



HAL
open science

Estimation of Analog/RF Parametric Test Metrics Based on a Multivariate Extreme Value Model

Ahcène Bounceur, Salvador Mir, R. Euler, K. Beznia

► **To cite this version:**

Ahcène Bounceur, Salvador Mir, R. Euler, K. Beznia. Estimation of Analog/RF Parametric Test Metrics Based on a Multivariate Extreme Value Model. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020, 39 (5), pp.966-976. 10.1109/TCAD.2019.2907923 . hal-02089332

HAL Id: hal-02089332

<https://hal.science/hal-02089332v1>

Submitted on 14 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Estimation of Analog/RF Parametric Test Metrics Based on a Multivariate Extreme Value Model

Ahcène Bounceur¹, Salvador Mir², Reinhardt Euler¹ and Kamel Beznia¹

¹CNRS - Lab-STICC Laboratory - Université de Bretagne Occidentale

20 Avenue Victor Le Gorgeu, 29238, Brest, France

Email: {Ahcene.Bounceur, Reinhardt.Euler}@univ-brest.fr

² Univ. Grenoble Alpes, CNRS, Grenoble INP, TIMA

38000, Grenoble, France

Email: Salvador.Mir@univ-grenoble-alpes.fr

Abstract—Analog/RF built-in test (BIT) techniques are essential for reducing the very high costs of specification-based tests and for high-safety applications. The adoption of a BIT technique needs to be decided at the design stage, and this can be facilitated by estimating the test quality in terms of errors such as Test Escapes (T_E) and Yield Loss (Y_L). Test quality estimation at the design stage has been traditionally very difficult for analog/RF circuits due to the lack of fault models that properly cover parametric faulty behavior. In recent years, statistical simulation has been considered in combination with learning techniques for the estimation of parametric test metrics. Extreme Value Theory (EVT) has provided a rigorous tool for the computation of parametric test metrics. However, test metrics estimation has been limited to the use of a univariate model. In this paper, we extend this approach by using a multivariate extreme value model. We illustrate this for the evaluation of a RF LNA BIT technique using a bivariate model.

I. INTRODUCTION

The design and development of embedded analog/RF test techniques is essential for tackling the large costs required for testing mixed-signal/RF blocks in today SoCs. Testing these blocks represents a disproportionate share of the total test cost, given the small area these circuits occupy in a SoC. Many different approaches have been presented for common blocks such as signal converters, frequency synthesizers or RF front-ends. However, the lack of structured integrated test solutions for these circuits has prevented widespread development. Indeed, each mixed-signal or RF block requires customized BIT solutions that are hard to evaluate before production. Without tools for an adequate evaluation at the design stage, it is not possible to explore the most convenient solutions, and it becomes often impossible to convince design and product managers to embed any of them.

The difficulties for test metrics estimation at the design stage for mixed-signal/RF circuits stem from the lack of adequate analog/RF fault models that properly cover parametric faulty behavior. This type of faulty behavior is in general due to multiple process and design parameters that are varying simultaneously. We do not consider in this work defects that modify the topology of the circuit since this is most often treated in the literature following an approach similar to catastrophic faults for digital circuits. For parametric faults, Monte Carlo circuit level simulation can be considered to generate a sample

of circuit instances under process variations. Each circuit instance is represented by a vector of its output parameters (the performances specified in the data sheet and the actual test measures). A circuit instance is functional if all the performances meet the specifications. A circuit instance passes the test if all the test measures are within their pre-defined limits. A test error occurs when either a faulty circuit escapes the test (i.e., Test Escapes, T_E) or a functional circuit fails the test (i.e., Yield Loss, Y_L). Parametric test metrics estimation with ppm precision requires a very large sample of circuits ($> 10^6$ circuits). Unfortunately, Monte Carlo circuit simulation is time consuming and can only generate a relatively small sample in reasonable time. In the literature, several methods have been proposed to overcome this limitation. They can be divided into two types [1].

