Fast Deployment of Alternate Analog Test Using Bayesian Model Fusion

John Liaperdos*¶, Haralampos-G. Stratigopoulos†‡, Louay Abdallah†‡,
Yiorgos Tsiatouhas§, Angela Arapoyanni¶, and Xin Li‖
*Technological Educational Institute of Peloponnese, Department of Computer Engineering, Sparta 23100, Greece
†Université Grenoble Alpes, TIMA, F-38000 Grenoble, France
‡CNRS, TIMA, F-38000 Grenoble, France
§University of Ioannina, Department of Computer Science & Engineering, Ioannina 45110, Greece
¶National and Kapodistrian University of Athens, Department of Informatics and Telecommunications, Athens 15784, Greece
‖Carnegie Mellon University, ECE Department, Pittsburgh, PA 15213, USA

Abstract— **In this paper, we address the problem of limited training sets for learning the regression functions in alternate analog test. Typically, a large volume of real data needs to be collected from different wafers and lots over a long period of time to be able to train the regression functions with accuracy across the whole design space and apply alternate test with high confidence. To avoid this delay and achieve a fast deployment of alternate test, we propose to use the Bayesian model fusion technique that leverages prior knowledge from simulation data and fuses this information with data from few real circuits to draw accurate regression functions across the whole design space. The technique is demonstrated for an alternate test designed for RF low noise amplifiers.**

I. Introduction

The current industry practice for production testing of analog circuits is specification testing, wherein all the performances promised in the datasheet are measured one by one and, subsequently, are compared to their specification limits to reach a pass or fail decision. This approach involves a high cost since it requires automatic test equipment with advanced features and it takes up significantly long test times. For many Systems-on-Chip and Systems-in-Package it has been reported that testing the analog functions is responsible for up to 50% of the total test cost, despite the fact that the analog circuits occupy less than 5% of the total die area.

The alternate test paradigm has been proposed as a replacement of the conventional specification tests with the aim to largely simplify the test procedure and reduce the test cost [1], [2]. The underlying idea is to apply appropriate test stimuli and extract alternate measurements from which we can infer implicitly the performances. The low cost stems from the fact that a single test configuration that employs DC or low-frequency test stimuli and a single acquisition of alternate DC or low-frequency alternate measurements suffice to predict accurately all the performances. This has been demonstrated for different types of analog circuits, including baseband analog [1], [3], RF [1], [2], [4]–[6], data converters [7], [8], and PLLs [9].

The feasibility of alternate test lies in the fact that both the performances and alternate measurements are subject to the same process variations. This implies that alternate measurements can be selected such that they are highly correlated to the performances. In this way, variations in the performances and, specifically, excursions outside their specification limits, can be tracked implicitly through the variations in the alternate measurements. With this in mind, the alternate test can only be applied to circuits that exhibit process variations and a defect filter is required to screen out circuits that fail due to defects, before they are actually forwarded to alternate test [10].

The first step in alternate test is to identify information-rich alternate measurements. This is a circuit-specific problem since the input, output, frequency band, transfer function, etc., depend on the type of the analog circuit, as well as on its architecture. It is also a very challenging problem since the large number of process parameters and their intricate interactions make it impossible to argue qualitatively that an alternate measurement captures all variation scenarios that can occur. The best practice is to select a test stimulus and associated alternate measurements and optimize directly the test stimulus until the accuracy of the alternate test becomes satisfactory [1], [11]. Very often, however, the alternate measurements are extracted *ad hoc* without a specific rationale and are shown to be effective only experimentally. A typical approach is to identify as many alternate measurements as possible and then compact this large set using feature selection algorithms [12]–[15].

The intricate relationship among performances and alternate measurements makes it impossible to build the mapping in the form of a closed-form mathematical relationship. For this reason, the mapping is built through statistical regression. Different regression tools can be employed for this purpose, including polynomial regression, multivariate adaptive regression splines (MARS), feed-forward neural networks, support vector machines (SVMs), etc. [16], [17].

