Fast Stress Analysis for Runtime Reliability Enhancement of 3D IC Using Artificial Neural Network

Lang Zhang*, Hai Wang*, and Sheldon X.-D. Tan†
*School of Microelectronics & Solid-State Electronics,
University of Electronic Science & Technology of China, Chengdu, Sichuan, 610054 China
†Department of Electrical Engineering, University of California, Riverside, CA 92521 USA

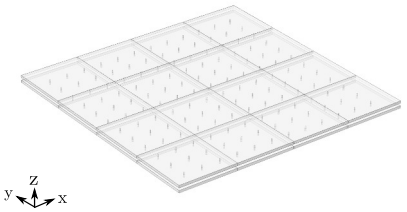
***Abstract*—Heat dissipation and the related thermal-mechanical stress problems are the major obstacles in the development and commercializing process of 3D ICs. Dynamic thermal management (DTM) techniques can be used to alleviate such problems and enhance the reliability of 3D ICs. However, the time varying stress information is hard to obtain at runtime which limites the effectiveness of DTM. In this paper, we propose a fast stress analysis method for runtime usage. The new method builds artificial neural network (ANN) model by training offline using thermal and stress data. Next, the ANN model is used to generate important stress information, such as maximum stress around each TSV, for DTM methods at runtime. In order to improve the stress estimation accuracy and speed, specially designed input selection plans are proposed and implemented for ANN model generation. Experiments with different configurations of ANN models show that the new method is able to estimate important stress information at extremely fast speed with good accuracy for runtime 3D IC reliability enhancement usage.**

I. Introduction

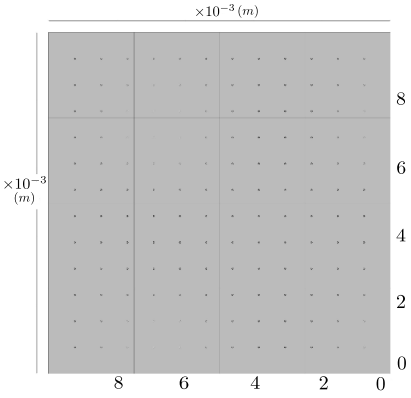
Three dimensional stacked IC (3D IC) is a promising technique to overcome the scaling problems encountered in traditional 2D IC. It provides benefits in enabling heterogeneous integration, alleviating interconnect barrier problem, and improving power performance, etc. The benefits are provided by integrating IC dies vertically, and connecting different die layers by through-silicon vias (TSV) [1]. With so many advantages against traditional 2D IC, the 3D IC suffers from severe thermal induced reliability problems, due to its poor heat removing ability in the vertical dimension [2], [3]. The TSV structure makes the thermal induced reliability even worse: caused by different materials used in TSV and die, thermal variations in both space and time lead to stress variations around TSV, which shortens the lifetime of 3D IC [4].

Dynamic thermal management (DTM) techniques [5], [6], [7] can be applied to 3D IC in order to enhance the thermal induced reliability. However, most of the existing DTM methods aim at lowering the average temperature of the chip and minimizing temperature variation across different parts of the 3D IC body, due to the lack of information on the reliability side. Although the thermal induced stress value is highly dependent on the temperature value, a higher temperature at one position does not indicate a larger stress value there as will be shown in this paper. The major reason is that the stress value is not only effected by temperature but also significantly effected by its distance to the nearby TSV structures.

The TSV stress can be simulated by finite element methods (FEM) [8], [9]. However, being relatively accurate, FEM based



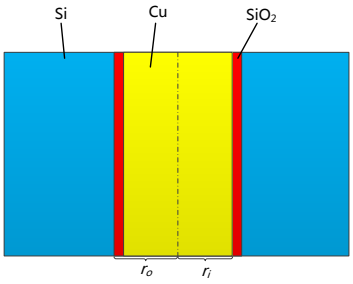
(a) Side view showing the two die layers connected with 144 TSVs.



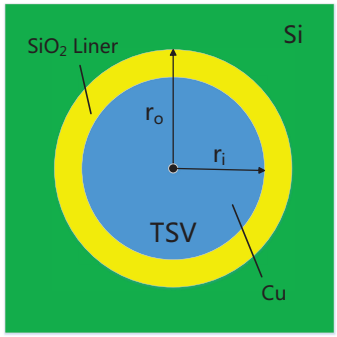
(b) Bottom view of chip showing the 16-core architecture for each layer.

