Combining Adaptive Alternate Test and Multi-Site

Gildas Leger
Instituto de Microlectrónica de Sevilla, Centro Nacional de Microelectrónica
Consejo Superior de Investigaciones Científicas (CSIC) and Universidad de Sevilla
Av. Américo Vespucio s/n, 41092 Sevilla, Spain.
e-mail: leger@imse-cnm.csic.es

***Abstract*—Testing analog, mixed-signal and RF circuits represents one of the main cost components for complex SoCs. Multi-site Testing is widely accepted as a straightforward technique to reduce the effective test time. This paper shows that an adaptive Alternate Test approach can be compatible with a multi-site strategy. The proposed solution consists in ordering off-line the signatures acquisition sequence and training incremental regression models for each new feature. These models can be used to diagnose the circuit as good, provided that the estimate of the performance is larger than the specification plus a guard-band related to the model error. If all the sites are diagnosed as good, the test program can be halted before completion. This decision is taken on-line and makes this scheme adaptive. We provide an analytical study of the expected test time reduction and of the test escape penalty that is incurred. Results obtained from post-layout MonteCarlo simulations of an LNA demonstrate the validity of the approach and show that significant test time improvements can be obtained, even for large number of sites, whenever the manufacturing yield is sufficiently high.**

I. Introduction

The test of Analog, Mixed-Signal and RF (AMS-RF) circuits is recognized as a one of the major bottlenecks for the reduction of manufacturing costs. As a matter of fact, AMS-RF test is often quoted to represent 50% of total test time in complex SoCs.

While some researchers regard defect-oriented Design-for-Test and BIST as way to reduce test time, others have taken the pragmatic path of optimizing what is already accepted: specification-based test. In this sense, test ordering is definitely not a new topic in the field of AMS-RF testing [1], [2]. The underlying question is utterly simple: can the test plan order optimize test time? In the end, the optimization depends on the fault likelihoods, the duration and the coverage of the different test in the test plan. But it appears obvious that time will be saved if faulty parts can be detected early, since they do not need to go through the complete test plan.

Another well-accepted method to improve throughput is parallelism. Multi-site testing is now a common industrial practice and roadmaps show that this trend is expected to continue [3]. Unfortunately, multi-site testing turns test ordering almost useless. In order to stop the test program in a multi-site framework, all the sites should be declared faulty. For practical manufacturing yield values, it is not probable that all the sites be faulty. The test time for those sites is thus the test time of the best circuit, which will likely be the test time of the full test set. The probability of M circuits (sites) failing a given specification $spec_i$ is given by,

$$P_{all\ bad\ @\ spec_i} = (1 - Y_i)^M \qquad (1)$$

where Y_i is the yield associated to this specification. This exponential dependence on M makes the probability vanish.

What we propose is to use the mapping power of Alternate Test to retain the time reduction benefits of test ordering even in a Multi-Site framework.

II. Multi-site Adaptive Alternate Test

In the last decade, Alternate Test has been proposed as a way to avoid costly performance measurements [4]. The idea is to leverage the power of recent multi-dimensional non-linear machine-learning algorithms to build a model that links simple measurements to the performance. As a matter of fact, this approach can also be used to remove redundancy from a conventional performance-based test plan. Regression tools provide indirect estimates of the different performances while classification tools determine the pass-fail boundaries in the cheap measurement space. The accuracy of these models is characterized by their generalization error: the error committed on a set different from the one that was used to train the model. Selecting more or less features leads to models with different errors.

Let us suppose that we have a satisfying model that manages to regress/classify the samples using N measurements (features) with an accuracy similar to performance measurements. Notice that the features may be simple measurements but also a subset of the conventional performance measurements. This will be the starting point of our proposal. The determination of the model and its associated features is out of the scope of this paper.

A-priori, it may seem that test ordering makes no sense since all the features are required to operate the model. However, taking a closer look we can see that is is possible to train several models with an increasing number of features. It is to be expected that the first model, with only one feature, would have a poor accuracy while the last one would reach the specification-based level of the original model. Each of these incremental models has thus an associated test time and coverage. It appears possible to order the features (and thus the incremental models) to optimize test time. This is very similar to former test ordering.

