Indirect test of RF circuits using ensemble methods

Hassan El Badawi, Florence Azaïs, Serge Bernard, Mariane Comte, Vincent Kerzérho, François Lefèvre

Indirect test of RF circuits using ensemble methods

H. El Badawi [1,2], F. Azais [1], S. Bernard [1], M. Comte [1], V. Kerzérho [1], F. Lefevre [2]
[1] LIRMM, University of Montpellier, CNRS, 161 rue Ada, 34095 Montpellier Cedex, France
[2] NXP Semiconductors, 2 Esplanade Anton Philips, 14000 Caen, France

***Abstract* — The adoption of indirect test for analog and RF integrated circuits (ICs) can tackle the rising costs of the classical industrial testing of these circuits, hence relaxing the requirements on test equipment. Based on machine-learning techniques, the concept of indirect test is to create a mapping between an indirect and low-cost measurement and the performance of the circuit by training a regression model. In this work, we explore the potential benefit of using ensemble learning. Instead of using a single regression model to predict the performance, the use of ensemble learning consists of combining multiple regression models to enhance the model's generalization. Different ensemble methods based on bagging, boosting or stacking are investigated and compared to classical individual models. Results are illustrated and discussed on three RF performances of a LNA for which we have production test data.**

Keywords: indirect testing, RF integrated circuits, machine-learning algorithms, ensemble methods, test efficiency

I. INTRODUCTION

Process variation and manufacturing imperfections could lead to performance degradation or even compromise the device functionality, thus IC manufacturers have to test every fabricated device in order to guarantee its functionality and ensure product quality. This testing task has a significant impact on the total cost of the finished product. In case of analog and RF circuits, this impact is stronger than for digital circuits because the testing process relies on the use of expensive and sophisticated test equipment in order to verify the different device specifications. To reduce the testing cost, researchers have investigated the implementation of an indirect test strategy. Essentially, the aim is to replace the conventional specification measurements by some low-cost Indirect Measurements (IMs). The idea is then to use regression models that establish the correlation between the two types of measurements. These kinds of predictive models are generally built using machine learning algorithms. This concept of indirect testing was first introduced for analog circuits [1], then extended to RF circuits [2]. Different aspects have been analyzed, such as the choice of the prediction model, the test stimulus choice, the processing of complex signatures, the use of embedded sensors [2,3] and multi-Vdd test conditions [4], or the selection of appropriate IM [5]. In this work, the objective is to investigate novel prediction models built with ensemble learning methods, which have been previously implemented in other application domains and have shown improved generalization ability.

This abstract is organized as follows. Section II summarizes the basics of the indirect test approach. Section III gives an overview on the used algorithms and the metrics to evaluate prediction model performances. Finally, Section IV presents the case studies and a summary of obtained results.

II. INDIRECT TEST PRINCIPLE

The underlying idea of indirect testing is that process variations that affect the device performance also affect indirect parameters. If the correlation between the indirect parameter space and the specification space can be established, then specifications may be verified using only the low-cost indirect signatures. Unfortunately, the relation between these two sets of parameters is complex and cannot be simply identified with an analytic function. The solution commonly implemented uses machine-learning algorithms. The indirect test synopsis is split into two distinct phases, namely training and production testing, as illustrated in Figure 1.

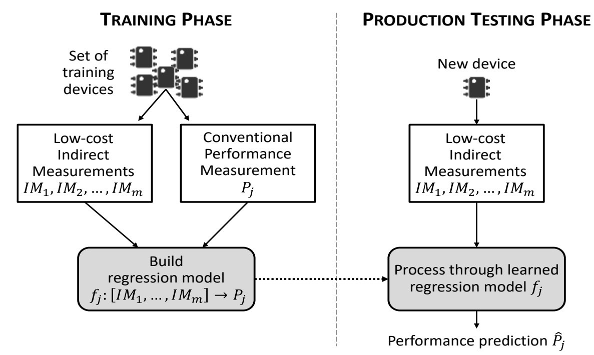

Fig.1. Indirect test synopsis.

The idea is to learn during the training phase the unknown dependency between the low-cost indirect measurements (IMi) and the conventional performance measurements (Pj). To achieve this, both the specification tests and the low-cost measurements are performed on a set of training devices and a machine-learning algorithm is trained to build regression models that map the indirect parameters space to the performance parameters space. During the production testing phase, only the low-cost indirect measurements are performed, and the specifications of every new device are predicted using the mapping learned in the initial training phase.

III. PREDITCION MODELS AND MODEL EVALUATION

Numerous regression models exist and are used in the context of indirect test. Traditionally, a single regression model is built to predict the performance on new instances. Multiple Linear Regression (MLR), Multi-Adaptive Regression Splines (MARS), and Support Vector Machine (SVM) are the mostly used algorithms. Nonetheless, some limitations or lack of generalization can arise from using a single regression model where performances can also differ with each case study.

To face the limitations of a single predictor, researchers have started to investigate the use multiple regression models and try to aggregate their outcomes to get the final prediction values. The idea is that with an appropriate combination of diverse individual models, it should be possible to exploit the strengths and overcome the weaknesses of the individual models and obtain better overall predictive performance. This approach is called ensemble learning, which refers to the procedures used to train multiple individual regression models (base learners) and combine their outputs in order to improve the stability and the predictive power of the ensemble model. Numerous methods for constructing ensemble models have been proposed in the literature [6], which includes parallel and sequential methods, based either on a single type of base learners (homogenous ensemble model) or learners of different types (heterogeneous ensemble model). The three most used methods are Bagging, Boosting, and Stacking. In this work, we have investigated one model built with Bagging, two models built with Boosting (AdaBoost and Gradient Boosting), and two models built with Stacking (Classical and Classical + Random Forest).

