Time-effective Fault Diagnosis Algorithms for Analog and Mixed-signal Circuits Using Sparsity-aware Multi-class Relevance Vector Machine

Qiwu Luo[1], *Member, IEEE*, Yigang He[1], *Member, IEEE*, Yichuang Sun[2], *Senior Member, IEEE*, and Lifen Yuan[1]
[1] School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, P. R. China
[2] School of Engineering and Technology, University of Hertfordshire, Hatfield, U.K.
E-mail: luoqiwu@hfut.edu.cn

***Abstract*—Except for the advantages of supporting arbitrary kernels, probabilistic predictions and automatic estimation of hyper-parameters, relevance vector machine (RVM) also encounters some of training time increase and classification accuracy recession, compared with SVM. In order to suppress such 'nuisance' imperfections, this paper proposed a sparsity-aware RVM model for multi-class classification (denoted as Sa-MRVM) by developing a configurable singular entropy decision mechanism. Multiple driven data sets captured from both emulational and actual circuits under test (CUTs) are involved to further improve the model's generalization ability and judging confidence. Experimental results carried out on two CUTs indicate that our proposed learning methodology is speedy and accurate enough for real world fault diagnosis tasks of analog and mixed-signal circuits.**

Keywords—Fault diagnosis; analog and mixed-signal circuits; singular entropy; RVM; supervised learning

I. INTRODUCTION

Fault diagnosis of analog and mixed-signal (AMS) circuits is going to face the daunting tasks of managing the increasingly large-scale integration and handling omnipresent component tolerance [1]. And reliable health monitoring and rapid fault disposal to electronic products will provide better user experience. Take retail industry for instance, if we can solve failures of electronic article surveillance (EAS) devices in time, much more time and effort will be saved for both EAS device suppliers and retailers [2, 7]. These latest situations bring challenges to the fault detection and diagnosis of AMS circuits.

At present, rapidly growing machine learning is addressing the issue of how to build self-improved computers (or embedded processors, reconfigurable units) through experience [3, 4]. Then, a great number of supervised learning algorithms and classification theories emerge as the times require. Zhang *et al.* [5] proposed an online evaluation strategy (MKALSSVR) for analog circuits. The simulation results carried out from a typical Sallen-Key low-pass filter circuit proved its impressive performance and fast speed contrasted with LSSVR and ε-SVR. However, it is required to carefully set three crucial model parameters: error insensitive zone ε, penalty factor C, and kernel specific parameter γ. To solve this issue, Chen *et al.* [6] imported double chains quantum genetic algorithm to improve the classical SVM model. The DCQGA-SVM achieved better classifying accuracy and runtime cost when applies on two small-scale analog CUTs, compared with SVM, PSO-SVM, and QGA-SVM. These SVM-based models are driven by purified simulation sample sets, and with a shared goal to develop cross-iterative algorithms for pursuing model parameters rapidly and aggressively.

In this paper, extended the supervised learning algorithms in [6] and diagnostic methodologies in [7], a sparsity-aware relevance vector machine model for multi-class classification is proposed based on configurable singular entropy decision theory (CSEDT), we denote this model with 'Sa-MRVM' for short. The main contributions are as follows.

- Compared with state-of-the-art SVM-based models, the universally recognized slight recession on accuracy of RVM model has been compensated to an acceptable extent, since the mixed data samples acquired from both simulation circuits and real-world CUTs possess higher information entropy.
- Compared with classical MRVM-based models, the Sa-MRVM requires smaller time overheads on both training and diagnosis, since the CSEDT-based principal component extraction process has dramatically reduced the data scale.

The rest of the paper is organized as follows. Section II declares the basic theory of the Sa-MRVM. Section III presents the diagnostic actualizing procedures. Section V demonstrates the experimental results for the CUTs illustrated in Section IV. Finally, Section VI concludes the paper.

