IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 3 December 2022, accepted 18 December 2022, date of publication 20 December 2022,
date of current version 29 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3230986

RESEARCH ARTICLE

A Deep-Learning Approach for Wideband Design of 3D TSV-Based Inductors

XIANGLIANG LI, PENG ZHAO, (Member, IEEE), SHICHANG CHEN, (Member, IEEE), KUIWEN XU, (Member, IEEE), AND GAOFENG WANG, (Senior Member, IEEE)
Engineering Research Center of Smart Microsensors and Microsystems, Ministry of Education, School of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018, China

Corresponding authors: Peng Zhao (pengzhao@hdu.edu.cn) and Gaofeng Wang (gaofeng@hdu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62141409, Grant 61971171, and Grant 61971174; and in part by the Zhejiang Provincial Key Research & Development Project under Grant 2021C01041.

ABSTRACT A high-efficient wideband through-silicon vias (TSVs) modeling method based on deep learning is proposed, and a compact three-dimensional (3D) spiral inductor is designed using the proposed method. By comparing different activation functions and loss functions, an adaptive deep neural network (DNN) based on Gaussian Error Linear Unit (GELU) and Huber functions for constructing parameterized TSV models is proposed. The model has much higher accuracy and better robustness than commonly used circuit equivalent models over a wide range of bandwidths. Moreover, a compact 3D spiral inductor using ground TSV is designed based on the DNN model. This 3D inductor greatly reduces the inductor area compared to planar inductors and has weak crosstalk between TSV pairs. The designed inductor is simulated by direct electromagnetic calculation to verify the proposed method and design.

INDEX TERMS Deep learning, equivalent-circuit model, through-silicon vias (TSVs), 3D spiral inductor.

I. INTRODUCTION

In three-dimensional integrated circuit (3D-IC), through-silicon via (TSV) is the key technology to realize the interconnection of upper and lower chips in 3D-ICs, which can greatly shorten the length of the interconnects between chips, reduce the occupied area, and achieve high performance and high density system on chips (SOCs) [1]. Moreover, by making use of TSV, three-dimensional (3D) inductor devices with higher inductance density and smaller occupancy area can also be realized. However, for the structure containing multiple TSVs, the crosstalk between adjacent TSVs seriously affects the signal integrity of 3D-ICs [2], [3], [4].

To suppress the TSV crosstalk, the methods of using a protection ring [5], coaxial TSV structure [6], active substrate coupling [7], coding algorithm [8] were proposed. In another hand, the existence of TSV makes the electromagnetic (EM) simulation more complex and time consuming. In this circumstance, not only the metal layer but also the dielectric material enclosed the TSVs should be meshed, even the layered medium Green's function is used. This is because oxide material isolating the metal and substrate has finite area in horizontal plane, which cannot be characterized using the planar Green's function. Therefore, some modeling methods of TSVs have been proposed to avoid the EM simulation of TSVs. The modeling of TSVs has been mainly focused on establishing RLCG equivalent circuit models by extracting relevant parasitic parameters through analytical formulas [9], [10], [11], [12], [13], [14]. In [9], the resistance, inductance, and capacitance parameters of the TSV were modeled. In [10], a modeling method based on integral equations to solve Maxwell's system of equations was proposed, but the solution was complex, and the effective applicable bandwidth was only 10 GHz. Moreover, the applicable frequency bands for the equivalent models proposed in [11] and [12] were only 20 GHz and 40 GHz, respectively. In [13], the impedances of the differential TSVs were extracted using the partial element equivalent circuit (PEEC) method, while the admittances were calculated analytically. Moreover, Kim proposed a wideband equivalent circuit model with consideration of the effective substrate current loop [14], but only adapted to round-signal type TSVs. The 3D EM simulation of TSVs is time consuming. Especially for TSV optimization, the EM simulation will be repeatedly used. Although the

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

equivalent circuit model of TSV can replace the EM simulation, accurate model is difficult to build for the complex structure. Therefore, a fast and accurate model that can efficiently predict the EM characteristics of the TSVs and achieve the optimization design is very necessary.

In this work, a method for modeling TSVs using deep neural network (DNN) is proposed. In the process of modeling, by comparing the effects of different activation functions and loss functions, the Gaussian Error Linear Unit (GELU) activation function and Huber loss function are found that have better performance than the traditional Rectified Linear Unit (ReLU) [15] activation function and mean squired error (MSE) loss [16] function for TSV modeling. It is also found that the proposed method has higher computational accuracy and can meet the accuracy requirements in a wide range of frequency bands comparing with the equivalent circuit TSV models [17]. In addition, using this accurate and highly efficient method, a compact TSV-based spiral inductor is designed, which has a smaller area than the conventional on-chip inductor and can be further stacked by multilayer chips to obtain a smaller footprint. The designed inductor is verified using full-wave simulations, which demonstrate the accuracy of the proposed model and the miniaturization of the inductor design.

II. TRAIN THE DNN-TSVs MODEL

A. DEEP NEURAL NETWORK MODEL

FIGURE 1 shows the flow chart of DNN-TSV training model. After building the parameterized TSV structures, some typical TSVs with the sampling parameters are simulated by EM solver. The geometric parameters of the TSV model and the corresponding S-parameters are used as the input dataset of the DNN-TSV model. Before training, all input variables are normalized to $[-1, 1]$, and the datasets are divided into training data and testing data according to the ratio of 7:3.

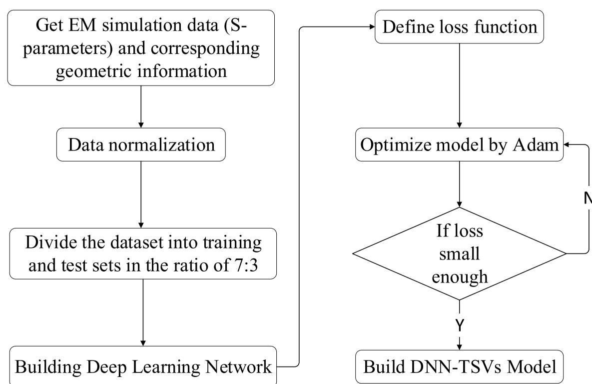

FIGURE 1. The flow chart of training DNN-TSVs model.