The regression models that map the alternate measurements to each of the performances are learned by employing a training set of circuits. At the initial phase of the alternate test deployment, the training set is not fully representative of the fabrication process and alternate test decisions are unavoidably prone to error for outlier circuits that lie towards the tails of the circuit distribution. In this paper, we employ the Bayesian model fusion (BMF) technique [18]–[21] with the aim to learn regression models that are valid across the circuit distribution

978-3-9815370-4-8/DATE15/©2015 EDAA

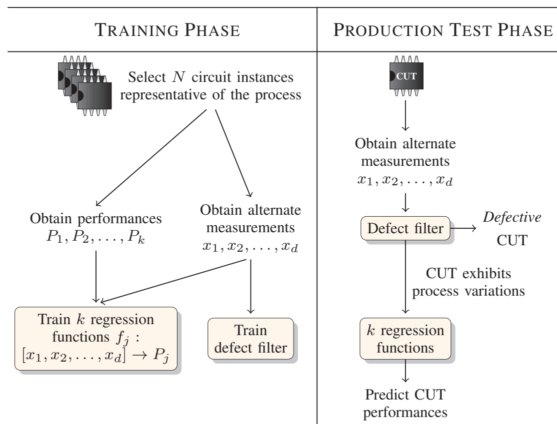

Fig. 1. Alternate test flow.

right at the onset of production without needing to hold off to collect a representative training set before we can fully trust the alternate test decisions. The underlying idea is to learn the regression functions by employing in addition to the real data prior information from post-layout simulation.

The rest of the paper is structured as follows. In Section II, we provide a brief and concise overview of alternate test. In Section III, we discuss the motivation behind this work. In Section IV, we discuss in detail the BMF learning procedure. In Section V, we demonstrate the BMF learning procedure in the case of an alternate test approach for a low noise amplifier (LNA) and we compare it to the conventional learning procedure. Finally, Section VI concludes the paper.

II. Brief Overview of Alternate Test

The alternate test flow is illustrated in Fig. 1. Let $P_1, P_2, \ldots, P_k$ denote the k performances of the circuit under test (CUT) that need to be determined during the conventional specification test approach. Let also $\boldsymbol{x} = [x_1, x_2, \ldots, x_d]$ denote a pattern of d low-cost alternate measurements. The mapping between $\boldsymbol{x}$ and performance P_j denoted by f_j

$$f_j \,:\, \boldsymbol{x} \rightarrow P_j,\; j = 1, 2, \ldots k \tag{1}$$

is learned through regression modeling in an off-line, preparatory training phase that employs N circuits collected from different lots, wafers, and sites on the same wafer, such that they are as representative as possible of the fabrication process. In the conventional learning procedure, the N circuits are split into training and validation sets. The training set is used to learn the regression model while the validation set is used as an independent set to assess the prediction error of the regression model that we would observe on previously unseen circuits that are not employed during training. If the prediction error on the validation set is deemed low enough, then the regression model can be readily used in production test to predict the performances of a CUT by relying solely on the alternate measurements and without needing to perform the conventional specification tests.

By definition, the alternate test can be applied to any CUT that exhibit process variations as long as it does not contain any defects. In the presence of a defect, the alternate measurement pattern will lie far away from the convex body that encloses the training set and any alternate test entails a risk since the regression model will extrapolate resulting in a somewhat random prediction. For this purpose, a defect filter is also learned during the training phase so as to screen out circuits with defects in the production test phase before they are forwarded to the regression models [10].

III. Motivation for this work

Alternate test can be accurate as long as the following criteria are satisfied: (a) the alternate measurements are information-rich and correlate well with the performances; (b) the regression models are accurately trained; and (c) the training set is fully representative of the fabrication process.

As mentioned in the introduction, criterion (a) is a circuit-specific problem and is outside the scope of this paper.

Criterion (b) can be satisfied by choosing an advanced regression modeling technique. Thereafter, any inaccuracies ascribed to regression modeling can be circumvented by training different regression models, for example, using different training sets and/or different sets of alternate measurements, and choosing to retest a circuit through the conventional specification tests in case there is a disagreement amongst the predictions of the regression models [22].

