Performance Prediction of Large-Scale 1S1R Resistive Memory Array Using Machine Learning

Zizhen Jiang[1*], Peng Huang[1, 2], Liang Zhao[1], Shahar Kvatinsky[3],
Shimeng Yu[4], Xiaoyan Liu[2], Jinfeng Kang[2], Yoshio Nishi[1], and H. -S. Philip Wong[1]
[1]Dept. of Electrical Engineering, Stanford University, CA 94305, U.S.A.;
[2]Peking University, Beijing, 100871, China; [3]Dept. of Computer Science, Stanford University, CA 94305, U.S.A.;
[4]Arizona State University, Tempe, AZ 85281, USA; Email: jiangzz@stanford.edu

***Abstract*—A methodology to analyze device-to-circuit characteristics and predict memory array performance is presented. With a five-parameter characterization of the selection device and a compact model of RRAM, we are able to capture the behaviors of reported selection devices and simulate 1S1R cell/array performance with RRAM compact modeling using HSPICE. To predict the performance of the memory array for a variety of selectors, machine-learning algorithms are employed, using device characteristics and circuit simulation results as the training data. The influence of selector parameters on the 1S1R cell and array behavior is investigated and projected to large Gbit arrays. The machine learning methods enable time-efficient and accurate estimates of 1S1R array performance to guide large-scale memory design.**

***Keywords*—RRAM, selector, 1S1R, crossbar array, prediction, machine learning**

I. Introduction

Two-terminal selection devices have been widely investigated to suppress the unwanted sneak path currents and boost the memory window of purely passive RRAM array by adding non-linearity to the cell [1-5]. Choosing an inappropriate selector-resistor match can result in poor 1S1R behavior (Fig. 1(a)), which leads to poor array characteristics. Physics-based device models are generally adopted to accurately capture device characteristics for simulations. However, the use of physics-based compact models is still infeasible for simulations of large arrays due to the enormous amount of required computing resources (Fig. 1(b)). For example, a 1 Mb RRAM 1S1R array simulation takes more than 24 hours of CPU time using a main memory of 20 GB for a 64-core Intel(R) Xeon(R) (@2.67 GHz) server. Rapid exploration of the 1S1R design space requires simulating of a large number of possible 1S1R options. To reduce the simulation time for Mbit to Gbit RRAM arrays, an efficient method to estimate various 1S1R array behaviors is required.

This work focuses on a new methodology that uses features extracted from device characteristics to predict 1S1R cell/array behaviors, based on machine learning regression algorithms. Despite the recent popularity of machine learning applications, such computer vision [7] and language processing [8], few researchers, such as machine learning and semiconductor domains, have considered using machine-learning algorithms to predict device impacts on circuit/system behaviors. The use of machine learning enables us to rapidly investigate (less than one second with an 1.3 GHz Intel Core i5 personal laptop) the 1S1R cell/array (Fig. 2) behaviors for different selector and RRAM device characteristics. The paper is organized as follows. In section II, the five-parameter characterization of selectors is described and the analysis methodology is presented. Section III investigates the impacts of the five-parameters of selectors on 1S1R cell behavior.

Validation of machine learning prediction results of 1S1R cell/array behaviors by circuit simulations is provided in section IV. Section V discusses the Pareto optima of 1S1R crossbar array, followed by conclusions and suggestions for future work in Section VI.

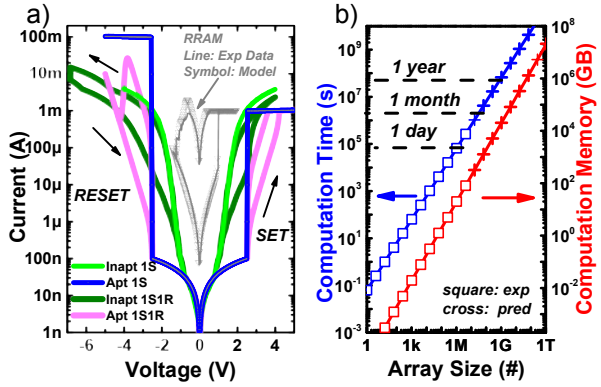

Fig. 1. (a) Inappropriate selector-RRAM match results in undesirable high RESET voltage, which optimized selector characteristics can solve. Light green curve, the inappropriate selector, inapt 1S [1][6]; green curve, the corresponded inappropriate selector-RRAM match 1S1R; blue curve, the appropriate (optimized) selector, apt 1S; pink curve, the corresponded appropriate selector-RRAM match, apt 1S1R. (b) Computation time and memory grow exponentially with the increase of array size. It takes a year to simulate a 400 Mb array.