The first type of methods requires an estimation of the joint probability density function (PDF) of the output parameters using an initial small sample of circuits generated via Monte Carlo simulation. Next, by sampling this statistical model, it is possible to numerically generate an arbitrarily large sample of circuits that follows the same distribution. This large sample will contain a representative set of parametric defective devices, and test metrics can be calculated using relative frequencies. This type of methods is simple to apply since only the output parameters of the circuit are considered and there are multiple techniques to estimate the joint PDF. In [2], the statistical model is assumed to be a multivariate Gaussian. As a general parametric approach, [3] proposes the use of copulas. A copula is a multivariate distribution that models the dependence between the output parameters, separately from the marginal distributions of these parameters. However, the copula model may be not be easy to find in all cases. When the copula model cannot be found, [4] proposes the use of a general non-parametric density estimation technique. The major problem with these techniques is that the estimated PDF of the output parameters may be inaccurate at the tails of the distribution, where parametric defective circuits are found.

The second type of methods use statistical learning to accelerate Monte Carlo simulation [5] so that devices in the distribution tail can be more easily generated. A multidimensional classifier is first trained to learn a boundary within

the input parameter space that separates devices with output parameters in the bulk of the distribution from devices with output parameters at the tails of this distribution. Next, during Monte Carlo electrical simulation, the simulation of a circuit instance is blocked if the classifier predicts from the input parameter vector that the output parameters will fall in the bulk of the distribution. Otherwise, the circuit instance is simulated. This type of methods is more difficult to apply since a multivariate classifier in the input parameter space must be trained. However, it provides a population of extreme circuits from which test metrics can be computed using an estimation of the tail density following a univariate Extreme Value Theory (EVT) model [5] [6] [7]. In [8], a multidimensional problem for test metrics estimation is considered using EVT although it is transformed to a one-dimensional problem and a univariate model is used.

In this paper, we consider the second type of approaches with a multivariate extreme value model. The paper is organized as follows. Section II formulates the computation of test metrics. The multivariate EVT model is presented in Section III. Next, Section IV describes two different procedures for test metrics estimation using extreme circuits: the first is based on the computation of relative frequencies and the second is based on the multivariate EVT model. For the validation of the proposed techniques, an RF LNA BIT technique is presented in Section V. The validation is facilitated by the fact that a database with 10^6 simulated instances has been made available. The results of the test metrics estimation with the new techniques are presented in Section VI. Finally, conclusions are presented in Section VII.

II. TEST METRICS

In this section we first present the definitions of the test metrics and next we describe the analytical equations to compute them.

A. Test metrics definition

In this paper, we will consider the test metrics Test Escapes (T_E) and Yield Loss (Y_L). T_E represents the proportion of the faulty circuits that pass the test and Y_L the proportion of the circuits failing the test within those that are functional. In terms of probabilities, test metrics are defined as follows:

$$\begin{aligned} T_E &= \Pr(\text{circuit is faulty} \mid \text{it passes the test}) \\ &= 1 - \Pr(\text{circuit is functional} \mid \text{it passes the test}), \end{aligned} \quad (1)$$

$$\begin{aligned} Y_L &= \Pr(\text{circuit fails the test} \mid \text{it is functional}) \\ &= 1 - \Pr(\text{circuit passes the test} \mid \text{it is functional}). \end{aligned} \quad (2)$$

Let us consider a circuit with n performances $\mathbf{P} = (P_1, P_2, \dots, P_n)$ and n specifications $s = (s_1, s_2, \dots, s_n)$. This circuit is designed such that each performance P_i satisfies the specification s_i (i.e. $P_i \leq s_i, i = 1, \dots, n$), considering without lack of generality a unilateral specification to simplify the notation. For this circuit we consider a set of m test measures $\mathbf{T} = (T_1, T_2, \dots, T_m)$ and m test limits