Finally, to be able to evaluate the different prediction models, we should base our judgment on evaluation metrics. The most commonly-used metric is the coefficient of determination R^2, which is a measure of the goodness of fit of a model. This score is computed as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

where y_i is the actual performance value of the ith instance, $\hat{y}_i$ is the predicted performance value of the ith instance, and n is the number of instances.

Another metric has been suggested in [5], which permits to quantify the prediction reliability of a model. This metric, called Failing Prediction Rate (FPR), expresses the percentage of circuits with a prediction error that exceeds the conventional measurement uncertainty ε_{meas}:

$$FPR = \frac{1}{n} \sum_{i=1}^{n}(|y_i - \hat{y}_i| > \varepsilon_{meas}) \quad (2)$$

with $(|y_i - \hat{y}_i| > \varepsilon_{meas}) = 1$ if true
$(|y_i - \hat{y}_i| > \varepsilon_{meas}) = 0$ otherwise

Lastly, if the test limits are available, we can compute another metric called the Misclassification Rate (*MR*). This metric simply expresses the ratio of misclassified circuits with respect to the total number of circuits.

IV. Case Study and Results

The test vehicle is a Low-Noise Amplifier (LNA) for which we have production test data on 3,850 devices. More precisely, test data include the conventional measurements of three RF performances, namely the gain, the output power at 1dB compression point (P1dB) and the third-order intercept point (IP3). Test data also include 79 low-cost indirect measurements which correspond to DC voltages on internal nodes (the device is equipped with an internal DC bus and internal DC probes) and DC signatures delivered by built-in process monitors. To perform the training phase, we have sampled 2000 of the initial test data by using Latin Hypercube Sampling (LHS), and the remaining circuits (1850) were used for result evaluation.

Eight types of model have been trained and evaluated, varying the number of selected IMs between 1 to 15. Three models are single ones built with classical methods (MLR, MARS, SVM); Five models have been trained using an ensemble method (as specified in section III). A summary of results is presented in Table I, which shows for each RF performance the best model obtained using either a classical method or an ensemble method.

TABLE I. Comparison between classical and ensemble methods: Summary of best results for the three RF performances

	Best solution selected from max(R^2) on validation set					
	RF Perf	***Model***	***R^2 (*)***	***FPR (*)***	***MR (*)***	***# feat.***
Classical method	Gain	MARS	0.65	2.86%	0%	9
	P1dB	SVM	0.85	12.32%	0.1%	8
	IP3	SVM	0.93	0.59%	4.2%	14
Ensemble method	Gain	Stack+RF	0.72	1.51%	0%	9
	P1dB	Stack+RF	0.87	11.24%	0.1%	12
	IP3	Stack+RF	0.94	0.70%	4.2%	14

A first comment is that the type of model that gives the best solution differs depending on the RF performance in case of classical methods (MARS or SVM), while it is always the same type of model that leads to the best solution in case of ensemble methods (Stack + Random Forest). It is an interesting feature to have a solution able to handle a variety of different situations. Then the second comment is that this type of ensemble model leads to an accuracy improvement for the three RF performances compared to the best classical models. However, the level of improvement is different for each RF performance. In particular, it can be observed that the lower the accuracy reached by the best classical model on a given RF performance, the stronger the improvement obtained with the ensemble model. Globally, the use of ensemble models built with stacking appears to be an interesting option.

V. Conclusion

In this work, we have explored the use of ensemble methods for indirect test of RF circuits. Different ensemble methods based on bagging, boosting and stacking have been investigated and compared to classical individual models. Results have shown the superiority of ensemble models built with stacking compared to other ensemble models. Results have also shown that such models can outperform the classical individual models, both in terms of accuracy and reliability, and that they offer a superior predictive power over a variety of different situations.

Acknowledgment

This work has been carried out under the framework of PENTA-EUREKA project 16003 “HADES”.

References

[1] P. N. Variyam et al., “Prediction of analog performance parameters using fast transient testing,” IEEE Trans. On Computer-Aided Design of Integrated Circuits and Systems, Vol. 21, no. 3, pp. 349–361, 2002.

[2] S. Ellouz et al., “Combining internal probing with artficial neural networks for optimal RFIC testing”, Proc. IEEE Int’l Test Conference (ITC), p.9, 2006.

[3] L. Abdallah et al., “Sensors for built-in alternate RF test,” Proc. IEEE European Test Symposium (ETS), pp. 49-54, 2010.

[4] M.J. Barragan, et al., “Improving the Accuracy of RF Alternate Test Using Multi-VDD Conditions: Application to Envelope-Based Test of LNAs”, Proc. IEEE Asian Test Symposium (ATS), pp. 359-364, 2011.

[5] S. Larguech, et al., “Evaluation of indirect measurement selection strategies in the context of analog/RF alternate testing,” Proc. Latin American Test Workshop, p.6, 2014

[6] Z-H. Zhou, “Ensemble Methods: Foundations and Algorithms”, Chapman and Hall/CRC, Machine Learning & Pattern Recognition Series, 2012.