set of nonlinear equations more efficiently. For example, it is expected that a trained deep neural network can be used to provide a good initial guess for the electrostatic potential. With a better initial solution, the number of the Newton-Raphson iterations required to solve the nonlinear Poisson equation can be reduced.

In this preliminary report, a deep neural network which can predict the electrostatic potential profiles of pn diodes is introduced. The structure of this extended abstract is as follows: In Section II, the neural network proposed in this work is briefly introduced. The numerical results for pn diodes at equilibrium are shown in Section III. It is clearly demonstrated that the number of the Newton-Raphson iterations can be considerably reduced by adopting the neural network. Moreover, a brief discussion on the future research direction is provided. Finally, the conclusion is made in Section IV.

II. Neural Network

The conceptual diagram for the proposed neural network is shown in Fig. 1. First, the neural network is trained as shown in Fig. 1(a). The training data set contains a list of the device specifications and the resultant electrostatic potential profiles. In this preliminary work, the device specifications are simplified as two scalar numbers, which represent the doping densities of a pn diode. It is prepared before starting the training phase. The supervised learning [4] is performed with the backpropagation algorithm. In addition to the training data set, the validation data set is prepared. Since the neural network is not trained with the validation data set, it is used to check the validity of the trained neural network. After trained with the predefined training data set, the neural network can be used to generate the initial potential profile in the semiconductor device as shown in Fig. 1(b).

The configuration of the proposed neural network is shown in Fig. 2. The doping densities of two regions (p-type and n-type regions) are used as the input parameters. Our goal is to generate the electrostatic potential profile as results of the last, output layer. For that purpose, two hidden layers have been introduced. Each hidden layer is fully connected

978-1-7281-0940-4/19/$31.00

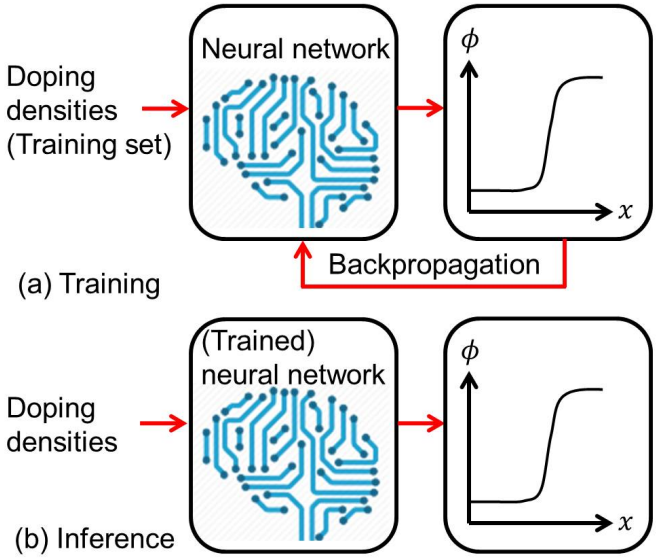

Fig. 1. Conceptual diagram for the proposed neural network. (a) Training and (b) inference phases.

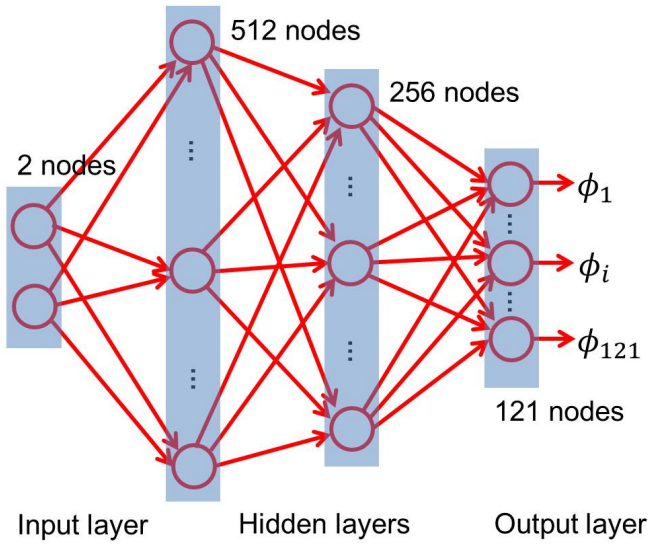

Fig. 2. Layer structure of the proposed neural network. Two hidden layers are introduced. Each hidden layer is fully connected with its neighboring layers.

with its neighboring layers. The first layer consists of 512 artificial neurons and uses the rectified linear unit (ReLU) as the activation function,

$$ReLU(x) = max(0, x). \quad (1)$$

The dropout parameter is set to be 0.2. The second layer has the same structure with the first one, except for the reduced number of artificial neurons, 256. The last, output layer has 121 artificial neurons. No activation function is used in the output layer.