$l = (l_1, l_2, \dots, l_m)$. The circuit passes the test if each test measure T_j satisfies the test limit l_j (i.e. $T_j \leq l_j, j = 1, \dots, m$), considering again a unilateral test limit to simplify the notation. The test metrics are then written as

$$\begin{aligned} T_E &= 1 - \frac{\Pr(P_1 \leq s_1, \dots, P_n \leq s_n, T_1 \leq l_1, \dots, T_m \leq l_m)}{\Pr(T_1 \leq l_1, \dots, T_m \leq l_m)} \\ &= 1 - \frac{F_{PT}(s_1, s_2, \dots, s_n, l_1, l_2, \dots, l_m)}{F_T(l_1, l_2, \dots, l_m)}, \end{aligned} \quad (3)$$

$$\begin{aligned} Y_L &= 1 - \frac{\Pr(P_1 \leq s_1, \dots, P_n \leq s_n, T_1 \leq l_1, \dots, T_m \leq l_m)}{\Pr(P_1 \leq s_1, \dots, P_n \leq s_n)} \\ &= 1 - \frac{F_{PT}(s_1, s_2, \dots, s_n, l_1, l_2, \dots, l_m)}{F_P(s_1, s_1, \dots, s_n)} \end{aligned} \quad (4)$$

where $F_P(s_1, s_2, \dots, s_n)$ is the joint Cumulative Distribution Function (CDF) of the performances, $F_T(t_1, t_2, \dots, t_m)$ is the joint CDF of the test measures, and $F_{PT}(s_1, s_2, \dots, s_n, t_1, t_2, \dots, t_m)$ is the joint CDF of the performances and the test measures.

B. Definition of test metrics for extreme data analysis

Let us now consider circuit instances that have output parameters exceeding a large threshold for the performances $\mathbf{u} = (u_1, u_2, \dots, u_n)$, where $u_i < s_i, (i = 1, \dots, n)$, for the case of n specifications, and/or a large threshold for the test measures $\mathbf{v} = (v_1, v_2, \dots, v_m)$ where $v_j < l_j, (j = 1, \dots, m)$, for the case of m test limits. We can then define test metrics considering threshold exceedances, so that extreme value data analysis can be applied [6].

We consider first the case of test escapes T_E as defined by Equation (1). If we only consider the population of circuits that pass the test, T_E becomes the probability that a circuit is faulty, i.e., the probability to violate at least one specification as follows:

$$T_E = \Pr(P_1 > s_1 \vee P_2 > s_2 \vee \dots \vee P_n > s_n). \quad (5)$$

Let us define condition $A = (P_1 > s_1 \vee P_2 > s_2 \vee \dots \vee P_n > s_n)$ and condition $B = (P_1 > u_1 \vee P_2 > u_2 \vee \dots \vee P_n > u_n)$ such that A implies B ($A \subseteq B$), and where $\Pr(A)$ is the probability that the circuit is faulty and $\Pr(B)$ the probability that the circuit is extreme. Then T_E can be formulated as follows:

$$\begin{aligned} T_E &= \Pr(A) \\ &= \Pr(A \wedge B) \\ &= \Pr(A \mid B) \times \Pr(B). \end{aligned} \quad (6)$$

We replace A and B in Equation (6) and obtain:

$$\begin{aligned} T_E &= \Pr((P_1 > s_1 \vee P_2 > s_2 \vee \dots \vee P_n > s_n) \mid \\ &\quad (P_1 > u_1 \vee P_2 > u_2 \vee \dots \vee P_n > u_n)) \times \\ &\quad \Pr(P_1 > u_1 \vee P_2 > u_2 \vee \dots \vee P_n > u_n). \end{aligned} \quad (7)$$

EVT allows an estimation of the conditional probability that appears as a first term in Equation (7), while the second term

is estimated via standard probability theory. Similarly as for T_E , we can estimate Y_L considering only the population of circuits that are functional. We obtain in this case:

$$Y_L = \Pr((T_1 > l_1 \vee T_2 > l_2 \vee \dots \vee T_m > l_m) | (T_1 > v_1, T_2 > v_2 \vee \dots \vee T_m > v_m)) \times \Pr(T_1 > v_1 \vee T_2 > v_2 \vee \dots \vee T_m > v_m). \quad (8)$$