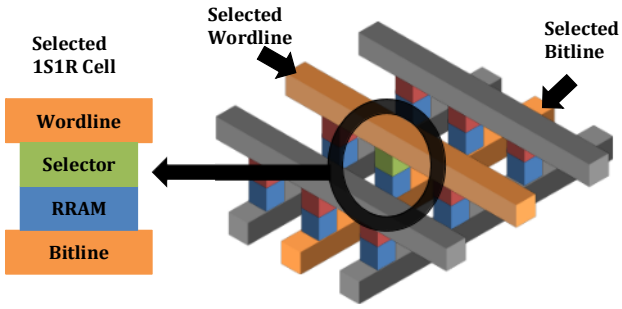

Fig.2. Schematic of a 1S1R crossbar array.

II. Analysis Methodology

Fig. 3 depicts the work flow of our simulation and prediction methodology. We first extract the key features (inputs, x) from device characteristics and simulate the circuit behavior (outputs, y) in a circuit simulator. As an example in this paper, the characteristics of individual selector (1S) and RRAM (1R) devices are extracted as input, while the performance of 1S1R cell characteristics and 1S1R crossbar array (up to 1 Mbit) are evaluated by HSPICE simulations and act as the output. Using a group of input and output results (the training set), we train the machine learning regression algorithms to predict the circuit performance for other inputs (test inputs) and for array sizes larger than the training set. The prediction accuracy is verified by circuit simulation (using HSPICE) of the test inputs that are not part of the training set.

978-1-4673-6933-6/15/$31.00 ©2015 IEEE

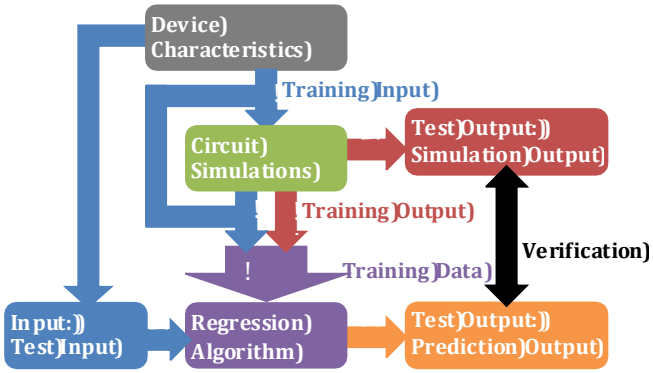

Fig. 3. Schematic of the work flow of our simulation and prediction methodology.

A. Device Characteristics – Input (Features)

A physics-based Verilog-A compact model is used to describe the resistive switching of RRAM devices [9]. The grey symbols in Fig. 1(a) represents the simulation data, which reproduce the SET and RESET characteristics measured in experiments (Fig. 1(a), grey line) [9].

To capture the reported I-V characteristics of selection devices, a five-parameter characterization is used, as depicted in Fig. 4(a). Reported selection devices can be divided into three regions: low conduction (*L*), tunneling (*T*) and saturation (*S*). We use V_{th} (V) and I_{th} (dec) to capture the typical resistance feature in the L-region, and S_{th} (V/dec) to depict the *T-region* slope, while I_{on} (dec) and S_{on} (V/dec) to describes *S-region* behaviors. We use the logarithm of the current to base 10 for the values of I_{th} and I_{on}. *L-region* normally defines the selection devices' intrinsic leakage before the onset of the tunneling region (*T-region*). Characteristics of recent-reported selectors are decoupled into the five-parameter selector space (Fig. 4). The selectors, as reported by J. Huang, E. Cha and S. Jo [1][2][4], exhibit I-V characteristics described by three regions. Some selectors may not have all three regions, such as the selector L. Zhang reported in [5], which only has the *T-region* (blue curve with triangle symbols).

In this paper, we consider selector characteristics as the main design variables. V_{th}, I_{th}, S_{th}, I_{on}, and S_{on} are the inputs for 1S1R cell analysis. Besides the five-parameter inputs, the array size and power supply (V_{DD}) are additional inputs for 1S1R array analysis.