II. THEORY OF PROPOSED SA-MRVM

A. Parameter selection and feature extraction

Given an AMS circuit under a certain frequency, accessible node voltages, amplitude-frequency (A-F) responses and phase-frequency (P-F) responses constitute an $1\times N$ circuit parameter vector, when considering wide band circuit response under P acquisition frequency points, the raw data matrix is $A_{N\times P}$ [7]. The main challenge in this paper is how to efficiently diagnosis arbitrary circuit faults based on this huge three-dimensional raw matrix (if data collection is repeatedly issued for reducing systematic errors) with a considerable accuracy.

B. Multi-class Relevance Vector Machine (MRVM)

The RVM, a sparse Bayesian extension of SVM, was first proposed by Tipping in [8]. It then was applied for multi-class classification in [9]. Here we briefly review MRVM as follows.

978-1-5386-4881-0/18/$31.00 ©2018 IEEE

Given a raw input and its corresponding target data set pairs $\{\mathbf{x}_k, t_k\}_{k=1}^{K}$, if only consider scalar-valued target functions and follow the standard probabilistic formulation, the model of classical RVM is the same as that of SVM.

$$t_k = y(\mathbf{x}_k;\mathbf{w}) = \mathbf{w}^T \phi_k(\mathbf{x}) + \mathbf{e}_k \tag{1}$$

where $\mathbf{w}=[w_1, w_2, ..., w_K]^T$ is the adjustable weight vector, $\phi(\mathbf{x})=[\phi_1, \phi_1, ..., \phi_K]^T$ is a non-linear function which can map input data $\mathbf{x}$ into a higher dimensional space, and $\mathbf{e}_k=\{e_k\}_{k=1}^{K}$ is the independent noise vector. If consider RVM for classification, then $\mathbf{e}_k=\mathbf{0}$, the target data set is rewritten as

$$t_k = \mathbf{w}^T \phi_k(\mathbf{x}) = \sum_{k=1}^{K} w_k \kappa(\mathbf{x}, \mathbf{x}_k) \tag{2}$$

where $\kappa(\mathbf{x}, \mathbf{x}_k)$ is a materialized transformation of $\phi_k(\mathbf{x})$, and it is always called kernel function.

The likelihood function in the above analysis obeying Bernoulli distribution

$$P(\mathbf{t}|\mathbf{w}) = \prod_{k=1}^{K} \sigma\{y(\mathbf{x}_k;\mathbf{w})\}^{t_k} \left[1-\sigma\{y(\mathbf{x}_k;\mathbf{w})\}\right]^{1-t_k} \tag{3}$$

where t is defined as class labels, $\sigma\{y(\mathbf{x};\ \mathbf{w})\}$ is the sigmoid function $1/(1+e^{-y(x,w)})$. Then, several iterative operations are started to find the weight vector $\mathbf{w}$ and the hyper-parameters $\boldsymbol{\alpha}$, the set of non-zero weights in $\mathbf{w}$ are defined as the relevance vectors in final. This methodology can be extended to multi-class classification through generalizing (3) to

$$P(\mathbf{t}|\mathbf{w}) = \prod_{k=1}^{K}\prod_{m=1}^{M} \sigma\{y_m(\mathbf{x}_k;\mathbf{w}_m)\}^{t_{km}} \tag{4}$$

where M is defined as the number of classes, t_{km} is the indicator variable for the case k to be one member of class m, and y_m is the predictor for the class m [8, 9]. One concern about this RVM-based multi-class classification needs a great quantities binary classifications to be undertaken. Imported the principles of multi-nominal logistic regression, (4) is then alternated as

$$P(\mathbf{t}|\mathbf{w}) = \prod_{k=1}^{K}\prod_{m=1}^{M} \sigma\{y_m; y1, y2, \cdots, y_M\}^{t_{km}} \tag{5}$$

where the class predictors are coupled in the multinomial Logit function $\sigma(y_m;\ y_1,\ y_2,\ ...,\ y_M) = e^{-y_m}(e^{y_1}+...+e^{y_M})$. Solving this problem requires the specification of the priors associated with the hyper-parameters, $\boldsymbol{\alpha}$, which have been obtained in [8]. The abovementioned theories form the basis of the MRVM focusing on multi-class classifications.