However, there is a difference that leads to a major change with respect to the impact on multi-site testing. Traditional

978-3-9815370-4-8/DATE15/©2015 EDAA

tests in a test program are seen as a succession of screenings: a circuit that fails a given test is a bad device, but a circuit that pass a given test cannot be diagnosed as a good device. On the contrary, an alternate test model can also be seen as a validity confirmation. Up to a certain boundary (related to the generalization error), the circuit can be diagnosed as good.

A direct consequence of this change is that yield becomes the ally of multi-site instead of its enemy. The test program (i.e. the acquisition of new features) is stopped when there is sufficient confidence that all the sites are good circuits. The higher the yield, the higher the probability.

Our proposal is as follows:

- Order the N features by stepwise addition
- Train the corresponding N incremental models
- For the first $N-1$ models, consider a guard-band on the specification to guarantee that the circuit is good. This guard-band can be defined as a multiple of the generalization error. Supposing, without loss of generality, a lower bound specification $spec$, the probability of a circuit to be diagnosed good by model i can be defined as,

$$P_{good_i} = P\left(\tilde{p} > spec + k \times \sigma_i\right) \quad (2)$$

  where $\tilde{p}$ is the estimated performance of the circuit, σ_i is the generalization error of model i and k is the guard-band factor.
- For the last model – which is the model that reaches higher accuracy – do not take guard-band (or adjust its value as a function of the Test Coverage vs Yield Loss trade-off [5])
- Perform the feature measurements sequentially and apply the corresponding models to the M sites
- If all the sites are diagnosed as good (taking into account the guard-band), then stop the test program

Though the test ordering is performed off-line, the decision to stop the test program is taken on-line depending on the cluster composition. In this sense, this approach can be seen as an adaptive multi-site strategy.

In a way, this is a similar philosophy to what was proposed in [6] with some fundamental differences: We do not restrict the features to specifications. Instead of using a correlation matrix between specifications, we use incremental models that can precisely capture non-linear relationships. And finally, we make it compatible with multi-site by setting the focus on acceptance instead of rejection.

Some similarities can also be found with the algorithm proposed in [7]. In that paper, the circuits are submitted to two defect filters, one strict and one lenient (the strictness is similar to our guard-band). The circuits in the uncertainty zone undergo thorough measurements, those that pass the strict filter are diagnosed by alternate test while those that fail the lenient filter are discarded. Our incremental models could also be viewed as a cascade of those defect filters: the circuits that are not diagnosed with sufficient confidence at step i are submitted to measurement $i+1$.

In [8], it is shown that test ordering is still interesting in a mutli-site framework because it allows to detect and suppress redundant tests. However, the goal of that paper is not to reduce test time by detecting early-failing devices but rather to eliminate some tests from the test program. In our opinion, this is closer to feature selection than to test ordering.

The manufacturing yield (in what follows, the yield Y), is the fraction of the manufactured devices that are actually good devices. The Yield Loss is related to the probability of good circuits to be diagnosed as bad by the test program. On the contrary, Test Escape (or defect-level) is related to the probability of bad circuits to be diagnosed as good by the test program. These metrics are sometimes defined with respect to the number of good devices for the former and with respect to the number of bad devices for the latter. In what follow we will consider the probabilities for any manufactured circuit. The relative metrics can easily be deduced dividing by Y or $1-Y$;

With the proposed method, the Yield Loss is that of the final model (which is supposed to be satisfying) since none but the last model is allowed to diagnose the circuits as bad. However, there could by a penalty on the Test Escape if the guard-band is not sufficiently high. This penalty should thus be made negligible with respect to conventional test.

For the sake of clarity, the study will consider a single performance (p) with a lower limit specification ($spec$) such that for good circuits $p > spec$. Regression models that draw an estimate of the performance $\tilde{p}$ will be used. The confidence in this estimate is directly related to the generalization error of the model (i.e. the rms error on a data set independent from the training set).

III. Analytical Study

A. Test ordering

The proposal starts from an appropriate model with its corresponding set of N features and orders these N features by stepwise addition, considering as a cost function the generalization error. It is assumed that proper feature selection has been carried out to remove useless or redundant features.

- start the feature set from the void set
- generate all the possible feature sets by adding only one feature to the actual feature set
- train the corresponding models
- select the feature set that gives the lowest generalization error and save the corresponding model
- iterate until all the features are ordered

At the end of this phase, the order in which the features must be acquired is obtained as well as a set of N associated "incremental" models. Together with the models we also get a decreasing error profile which will be used to properly set the guard-band.