B. Circuit Simulation Results – Output (Targets)

Applying this five-parameter characterization for the selectors with the Verilog-A compact model for RRAM devices, we generate HSPICE circuit netlists to simulate 1S1R cell and crossbar arrays (up to the size of 1 Mb) behaviors to illustrate the concept of using machine learning algorithms to estimate array behaviors efficiently.

1) For a 1S1R cell, minimum write voltage, V_{MIN} (SET, V_{MINS} and RESET, V_{MINR}), minimum read voltage and the read window are the outputs, for a given set characteristics for the resistive switching devices. Considering that READ should not interrupt the resistance state, the read window of a 1S1R cell is defined as

$$Read\ Window = V_{MIN} - V_R; \tag{1}$$

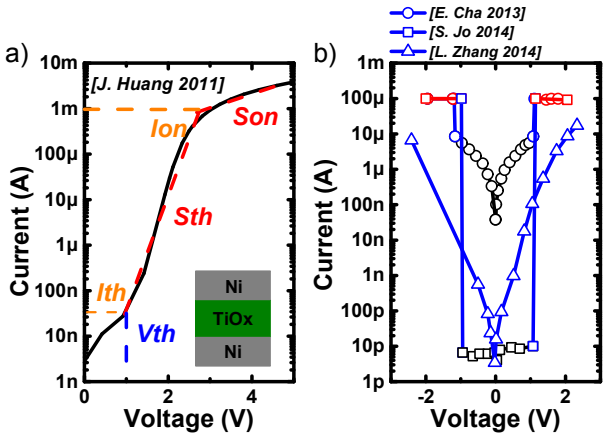

Fig. 4. (a) Five-parameter selector characteristics illustration, where V_{th} and I_{th} imply L-region selector behavior, S_{th} captures T-region behavior and I_{on} and S_{on} indicates S-region behavior. (b) Reported selectors can be described using this characterization. The black line is L-region, blue line is T-region and red line is S-region.

where V_{MIN} is the minimum write voltage and V_R is the minimum read voltage when the current difference between HRS and LRS is more than 100 nA required by the read sense amplifier.

2) For 1S1R crossbar arrays, the write margin, read margin (>100 nA), and energy consumptions are the outputs. Different selector and RRAM device characteristics have different V_{MIN}. The write margin is re-defined as

$$Write\ Margin = \log\left(\frac{|V_{ACC}|}{|V_{MIN}|}\right); \tag{2}$$

where V_{ACC} is the access voltage on the worst-case cell. Write margin > 0 indicates the worst-case cell can be programmed successfully.

C. Data Processing

We generate 105 random generated five-parameter sets with different array size and power supplies as training data (using *a patterning half-pitch of 22 nm*); another 45 random sets are used as test data. 'Adaptive Boosting' ('AdaBoost') [10] and 'Support Vector Regression' ('SVR') [11] are adopted for different output predication, which are chosen among different tested machine learning regression algorithms [12] ('Nearest neighbor', 'Decision tree', 'AdaBoost' and 'SVR') to exhibit the smallest error. 'AdaBoost' is a linear combination of step-wise functions. In 'SVR', the inputs are first mapped onto an m-dimensional feature space using a fixed nonlinear mapping (we use the 'Radial Basis Function Kernel'). Then a linear model is constructed from this mapped feature space. The flexibility of 'AdaBoost' and 'SVR' models' nonlinearity allows fitting the nonlinear behavior of 1S1R cell/crossbar arrays effectively.

III. 1S1R Cell Level Analysis

The estimation of a 1S1R cell performance helps identify the selector characteristics that reduce the minimum write voltages and expand the read windows. To estimate the 1S1R cell performance, a single cell is simulated using HSPICE.