If criterion (c) is not satisfied, then there will be outlier CUT towards the tails of the circuit distribution (e.g. process corners) whose alternate measurements lie outside the convex body that encloses the alternate measurements of the circuits that are employed for training the regression models. For those outlier CUT the alternate test entails a risk since, as in the case of CUT with defects, the regression model will extrapolate, resulting in a somewhat random prediction. Therefore, a large volume of data need to be collected over a large period of time until we have at hand a training set that is fully representative of the fabrication process to train the regression models accurately and start applying alternate test with high confidence. Alternatively, the defect filter can also be used to screen out the outlier circuits and forward them instead to the conventional specification tests. The defect filter and the regression models can be calibrated as we keep retesting CUT and enhancing the training set, such that after a point in time the alternate test can be applied blindly for every circuit resulting in accurate decisions. However, in addition to increasing test cost for a fraction of CUT for a period of time, this two-tier test approach poses certain difficulties in the test floor. Another approach is to generate large synthetic data sets from real data to train the regression models [23]. However, this approach is unlikely to succeed since the synthetic data are generated without having any information about the tails of the circuit distribution.

In this paper, the “incompleteness” of a small training set collected from the first production wafers and lots is

circumvented by employing in the analysis post-layout simulations. The real data are not combined with the post-layout simulation data in a naive manner. Instead, we employ the more sophisticated BMF technique which is explained in detail in the next section.

IV. Bayesian Model Fusion

We assume without loss of generality a single performance P. Our objective is to learn the regression function $f : \boldsymbol{x} \rightarrow P$. Let us assume that we have at hand data from N real circuits. We define the vectors $\boldsymbol{P}_L = [P^{(1)}, \cdots, P^{(N)}]$ and $\boldsymbol{x}_L = [\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}]$, where $P^{(n)}$ and $\boldsymbol{x}^{(n)}$ denote the performance and alternate measurement pattern, respectively, for the n-th circuit $n = 1, \cdots, N$.

The conventional learning procedure is to use a fraction of the real data for training and the rest of the real data for validating the generalization ability on previously unseen circuits. However, as explained above, in a practical scenario this real data set contains very limited information about the process corners and the regression function will be valid mainly around the nominal point. The aim of the BMF technique is to learn the regression function by leveraging information about the process corners from a large volume of post-layout simulation data that is readily available and combining this information with the real data. We refer to the post-layout simulation data as *early-stage* data and to the real data as *late-stage* data.

Formally, we consider two versions of the regression function f, namely an early-stage regression function, denoted by f_E, that is trained using only early-stage data and a late-stage regression function, denoted by f_L, that is trained using the BMF learning procedure. We use the following general forms

$$f_E(\boldsymbol{x}) = \sum_{m=1}^{M} a_{E,m} \cdot b_m(\boldsymbol{x}) \qquad (2)$$

$$f_L(\boldsymbol{x}) = \sum_{m=1}^{M} a_{L,m} \cdot b_m(\boldsymbol{x}), \qquad (3)$$

where $b_m(\boldsymbol{x})$ is the m-th basis function and $a_{E,m}$, $a_{L,m}$ correspond to the m-th coefficient of the early-stage and late-stage regression function, respectively, $m = 1, \cdots, M$.

The BMF learning procedure consists of solving for the late-stage model coefficients that maximize the *posterior* distribution $\text{pdf}(\boldsymbol{a}_L | \boldsymbol{P}_L, \boldsymbol{x}_L)$, that is,

$$\max_{\boldsymbol{a}_L} \text{pdf}(\boldsymbol{a}_L | \boldsymbol{P}_L, \boldsymbol{x}_L), \qquad (4)$$

where $\boldsymbol{a}_L = [a_{L,1}, \cdots, a_{L,M}]$. By applying Bayes' theorem, we can write

$$\text{pdf}(\boldsymbol{a}_L | \boldsymbol{P}_L, \boldsymbol{x}_L) \propto \text{pdf}(\boldsymbol{a}_L) \cdot \text{pdf}(\boldsymbol{P}_L, \boldsymbol{x}_L | \boldsymbol{a}_L). \qquad (5)$$

Thus, the problem boils down to

$$\max_{\boldsymbol{a}_L} \text{pdf}(\boldsymbol{a}_L) \cdot \text{pdf}(\boldsymbol{P}_L, \boldsymbol{x}_L | \boldsymbol{a}_L). \qquad (6)$$

Assuming that the late-stage model coefficients are independent, we can write

$$\text{pdf}(\boldsymbol{a}_L) = \prod_{m=1}^{M} \text{pdf}(a_{L,m}). \qquad (7)$$