B. Expected test time

Let us consider a device for which has been designed an alternate test that involves N features. It is supposed that the quality of this alternate test would be sufficient to replace the specification-based test. This device will be tested in an M-site configuration.

The probability that the M devices from the M sites are diagnosed as good at the same time by model i is,

$$P\left(allOK\ @_i\right) = P_{good_i}^M \tag{3}$$

The probability for the test plan to be stopped at step i is,

$$\begin{aligned} &P\left(stop\ @_i\right) = \cdots \\ &P\left(allOK\ @_i, \overline{allOK\ @_{i-1}}, \cdots, \overline{allOK\ @_1}\right) \end{aligned} \tag{4}$$

Indeed, some circuits that could be diagnosed at step i have already been diagnosed in previous steps.

In order to calculate this probability, we can sum over all the possible combinations of results at previous steps,

$$\begin{aligned} &P\left(allOK\ @_i\right) = \cdots \\ &P\left(allOK\ @_i, \overline{allOK\ @_{i-1}}, \cdots, \overline{allOK\ @_1}\right) \\ &+ \cdots \\ &+ P\left(allOK\ @_i, allOK\ @_{i-1}, \cdots, \overline{allOK\ @_1}\right) \end{aligned} \tag{5}$$

There are 2^{i-1} combinations. Fortunately, many of them make no sense and have thus almost zero probability. Indeed, we have,

$$P\left(\overline{allOK\ @_k}, allOK\ @_l\right) = 0\ \forall\ k > l \tag{6}$$

It comes that the only non-null probabilities are the $P\left(stop\ @_k\right)$ for all $k < i$. Hence, we finally obtain,

$$P\left(stop\ @_i\right) = P_{good_i}^M - \sum_{j=1}^{i-1} P\left(stop\ @_j\right) \tag{7}$$

This result is not surprising. The devices that are actually diagnosed at step i are those that would be diagnosed by model i minus those that would have been diagnosed earlier. This recursive definition reduces to,

$$P\left(stop\ @_i\right) = P_{good_i}^M - P_{good_{i-1}}^M \tag{8}$$

This holds for any i superior to 1. For $i = 1$, the equation is simply,

$$P\left(stop\ @_1\right) = P_{good_1}^M \tag{9}$$

In order to evaluate the test time, take into account that the devices that reach the final model (i.e. model N) will all see the same test time, should they fail or pass the test. With a test time T_i associated to each feature, the test time associated to model j is,

$$T_{model}\left(j\right) = \sum_{i=1}^{j} T_i \tag{10}$$

And the expected test time can be written as,

$$\begin{aligned} TestTime = {} & T_{model}\left(N\right) \times \left(1 - \sum_{i=1}^{N-1} P\left(stop\ @_i\right)\right) \\ & + \sum_{j=1}^{N-1} T_{model}\left(j\right) \times P\left(stop\ @_j\right) \end{aligned} \tag{11}$$

Using the expressions (8), (9) and (10) this expression simplifies to,

$$TestTime = T_1 + \sum_{j=1}^{N-1} T_{j+1} \times \left(1 - P_{good_j}^M\right) \tag{12}$$

This expected test time shall be compared with the test time without ordering. As all the circuits would be tested with the best model only, the test time without ordering would be that of model N: the sum of the test times associated to all the features (10). The expected test time is thus lower than the original model test time, since all the feature test times except the first one (which all the circuits must pass) are weighted by a factor $\left(1 - P_{good_i}^M\right)$ lower than one. This factor actually corresponds to the fraction of circuits detected by previous steps. To go further in the study, some assumptions on the performance distribution should be made. Let us assume a Normal distribution with mean value μ_{perf} and standard deviation σ_{perf}. In this way, we can explicitly write

$$P_{good_j} = \frac{1}{2}\left(1 - erf\left(\frac{spec - \mu_{perf} + k\sigma_j}{\sigma_{perf}\sqrt{2}}\right)\right) \tag{13}$$

C. Guard-band selection and Test Escape

As explained in Section II, the proposed approach does not impact Yield Loss but may cause an increase of the Test Escape. Let us try to evaluate the penalty on the test escape, as a function of the guard-band.

We can write the probability of a test escape occurring at any step before the last one as,

$$\begin{aligned} Penalty = P\,(&(p < spec)\ AND \\ &(allOK\ @_1 OR \cdots OR\ allOK\ @_{N-1})) \end{aligned} \tag{14}$$

By developing the expression in (14), some crossed terms appear, since $P\left(A\ OR\ B\right) = P\left(A\right) + P\left(B\right) - P\left(A\ AND\ B\right)$. However, we can make the approximation that the probability that an escape would occur at two or more steps is negligible.