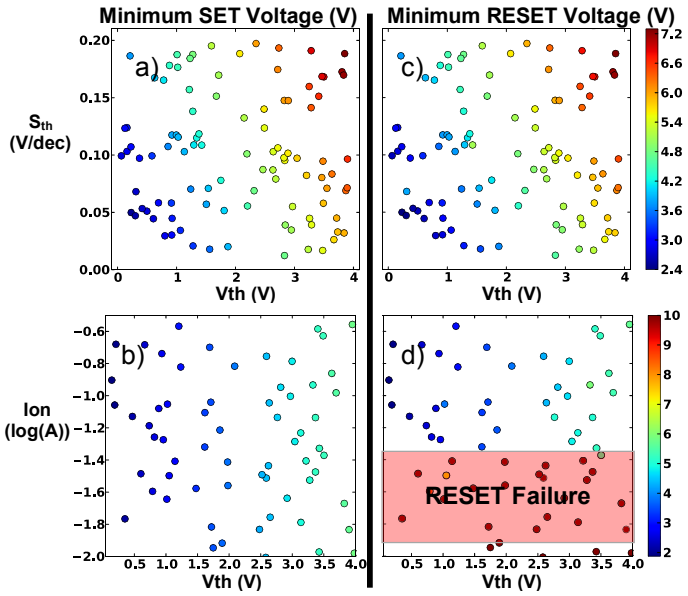

Fig. 5. Simulation results of minimum SET and RESET voltages for 1S1R cells (absolute value shown by the color scale (V)): (a-b) V_{MINS} and (c-d) V_{MINR}. For (a)(c), I_{th} = 1 pA, I_{on} = 100 mA, and S_{on} = ∞ V/dec; For (b)(d) I_{th} = 1 pA, S_{th} = 0.01 V/dec, and Son = ∞ V/dec.

As seen from Fig. 5(a-b), the SET minimum write voltage (V_{MINS}) depends primarily on V_{th} and S_{th}, where the dominant factor is V_{th} due to the intrinsic abrupt RRAM SET process requiring a certain voltage level on RRAM device to SET the device. I_{on} and S_{on} can limit low resistance state after SET and thereby the resistance window.

RESET minimum write voltage (V_{MINR}) depends on I_{on}, V_{th}, and S_{th}, as shown in Fig. 5(c-d), but it is comparably more dependent on I_{on} than the SET process. When the RRAM device is SET to deep LRS, the I_{on} of the selector can limit the voltage applied across the RRAM device and result in RESET failure (Fig. 5(d)), if the equivalent resistance of selector at that voltage is equivalent to or lower than the RRAM resistance. Thus, minimum RESET voltage has to increase.

The results of the SET and RESET minimum voltages indicate that one can trade off V_{th} with S_{th} (have larger V_{th} but smaller S_{th}, or smaller V_{th} but larger S_{th}) to reach the same minimum write voltage.

Minimum read voltage depends on V_{th}, I_{th} and S_{th}. Thus read window, as defined in (1), is determined by V_{MIN} and V_R.

IV. Prediction of 1S1R Cell/Array Behaviors

To predict the performance of large memory arrays, it is useful to have simple, analytical description of the 1S1R cell and array behavior based on the characteristics of the selectors and the RRAMs. However, this is not possible even in the simple case of a resistive network and gross simplification has to be made [13]. In this paper, we demonstrate the use of machine learning regression algorithms for projecting: 1) other selector and RRAM pair characteristics through the use of a small set of selector and RRAM pair characteristics; 2) large memory array behavior through the use of a small number of full circuit simulations as training input data. Machine learning

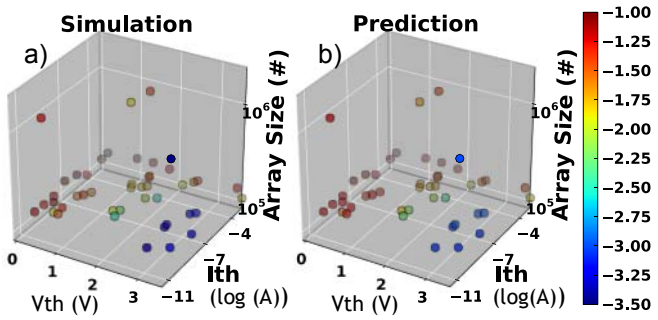

Fig. 6. Simulation (a) and prediction (b) results of 1S1R array (512×512 and 1024×1024) RESET power consumption for test data. The color scale represents the logarithm of power consumption (log (W)). The unit of V_{th} and I_{th} is V and log10(A) respectively.

analysis explores the construction and study of algorithms that can learn from data and then makes predictions. Once the algorithms are trained from training data, it can predict 1S1R cell and array behavior expeditiously without performing time-consuming full circuit simulations.