Fig. 1. A 3D IC model built in COMSOL.

This work is supported in part by National Natural Science Foundation of China, under grant No. 61404024, in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, in part by the Open Foundation of State Key Laboratory of Electronic Thin Films and Integrated Devices, under grant No. KFJJ201409.
Corresponding author: Hai Wang (wanghai@uestc.edu.cn).



(a) Cross-section view.



(b) Longitudinal-section view.

Fig. 2. TSV filled with Cu with a SiO_2 liner.

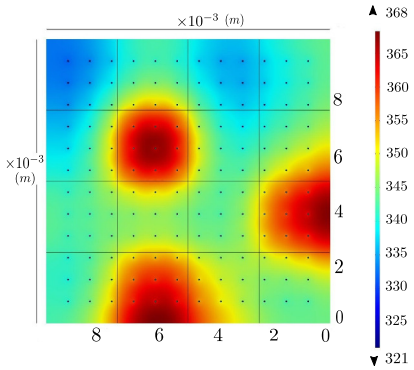


Fig. 3. Temperature (K) distribution of the bottom surface of the 3D IC model in COMSOL.

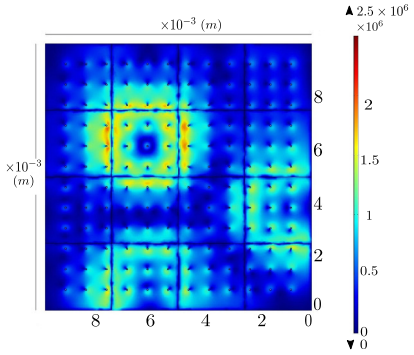


Fig. 4. Von Mises thermal stress (N/m^2) distribution of the bottom surface of the 3D IC model in COMSOL.

methods cannot be used for DTM due to their large computing costs. The analytical stress models have also been developed [8], [4], [10] to analyze thermal stress behaviors around TSV structures. However, these methods assume simple TSV structures in order to simplify the mathematical equation derivation steps, and generates model with relatively large error.

In this work, a new fast stress analysis method is developed for runtime reliability enhancement of 3D IC. A 3D IC model with TSVs is built first in COMSOL. The TSV structure is copper filled, with silicon dioxide liner between the copper and silicon. Then, stress and thermal data are obtained by simulating the 3D IC model with different power distributions. An artificial neural network (ANN) with backpropagation (BP) algorithm is trained with the temperature distribution data around each TSV as model input and important stress information (such as maximum stress) around each TSV as model output. The input temperature data is carefully selected in order to remove the redundancy and avoid the rotational symmetry problem. The trained ANN model is used for each TSV as the maximum stress detector and can be in cooperated with DTM methods for reliability enhancement of 3D IC.

II. Stress analysis for 3D IC with TSV structures using FEM based method

In this section, we introduce some preliminaries for stress modeling in 3D IC with the appearance of TSVs.

TSVs are introduced in 3D IC as the vertical communication devices to solve the interconnect scaling problem. As a necessary structure in 3D IC, TSVs, however, lead to thermal induced stress problem, which harms the reliability of the chip. There are two major reasons for the problem. For the first reason, TSV usually has a much higher thermal conductivity than silicon wafers because of the materials it used. As a result, large temperature gradient may appear in the area close to TSV, which usually leads to larger thermal stress. For the second reason, mismatch in coefficient of thermal expansion (CTE) also brings significant stress increase. The CTE of silicon is $4.15 \times 10^{-6}K^{-1}$, while copper's CTE is $17 \times 10^{-6}K^{-1}$, which is nearly four times bigger than silicon. When temperature increases with the same degree, copper expansion will be much more significant than silicon, resulting in considerable stress.

A two-layer 3D IC chip structure with TSVs is shown in Fig. 1. This is a chip with multi-core architecture, with TSVs uniformly distributed across the chip. A popularly used TSV structure [8] with full copper filling and a silicon dioxide liner between copper and silicon applied is shown in Fig. 2.