We define the *prior* distribution $\text{pdf}(a_{L,m})$ by involving the prior knowledge from the early-stage data. Specifically, $\text{pdf}(a_{L,m})$ is assumed to follow a Gaussian distribution with mean $a_{E,m}$ and standard deviation $\lambda |a_{E,m}|$

$$\text{pdf}(a_{L,m}) = \frac{1}{\sqrt{2\pi}\lambda |a_{E,m}|} \cdot \exp\left[-\frac{(a_{L,m} - a_{E,m})^2}{2\lambda^2 a_{E,m}^2} \right]. \qquad (8)$$

This approach accounts for the fact that $a_{L,m}$ is expected to be similar to $a_{E,m}$ and deviate from $a_{E,m}$ according to the absolute magnitude of $a_{E,m}$.

The *likelihood function* $\text{pdf}(\boldsymbol{P}_L, \boldsymbol{x}_L | \boldsymbol{a}_L)$ is expressed in terms of the real data. Specifically, since the real data are obtained independently, we can write

$$\text{pdf}\left(\boldsymbol{P}_L, \boldsymbol{x}_L | \boldsymbol{a}_L\right) = \prod_{n=1}^{N} \text{pdf}\left(P^{(n)}, \boldsymbol{x}^{(n)} | \boldsymbol{a}_L\right). \qquad (9)$$

Furthermore,

$$\text{pdf}\left(P^{(n)}, \boldsymbol{x}^{(n)} | \boldsymbol{a}_L\right) = \text{pdf}(\varepsilon^{(n)}), \qquad (10)$$

where $\varepsilon^{(n)}$ is the prediction error introduced by the late-stage regression for the n-th real circuit

$$\varepsilon^{(n)} = P^{(n)} - f_L(\boldsymbol{x}^{(n)}). \qquad (11)$$

This error is a random variable that is assumed to follow a zero-mean Gaussian distribution with some standard deviation σ_0

$$\text{pdf}(\varepsilon^{(n)}) = \frac{1}{\sqrt{2\pi}\sigma_0} \cdot \exp\left(-\frac{\left(\varepsilon^{(n)}\right)^2}{2\sigma_0^2} \right). \qquad (12)$$

Therefore, combining (10), (11), (12), and (3), we can write

$$\text{pdf}\left(P^{(n)}, \boldsymbol{x}^{(n)} | \boldsymbol{a}_L\right) = \frac{1}{\sqrt{2\pi}\sigma_0} \cdot \exp\left\{ -\frac{1}{2\sigma_0^2} \cdot \left[P^{(n)} - \sum_{m=1}^{M} a_{L,m} \cdot b_m\left(\boldsymbol{x}^{(n)}\right) \right]^2 \right\}. \qquad (13)$$

By combining (7), (8), (9), and (13) and taking the natural logarithm, the maximization problem in (6) becomes

$$\max_{\boldsymbol{a}_L} \sum_{m=1}^{M} \log[\text{pdf}(a_{L,m})] - \frac{1}{2\sigma_0^2} \cdot \sum_{n=1}^{N} \left[P_L^{(n)} - \sum_{m=1}^{M} a_{L,m} \cdot b_m\left(\boldsymbol{x}^{(n)}\right) \right]^2. \qquad (14)$$

The optimal values of σ_0 and λ are determined by k-fold cross-validation [16], [17].

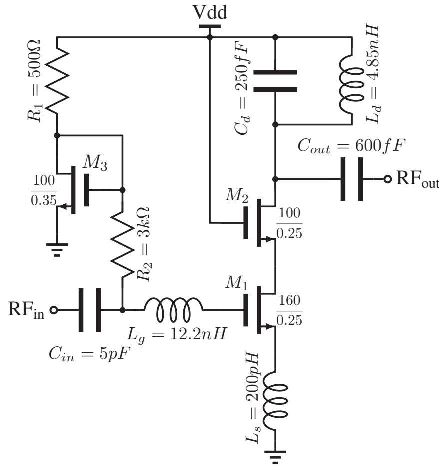

Fig. 2. CMOS inductively degenerated cascode LNA.