C. Sparsity-aware Multi-class Relevance Vector Machine

Supposing that the raw sample vector set $A_{N\times P}$ contains P input vectors $\varphi(x_j)$, and each input vector $\varphi(x_j)$ contains N elements, this raw sample set $A_{N\times P}$ can be expressed as $A_{N\times P} = [\varphi(x_1), \varphi(x_2),..., \varphi(x_N)]$. Hence there are r base vectors $\{\varphi(v_1), \varphi(v_2),..., \varphi(v_r)\}$ in the sample vector set. Any input vector $\varphi(x_j)$ which belongs to the sample vector $A_{N\times P}$ can be linearly represented by the base vector set. However, the raw base vectors inevitably involve redundant information, they would trigger unwanted computation overheads. In this work, the configurable singular entropy decision theory (CSEDT) [10] is imported to obtain the principal component of base vector set. The vector set $A_{N\times P}$ then can be decomposed by singular value decomposition (SVD) as

$$A_{N\times P} = U_{N\times r} M_{r\times r} V_{r\times P}^T \ ,\ M_{r\times r} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{pmatrix} \tag{6}$$

where $U_{N\times r}$ and $V_{r\times P}$ are the orthogonal left and right singular matrix of $A_{N\times P}$, respectively. The descending singular values λ_i, i=1, 2, ..., r, ($\lambda_1 > \lambda_2 > ... > \lambda_i$) of $A_{N\times P}$ constitute the main diagonal elements of $M_{r\times r}$.

As a basic concept of information theory, entropy is employed to evaluate the mean information content, the singular entropy can be calculated as

$$E_T = \sum_{t=1}^{T} \Delta E_t, T \le r \tag{7}$$

where T is the order of singular entropy, the ΔE_t is the incremental value of that singular entropy at the order of t. It can be calculated as

$$\Delta E_t = -(\lambda_t / \sum_{k=1}^{r} \lambda_k)\log_2(\lambda_t / \sum_{k=1}^{r} \lambda_k) \tag{8}$$

where λ_t is the t^{th} singular value of raw matrix $A_{N\times P}$.

Consequently, all the raw training data can be represented by those base vector sets in a sparser manner. In addition, the vector set size can be flexibly adjusted through determining the order of singular entropy E_T. Compared with the original MRVM, our improved Sa-MRVM possesses better sparsity, selectivity, and adaptability.

III. FAULT DIAGNOSIS PROCEDURES

The Sa-MRVM-based fault diagnosis procedures of AMS circuits are demonstrated as follows, and the corresponding flow-chart is given in Fig. 1.

Step 1. Data acquisition. Collect massive raw data involving abundant fault information of CUTs. For a particular kind of fault in $\{C_m\}_{m=1}^{M}$, $N\times P$ circuit parameters of accessible node voltages, A-F responses and P-F responses under P discrete scanning frequency points are gathered by our previous portable diagnostic equipment designed in [7], and then be transmitted to host PC merging with the circuit simulation data with the same dimensions. In order to reduce testing error, each group of data are acquired for repeated Z times.

Step 2. Normalization pre-processing. In order to ensure that the fault feature parameters with different circuit attributes fairly participate in classifier training, all of them are normalized to an equivalent dimension.

Step 3. CSEDT-based principal component extraction. This is the *first* round of sparsification aiming at data input of classifier. Most of the redundant data unconsciously expanded in Step 1 will be discriminatively eliminated through the singular entropy decision theory described in Sec. II.C.

Step 4. MRVM-based classification training. This is the *second* round of sparsification supported by classifier model itself. Referred to the classical RVM theory in [8], the sparse Bayesian learning methodology is adopted for M multi-fault-class classifications in this paper, and the hyper-parameters are obtained from the iterative formula: $\alpha_k^{new} = (1 - \alpha_k \Sigma_{k,k}) / \arg\max_{w} p(w \mid t, \alpha_k)$.