The stress in solid in cartesian coordinate can be expressed

as [9]:

$$\begin{cases} -f_x = k_x - \dfrac{E\alpha}{1-2\nu} \cdot \dfrac{\partial T}{\partial x} \\ -f_y = k_y - \dfrac{E\alpha}{1-2\nu} \cdot \dfrac{\partial T}{\partial y} \\ -f_z = k_z - \dfrac{E\alpha}{1-2\nu} \cdot \dfrac{\partial T}{\partial z} \end{cases} \quad (1)$$

where

$$\begin{cases} k_x = \mu(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}) + (\mu+\lambda)(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 v}{\partial y \partial x} + \dfrac{\partial^2 w}{\partial z \partial x}), \\ k_y = \mu(\dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 v}{\partial z^2}) + (\mu+\lambda)(\dfrac{\partial^2 u}{\partial x \partial y} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 w}{\partial z \partial y}), \\ k_z = \mu(\dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2}) + (\mu+\lambda)(\dfrac{\partial^2 u}{\partial x \partial z} + \dfrac{\partial^2 v}{\partial y \partial z} + \dfrac{\partial^2 w}{\partial z^2}), \\ \mu = \dfrac{E}{2(1+\nu)}, \\ \lambda = \dfrac{E\nu}{(1+\nu)(1-2\nu)}. \end{cases}$$

The terms (f_x, f_y, f_z) are forces in x, y, and z directions, (u, v, w) are displacements in the three directions, E is the elastic modulus, ν is the Poisson ratio, α is the thermal expansion coefficient, T is the temperature, μ and λ are the Lamé coefficients. From (1), it can be observed that (f_x, f_y, f_z) changes with temperature T.

As a powerful method for the analysis of thermo-mechanical stress in a complex structure where experimental investigation is quit difficult, FEM method can be used to build the 3D IC model with TSV structures based on (1). We have built a two-layer 3D IC model with 12×12 TSVs uniformly placed in the whole chip using the FEM based software COMSOL. The size of whole chip is $1cm \times 1cm \times 300\mu m$, and it is divided into 4×4 same sized blocks to represent 16 cores, both of the two layers are $1cm \times 1cm \times 100\mu m$ as shown in Fig. 1. For the TSV structure, we set the values of r_i and r_o in Fig. 2 as $20\mu m$ and $24\mu m$, respectively. We also couple the solid heat conduction field and the solid mechanical field. All surfaces of chip are fixed and other boundary conditions can be adjusted. The structure is assumed to be stress free at the room temperature $300K$. Power is applied to the bottom layer. Using the FEM model, temperature and Von Mises stress information can be extracted with different power distributions as model input.

Take the bottom layer as an example. From [10], it is known that the stress in the Z direction is 0, and as a result, we only need to analyze the stress in the 2-D plane. Fig. 3 and Fig. 4 show the temperature and Von Mises stress of bottom surface with a given power distribution, respectively. By observation and comparison of these two figures, it is easy to find out that stress depends not only on temperature, but also on TSV positions: there are many places which have *higher* temperatures but turns out to have *smaller* stresses than that of some lower temperature places, because the previous places are much further away from TSVs than the latter ones. This is an important observation which means that taking only the temperature as the optimization consideration is insufficient for DTM methods.

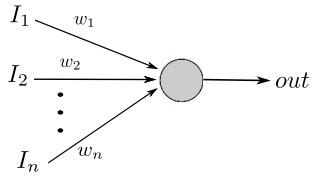


Fig. 5. The function of one neuron.

III. Fast stress analysis for 3D IC

In the previous section, we have shown that stress distribution in 3D IC does not only relates to the temperature distribution, but also relates to the TSV distribution. As a result, it is important to take the stress information directly into account in DTM methods. Unfortunately, the FEM based stress analysis method is too expensive to be used for DTM, which requires stress information to be updated at runtime. In order to solve this problem, we propose an ANN based method providing a viable way for runtime enhancement of the reliability of 3D IC.

The new method aims at calculating the maximum stress around each TSV on chip rapidly by taking the temperature distribution around the TSV as input. The maximum stress around a TSV is chosen to save the computing cost of the ANN model as it is the most important factor to affect the reliability. Please note that other output other than maximum stress or even more outputs can also be chosen if necessary.