In the DNN, as the basic components of neural networks, a neuron is mathematically expressed as

$$Y = \varphi\left(\sum_{i}^{n} a_i X_i + b\right) \tag{1}$$

where X_i denotes the i-th element of the input matrix, a_i represents the weight factor of the i-th element, b is the offset, and φ represents the activation function, which can introduce nonlinear factors into neurons and be used to fit various nonlinear models. In practical applications, the ReLU activation function is the most used activation function. However, it is found that for the training of TSV S-parameters, the performance using the ReLU function is not good.

As the recently proposed activation functions, the Swish and the GELU activation functions have some advantages in modeling [18]. The Swish activation function has the properties of no upper bound and lower bound, smooth and non-monotonic, and all of which can improve the performance of the activation function. The mathematical expression of Swish activation function can be expressed as:

$$Swish\,(x) = x \cdot sigmoid\,(x) = x \cdot \frac{1}{1+e^{-x}} \tag{2}$$

Moreover, when a deterministic nonlinear activation function such as ReLU is used as the activation function of a neural network, stochastic regularization needs to be added to the network to improve the generalization ability of the model, but the stochastic regularization and the activation function are separated, making a network look like a patchwork of sub-networks [19]. Unlike the ReLU function, the GELU function, which combines properties of the ReLU function and dropout as well as adds stochastic regularization to the activation function, can provide both nonlinearity and stochastic regularity to the neural network. The GELU activation function can be expressed as

$$GELU = x \cdot P\,(X \leq x) = x \cdot \Phi\,(x) \tag{3}$$

where $\Phi\,(x)$ denotes the cumulative function of the Gaussian normal distribution of x. Since the above equation is not directly computable, it can be approximated as

$$GELU\,(x) = 0.5x\left(1 + tanh\left[\sqrt{\frac{2}{\pi}}\left(x + 0.044715x^3\right)\right]\right) \tag{4}$$

In this work, the ReLU, Swish and GELU activation functions are all used for the training of DNN-TSV model, and the performance of these functions will be compared in section III.

After choosing the activation function, a deep neural network model is made up of multiple layers of neurons connected in pairs between adjacent layers. A complete neural network consists of one input layer, several intermediate layers and one output layer. The mathematical expression of a complete neural network is shown in (5).

$$Y = \varphi\left(A^n \cdots \varphi\left(A^2\left(\varphi\left(A^1 X + b^1\right)\right) + b^2\right) \cdots + b^n\right) \tag{5}$$

where X denotes the geometric information of the TSV model (inputs to the neural network) and Y denotes the S-parameters of the TSV (outputs to the neural network).

B. LOSS FUNCTION AND OPTIMIZATION ALGORITHM

Owing to the gradient descent convergence, the MSE loss function is widely used in general regression problems [20]. However, for the DNN-TSV modeling, due to the existence of a certain number of outlier samples, the MSE squares the error and makes the error of the outliers too large, which resulting in the slow decrease of test loss of training. As a consequence, it is difficult or even impossible to reach the desired loss threshold.

Huber loss combines the advantages of MSE and mean absolute error (MAE) [21]. It enhances the outlier robustness of MSE, reduces the sensitivity problem to outliers, and compensates the slow decreasing speed of MAE loss. Therefore, the Huber loss function is used in the DNN-TSV modeling. Specifically, the Huber loss is defined as follows:

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y-f(x))^2, & \text{if } |y-f(x)| \le \delta \\ \delta|y-f(x)| - \frac{1}{2}\delta^2, & \text{if } |y-f(x)| > \delta \end{cases} \tag{6}$$

where y is the true value, $f(x)$ is the predicted value, and δ is a hyperparameter of Huber loss which is set to 1 in this paper. In training the DNN-TSV model, y represents the real S-parameters in the dataset, and $f(x)$ represents the S-parameters obtained from the DNN-TSV model. By setting δ, the loss function is derivable, and the gradient is more stable in the interval $[-\delta, \delta]$, and the influence of outliers on the model is attenuated in the interval $[-\infty, -\delta]$ and $[\delta, +\infty]$. Therefore, the robustness of the loss function is improved.

Next, the loss function is minimized using an optimization algorithm, and the loss function is used as the objective function for optimization in the optimization process, and the training error of the loss function is reduced by continuously adjusting the weight factors and offset in the neural network. Considering the advantage of adaptive moment estimation (Adam) [22] algorithm that it can design independent adaptive learning rates for different parameters by computing the first- and second-order moment estimates of gradients, the Adam function is selected as the optimization function. The mathematical expression of Adam function is as follows:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\cdot g_t \tag{7}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)\cdot g_t^2 \tag{8}$$

where m_t and v_t are the first- and second-order moment estimators of gradient, respectively, g_t is the gradient of the loss function after the t step iterations, and β_1 and β_2 are constant delay factors, which are usually set at 0.9 and 0.999, respectively.

After bias correction, the first- and second-order moment estimators are respectively expressed as

$$m_t' = \frac{m_t}{1-\beta_1^t} \tag{9}$$

$$v_t' = \frac{v_t}{1-\beta_2^t} \tag{10}$$

$$\theta_{t+1} = \theta_t - \frac{\mu}{\sqrt{v_t'} + \varepsilon} m_t' \tag{11}$$

where θ_{t+1} is the updated parameter of θ_t, and μ and ε are two constants, which are usually set as 0.001 and 1, respectively.

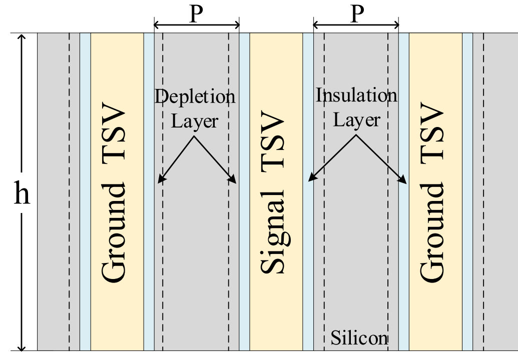

FIGURE 2. Schematic of GSG-type TSVs.