Step 5. Sa-MRVM-based classification simulation. Various training samples collected from the current classical diagnostic circuit cases will be applied to train the Sa-MRVM model for its performance evaluation and optimization. The Step 3 and Step 4 would be accordingly repeated to search an appropriate tightness of singular entropy order T for a considerable classification performance.

Step 6. Sa-MRVM-based diagnosis testing. The classification learning model with auto-determined parameters $\boldsymbol{\alpha}$ and T is implemented as IP core then embedded on FPGA for realistic fault diagnosis task.

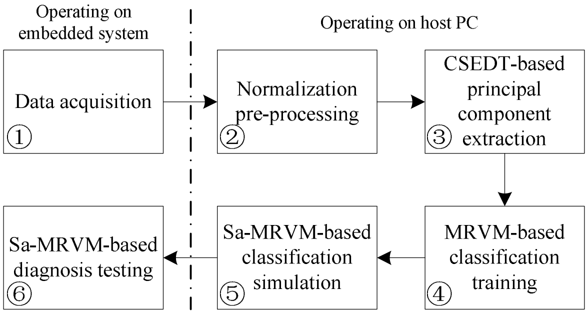

Fig. 1. The flow chart of fault diagnosis based on Sa-MRVM.

It is worth mentioning that the first step and final step are running on the embedded diagnosis system with real-time requirement, while the intermediate four steps are operating on a mobile personal computer and the time-effectiveness could be loosely required.

IV. ILLUSTRATIVE CIRCUIT UNDER TEST AND FAULT SET

A. Universal Analog CUT 1: Salen-Key Band-pass Filter

The CUT 1, a purified analog circuit, is a simple two-port circuit network. As shown in Fig. 2, the input and output signal can be easily gathered on test points TP1 and TP2, the nominal values and chip model of the components are also given. Suppose only one of the components of C1, C2, R2 and R3 degrades to erroneous mode in each faulty situation, the corresponding fault classes are summarized in TABLE I, where the symbols of ↓ and ↑ imply significantly higher and lower than the nominal values by 20%, respectively.

Apart from the Monte Carlo analysis on PSPICE simulation in [1, 5, 6], our proposed methodology supports the post-testing sample acquisition from the realistic CUT. Consequently, the training data set will be more abundant and objective under this approach combining pre-simulation and post-testing. However, it will take more time to collect the actual sample. Fortunately, the automatic fast frequency-sweeping method set up in [7] can greatly improve the efficiency of sample collection.

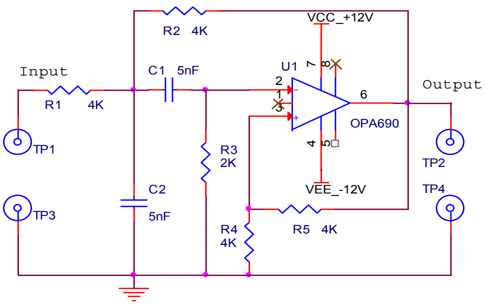

Fig. 2. Analog CUT of the Sallen-Key bandpass filter.

TABLE I. FAULT CLASSES OF CUT 1

Fault Code: *Fi*	**Component**	**Standard Value**	**Tolerance (%)**	**Fault Value**
F0	Normal	/	/	/
F1	C1 ↓	5.00 nF	5	2.50 nF
F2	C1 ↑	5.00 nF	5	10.00 nF
F3	C2 ↓	5.00 nF	5	2.50 nF
F4	C2 ↑	5.00 nF	5	10.00 nF
F5	R2 ↓	3.00 kΩ	1	1.50 kΩ
F6	R2 ↑	3.00 kΩ	1	6.00 kΩ
F7	R3 ↓	2.00 kΩ	1	1.00 kΩ
F8	R3 ↑	2.00 kΩ	1	4.00 kΩ

B. Actual AMS CUT 2: Multi-Order Band-pass Filter

The DSP controller is the most crucial part in a flagship AM-EAS system, which belongs to a typical analog/mixed-signal circuit. For further investigation of the performance and fair comparison, the previous quad-opamp 4th order band-pass filter in [7] is selected as the second CUT with the same fault settings. This CUT contains 32 OPAMPs, 104 resistances, 56 capacitances, 8 digital control switches, and 8 Schottky diodes.