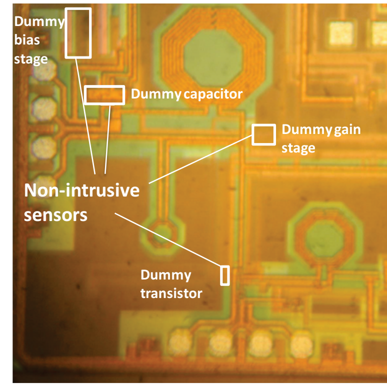

Fig. 3. Photo of fabricated chip [24].

V. Case Study

A. CUT and alternate test

Our test vehicle is a 2.4 GHz inductively degenerated cascode LNA shown in Fig. 2. We consider the alternate approach based on non-intrusive sensors proposed in [24]. The non-intrusive sensors consist of dummy analog stages and single components extracted directly from the topology of the LNA. In particular, we consider a dummy bias stage identical to the bias stage of the LNA formed by transistor M_3 and resistor R_1, a dummy gain stage identical to the gain stage of the LNA formed by transistors M_1 and M_2, a dummy diode-connected transistor identical to the transistor $M1$, and a dummy capacitor identical to the capacitor C_{in}. These dummy analog stages and single components are placed on the same die in close physical proximity to the analog stages and components of the LNA that they are mimicking without being electrically connected to the LNA. The alternate measurement pattern provided by these sensors includes the DC bias, DC voltage gain, transistor transconductance, and capacitance value. The underlying idea is that the sensors "witness" the same process variations as the LNA, thus the alternate measurements are correlated to the performances of the LNA and can be used to predict the variations in the LNA performances following the alternate test paradigm. We consider the four main performances of the LNA typically measured in production testing, namely the gain (S21), noise figure (NF), 1 dB-compression point (1-dB CP), and input third-order intercept point (IIP3). This built-in alternate test approach, in addition to being low-cost and incurring low area overhead, has the important property that is totally transparent to the LNA. The built-in sensors are non-intrusive, thus the performances of the LNA are unaffected by the test and the design and test are dissociated. For more details, the interested reader is referred to [24].

The LNA and the non-intrusive sensors are designed using the 0.25 μm Qubic4+ BiCMOS technology by NXP Semiconductors. The photo of the fabricated chip is shown in Fig. 3. In total, 140 chips were fabricated in a Multi-Project-Wafer (MPW) run. The chips are scattered across different sites on a wafer. The four LNA performances and the alternate measurement pattern were obtained on each chip using benchtop equipment. There is sufficient measurement dispersion due to process variations, larger than the measurement error, which allows us to formulate appropriately the learning problem. Again, for more details, the interested reader is referred to [24].

B. Experiment design

The late-stage real data in our case study come from the 140 fabricated circuits. The early-stage simulation data are generated through a Monte Carlo post-layout simulation with 1000 runs. The simulation takes into account the complete signal path, including the circuit, pins, package, test board, etc. Before proceeding with our experiment, we first confirmed the basic underlying assumption of BMF that the early-stage and late-stage real data distributions are similar.

The course of the experiment is as follows:

Step 1: Select randomly 100 real circuits out of the 140 available real circuits.

Step 2: Build the sets of late-stage real *inliers* and late-stage real *outliers*. The set of late-stage real inliers contains $N_{\text{tr}} = 100 - N_{\text{val}}$ circuits and the set of late-stage real outliers contains the rest N_{val} circuits. N_{val} is selected in the range $[30, 70]$ (see step 5). The late-stage real outliers are the N_{val} most distant circuits from the sample mean in an Euclidian sense when all circuits are projected onto the space of alternate measurements. For the purpose of our experiment, we consider that the late-stage real inliers are the circuits that are typically available at the onset of production for training the regression functions in alternate test. The late-stage real outliers correspond to "extreme" circuits that are met much more rarely and are unlikely to be available at the onset of production.