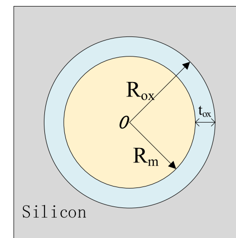

FIGURE 3. Top view of the TSV.

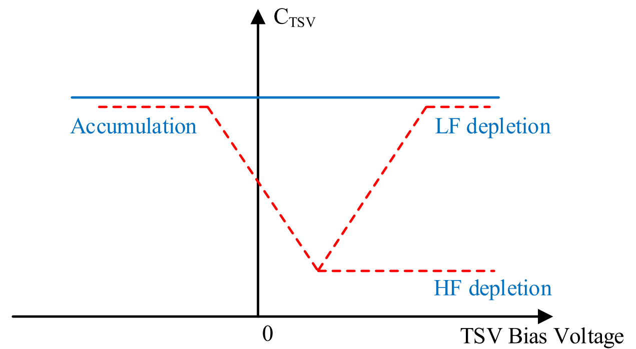

FIGURE 4. The total capacitance of TSV varies with the DC bias voltage.

Since the TSV structure contains a metal-oxide-semiconductor (MOS) structure, the silicon doping forms a depletion layer with the voltage applied to the TSV, as shown in FIGURE 2. The MOS structure yields a nonlinear voltage-dependent capacitive effect due to the formation of the depletion layer. For a P-type substrate, as shown in FIGURE 4, when large negative bias voltage is applied to the TSV, most of the carriers will accumulate on the TSV surface, and the total TSV capacitance approaches the oxide capacitance.

As the negative voltage decreases, a depletion layer is formed and thus the depletion layer capacitance, which is equivalent to the series connection of the oxide and depletion layer capacitances, so the total TSV capacitance decreases. Continuing to increase the voltage, the capacitance will pass through a minimum point at low frequencies (below 100 Hz) and then increases again. Due to the formation of a minority carrier inversion layer on the TSV surface, and the capacitance eventually remains stable at high frequencies as the voltage increases. In [23], the C-V characteristics of the TSV at low and high frequencies are discussed in detail. In [24], the MOS capacitance was modeled according to the size and bias voltage of the TSV. Both the DNN-TSV modeling method proposed in this paper and the TSV-based inductor designed later in Section IV are used without DC bias voltage, so the MOS effect of TSV is weak and can be not considered.

The performance of various combinations of aforementioned activation and loss functions in training ground-signal-ground (GSG) typed DNN-TSV models (as shown in FIGURE 2) is compared, and the final test loss achieved are given. In this paper, the final test loss is defined as when training the DNN-TSV model, the number of hidden layers of the neural network starts from 3, and the loss value is calculated after every 10,000 iterations using the test set according to the definition of the loss function. And then, the hidden layers are added and retrained until the number of hidden layers increases to 7, and the final minimum loss value is taken as the final test loss.

In FIGURE 2, there are three TSVs, the signal TSV is as the signal transmission channel of upper and lower chips, and the ground TSV provides a reference ground for signal transmission. The top view of each TSV is shown in FIGURE 3. In Figs. 2 and 3, p is the distance between two adjacent TSVs, h is the height of the TSV, R_m is the radius of the metal layer in the middle of the TSV, t_{ox} is the thickness of the oxide layer in the outer TSV, and R_{ox} is the radius of the TSV. The parameters of metal layer and dielectric layer are $\sigma_{si} = 10\,S/m$, $\rho_{tsv} = 1.68 \times 10^{-8}\Omega \cdot \mathrm{m}$, $\mu_r = 1$, $\varepsilon_{r,si} = 11.9$, and $\varepsilon_{r,Insulator} = 4$. The sweeping range of geometric parameters are shown in Table 1.

TABLE 1. Sweep parameters.

Input Parameters	Starting Value	End Value	Step
h (μm)	10	50	5
p (μm)	5	20	5
R_m (μm)	2	5	1
t_{ox} (μm)	0.05	0.5	0.1
Frequency (GHz)	0	100	1

For multi-parameter modeling of DNN-TSVs, e.g., when h, p, R_m and t_{ox} all vary as the value shown in Table 1, the performance of various function combinations is tested and shown in Table 2. For the modeling, the number of hidden layers is set to 7, the number of neurons in each hidden layer is set to 50, the hyperparameter δ in the Huber loss

TABLE 2. Train GSG type DNN-TSVs model with TSV height, spacing, metal layer radius, oxide layer thickness as input variables (the number of datasets = 864, the number of hidden layers = 7, the number of neurons in each hidden layer = 50, the hyperparameter δ in the Huber loss function = 1, the learning rate = 0.001).

Test Loss	MSE	Huber
ReLU	1.91×10^{-6}	7.14×10^{-7}
GELU	4.1×10^{-7}	$\mathbf{2.95 \times 10^{-7}}$
Swish	9.91×10^{-7}	7.17×10^{-7}

TABLE 3. Train GSSG differential type DNN-TSVs model with TSV height, spacing, metal layer radius, oxide layer thickness as input variables (the number of datasets = 864, the number of hidden layers = 7, the number of neurons in each hidden layer = 50, the hyperparameter δ in the Huber loss function = 1, the learning rate = 0.001).

Test Loss	MSE	Huber
ReLU	4.1×10^{-6}	1.24×10^{-6}
GELU	1.23×10^{-6}	$\mathbf{3.34 \times 10^{-7}}$
Swish	9.59×10^{-7}	7.83×10^{-7}

function is set to 1, and the learning rate is set to 0.001. As can be seen from this table, the final test loss using the Huber loss function is smaller than that using the MSE loss function. Moreover, the final test loss using the GELU activation function is smaller than that using the ReLU and Swish activation functions. Among all the combinations, the final test loss using the combination of GELU activation function and Huber loss function is the smallest.

In addition, the S-parameters generated by the DNN-TSV model depending on different activation functions and Huber loss function are compared to demonstrate the actual effect of the activation functions. The S-parameters of the GSG-type TSV with geometric parameters of $h = 42\mu m$, $p = 6\ \mu m$, $R_m = 2.5\ \mu m$, $t_{ox} = 0.23\mu m$ are shown in FIGURE 5. It can be observed that the S-parameters generated by the DNN-TSV model with GELU activation function are the closest one to the results of the EM simulation. And the combination of GELU activation function and Huber loss function is chosen for modeling of DNN-TSV.