In machine learning and cognitive science, ANNs are a family of statistical learning models inspired by biological neural networks (the central nervous systems of animals, in particular the brain) and are used to estimate or approximate functions that can depend on a large number of inputs. ANNs are generally presented as systems of interconnected "neurons", which send messages to each other. For neurons shown in Fig. 5, they have the same function:

$$out = \sum_{i=1}^{n} I_i w_i, \quad (2)$$

where the terms I_i, $i = 1, 2, \ldots, n$ are inputs. (2) means the output of a neuron depends on the weighted sum of all its inputs. The connections have weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning. A large variety of ANNs have been developed and widely used in many fields, such as function approximation, classification, and data processing. Considering the generalization ability, we select backpropagation (BP) network as our neural network model.

Backpropagation is a common method of training ANNs used in conjunction with an optimization method such as gradient descent. It calculates the gradient of a loss function with respect to all the weights in the network. The gradient is fed to the optimization method which in turn uses it to update the weights, in an attempt to minimize the loss function. This ANN can be simply explained in Fig. 6 which has a hidden layer. The network usually has one or several hidden layers, with the corresponding nodes in such layers

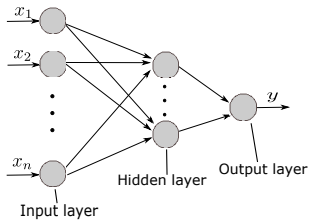


Fig. 6. A BP net with one hidden layer (n inputs and one output).

called hidden neurons. With one or more hidden layers, the network is able to model higher-order statistical properties. Initially, before training, the weights will be set randomly. Then the neurons learn from training samples. Each sample, taken from the system to be modeled, is an input-output set $\{x_1, x_2, \ldots, x_n, t\}$, where $\{x_1, x_2, \ldots, x_n\}$ are the inputs and t is the corresponding output of the system to be modeled. Since the ANN is not exactly the same as the system to be modeled, given the sample inputs $\{x_1, x_2, \ldots, x_n\}$ to the ANN will generate an output y which may differ from t (especially when the weights are initially random). During the training process, the weights will be tuned automatically to make y close to t.

In our BP neural network model, the temperatures and maximum thermal stress around TSV are set as input and output. So in Fig. 6, $x_1, x_2, \ldots, x_n$ are spatially distributed temperature values around one certain TSV and y represents the maximum stress around that TSV. For the k-th TSV, we will have one temperature vector $X_k = \{x_{k1}, x_{k2}, \ldots, x_{kn}\}$, and one maximum stress scalar y_k. We call such data $S_k = (X_k, y_k)$ a *sample*. Considering the large scale of samples and the resilient backpropagation's (RPROP) advantage in nonlinear mapping, processing efficiency and less resources occupying [11], we select RPROP as the learning algorithm. Since the irregularly shaped finite element grid in COMSOL prevents us from taking sample x_i from the same location (relative location to the TSV center) for different TSVs, we first formulate a common square shaped grid for all TSVs. Then, we do interpolation according to the common grid to generate p unified samples: $S_1, S_2, \ldots, S_p$.

A. Input selection in ANN

By far, we have the basic framework established for ANN modeling. But the input selection has huge impact on the performance and efficiency of ANN: how to select temperature elements and form the input vector X_k for the k-th sample? At the beginning, we put temperatures of all the square shaped grid elements around the TSV into X_k as a input vector and build a BP neural network which has two hidden layers. When training this network, we find it consumes a lot of time in training because of the big number of inputs. In addition, validation process shows the mean squared error (MSE) of the resulted model is quit large.

With a lot of experiments and verifications, it is observed that the maximum stress always appears within a certain distance from the TSV center. As a result, there are a lot

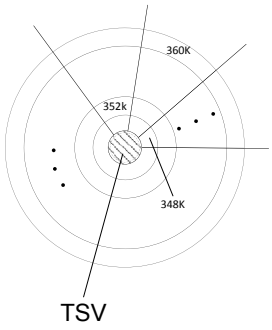


Fig. 7. New circular shaped common grid to optimize input selection. Only temperatures of the grid elements close enough to the center are selected into the ANN training process.