Let us consider the marginal test escape at step i,

$$TE_i = P\left(\left(p < spec\right), \left(allOK\ @_i\right)\right) \tag{15}$$

It comes that,

$$Penalty \approx \sum_{j=1}^{N-1} TE_j \tag{16}$$

In order to justify this approximation, let us consider the generalization error of the prediction as a random variable r that is independent of the actual performance, so that,

$$\tilde{p} = p + r \tag{17}$$

For a test escape to occur, it is thus necessary (though not sufficient) that the model error be superior to the guard-band (i.e. $k \times \sigma_i$). This leads to the following inequality for the crossed terms:

$$\begin{aligned} &P\left(p < spec, \tilde{p} > spec + k\sigma_i, \tilde{p} > spec + k\sigma_j\right) \\ &\cdots < P\left(r_i > k\sigma_i\right) P\left(r_j > k\sigma_j\right) \end{aligned} \tag{18}$$

Considering a normal unbiased distribution for the generalization error it comes,

$$P(r_i > k\sigma_i) = \frac{1}{2}\left(1 - erf\left(\frac{k}{\sqrt{2}}\right)\right) \tag{19}$$

The crossed terms in (14) would thus have a probability lower than $\left(\frac{1}{2}\left(1 - erf\left(\frac{k}{\sqrt{2}}\right)\right)\right)^{n_c}$, where n_c is the number of events involved in the crossed term. For any significant guard-band this probability would rapidly vanish.

It is difficult to evaluate the individual test escapes, since they will depend on both the probability density function of the performance and the generalization error of the models. For all but the last model, a test escape occurs only if all the M sites are diagnosed as good. Considering independent sites we can write,

$$\begin{aligned} TE_i &= P(p_1 < spec, \tilde{p_1} > spec + k\sigma_i, \cdots \\ &\quad \tilde{p_2} > spec + k\sigma_i, \cdots, \tilde{p_M} > spec + k\sigma_i) \\ &= P(p_1 < spec, \tilde{p_1} > spec + k\sigma_i) \times P_{good_i}^{M-1} \end{aligned} \tag{20}$$

The conditional probability on the generalization error leads to:

$$\begin{aligned} &P(p < spec, \tilde{p} > spec + k\sigma_i) \\ &= \int P(p < spec, p + r > spec + k\sigma_i \,|r) P(r)\, dr \\ &= \int_{r=k\sigma_i}^{\infty} \left(\int_{p=spec-r+k\sigma_i}^{spec} P(p)\, dp\right) P(r)\, dr \end{aligned} \tag{21}$$

We can hardly make any assumption about the shape of the performance distribution $P(p)$. However, there is an upper bound on the inner integral in (21). Actually,

$$\int_{p=spec-r+k\sigma_i}^{spec} P(p)\, dp < \int_{-\infty}^{spec} P(p) = 1 - Y \tag{22}$$

Then, if the generalization error follows a Gaussian distribution, we have

$$TE_i < \frac{1-Y}{2}\left(1 - erf\left(\frac{k}{\sqrt{2}}\right)\right) P_{good_i}^{M-1} \tag{23}$$

Considering that $0 \leq P_{good_i} < Y$ (which should always be the case for significant guard-banding), a bound on the test escape penalty can be set as:

$$Penalty < (N-1)\frac{1-Y}{2}\left(1 - erf\left(\frac{k}{\sqrt{2}}\right)\right) Y^{M-1} \tag{24}$$

For practical values of Yield (i.e. $Y > 0.9$), number of sites (i.e. $M \in [2, 64]$) and number of models ($N < 15$), it comes that a guard-banding of $k = 4.8$ would guarantee a test escape impact below 1ppm. Keep in mind that this is an upper bound and the minimum affordable value of guard-banding may be significantly lower.

In order to get some insight on the quality of this bound, we have performed a MonteCarlo simulation on synthetic data, considering a Gaussian distribution for the performance ($N(\mu = 13.2, \sigma = 1.3)$) and linearly decreasing Gaussian generalization errors to mimic the different models ($N(\mu = 0, \sigma \in [0.03, 0.2])$). It comes out from these simulations that a guard-band of $k = 4.2$ would be sufficient to obtain a test escape penalty below 1ppm.