Step 3: Select learning procedure and build the corresponding training set. We consider the four learning procedures listed in Table I. In particular, the BMF learning procedure is compared to the conventional learning procedure that uses as training set (a) only the late-stage real inliers, (b) a "raw" combination of early-stage simulation data and late-stage real

TABLE I
ALTERNATE TEST LEARNING PROCEDURES

learning method	training set	validation set
BMF	"intelligent" mixture of early-stage simulation data and late-stage real inliers	late-stage real outliers
standard	late-stage real inliers	
straightforward combination	"raw" mixture of early-stage simulation data and late-stage real inliers	
simulation-based	early-stage simulation data	

inliers, and (c) only the early-stage simulation data. Case (a) is the standard learning procedure used today, case (b) combines together simulation with real data in a straightforward manner, unlike the BMF learning procedure that relies on an "intelligent" combination, and case (c) is a "naive" simulation-based approach. For all learning procedures, we use polynomial regression with pure quadratic polynomial basis functions and no interaction terms.

Step 4: Compute the error of the learned regression function on the validation set. For all learning procedures, the late-stage real outliers are used as the validation set. We employ two different accuracy metrics to express the error, namely the root-mean-square (RMS) error (ε_{RMS}) and the maximum absolute error ($|\varepsilon|_{max}$).

Step 5: Repeat steps 2-4 for different values of N_{val} in the range $[30, 70]$, in order to study how each learning procedure performs with respect to the size of the late-stage real data set that is employed for training.

Step 6: Repeat steps 1-5 10 times and report the average value of the two accuracy metrics observed for each performance, learning method, and value of N_{val}. This step applies the bootstrapping idea so to report as trustworthy accuracy metrics as possible given the small real data set that we have at hand.

C. Results

The results are shown in Fig. 4, where columns correspond to the four performances and rows correspond to the two accuracy metrics. In each plot, we show the average accuracy metric for each of the four learning procedures listed in Table I versus N_{val}. The markers indicate the exact average values of the accuracy metrics. To enhance the readability of the plots, we fit quadratic lines to the markers, in order to smooth out the statistical variations that are due to the small sample.

The main conclusion drawn by studying carefully the plots in Fig. 4 is that, for any performance and any accuracy metric and regardless the number of the late-stage real inlier circuits used for training, the BMF learning procedure either performs better than the other three conventional learning procedures or, at worst, it is equivalent to one of them in statistical terms. In particular, the simulation-based learning procedure shows consistently the worst performance. This is expected since the distributions of simulation and real data inevitably show some discrepancy. The performance of the standard learning procedure deteriorates monotonically as N_{tr} becomes smaller or, equivalently, N_{val} becomes larger. This is expected since the information available for training is weakened and our ability to extrapolate the regression towards the tails of the distribution deteriorates, resulting in large prediction error on the validation set. In some cases, for very small training set sizes, it turns out that the standard learning procedure presents even a worst performance compared to the "naive" simulation-based approach. The straightforward combination learning procedure performs better than the standard and simulation-based learning procedures since information about outliers is included during training. However, it is observed to be less effective than the BMF learning procedure. This is explained by the fact that the straightforward combination learning procedure combines simulation and real data with equal weight, while the BMF learning procedure appropriately assigns the optimal weight through cross-validation. The improvement that the BMF learning procedure offers as compared to the straightforward combination learning procedure is significant if we project it to parts-per-million. It should also be noticed that the BMF learning procedure shows a remarkably stable behaviour, maintaining nearly constant accuracy metrics even when the number of the late-stage real inliers used for training is small. This implies that the BMF learning procedure, by statistically extracting prior knowledge from simulation data, is capable of generating accurate regression functions across the design space based only on few real circuits. Thus, the BMF learning procedure can be used to start deploying the alternate test right at the onset of production, without needing to wait to collect beforehand a large volume of data, as is the standard practice today. This result showing that the BMF learning procedure reduces the burden of collecting late-stage data is consistent with the outcome of other studies that employ the BMF learning procedure in the context of pre-silicon validation, yield learning, and post-manufacturing tuning [18]–[21].

VI. CONCLUSION

We demonstrated the use of the BMF learning procedure in the context of alternate test with the aim to reduce the real data that need to be collected before the alternate test is deployed with high confidence. We demonstrated that the BMF learning procedure is very stable providing accurate performance predictions even in the case where the available real data is really limited. This is achieved by reusing readily available simulation data as prior knowledge when learning the regression functions based on the limited real data. In our example case study, the BMF learning procedure outperformed the standard learning procedure for a given small training set of real data. It also outperformed two other more simplistic learning procedures, one that uses as a training set a straightforward combination of simulation and real data and one that uses as a training set only simulation data.