V. SIMULATION, APPLICATION AND DISCUSSION

A. Algorithm Simulation for CUT 1 and Result Discussion

Applying the evaluating terminologies in [11], Table II summarizes the true-positive rate (TPR), model training time and average diagnosis time. It is observed that the TPRs of classifier models based on DCQGA-SVM, MRVM and Sa-MRVM are better than that of interval math model, which illustrate that learning models are more intelligent than the latter one relying only on direct data analysis and feature alignment. In general, the diagnosis accuracy of MRVM should be a little bit lower than that of DCQGA-SVM due to its slightly contracted data integrity resulted from the sparse Bayesian model [4]. The fact that the figure of Sa-MRVM driven with more sparse data input is lower than that of MRVM has proved this standpoint. However, and surprisingly, the TPR of MRVM is slightly improved than that of DCQGA-SVM, the newly introduced 50 of actual samples are the main contributors to this deteriorated-accuracy compensation, while DCQGA-SVM in [6] was only driven by the simulation data.

The system time consumption contains two parts: training time on PC and diagnosis time on FPGA. The former is mainly consumed on the iterative computation to obtain the parameters

978-1-5386-4881-0/18/$31.00 ©2018 IEEE

of learning model. The latter pays for error-cost assessment and data migration among different memories. Because the training time for pursuing model parameters is cost only once on PC while the classifying time is repeatedly required in fault diagnosis applications, we consider more on the latter time overhead. Benefited from the ultra-sparse data input, our Sa-MRVM performs better than DCQGA-SVM and MRVM, even tends to catch up with the simplest interval math model.

TABLE II. PERFORMANCE COMPARISON OF FOUR DIFFERENT METHODS APPLYING ON CUT 1

Classifier Model	TPR (%)	Training Time [a] (s)	Average Diagnosis Time (s)
DCQGA-SVM [6]	97.4074	3840 [b]	4.12 [d]
Interval Math [7]	96.6862	1330 [b]	1.68 [e]
MRVM	97.9623	14860 [c]	2.39 [e]
Sa-MRVM [f]	97.1039	2640 [c]	1.82 [e]

[a] PC: MacBook Pro with Intel Core i7 (2.9 GHz), 16GB 2133MHz LPDDR3.

[b] Only consider time consumptions based on Monte-Carlo analysis (1000 samples) from PC.

[c] Consider time consumptions based on both Monte-Carlo analysis (1000 samples) from PC and actual acquired data (50 samples) from FPGA.

[d] Time consumption on PC.

[e] Time consumption on FPGA.

[f] The order of singular entropy T was determined as 0.86.

B. *Algorithm Application on CUT 2 and Result Evaluation*

Due to the cross-iteration of hyper-parameters, DCQGA-SVM model involves biggish diagnosis time. Thus, we omit it in following application tests. As illustrated in TABLE III, compared with the interval math theory, the extra 1.2 s of time cost has obtained about 1 percent of TPR increase. Compared with MRVM, although the TPR of our Sa-MRVM is slightly lower, but it performs much more competitive records on time-effectiveness (2.15 s *vs*. 9.47 s), which is mainly benefited from the CSEDT scheme. Such performance could be accepted by average EAS device suppliers and retailers. Another advantage is that the average confidence interval of Sa-MRVM is more concentrated than that of MRVM, which indicates that the sparsity process also weakens the noise interferences implicitly. Since the confidence level is inherited from the original RVM model, the interval math model cannot support this function.