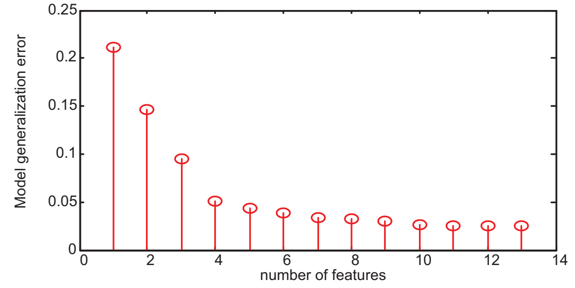

Fig. 1. Generalization error of the models with an increasing number of features.

IV. A PRACTICAL CASE OF STUDY

In order to validate this study and draw some conclusions on the utility of the approach, the following experiment is proposed on the data obtained by MonteCarlo simulation of an LNA complemented by an envelope detector [9]. The discussion is limited to the prediction of a single performance –the LNA's gain in this case– to facilitate the interpretation of the results, although the method is applicable to any number of target performances if a classifier is used.

A total of 46 signatures were acquired for each LNA instance, including DC voltages in all the nodes of the LNA, the output of the envelope detector, and all these signatures measured under power supply stress. A population of 2000 instances of the LNA was generated using Monte Carlo simulation [9], [10]. Using wrapper-based feature selection methods [11], a reduced set of 13 signatures (or features) is finally obtained, that lead to a model generalization error of 0.025dB for the LNA gain, for a Perceptron Neural Network. This precision is very high and compares with a standard functional measurement.

The 2000 instances are split in two groups of 1000 devices: the training set and the validation set. Using the training set, we perform feature ordering by stepwise addition as explained in section III.A. Notice that the generalization error is estimated by cross-validation techniques on the training set only. Up to that point, no circuit from the validation set has been involved. For the sake of simplicity, we assume that all the signatures take the same amount of time to be acquired. The test time associated to model 1 would be 1, while for the final model (i.e. model 13), it would be 13. After this operation, the feature list and the associated models are ordered with decreasing error. Fig.1 plots the generalization error versus the number of features in the case of the LNA gain.

It can be observed that the error reduction rate is not linear. The models rapidly reach a decent precision: with only 4 features the generalization error has already dropped from 0.21dB to 0.05dB.

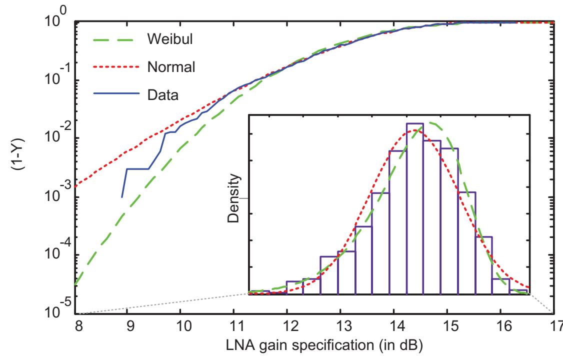

Fig. 2. Yield variation with the gain specification

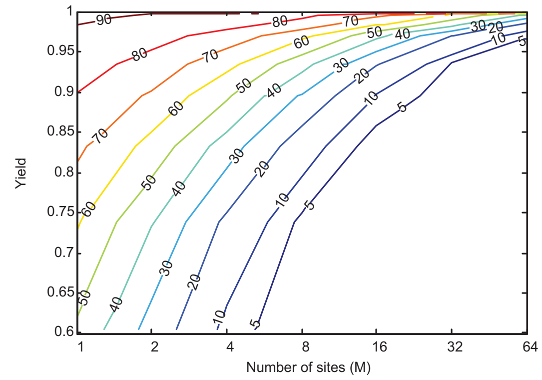

Fig. 3. Test time reduction (in %) obtained by Alternate Test Ordering, as a function of the Yield and the number of sites

With the model list and the associated generalization errors, the performance of circuits from the validation set can be predicted. If the prediction is superior to the specification by k times the generalization error, the circuit is diagnosed as good. Considering M sites, if the M circuits from the clusters are diagnosed as good, the test procedure is halted and the next cluster is tested. In this way, each circuit is associated to a given test time, depending on the step at which the test procedure was halted.