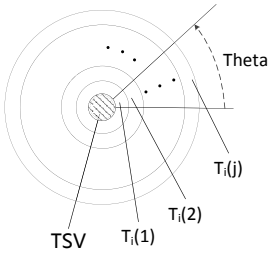
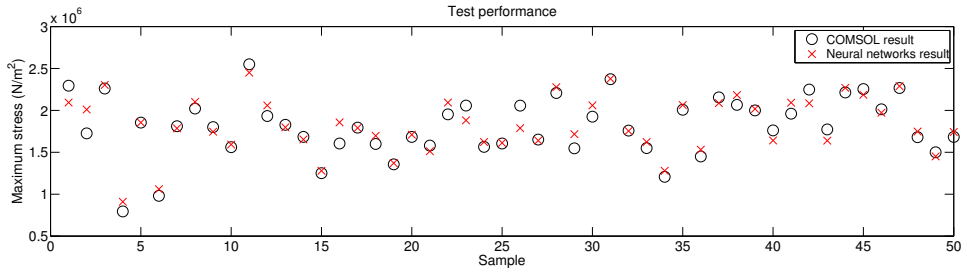


Fig. 8. Method to solve the rotational symmetry problem. Total temperatures are calculated for each piece along the radial lines by summing up $T_i(1), T_i(2), \ldots, T_i(j)$ for the i-th piece for example.

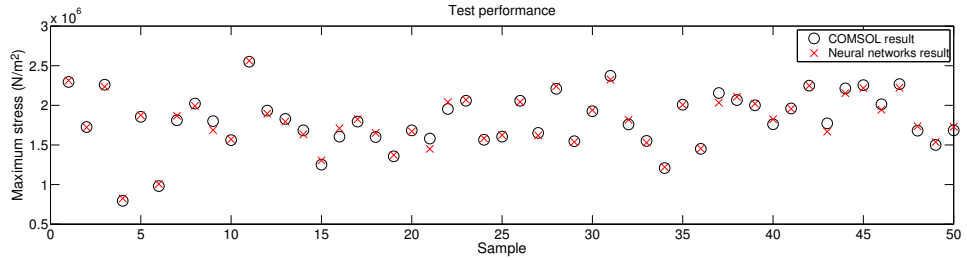
of redundant data at the far away positions from the center. In addition, it is quite obvious that having a round structure, the TSV and the areas around it have a *rotational symmetry property*: two different thermal/stress maps can be the exactly the same after a certain amount of rotation, which simply means redundancy there if we take them as different samples. With the two observations, we can improve the original simple input selection method, and propose the first input selection plan A by utilizing the first observation and the second input selection plan B by employing both observations.

1) Input selection plan A: In order to save training time and enhance the ANN model accuracy, we need to reduce the input number for each sample.

As discussed in Section II, TSV usually has a higher thermal conductivity than silicon wafers. As a result, large temperature gradient may appear in the area close to TSV, resulting in large stress. In addition, the stress also tends to appear at the positions where different materials attach, that is, near TSV. Because of the reasons above, the first observation verifies that the largest stress always appears *not far away* from the TSV center, thus taking all temperature data around one TSV as inputs introduces a lot of redundancy data, and can harm the ANN training/validation speed as well as accuracy. As a result, input selection plan A is proposed by first generating common grids in a circular shaped way as shown in Fig. 7, instead of the previous square shaped way. Then, we simply ignore the temperature information for the grid elements which



(a) Accuracy result of the ANN model with input selection Plan A. The ANN model has two hidden layers, with 60 neurons in the first layer and 15 neurons in the second layer.



(b) Accuracy result of the ANN model with input selection Plan B. The ANN model has two hidden layers, with 55 neurons in the first layer and 15 neurons in the second layer.

Fig. 9. Accuracy results of the ANN models with input selection plan A and plan B, respectively, showing 50 validation samples.

are further than the threshold distance. It results in significant redundancy reduction.

2) Input selection plan B: According to the second observation, the input selection plan A does not solve all the problems: since the cross section of TSV is a circle, the temperatures around the TSV can be *rotational symmetry*. Two or more samples, which are actually the same with respect to a rotation angle, are taken into consideration as different in the training process, resulting in less accurate ANN model.

In order to solve this problem in an extremely fast way, input selection plan B is proposed with an improved method as demonstrated in Fig. 8. We start from the input selection plan A: assume the plane around the TSV are divided into k pieces (each piece with an angle $\theta = 2\pi/k$ in the figure) around the full circle by the radial lines, and j rings divided by the circles, making a total of $k \times j$ sections. Then we can calculate the total temperature value of all sections in each piece, for example, there is $M_i = T_i(1) + T_i(2) + \cdots + T_i(j)$ for the i-th piece. Then, we find the largest total temperature value, and record the corresponding piece's angle, with a respect to a reference line. Finally, we are able to rotate the largest total temperature piece back to the reference line for all samples, and solve the rotational symmetry problem.