It is noted that the layer structure shown in Fig. 2 has been obtained by performing several numerical experiments with various layer structures. Depending on the problem, other layer structures may exhibit better results. Unfortunately, we have not found a general rule to build the best layer structure without intensive numerical experiments. It would be an interesting research topic to establish an efficient method to find a sufficiently good layer structure.

As far as the actual implementation is concerned, we have used the sequential model among the Keras layer models [5]. We use the RMSprop (with the learning rate of 0.00005) as the optimizer and the mean square error as the loss function. The number of epochs is 1000.

III. Numerical Results

One-dimensional, abrupt pn diodes at equilibrium are considered. The material is silicon and the temperature is 300 K. The doping density of each region varies from 1×10^{15} cm^{-3} to 1.3×10^{17} cm^{-3}. The p-type and n-type doping densities vary independently. By adopting the depletion approximation, the depletion width varies from 0.13 μm (both regions with 1.3×10^{17} cm^{-3}) to 1.24 μm (both regions with 1×10^{15} cm^{-3}). A sufficiently long structure, whose length is 3 μm, is used for all the cases to ensure that the entire depletion region is included in the simulation domain. The real space is discretized with a uniform grid, whose spacing is 25 nm. The number of nodes in the output layer is matched to that of the grid points.

The difference between the predicted electrostatic potential and the numerical simulation result is regarded as an error. The numerical solution of the nonlinear Poisson equation is obtained by using our in-house tool and the intrinsic carrier density of Si at 300 K is set to be 1×10^{10} cm^{-3}. In Fig. 3, the mean absolute error of the electrostatic potential is shown as a function of the learning epoch. It is estimated after each learning epoch is finished. Throughout the training phase, the error is reduced for both the training data set and the validation data set. At the initial phase, the mean absolute error is rapidly decreasing with the epoch number. However, after a few hundred epochs, the error does not decrease any more. With the given layer structure shown in Fig. 2, it seems that the neural network is sufficiently trained after a few hundred epochs.

After the training phase is finished, the neural network can be used to predict the electrostatic potential profile. Fig. 4 shows the electrostatic potential profile generated by the trained neural network. Various structures with different doping densities are tested. The test structures shown in Fig. 4 are not included in the original training set. The doping densities are randomly selected in a range of (1×10^{15} cm^{-3}, 1.3×10^{17} cm^{-3}). Nevertheless, the generated profile agrees with the numerical simulation result reasonably well for each test structure. Boundary values at both ends are well reproduced and the depletion layers are accurately predicted.

Of course, it should be noted that the profile generated by the deep neural network cannot be perfectly matched to the numerical simulation result. In Fig. 4, slight differences between the symbols (the predicted profiles) and the solid lines

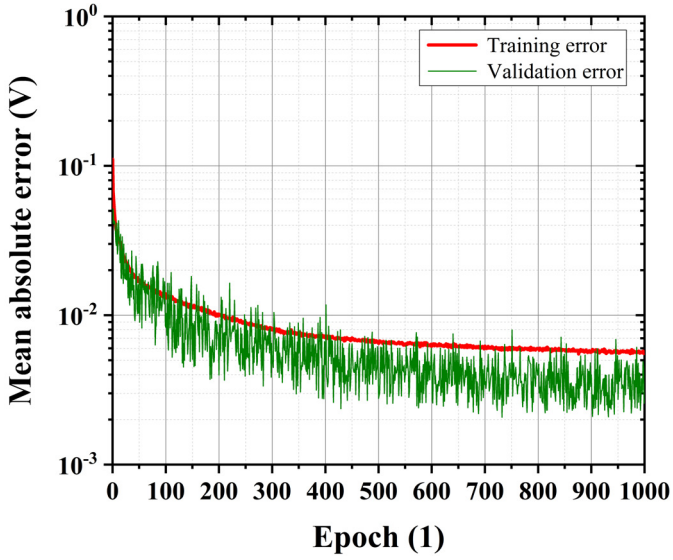

Fig. 3. Training and validation errors as functions of the learning epoch. The errors are rapidly decreasing with the epoch number.