For different TSV structures, the modeling method can be still adopted. FIGURE 6 shows the configuration of a ground-signal-single-ground GSSG-typed differential-TSVs (D-TSVs). The performance of various function combinations is tested and shown in

Table 3. The combination of the GELU activation function and the Huber loss function is also the optimal combination.

III. VALIDATION OF DNN-TSV MODELS

The EM simulation results are used to verify the results obtained by the DNN-TSV models. The structures of GSG- and GSSG-typed TSVs are calculated, and the final S-parameters are compared for different methods. Moreover, after building the DNN-TSV model, the EM computation of the S-parameters are replaced by the DNN model,

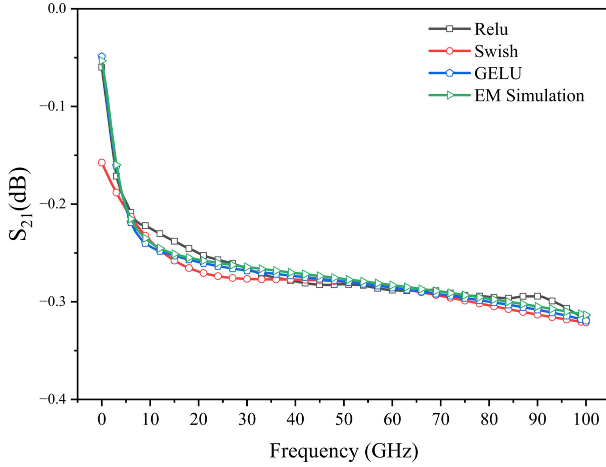

FIGURE 5. Comparison of S-parameters (S_{21}) generated by DNN-TSV models constructed from three activation functions respectively.

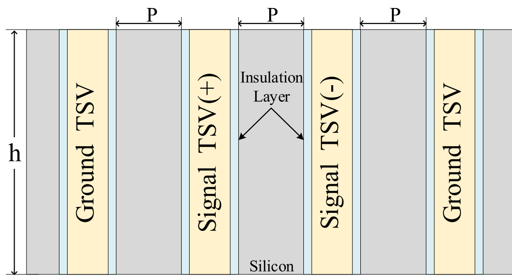

FIGURE 6. Schematic of GSSG-type differential TSVs.

so the efficiency is very high. For calculating the GSSG-type TSV structure with the same geometric parameters, the S-parameters predicted by the DNN-TSV model require only about 3 seconds, and occupy only 1.4 MB of memory, while the EM simulation will take about 474 seconds, and occupies 2.28 GB. In addition, although efficiency of applying the circuit model of TSVs is comparable to the DNN model, the accuracy cannot be guaranteed, so in the following, the accuracy of DNN-TSV model will be verifies.

Three methods are compared, i.e., the EM simulation, the DNN-TSV model and the equivalent circuit model based on [17]. The equivalent circuit model is shown in FIGURE 7, and the equivalent circuit models for both GSG-type and GSSG-type are given. FIGURE 8 and FIGURE 9 show S-parameters results for the GSG-type TSV. The geometric parameters of GSG-type TSVs are $h = 44$ μm, $p = 5.5$ μm, $R_m = 2.5$ μm, $t_{ox} = 0.1$ μm. The parameters of the circuit model are shown in the second column of Table 4, the DC resistance of TSV is listed in Table 4, and the AC resistance varies with frequency and is calculated according to Equation (12) in [17]. It can be seen from the figure that the deviation between the DNN-TSV model and the EM simulation is much smaller than the deviation between the equivalent circuit model and the EM simulation. FIGURE 10 and FIGURE 11 show S-parameters results for the differential GSSG-type TSV, the geometric parameters are set as $h = 42$ μm, $p = 5.2$ μm, $R_m = 2.5$ μm, $t_{ox} = 0.18$ μm.

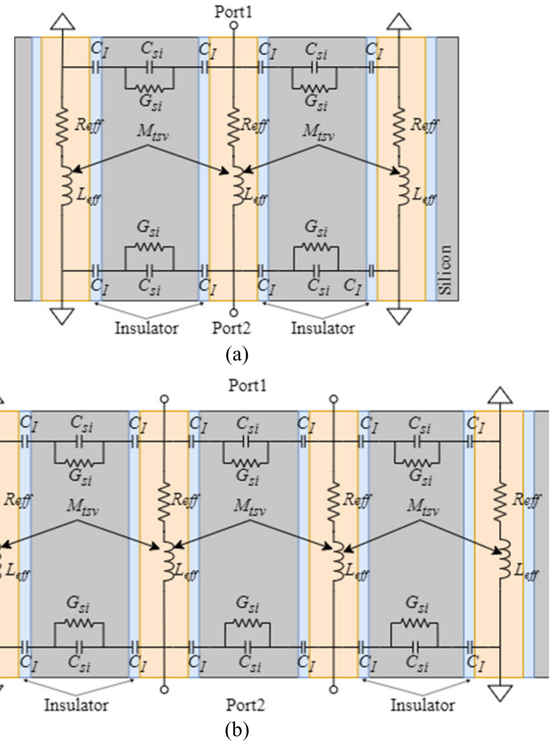

FIGURE 7. Equivalent circuit model base on [17], where (a) denotes the equivalent circuit model for GSG-type and (b) denotes the equivalent circuit model for differential GSSG-type.

TABLE 4. Circuit model parameters for GSG-type and differential GSSG-type.