TABLE III. PERFORMANCE COMPARISON OF THREE DIFFERENT METHODS APPLYING ON CUT 2

Classifier Model	TPR (%)	Average Diagnosis Time (s)	Average Posterior Probability [g] (%)	
			Worst	*Best*
Interval Math [7]	95.7800	0.98	/	/
MRVM	97.0746	9.47	67.6941	98.6612
Sa-MRVM	96.8942	2.15	87.8587	95.3447

[g] Record the posterior probability of each considered circuit fault then calculate their average value.

VI. CONCLUSION

Online fault diagnosis of large-scale analog and mixed-signal circuits is an art of pursuing performance balance between classifying accuracy, time efficiency, resource overhead, and serviceability. This paper proposed a sparsity-aware MRVM classification model improved by singular entropy technology, which is named as Sa-MRVM. Experimental results from two CUTs with incremental circuit scale show that the proposed Sa-MRVM diagnostic algorithm has balanced performances among classifying accuracy, convergence rate of training, and diagnosis time cost. Furthermore, the Sa-MRVM can provide posterior probability estimation of class membership. Even so, this work is only a preliminary study, more practical tests are waiting to be executed for further improvements of Sa-MRVM classification algorithms.

ACKNOWLEDGMENT

This work was jointly supported by the National Natural Science Foundation of China under Grant No. 51704089, 51577046, 51637004, the Anhui Provincial Natural Science Foundation of China under Grant 1808085QF190, the China Postdoctoral Science Foundation under Grant 2017M621996, the Ph.D Special Research Fund of HFUT under Grant JZ2016HGBZ1030, and in part by the national key research and development plan "Important Scientific Instruments and Equipment Development" under Grant 2016YFF0102200.

REFERENCES

[1] L. Yuan, Y. He, and Y. Sun, "A new neural-network-based fault diagnosis approach for analog circuits by using kurtosis and entropy as a preprocessor," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 3, pp. 586-595, Mar. 2010.

[2] E. Bottani, G. Ferretti, R. Montanari, A. Rizzi, and A. Volpi, "Performances of RFID, acousto-magnetic and radio frequency technologies for electronic article surveillance in the apparel industry in Europe: A quantitative study," *I. J. RF Technol.: Res. and Appl.*, vol. 3, no. 2, pp. 137-158, 2012.

[3] MI. Jordan and TM. Mitchell, "Machine learning: trends, perspectives, and prospects," *Science*, vol.349, no. 6245, pp. 255-260, Jul. 2015.

[4] J. Luo, CM. Vong, and PK.Wong, "Sparse bayesian extreme learning machine for multi-classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 4, pp. 836-843, Apr. 2014.

[5] A. Zhang, C. Chen, and HR. Karimi, "A new adaptive LSSVR with online multikernel RBF tuning to evaluate analog circuit performance," *Abstr. Appl. Anal.*, vol. 2013, no. 9, pp. 1-7, Dec. 2013.

[6] P. Chen, L. Yuan, Y. He, and S. Luo, "An improved SVM classifier based on double chains quantum genetic algorithm and its application in analogue circuit diagnosis," *Neurocomputing*, vol. 211, pp. 202-211, Oct. 2016.

[7] Q. Luo, Y. He, and Y.Sun, "Real-time fault detection and diagnosis system for analog and mixed-signal circuits of acousto–magnetic EAS devices," *IEEE Des. Test*, vol. 33, no. 3, pp. 77-90, Jun. 2016.

[8] ME. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, no. 3, pp. 211-244, Apr. 2001.

[9] H. Zhang and J. Malik, "Selecting shape features using multi-class relevance vector machine," Technical Report in UCB, No. UCB/EECS-2005-6, pp. 1-11, Oct. 2005.

[10] H. Yang, J. Chen, and S. Su, "Noise reduction for chaotic time series based on singular entropy," in *IEEE ISCID 2013*, pp. 216-218, Hangzhou, China, Oct. 2013.

[11] ASS. Vasan, B. Long, and M. Pecht, "Diagnostics and prognostics method for analog electronic circuits," *IEEE Trans. Ind. Electron.*, vol. 60, no. 11, pp. 5277-5291, Nov. 2013.