Provided that our validation set is only 1000 circuits, we have performed the experiment several times (actually 30 times) by randomizing the circuit indexes. The obtained results are averaged to mitigate the quantization effects on the test time, related to the cluster size. Since the true value of the specification is known for the circuits in the validation set, we can also check the test escape level. These post processing operations have a low computational cost, so the different parameters can be easily varied: The number of sites M, the error guard-band k, and also the specification value which is a proxy of the yield Y.

Fig.2 shows the yield evolution as we change the performance specification (what is represented is the Cumulative Density Function, which is equal to $1-Y$). The inset displays the distribution of the LNA gain and its best fit to Normal and Weibull Probability Density Functions (PDFs). The actual data seems to be bounded by the two fits, and does not appear reliable (at least for quantitative purpose) for yield values superior to 99% with such a dataset.

As a first experiment, we compute the test time reduction obtained by Alternate Test ordering as a function of the Yield and the number of sites. For this simulation, a guard-band of $k = 4.2$ has been taken which should guarantee a Test Escape penalty minor than 1ppm. Notice that we do not include the direct contribution of multi-site in this computation: the test time reduction is a relative to the same multi-site configuration without ordering. The purpose is to isolate the benefits of test ordering only.

Fig.3 shows the obtained contour plot, where the iso-levels are tagged with the associated test time reduction in percent. It can be seen that the test time improvement is greater for higher yield and for lower number of sites. This trend was expected: for both a low yield and a high number of sites, it is less likely that all the circuits in a cluster be diagnosed good, and hence that the test program be stopped early. Anyhow, for practical values like 8 sites and a Yield of 0.95, the test time can be cut by a factor two with respect to the same number of sites without ordering.

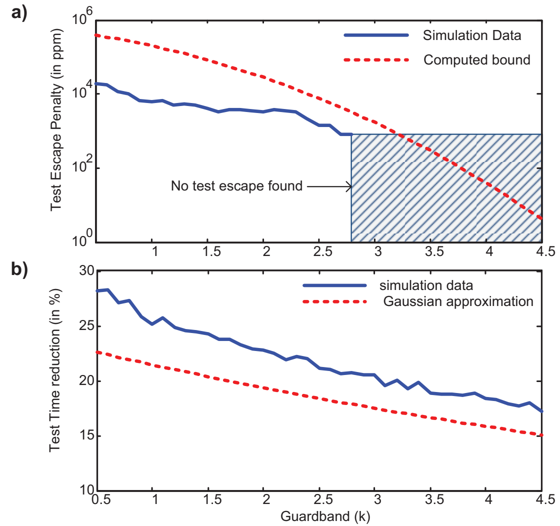

Fig. 4. a) Test Escape penalty,and b) Test Time reduction, versus guard-band for 4 sites and a yield of 0.74

For the used guard-band, the test escape penalty appears to be zero. However, this result is not statistically significant since we only have 1000 circuits in the validation set. Only test escapes levels higher than 0.1% might be detected. In order to study the influence of the guard-band, we thus varied k from 0.5 to 4.5 for a multi-site configuration of 4 sites and a Yield level of 0.74 (which sets the specification to 12.5dB) which should guarantee the occurence of test escapes.

On Fig.4-a) we can see that the test escape penalty effectively decreases with the guard-band, reaching zero slightly before $k = 3$. We have also represented the Test Escape Bound given by (23) and it appears to be much larger than the real test escape penalty. This high difference is understandable since

the substitution that was made with (22) is worse for lower yield values and smaller guard-bands. Actually, the bound appears to get closer to the real value as k increases.

In Fig.4-b), as expected, the Test Time reduction decreases with guard-band, since less circuits will actually have a specification beyond this limit. However, the reduction only varies in 10% over the simulated guard-band range. On the same plot the expected Test Time reduction for a Gaussian approximation of the performance distribution is shown. This expected reduction is computed using (12) and (13) and taking for the model errors the generalization errors plotted in Fig.1. The Test time reduction is slightly underestimated by the Gaussian approximation but is in the same order of magnitude.

V. Discussion

Let us briefly discuss the main limitations or open problems of this work. The first concern may be about the generalization of the results. Indeed, our validation example is a low-complexity circuit (an LNA) and we have considered a single simple performance, the gain. The single performance consideration allowed us to use a regression model and handle the guard-banding interpretation in the performance domain. For a multi-performance case, it would be wiser to use classifiers rather than regressors. Fortunately most classifiers also output a measure of the confidence in the classification that is conceptually similar to the generalization error and that could be used for feature ordering and for setting the appropriate guard-band. This is particularly obvious if this metric is the posterior class probability.