Parameters of the circuit model	GSG-type	Differential GSSG-type
C_{si} (F)	5.45991×10^{-15}	5.34899×10^{-15}
C_I (F)	8.07537×10^{-14}	6.72141×10^{-14}
G_{si} (Ω^{-1})	0.000518191	0.000507664
M_{tsv} (H)	1.74914×10^{-11}	1.92919×10^{-11}
R_{dc} (Ω)	0.022978	0.0359359
L_{eff} (H)	4.0326×10^{-11}	4.11915×10^{-11}

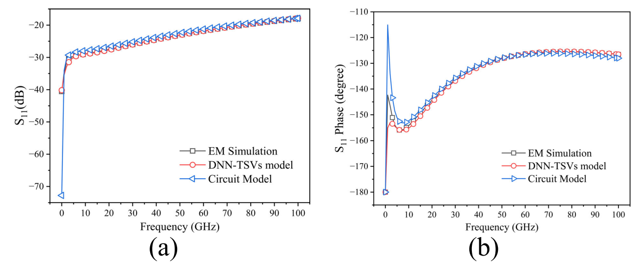

FIGURE 8. Comparison of S_{11} for GSG-type TSV. (a) S_{11} magnitude and (b) S_{11} phase.

The parameters of the circuit model are shown in the third column of Table 4, as above only the DC resistance of TSV is listed in Table 4. Similarly, the DNN-TSV model is also more accurate than the equivalent circuit method for the GSSG case.

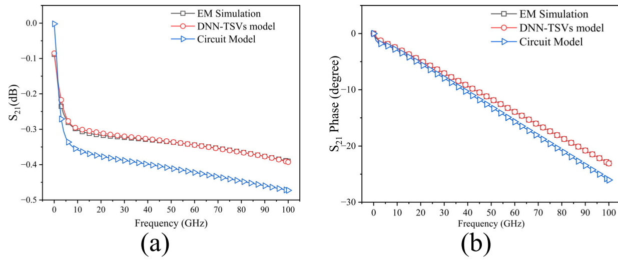

FIGURE 9. Comparison of S_{21} for GSG-type TSV. (a) S_{21} magnitude and (b) S_{21} phase.

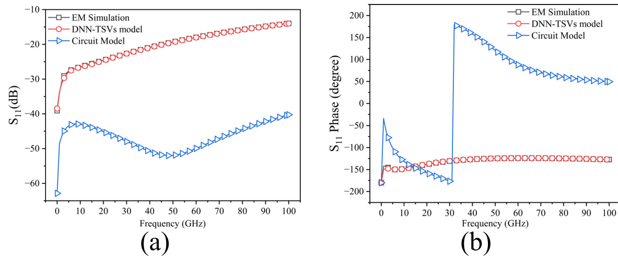

FIGURE 10. Comparison of S_{11} for differential GSSG-type TSV. (a) S_{11} magnitude and (b) S_{11} phase.

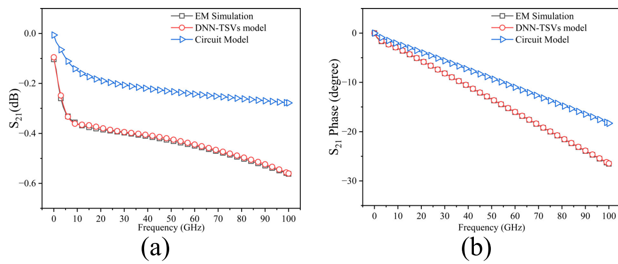

FIGURE 11. Comparison of S_{21} for differential GSSG-type TSV. (a) S_{21} magnitude and (b) S_{21} phase.

IV. COMPACT SPIRAL INDUCTOR BASED ON TSV

In this section, a compact spiral inductor based on TSVs is designed by virtue of the DNN-TSV models.

A. 3D SPIRAL INDUCTOR

The conventional 3D spiral inductor without ground TSV inserted between adjacent signal TSVs is shown in FIGURE 12. To suppress the crosstalk, a compact 3D spiral inductor with ground TSV is herein designed. To further increase the inductance density, the inductor is designed by stacking multilayer chips, as shown in FIGURE 13. Ground TSVs are inserted between signal TSVs to reduce crosstalk between signal TSVs and provide reference ground for signal TSVs.

The front view on the *yz* plane of this design is shown in FIGURE 14. The whole conductor structure consists of redistribution layers (RDLs), TSVs and bumps. The TSVs penetrate through a silicon substrate, with a central metal layer of copper and an oxide layer of SiO_2. The two TSV

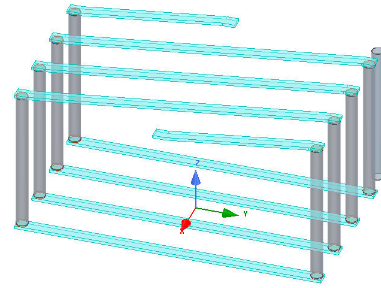

FIGURE 12. Schematic of the conventional 3D spiral inductor.

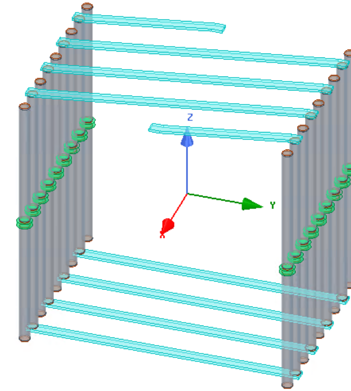

FIGURE 13. Schematic of the compact 3D spiral inductor.

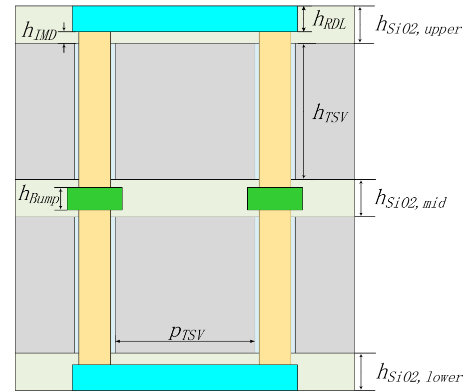

FIGURE 14. Front view along the yz plane.

layers are connected with copper bumps, with the top and bottom RDLs embedded in the SiO_2 layer.

The design parameters of the compact 3D spiral inductor are shown in Table 5, where p_{tsv} represents the spacing of TSVs on the same side and S_{tsv} denotes the spacing of TSVs on the other side.