IV. Experimental Results

The experiments are performed on a Linux server with two 2.90GHz 8-core 16-thread CPUs and 64GB memory. A 3D IC model with 144 TSVs is built in COMSOL 5.0 to obtain the thermal and stress data. The ANN based fast stress analysis method is implemented in MATLAB R2010a.

First, we have performed 19 simulations using the 3D IC model in COMSOL for 19 different power distributions covering a large variety of the 3D IC operating conditions. Because there are 144 TSVs on the chip, by running the 19 simulations, we have obtained around 2700 samples in total. The samples are divided into two parts, 2400 samples are used for training and building the ANN models, and the rest 300 samples are used for validating the accuracy of the ANN models. k and j in the input selection method are chosen to be 16 and 20, respectively. In the ANN training routine, the maximum iteration number is set as 1000. For both of the input selection plan A and plan B, the ANN models are set to having two hidden layers.

First, we test the accuracy of ANN models generated by the two input selection plans as shown in Fig. 9. For plan A, where the common circular shaped grid is used, the maximum stress from COMSOL data and from the ANN model matches with each other well in most validation samples as can be seem in Fig. 9 (a). However, there are still occasions when relatively large errors appear, such as for the validation sample 2, 16, and 26, etc. Through analyzing the results with input selection plan A, we have deleted such large errors are caused by the rotational symmetry property of the TSV model. The result of the rotational symmetry enhanced input selection plan B is shown in Fig. 9 (b). With even a smaller ANN model (only 55 neurons in the first hidden layer are used in plan B comparing

TABLE I
ACCURACY AND COMPUTING TIME RESULTS OF INPUT SELECTION PLAN A.

# nodes in 2 layers	t_1 (s)	t_2 (s)	0<err<0.01		0.01<err<0.05		0.05<err<0.1		0.1<err<0.2		0.2<err<0.3		0.3<err<0.4		0.4<err<0.5		0.5<err		maximum err	mean err
			n	p	n	p	n	p	n	p	n	p	n	p	n	p	n	p		
40, 10	60	0.051	13	4.3%	60	20.0%	55	18.3%	82	27.3%	52	17.4%	20	6.7%	10	3.3%	8	2.7%	0.742	0.1521
50, 15	62	0.058	17	5.7%	57	19.0%	60	20.0%	90	30.0%	48	16.0%	17	5.7%	4	1.3%	7	2.3%	0.722	0.1502
55, 15	65	0.060	16	5.3%	64	21.3%	58	19.3%	80	26.7%	44	14.7%	26	8.7%	6	2.0%	6	2.0%	0.693	0.1413
60, 15	67	0.059	18	6.0%	63	21.0%	63	21.0%	74	24.7%	46	15.3%	27	9.0%	5	1.7%	4	1.3%	0.792	0.1385
70, 20	68	0.061	15	5.0%	59	19.7%	62	20.6%	88	29.3%	38	12.7%	24	8.0%	8	2.7%	6	2.0%	0.745	0.1445

TABLE II
ACCURACY AND COMPUTING TIME RESULTS OF INPUT SELECTION PLAN B.

# nodes in 2 layers	t_1 (s)	t_2 (s)	0<err<0.01		0.01<err<0.05		0.05<err<0.1		0.1<err<0.2		0.2<err<0.3		0.3<err<0.4		0.4<err<0.5		0.5<err		maximum err	mean err
			n	p	n	p	n	p	n	p	n	p	n	p	n	p	n	p		
40, 10	57	0.041	30	10.0%	96	32.0%	73	23.7%	74	24.8%	21	7.0%	4	1.3%	2	0.6%	2	0.6%	0.759	0.0938
50, 15	58	0.045	31	10.4%	104	34.7%	65	21.7%	73	24.3%	19	6.4%	2	0.6%	1	0.3%	2	0.6%	0.704	0.0832
55, 15	60	0.056	41	13.7%	105	35.0%	83	27.7%	53	17.7%	13	4.3%	3	1.0%	0	0%	2	0.6%	0.693	0.0779
60, 15	63	0.054	30	10.0%	106	35.3%	83	27.7%	62	20.7%	13	4.4%	3	1.0%	1	0.3%	2	0.6%	0.792	0.0799
70, 20	65	0.055	31	10.4%	99	33.0%	78	26.0%	66	22.0%	18	6.0%	4	1.4%	2	0.6%	2	0.6%	0.745	0.0802

with 60 neurons used in plan A), it is very clear that the new input selection plan resolved the rotational symmetry problem and drastically improved the accuracy from the previous ANN model showing no validation sample with large error at all.