In addition, the equation of the expected test time reduction (12) has been derived in a generic case. The definition of the probability of a circuit to be diagnosed as good (2) is what introduces the link with the single performance and the guard-band. Whenever the incremental models exhibit a decreasing error profile – adding a new feature effectively improves the classification – the test time equation will hold. The only reason why we would not obtain test time reduction is that the P_{good_j} tend to zero, which is equivalent to saying that the classification errors is very high for all the incremental models but the last one. This is very improbable. We cannot guarantee that a convex (i.e. a rapidly decreasing) error profile is a general trend for any circuit, but it coincides with our personal experience and with the literature, where many models are built upon a small number of signatures. By the way it seems logical that the most relevant degradation mechanisms be identified easily (with few features) while higher-order subtle ones require more information. Anyhow, other error profiles would also provide test time reduction.

Further work may also focus on global optimization strategies that would combine feature selection and ordering to minimize the total test time while ensuring an adequate level of Yield Loss and Test Escape. Provided a set of features that lead to a high accuracy model, it may occur that a feature not present in this set provide strong diagnosing capabilities. Indeed the information contained in such a feature may be redundant with the information contained in a couple of features of the selected set. While for the final accuracy of the test the feature is useless, from a test time perspective it may be beneficial allowing to diagnose rapidly a high number of devices. Similarly, it would be interesting to develop the test time model for the different features, including shared setup costs. From a cost viewpoint, it may be better to group some features even if they do not provide the best error profile.

VI. Conclusion

This work demonstrates that test ordering is compatible with Alternate Test and may provide significant test time savings. The particular philosophy of the Alternate approach allows reversing the conventional way of thinking about test: instead of searching for failing devices we search for good devices. This fundamental change paves the way to efficient multi-site testing: while for the former approach high manufacturing yield is detrimental, for the latter it is beneficial.

Together with the proposed methodology, an analytical study permits to assess the test time gains and set the desired bound on the test escape penalty. Results obtained from post-layout simulation data from an LNA confirm the benefits of the proposed approach, which can reach very high Test Time reduction (up to 90% in this case). Though the benefits of test ordering decrease with high number of sites, they increase for high yields, providing additional cost savings to the direct multi-site gain.

Acknowledgment

This work has been partially funded by the Spanish Government project TEC2011-28302, co-funded with FEDER program.

References

[1] L. Milor and A. L. Sangiovanni-Vincentelli, "Minimizing production test time to detect faults in analog circuits," *IEEE Trans. on CAD*, vol. 13, no. 6, pp. 796–813, 1994.

[2] W. Jiang and B. Vinnakota, "Defect-oriented test scheduling," in *IEEE VLSI Test Symp.*, 1999, pp. 433–438.

[3] "International roadmap for semiconductors (ITRS) http://www.itrs.net," Tech. Rep., 2009.

[4] P. N. Variyam, S. Cherubal, and Chatterjee, "Prediction of analog performance parameters using fast transient testing," *IEEE Trans. on CAD*, vol. 21, no. 3, pp. 349–361, 2002.

[5] G. Roberts and S. Aouini, "An overview of mixed-signal production test from a measurement principle perspective," *IEEE Design & Test*, pp. 1–1, 2013.

[6] E. Yilmaz and S. Ozev, "Adaptive test elimination for analog/RF circuits," in *ACM/IEEE Design Automation Conf.*, Jul. 2009, pp. 720–725.

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

[8] N. Akkouche, S. Mir, E. Simeu, and M. Slamani, "Analog/RF test ordering in the early stages of production testing," in *IEEE VLSI Test Symp.*, 2012, pp. 25–30.

[9] M. J. Barragan, R. Fiorelli, G. Leger, A. Rueda, and J. L. Huertas, "Alternate test of LNAs through ensemble learning of on-chip digital envelope signatures," *Journal of Electronic Testing*, vol. 27, no. 3, pp. 277–288, Jan. 2011.

[10] M. Barragan, R. Fiorelli, G. Leger, A. Rueda, and J. Huertas, "Improving the accuracy of RF alternate test using multi-VDD conditions: Application to envelope-based test of LNAs," in *IEEE Asian Test Symp.*, Nov. 2011, pp. 359 –364.

[11] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, no. 1, 1997.