B. S PARAMETER-CIRCUIT EQUIVALENCE MODEL BASED ON DNN-TSV

For fast simulation and rapid design of this inductor, the DNN-TSV model is used. The specific process is as follows: Firstly, the designed TSV array is split into multiple substructures. For example, the TSV array on the same side

TABLE 5. Design parameters.

Parameters	Value (μm)
R_m	2
t_{ox}	0.15
h_{TSV}	50
p_{TSV}	5
S_{tsv}	20
w_{RDL}	4.3
h_{RDL}	1
h_{Bump}	2
$h_{sio2,upper}$	1.5
$h_{sio2,mid}$	3
$h_{sio2,lower}$	1.5
h_{IMD}	0.5

of the inductor designed in this section can be split into multiple GSG substructures. Secondly, the DNN-TSV model proposed above is utilized to obtain the S-parameter of GSG substructure. Thirdly, the RDL layer is equivalently modeled as a circuit model by using resistance, inductance, and capacitance, in which the calculation formulas are expressed as [25]. Moreover, TSVs with the same direction of current flow contribute to positive mutual inductance, whereas the opposite direction current flow would contribute to negative mutual inductance [26]. Therefore, the total inductance can be improved by placing same-direction TSVs closer and opposite-direction TSVs farther away [27], and if the spacing between the TSVs is much greater than six times the radius of the TSVs can be considered a weakly coupled interconnection [28], so that the mutual inductance between TSVs along the y-direction can be ignored in this circumstance.

$$R_{RDL,Bump} = \frac{\rho l}{s} \tag{12}$$

$$L_{RDL} = \frac{\mu_0 l}{2\pi}\left[ln\left(\frac{l}{w}\right) + 0.5 + 0.447\frac{w}{l}\right] \tag{13}$$

$$L_{Bump} = \frac{1}{2}\left[\frac{\mu_0}{2\pi}h_{Bump}\ln\left(\frac{p_{tsv}}{r_{Bump}}\right)\right] \tag{14}$$

$$M_{RDL} = \frac{\mu_0 l}{2\pi}\left[ln\left(\frac{l}{p_{RDL}}\right) + \sqrt{1+\frac{l^2}{p_{RDL}^2}} + \frac{p_{RDL}}{l} + \sqrt{1+\frac{p_{RDL}^2}{l^2}}\right] \tag{15}$$

where ρ is the conductor resistivity, l is the length of RDL, s represents the cross-sectional area of RDL, w is the width of RDL, p_{RDL} is the distance between two adjacent RDL, p_{tsv} is the distance between two adjacent TSVs, h_{Bump} is the height of bump, and r_{Bump} denotes the radius of bump. Finally, the S parameters from the DNN-TSV model and the equivalent circuit of RDL layer are co-simulated.

The S parameter-equivalent circuit model of the designed TSV inductor is shown in FIGURE 15. The results of the EM simulations are provided as reference to verify the accuracy of this S-parameter-circuit equivalence model, and the

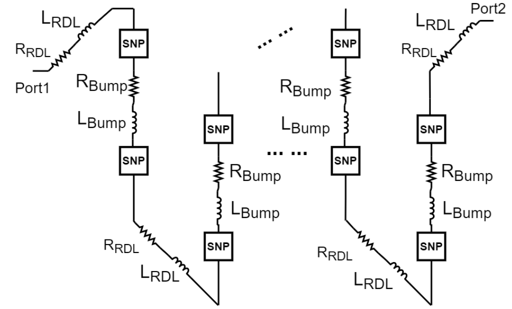

FIGURE 15. Schematic of S parameter-circuit equivalent model constructed based on the DNN-TSV model.

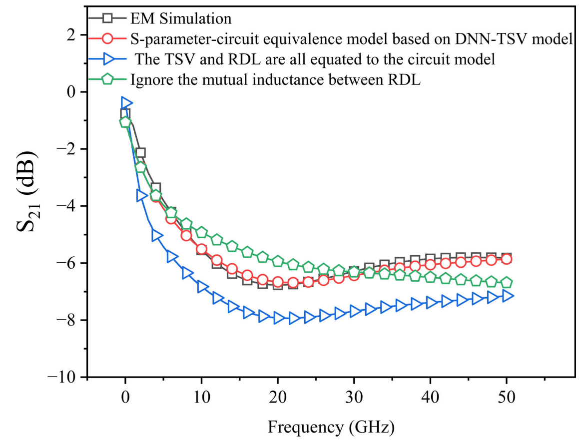

FIGURE 16. Validation of the S parameter-circuit equivalent model based on the DNN-TSV model using EM simulation.

TABLE 6. Efficiency comparison of DNN-TSV based S-parameter-circuit model and EM simulation.

Method	Time	Memory
This work	3 minutes	710MB
EM simulation	417 minutes	50GB

results using the equivalence circuit models of both TSV and RDL are also given for the comparison. From FIGURE 16, it is observed that the proposed method is much accurate than the equivalence circuit method, and its result is very closed to the EM simulation, and the test results show that ignoring the mutual inductance between RDLs would produce a large error.

In addition, because the computational requirements of the DNN-TSV model are small and the Bump and RDL in the inductor are equivalent to a simple RLC circuit model, the simulation efficiency of the inductor designed based on the DNN-TSV model is very high. A comparison of the efficiency of the DNN-TSV based design inductor and EM simulation is shown in Table 6.

The inductance values of the conventional 3D inductor and the proposed inductor are compared in FIGURE 17. The compact 3D inductor designed in this work can further increase

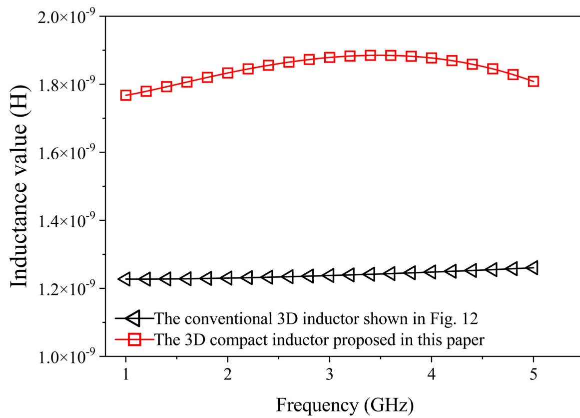

FIGURE 17. Comparison of inductance values of conventional 3D inductors and the compact inductors proposed in this paper.