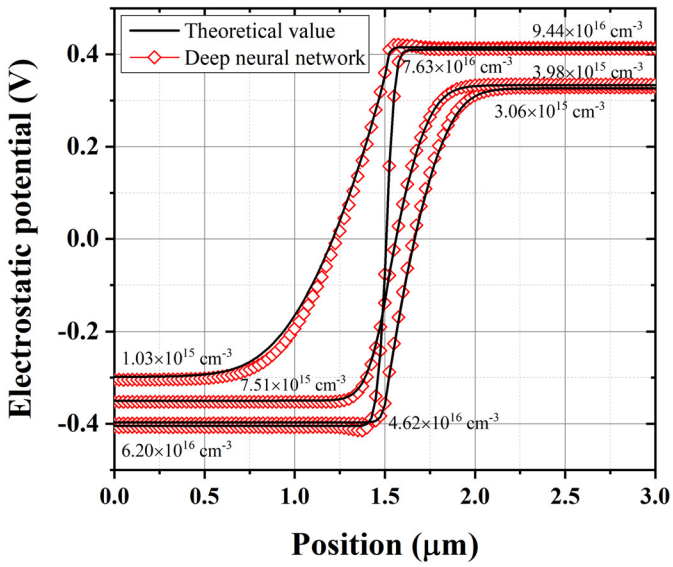

Fig. 4. Electrostatic potential profile generated by the trained neural network. Various structures with different doping densities are tested. The n-type region is located at the right side. For comparison, the numerical solutions are also drawn as solid lines.

(the numerical solutions) are visible. Therefore, the predicted electrostatic profile cannot be treated as the converged solution.

In order to understand the internal procedure for the neural network to predict the electrostatic potential profile, results of three layers are shown in Fig. 5. Two symmetric pn diodes are selected as representative examples. One has a relatively high doping density of 1.3×10^{17} cm^{-3} and the other has a low doping density of 1×10^{15} cm^{-3}. It can be observed that the high doping concentration excites the nodes strongly. Although it is difficult to interpret the physical meaning of

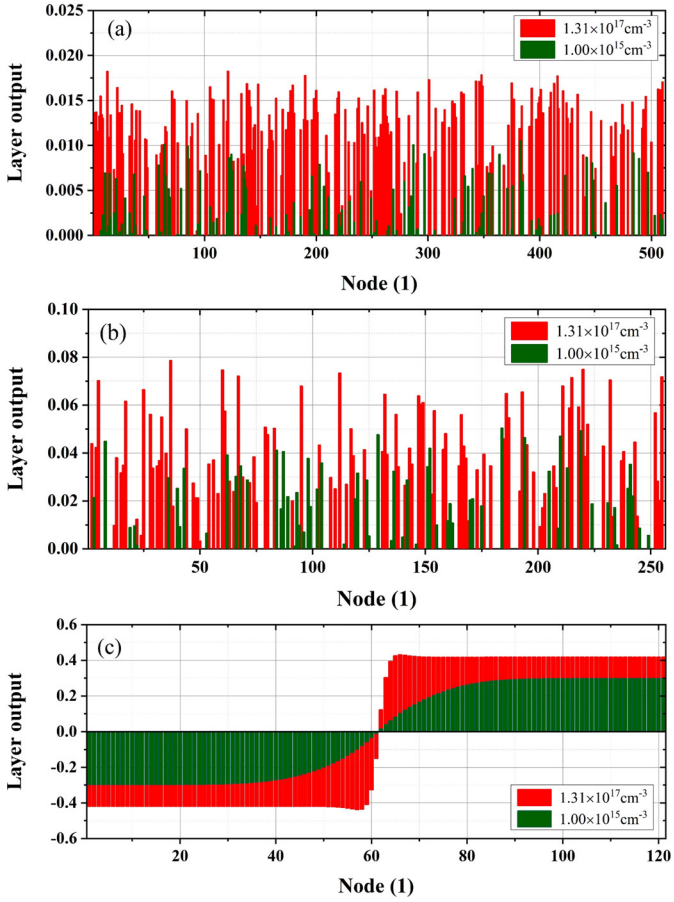

Fig. 5. Results of three layers. (a) The first hidden layer, (b) the second hidden layer, and (c) the output layer for two different symmetric pn diodes.

node outputs in the first and second layers (Figs. 5(a) and 5(b)), the output layer gives appropriate potential profiles for both structures as shown in Fig. 5(c). The sensitivity analysis on the trained neural network would be an interesting research topic.

Up to now, it has been demonstrated that the trained neural network can be used to guess the electrostatic potential profile at equilibrium. The predicted electrostatic potential can be used as the initial solution for the nonlinear Poisson equation. Since the initial guess for the electrostatic potential is quite similar to the numerical solution, it is expected that the convergence behavior can be improved significantly.