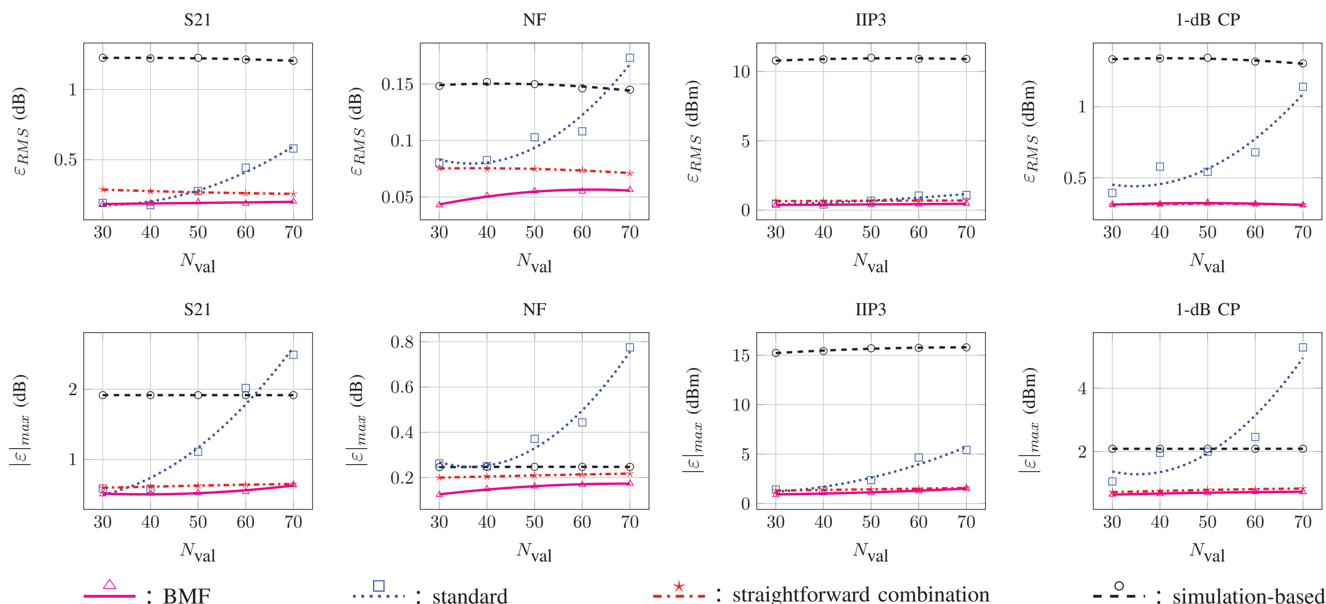

Fig. 4. Alternate test accuracy metrics for various performances and model construction cases.

REFERENCES

[1] P. N. Variyam, S. Cherubal, and A. Chatterjee, "Prediction of analog performance parameters using fast transient testing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 3, pp. 349–361, 2002.

[2] S. S. Akbay, A. Halder, A. Chatterjee, and D. Keezer, "Low-cost test of embedded RF/Analog/Mixed-signal circuits in SOPs," *IEEE Transactions on Advanced Packaging*, vol. 27, no. 2, pp. 352–363, 2004.

[3] R. Voorakaranam, R. Newby, S. Cherubal, B. Cometta, T. Kuehl, D. Majernik, and A. Chatterjee, "Production deployment of a fast transient testing methodology for analog circuits: Case study and results," in *IEEE International Test Conference*, 2003, pp. 1174–1181.

[4] S. Cherubal, R. Voorakaranam, A. Chatterjee, J. Mclaughlin, J. L. Smith, and D. M. Majernik, "Concurrent RF test using optimized modulated RF stimuli," in *IEEE International Conference on VLSI Design*, 2004, pp. 1017–1022.

[5] A. Halder and A. Chatterjee, "Low-cost alternate EVM test for wireless receiver systems," in *Proc. IEEE VLSI Test Symposium*, 2005, pp. 255–260.

[6] S. Ellouz, P. Gamand, C. Kelma, B. Vandewiele, and B. Allard, "Combining internal probing with artficial neural networks for optimal RFIC testing," in *Proc. IEEE International Test Conference*, 2006, pp. 4.3.1–4.3.9.