TABLE 7. Comparison of inductance value, area and inductance density.

Reference	Inductance (nH)	Area (mm^2)	Inductance density (nH/mm^2)
This work	1.8	6.8×10^{-3}	264.7
[29]	2.2	0.0649	33.85
[30]	5	0.0489	102.12
[31]	53	25	2.12

the inductance value without increasing the occupied area compared to the conventional 3D spiral inductor. Moreover, the comparison of the inductance value, occupied area, and inductance density of this design with the on-chip inductors in [29], [30], and [31] is shown in Table 7. It can be seen that the 3D compact inductor designed in this work can greatly improve the inductance density.

V. CONCLUSION

In this work, an efficient deep-learning approach for wideband design of 3D TSV-Based Inductors was proposed. By comparing various activation functions and loss functions, the GELU activation function and the Huber loss function were employed to build robust DNN-TSVs model. Using this model, the frequency responses of various TSVs structures including GSG-type TSVs and GSSG-type D-TSVs can be accurately analyzed in a wide frequency band. Based on this model, a compact 3D spiral inductor with ground TSV has been designed, which can greatly reduce not only the occupied area by comparison with the conventional inductors but also the crosstalk between the signal TSV pairs.

REFERENCES

[1] A. Rahman and R. Reif, "System-level performance evaluation of three-dimensional integrated circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 6, pp. 671–678, Dec. 2000, doi: 10.1109/92.902261.

[2] C. Liu, T. Song, J. Cho, J. Kim, J. Kim, and S. K. Lim, "Full-chip TSV-to-TSV coupling analysis and optimization in 3D IC," in *Proc. 48th Design Autom. Conf. (DAC)*, 2011, pp. 783–788.

[3] T. Song, C. Liu, Y. Peng, and S. K. Lim, "Full-chip signal integrity analysis and optimization of 3-D ICs," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 5, pp. 1636–1648, May 2016, doi: 10.1109/TVLSI.2015.2471098.

[4] C. Duan and B. LaMeres, *On and Off-Chip Crosstalk Avoidance in VLSI Design*. New York, NY, USA: Springer, 2010.

[5] J. Cho, "Modeling and analysis of through-silicon via (TSV) noise coupling and suppression using a guard ring," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 2, pp. 220–233, Feb. 2011, doi: 10.1109/TCPMT.2010.2101892.

[6] K. Salah, "Performance comparison between air-gap based coaxial TSV and conventional circular TSV in 3D-ICs," in *Proc. 8th IEEE Design Test Symp.*, Dec. 2013, pp. 1–3, doi: 10.1109/IDT.2013.6727102.

[7] P.-J. Yang, P.-T. Huang, and W. Hwang, "Substrate noise suppression technique for power integrity of TSV 3D integration," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2012, pp. 3274–3277, doi: 10.1109/ISCAS.2012.6272024.

[8] P. M. Yaghini, A. Eghbal, M. Khayambashi, and N. Bagherzadeh, "Coupling mitigation in 3-D multiple-stacked devices," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 12, pp. 2931–2944, Dec. 2015, doi: 10.1109/TVLSI.2014.2379263.

[9] G. Katti, M. Stucchi, K. D. Meyer, and W. Dehaene, "Electrical modeling and characterization of through silicon via for three-dimensional ICs," *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 256–262, Jan. 2010, doi: 10.1109/TED.2009.2034508.

[10] K. J. Han, M. Swaminathan, and T. Bandyopadhyay, "Electromagnetic modeling of through-silicon via (TSV) interconnections using cylindrical modal basis functions," *IEEE Trans. Adv. Packag.*, vol. 33, no. 4, pp. 804–817, Nov. 2010, doi: 10.1109/TADVP.2010.2050769.

[11] J. Kim, "High-frequency scalable electrical model and analysis of a through silicon via (TSV)," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 2, pp. 181–195, Feb. 2011, doi: 10.1109/TCPMT.2010.2101890.

[12] T.-Y. Cheng, C.-D. Wang, Y.-P. Chiou, and T.-L. Wu, "A new model for through-silicon vias on 3-D IC using conformal mapping method," *IEEE Microw. Wireless Compon. Lett.*, vol. 22, no. 6, pp. 303–305, Jun. 2012, doi: 10.1109/LMWC.2012.2195776.

[13] W. S. Zhao, J. Zheng, F. Liang, K. Xu, X. Chen, and G. Wang, "Wideband modeling and characterization of differential through-silicon vias for 3-D ICs," *IEEE Trans. Electron Devices*, vol. 63, no. 3, pp. 1168–1175, Mar. 2016, doi: 10.1109/TED.2016.2516345.

[14] K. Kim, K. Hwang, and S. Ahn, "Wideband equivalent circuit model for a through silicon via with effective substrate current loop," in *Proc. IEEE Electr. Design Adv. Packag. Syst. Symp. (EDAPS)*, Dec. 2015, pp. 12–14, doi: 10.1109/EDAPS.2015.7383663.

[15] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*.

[16] Z. Guan, P. Zhao, X. Wang, and G. Wang, "Modeling radio-frequency devices based on deep learning technique," *Electronics*, vol. 10, no. 14, p. 1710, Jul. 2021, doi: 10.3390/electronics10141710.

[17] K. Kim, K. Hwang, and S. Ahn, "An improved 100 GHz equivalent circuit model of a through silicon via with substrate current loop," *IEEE Microw. Compon. Lett.*, vol. 26, no. 6, pp. 425–427, Jun. 2016, doi: 10.1109/LMWC.2016.2562113.

[18] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017, *arXiv:1710.05941*.

[19] D. Hendrycks and K. Gimpel, "Gaussian error linear units (GELUs)," 2016, *arXiv:1606.08415*.

[20] B. Girod, "Psychovisual aspects of image processing: What's wrong with mean squared error?" in *Proc. 7th Workshop Multidimensional Signal Process.*, Sep. 1991, p. P-2, doi: 10.1109/MDSP.1991.639240.