Fig. 6 shows the convergence behavior of the nonlinear Poisson equation. Several structures with randomly selected doping densities are simulated. The maximum potential correction in the first Newton-Raphson iteration is much smaller than 0.1 V. Since the initial error is small, the converged solution can be obtained quickly. For example, in every case simulated in Fig. 6, the maximum potential correction in the fourth Newton-Raphson iteration is smaller than 10^{-5} V. Compared with the conventional method assuming the local charge balance, the convergence acceleration by the deep neural network is obvious. In order to obtain the maximum potential correction

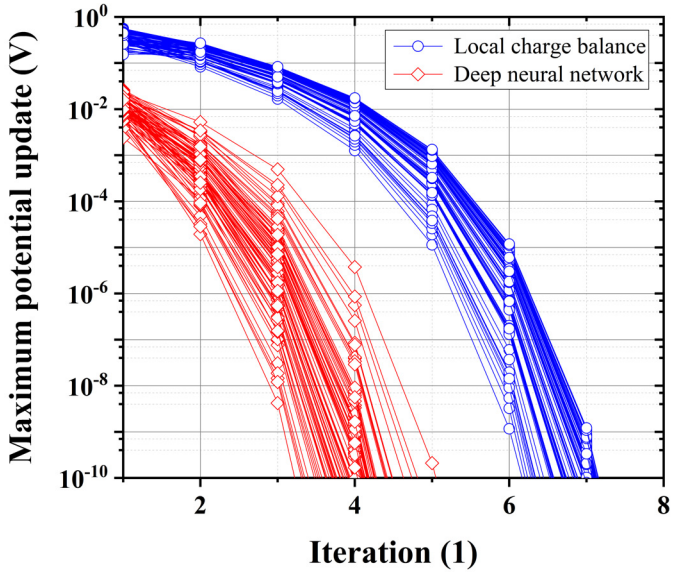

Fig. 6. Convergence behaviour of the equilibrium nonlinear Poisson equation. Two methods for the initial potential profiles (the local charge balance and the generated value by the neural network) are compared for several test structures.

smaller than 10^{-5} V, six or seven Newton-Raphson iterations are needed when the local charge balance approximation is adopted. Therefore, at least two or three iterations can be skipped without sacrificing the numerical accuracy. Since the neural network is trained only in the training phase, the computational burden in the inference phase is negligible.

It is noted that skipping two or three Newton-Raphson iterations for the nonlinear Poisson equation does not save the computation time significantly. However, the present approach has great potential for realistic situations. Although we have considered only the equilibrium case in this feasibility study, it can be extended to non-equilibrium cases [6]. By training the deep neural network with the simulated electrostatic potential profiles even at non-equilibrium cases, the neural network can provide an appropriate potential profile for a biased pn diode. With the predicted potential profile under the given bias point, the numerical simulation can be performed directly. On the other hand, in the conventional device simulation, several intermediate bias points should be simulated to cope with the system nonlinearity.

Therefore, the speed-up ratio for the Newton-Raphson method can be expressed as follows:

$$SpeedUp \approx \frac{N_{Bias}^{Conv} \times N_{Newton}^{Conv}}{N_{Newton}^{NN}}, \tag{2}$$

where N_{Bias}^{Conv} is the number of bias points to be simulated and N_{Newton}^{Conv} is the average number of the Newton-Raphson iterations per bias point. These quantities are for the conventional bias ramping scheme. On the other hand, N_{Newton}^{NN} is the number of the Newton-Raphson iterations when the predicted profile is used as the initial guess. Two numbers, N_{Newton}^{Conv} and N_{Newton}^{NN}, may be comparable. In such a case, the simulation speed-up is mainly achieved by skipping intermediate bias points. Further results will be reported elsewhere [6].

IV. Conclusion

In conclusion, the deep neural network can generate the initial electrostatic potential profile reasonably well. Compared with the local charge balance approximation, the number of the Newton-Raphson iterations required to get the converged solution is considerably reduced. It is expected that this work can be expanded to non-equilibrium cases with additional efforts.

Acknowledgment

This work was supported by GIST Research Institute (GRI) grant funded by the GIST in 2019.

References

[1] F. Benistant, "Future of TCAD: A foundry perspective," International Conference on Simulation of Semiconductor Processes and Devices, 2018.
[2] U. Kwon, T. Okagaki, S. Ahn, J. Shin, A.-y. Kim, Y.-s. Song, S. Kim, J. Kim, D. S. Kim, W. Qi, Y. Lu, H.-H. Park, W. Choi, "Intelligent DTCO (iDTCO) for next generation logic path-finding," International Conference on Simulation of Semiconductor Processes and Devices, pp. 45–48, 2018.
[3] G. Hegde and R. C. Bowen, "Machine-learned approximations to density-function theory Hamiltonians," Scientific Reports, vol. 7, p. 42669, 2017.
[4] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall, 2010.
[5] Keras library. [Online]. Available: https://keras.io
[6] S.-C. Han and S.-M. Hong, unpublished.