We have also changed the number of neuron nodes in each hidden layer and record all results in Table I and Table II, where "n" in the tables means number of nodes, and "p" represents percentage. It is noted that the accuracy does not always improve as we increase the ANN model size due to overfitting problems, and the best accuracy result appears for around 55 to 60 nodes in the first layer and 15 nodes in the second layer. For all cases, ANN model built by input selection plan B outperforms the one built by plan A.

The major purpose to build such ANN model is for runtime usage in cooperating with a dynamic thermal/power management method. As a result, the speed of stress analysis using the ANN based model is very important. In order to test the speed accurately, we have disabled the parallel computing ability of MATLAB, and use only single thread to do the model training and stress analysis using the ANN model. In Table I and Table II, we have recorded the training time (t_1 in both tables) and stress analysis time (t_2 in both tables). Please note that the stress analysis time is the total time of all the 300 TSV structures. The training process usually takes around one minute to finish. But it is still considered to be really fast because it only needs to be done once *offline*, and its computing time will not be counted in the runtime part. The only runtime part comes from the stress analysis, and it is very clear that such time is really small: only around 0.06 seconds for 300 TSV structures. The runtime overhead can be further reduced by using parallel computing techniques, because the stress around each TSV can be analyzed independently.

V. CONCLUSION

In this paper, we have proposed a fast stress analysis method for runtime usage. The new method builds artificial neural network (ANN) model by training offline using thermal and stress data. Next, the ANN model is used to generate important stress information, such as maximum stress around each TSV, for DTM methods at runtime. In order to improve the stress estimation accuracy and speed, specially designed input selection plans are proposed and implemented for ANN model generation. Experiments with different configurations of ANN models show that the new method is able to estimate important stress information at extremely fast speed with good accuracy for runtime 3D IC reliability enhancement usage.

REFERENCES

[1] R. S. Patti, "Three-dimensional integrated circuits and the future of system-on-chip designs," *Proc. of the IEEE*, vol. 94, no. 6, pp. 1214–1224, June 2006.

[2] F. Li *et al.*, "Design and management of 3D chip multiprocessors using network-in-memory," in *Proc. Int. Symp. on Computer Architecture (ISCA)*, 2006, pp. 130–141.

[3] A. Sridhar *et al.*, "3D-ICE: Fast compact transient thermal modeling for 3D ICs with inter-tier liquid cooling," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2010, pp. 463–470.

[4] M. Jung *et al.*, "TSV stress-aware full-chip mechanical reliability analysis and optimization for 3D IC," *Communications of the ACM*, vol. 57, no. 1, pp. 107–115, January 2014.

[5] J. Donald and M. Martonosi, "Techniques for multicore thermal management: Classification and new exploration," in *Proc. Int. Symp. on Computer Architecture (ISCA)*, June 2006, pp. 78–88.

[6] J. H. Lau and T. G. Yue, "Thermal management of 3D IC integration with TSV (through silicon via)," in *Proc. Electronic Components and Technology Conference (ECTC)*, 2009, pp. 635–640.

[7] Q. Zou *et al.*, "Thermomechanical stress-aware management for 3D IC designs," in *Proc. European Design and Test Conf. (DATE)*, 2013, pp. 1255–1258.

[8] K. H. Lu *et al.*, "Thermo-mechanical reliability of 3-D ICs containing through silicon vias," in *Proc. Electronic Components and Technology Conference (ECTC)*, 2009, pp. 630–634.

[9] T. Chen, D. M. Liao, and J. X. Zhou, "Numerical simulation of casting thermal stress and deformation based on finite difference method," *Materials Science Forum (MSF)*, vol. 762, pp. 224–229, 2013.

[10] S. K. Marella and S. S. Sapatnekar, "A holistic analysis of circuit performance variations in 3-D ICs with thermal and TSV-induced stress considerations," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, vol. 23, pp. 1308–1321, July 2014.

[11] Z. Xiong *et al.*, "Two-phases parallel neural network algorithm based on RPROP," in *Proc. IEEE TENCON*, 2005, pp. 1–6.