We are able to predict a variety of 1S1R cell and array properties: SET minimum voltage V_{MINS}, RESET minimum voltage V_{MINR}, Read minimum voltage V_R and Read window (*RW*); write margin, SET energy E_S, RESET energy E_R, and read margin δI_R. The relative error and the algorithm employed for each output of 45 randomly chosen test data are listed in Table I.

One significant observation is that we can use an array size of up to 256×256 as training data to predict the test data results for array size of 512×512 (4×) and 1024×1024 (16×) test data (Fig. 6). The relative error between the prediction and simulation results is 8.43% using 'AdaBoost'. It suggests that we can use training algorithms with features extracted from device characteristics and smaller array simulation results to predict larger array simulation results.

V. Pareto Optima of 1S1R Crossbar Array

Using this methodology, we predict and analyze the array behavior for RRAMs with different LRS biased using various power supply voltages (V_{DD}). For each RRAM *(R_{ON} = 10 kΩ, 100 kΩ, 1 MΩ and 10 MΩ)*, there are envelopes of optimal performance, attainable by using the proper selector characteristics (Fig. 7-8), - in other words, a Pareto optimum is reached. The LRS and series resistance limit the voltage drop on the worst-case cell, leading to this optimum. The ranges of selector parameters are listed in Table II.

For each RRAM device with different LRS using the same V_{DD},

a) Larger read margin than each curve is not attainable (Fig. 7), assuming V_{READ} = *3 V* and *RRAM on/off ratio = 10*. With an ideal selector, read margin (>100 nA) is always reachable for Gbit array.

b) Larger write margin than each curve is not attainable (Fig. 8(a)), assuming V_{DD} = *5 V* and V_{MIN} = *4 V*. Maximum write margin for this case is $log(V_{DD}/V_{MIN}) = log(5/4) \approx 0.1$.

For each RRAM with different LRS under same access voltage V_{ACC} on the worst case cell,

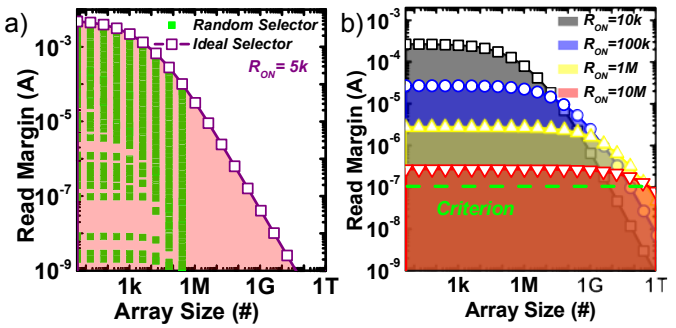

Fig. 7. (a) Envelope of read margin attainable for R_{ON} = 5 kΩ. The green symbols are the read margins of a range of randomly generated selectors. The purple symbols with line are the read margins of the optimal selectors. The pink-colored space marks the attainable read margin space by varying selector characteristics. (b) Envelope of read margin attainable by using optimal selector characteristics. For each RRAM with different LRS, the corresponding color space of read margin can be attainable by varying selector characteristics. R_{ON} is the LRS of RRAM.

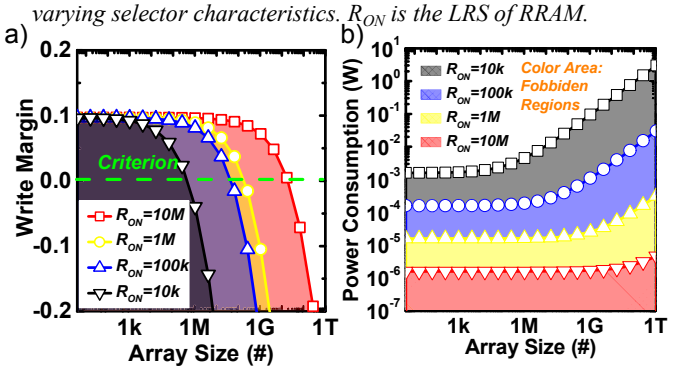

Fig. 8. (a) *Envelope of write margin attainable by using optimal selector characteristics. For each RRAM with different LRS, the corresponding color space of write margin can be attainable by varying selector characteristics.* (b) *Envelope of energy consumption not attainable by using optimal selector characteristics. For each RRAM with different LRS, the corresponding color space of energy consumption cannot be attainable by varying selector characteristics.*

c) Lower energy consumption than the curve is not attainable (Fig. 8(b)), assuming same access voltage on the worst-case cell V_{ACC} = *4 V*.