[21] C. Willmott, S. Ackleson, R. Davis, J. Feddema, K. Klink, D. Legates, J. O'donnell, and C. Rowe, "Statistics for the evaluation of model performance," *J. Geophys. Res.*, vol. 90, no. C5, pp. 8995–9005, Sep. 1985, doi: 10.1029/jc090ic05p08995.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[23] G. Katti, M. Stucchi, J. Van Olmen, K. de Meyer, and W. Dehaene, "Through-silicon-via capacitance reduction technique to benefit 3-D IC performance," *IEEE Electron Device Lett.*, vol. 31, no. 6, pp. 549–551, Jun. 2010, doi: 10.1109/LED.2010.2046712.

[24] D.-H. Kim, Y. Kim, J. Cho, B. Bae, J. Park, H. Lee, J. Lim, J. J. Kim, S. Piersanti, F. de Paulis, A. Orlandi, and J. Kim, "Through-silicon via capacitance–voltage hysteresis modeling for 2.5-D and 3-D IC," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 7, no. 6, pp. 925–935, Jun. 2017, doi: 10.1109/TCPMT.2017.2670063.

[25] C. R. Paul, "The concept of partial inductance," in *Inductance: Loop and Partial*. Hoboken, NJ, USA: Wiley, 2011, pp. 195–245.

[26] J. S. Pak, "PDN impedance modeling and analysis of 3D TSV IC by using proposed P/G TSV array model based on separated P/G TSV and chip-PDN models," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 1, no. 2, pp. 208–219, Feb. 2011, doi: 10.1109/TCPMT.2010.2101771.

[27] S. Mondal, S.-B. Cho, and B. C. Kim, "Modeling and crosstalk evaluation of 3-D TSV-based inductor with ground TSV shielding," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 1, pp. 308–318, Jan. 2017, doi: 10.1109/TVLSI.2016.2568755.

[28] E. Bogatin, *Signal Integrity: Simplified*. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

[29] P. Shen, W. Zhang, L. Huang, D. Jin, and H. Xie, "Improving the quality factor of an RF spiral inductor with non-uniform metal width and non-uniform coil spacing," *J. Semicond.*, vol. 32, no. 6, pp. 64–68, Jun. 2011, doi: 10.1088/1674-4926/32/6/064011.

[30] S. M. Chen, Y. K. Fang, C. S. Lin, Y. T. Chiang, F. R. Juang, and C. I. Lin, "A simple CMOS process compatible high performance parallel-stacked spiral inductor for advanced RF analog circuit applications," *Mater. Sci. Semicond. Process.*, vol. 13, no. 3, pp. 189–192, Sep. 2010, doi: 10.1016/j.mssp.2010.10.010.

[31] J.-J. Lee, Y.-K. Hong, S. Bae, J.-H. Park, J. Jalli, G. S. Abo, R. Syslo, B.-C. Choi, and G. W. Donohoe, "High-quality factor Ni–Zn ferrite planar inductor," *IEEE Trans. Magn.*, vol. 46, no. 6, pp. 2417–2420, Jun. 2010, doi: 10.1109/TMAG.2010.2044377.

XIANGLIANG LI received the B.S. degree in communication engineering from Yantai University, Yantai, China, in 2018. He is currently pursuing the M.E. degree with Hangzhou Dianzi University, Hangzhou, China. His research interests include computational electromagnetics and antennas.

PENG ZHAO (Member, IEEE) received the B.Eng. and M.Phil. degrees from the Electronic Engineering Department, Zhejiang University, China, in 2006 and 2008, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, China, in 2014. He is currently a Faculty Member of the Microelectronics CAD Center, Hangzhou Dianzi University. His research interests include computational electromagnetics and antennas.

SHICHANG CHEN (Member, IEEE) received the B.S. degree in electronic engineering from the Nanjing University of Science and Technology, in 2009, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, in 2013. From September 2013 to August 2014, he was with the City University of Hong Kong as a Postdoctoral Fellow. He is currently with Hangzhou Dianzi University as a Professor. His research interests include high-efficiency power amplifier, integrated circuits, and sensors.

KUIWEN XU (Member, IEEE) received the B.E. degree in electrical engineering from Hangzhou Dianzi University, Hangzhou, China, in 2009, and the Ph.D. degree in electrical engineering from Zhejiang University, Hangzhou, in 2014. From 2012 to 2013, he was a Visiting Ph.D. Student with the National University of Singapore, Singapore. From 2014 to 2015, he was a Senior Researcher with Huawei Technologies Company Ltd., Hangzhou. He was invited to the State Key Laboratory of Terahertz and Millimeter Waves, City University of Hong Kong, Hong Kong, China, as a Visiting Professor, in 2018. Since September 2015, he has been with Hangzhou Dianzi University, where he is currently a Professor. His research interests include electromagnetic inverse problems, microwave measurement, and novel antenna design. He has been an Associate Editor of the *IET Electronics Letters*, since 2018.

GAOFENG WANG (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Wisconsin-Milwaukee, Milwaukee, WI, USA, in 1993, and the Ph.D. degree in scientific computing from Stanford University, Stanford, CA, USA, in 2001.

He was a Scientist with Tanner Research Inc., Pasadena, CA, USA, from 1993 to 1996. He was a Principal Researcher and the Development Engineer with Synopsys Inc., Mountain View, CA, USA, from 1996 to 2001. In 1999, he worked as a Consultant with Bell Laboratories, Murray Hill, NJ, USA. He was the Chief Technology Officer (CTO) of Intpax Inc., San Jose, CA, USA, from 2001 to 2003. He was the CTO of Siargo Inc., Santa Clara, CA, USA, from 2004 to 2010. He was a Professor and the Head of the CJ Huang Information Technology Research Institute, Wuhan University, Wuhan, China, from 2004 to 2013. He was the Chief Scientist with Lorentz Solution Inc., Santa Clara, from 2010 to 2013. He is currently a Distinguished Professor with Hangzhou Dianzi University, Hangzhou, China. He has authored over 280 journal articles and holds 53 patents. His current research interests include integrated circuit and microelectromechanical system design and simulation, computational electromagnetics, electronic design automation, and wavelet applications in engineering.