[7] S. Goyal, A. Chatterjee, and M. Purtell, "A low-cost test methodology for dynamic specification testing of high-speed data converters," *Journal of Electronic Testing: Theory and Applications*, vol. 23, no. 1, pp. 95–106, 2006.

[8] S. Kook, A. Banerjee, and A. Chatterjee, "Dynamic specification testing and diagnosis of high-precision sigma-delta ADCs," *IEEE Design & Test of Computers*, vol. 30, no. 4, pp. 36–48, 2013.

[9] S.-W. Hsiao, X. Wang, and A. Chatterjee, "Analog sensor based testing of phase-locked loop dynamic performance parameters," in *Proc. IEEE Asian Test Symposium*, 2013, pp. 50–55.

[10] H.-G. Stratigopoulos and S. Mir, "Adaptive alternate analog test," *IEEE Design & Test of Computers*, vol. 29, no. 4, pp. 71–79, 2012.

[11] S. S. Akbay, J. L. Torres, J. M. Rumer, A. Chatterjee, and J. Amtsfield, "Alternate test of RF front ends with IP constraints: Frequency domain test generation and validation," in *Proc. IEEE International Test Conference*, 2006, pp. 4.4.1–4.4.10.

[12] H.-G. D. Stratigopoulos and Y. Makris, "Non-linear decision boundaries for testing analog circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 11, pp. 1760–1773, 2005.

[13] H. Ayari, F. Azais, S. Bernard, M. Compte, M. Renovell, V. Kerzerho, O. Potin, and C. Kelma, "Smart selection of indirect parameters for dc-based alternate RF IC testing," in *Proc. IEEE VLSI Test Symposium*, 2012, pp. 19–24.

[14] M. J. Barragan and G. Leger, "Efficient selection of signatures for analog/RF alternate test," in *Proc. IEEE European Test Symposium*, 2013.

[15] J. Liaperdos, A. Arapoyanni, and Y. Tsiatouhas, "Adjustable RF mixers' alternate test efficiency optimization by the reduction of test observables," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 9, pp. 1383–1394, 2013.

[16] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining , Inference, and Prediction*, Springer, 2001.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning*, Information Science and Statistics. Springer, 2007.

[18] X. Li, W. Zhang, F. Wang, S. Sun, and C. Gu, "Efficient parametric yield estimation of analog /mixed-signal circuits via Bayesian model fusion," in *Proc. IEEE/ACM International Conference on Computer-Aided Design*, 2012, pp. 627–634.

[19] F. Wang, W. Zhang, S. Sun, X. Li, and C. Gu, "Bayesian Model Fusion: Large-scale performance modeling of analog and mixed-signal circuits by reusing early-stage data," in *Proc. ACM/EDAC/IEEE International Conference on Computer-Aided Design*, 2013.

[20] C. Gu, E. Chiprout, and X. Li, "Efficient moment estimation with extremely small sample size via bayesian inference for analog/mixed-signal validation," in *Proc. ACM/EDAC/IEEE Design Automation Conference*, 2013.

[21] S. Sun, F. Wang, S. Yaldiz, X. Li, L. Pileggi, A. Natarajan, M. Ferriss, J. Plouchart, B. Sadhu, B. Parker, A. Valdes-Garcia, M. Sanduleanu, J. Tierno, and D. Friedman, "Indirect performance sensing for on-chip analog self-healing via bayesian model fusion," in *Proc. IEEE Custom Integrated Circuits Conference*, 2013.

[22] H. Ayari, F. Azais, S. Bernard, M. Compte, V. Kerzerho, O. Potin, and M. Renovell, "Making predictive analog/RF alternate test strategy independent of training set size," in *Proc. IEEE International Test Conference*, 2012, Paper 10.1.

[23] H.-G. Stratigopoulos, S. Mir, and Y. Makris, "Enrichment of limited training sets in machine-learning-based analog/RF test," in *Proc. Design, Automation & Test in Europe Conference*, 2009, pp. 1668–1673.

[24] L. Abdallah, H.-G. Stratigopoulos, S. Mir, and C. Kelma, "Experiences with non-intrusive sensors for RF built-in test," in *Proc. IEEE International Test Conference*, 2012, Paper 17.1.