The results imply that even with an ideal selector, we still need to increase RRAM LRS (>10 MΩ) to guarantee proper operation of memory arrays greater than 10Gbit. Increasing RRAM LRS increases the write margin, maintains the read margin and lowers the energy consumption.

VI. Conclusion

A general methodology to characterize the behavior of a 1S1R cell/array based on machine learning regression algorithms is presented. The five-parameter characterization of selection devices and Verilog-A model of RRAMs are implemented into a circuit simulator to demonstrate the validity of this methodology. This methodology enables us to rapidly predict (less than one second with an 1.3 GHz Intel Core i5 personal laptop) of the 1S1R cell/array behaviors without running full circuit simulations. The prediction accuracy is validated for test data for up to Mb array. Machine learning allows us to avoid the use of enormous computing resources and is able to the performances of arrays that are more than 16× larger arrays (up to Mb) based on training using data from smaller arrays (Kb). Based on this method, the Pareto optima of array performances, write margin, read margin and energy consumption, are analyzed for different RRAMs *(R_{ON} = 10 kΩ, 100 kΩ, 1 MΩ, 10 MΩ)* by thoroughly searching the entire five-parameter design space for the 1S1R. This thorough search is enabled by the rapid prediction of the machine learning approach. The reported methodology is not limited to specific RRAM and selection device models and can therefore guide the design of numerous different 1S1R arrays in the future.

Table I. Relative Error and Prediction Algorithm

	Output	*Relative Error*	*Prediction Algorithm*
1S1R Cell	V_{MINS}	8.15%	'SVR'
	V_{MINR}	13.04%	'SVR'
	V_R*	±0.27 V (48.03%)	'AdaBoost'
	RW	12.42%	'AdaBoost'
1S1R Array	V_{ACCS}*+	± 0.18 V (N.A.)	'SVR'
	V_{ACCR}*+	± 0.11 V (N.A.)	'SVR'
	E_S	13.79%	'AdaBoost'
	E_R	8.04%	'SVR'
	δI_R	14.84%	'AdaBoost'

** The relative error is shown by voltage difference. The relative error of V_R, V_{ACCS} and V_{ACCR} in terms of percentage (in parenthesis) is large or not applicable (N.A.). Because the prediction still gives similar voltage difference (~± 0.2) at a smaller output voltage result, e.g. when V_R = 0.1 V, the prediction may give 0.3, where the relative error in terms of percentage is 300%.*

+ We use V_{ACC} (V_{ACCS} for SET and V_{ACCR} for RESET) for the prediction and later calculate write margin by eqn. (2).

Table II. Ranges of Selector Parameters

	Unit	*Range*
V_{th}	V	[0, 4]
I_{th}	log(A)	[-18, -4]
S_{th}	V/dec	[0.01, 0.2]
I_{on}	log(A)	[log10(I_{COMP}*), -0.5]
S_{on}	V/dec	[2, ∞)

** I_{COMP} is the compliance current of RRAM during SET programming.*

Acknowledgement: This work is supported in part by STARnet SONIC, NSF NEEDS, IARPA TIC, and member companies of the Stanford Non-Volatile Memory Technology Research Initiative (NMTRI).

Reference: [1] Jiun-Jia Huang, *et al.,* IEDM, 2011; [2] Euijun Cha, *et al.,* IEDM, 2013; [3] Jiyong Woo, *et al.,* VLSI, 2014; [4] Sung Hyun Jo, *et al.,* IEDM, 2014; [5] Leqi Zhang, Leqi, *et al.,* IEDM, 2014; [6] Shimeng Yu, *et al.,* IEDM, 2012; [7] Pinar Duygulu, *et al.,* Computer Vision—ECCV, 2002; [8] Fabrizio Sebastiani, ACM computing surveys (CSUR), 34.1, 2002; [9] Zizhen Jiang, *et al.,* SISPAD, 2014; [10] Harris Drucker, ICML. Vol. 97. 1997; [11] Chih-Chung Chang, *et al.,* ACM Transactions on Intelligent Systems and Technology (TIST), 2011; [12] Christopher M. Bishop, Springer, 2006; [13] Jiale Liang, *et al.,* IMW 2012.