III. MULTIVARIATE EXTREME VALUE MODEL

Multivariate EVT provides a mathematical framework to deal with the case when the largest values of multiple variables of interest tend to occur simultaneously. The probability of multiple variables being simultaneously large are thus estimated by means of a Multivariate Extreme Value Distribution (MEVD). There are two main methods to obtain this distribution:

- The *Block Maxima* method considers large blocks of independent data samples and then estimates the distribution of the maxima of the blocks. These maxima fit a Generalized Extreme Value Distribution (GEVD).
- The *Peaks Over Threshold* method considers the exceedances of the data samples over a large threshold. These exceedances fit a Generalized Pareto Distribution (GPD).

We will use in this work the second type of methods since we are interested in modeling circuit output parameters that exceed a large threshold defined by a parameter specification or a test limit. In this Section, we first describe the estimation of the threshold of extreme exceedances. Next, we present the modeling of multivariate exceedances using a multivariate GPD. Finally, we show how to estimate the parameters of this model.

A. Estimating the threshold of exceedances

The choice of the threshold u that defines the exceedances of a random variable X is crucial in EVT. The threshold must be sufficiently high so that the distribution can be correctly approximated in the extreme by a GPD distribution with probability density function $h(x)$ given by

$$h(x) = \frac{1}{\sigma} \left(1 + \frac{\xi(x - \mu)}{\sigma} \right)^{\left(-\frac{1}{\xi} - 1\right)} \quad (9)$$

where μ, σ and ξ represent, respectively, the mean, the standard deviation and the shape parameter. It must be noted that if the threshold is too high, the data set of extreme exceedances will be too small, leading to an inaccurate estimation of the GPD model parameters.

There are several methods to estimate the threshold of exceedances. The method based on the *Mean Excess Function (MEF)* is very common. Assuming that the extreme data follows a GPD distribution with shape parameter $\xi < 1$, the *MEF* $e(u)$ follows the linear function

$$e(u) = \mathbb{E}[X - u | X > u] = \frac{\sigma + \xi u}{1 - \xi}. \quad (10)$$

For a set of n exceedances, the function $e(u)$ is estimated by

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n (X_i - u) \mathbb{I}_{X_i > u}}{\sum_{i=1}^n \mathbb{I}_{X_i > u}} \quad (11)$$

which must be linear beyond the selected threshold u , with u being as low as possible. Figure 1 shows an example for the selection of a threshold for the case study that will be described later.

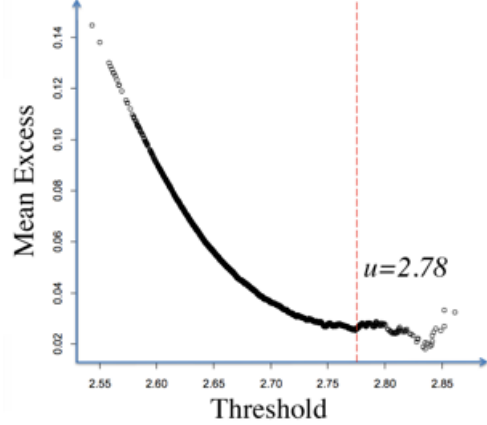


Fig. 1. Estimation of the threshold of exceedances by the Mean Excess Function.

B. Modeling of multivariate threshold exceedances

For the sake of simplicity, we consider a bivariate case. Let (Z_1, Z_2) be the observed random vector, (u_1, u_2) a given threshold vector and $(X, Y) = (Z_1 - u_1, Z_2 - u_2)$ the random vector of the exceedances of the random vector beyond the threshold. There exist two definitions of Bivariate Generalized Pareto Distributions (BGPD) for the exceedances: BGPD-I and BGPD-II. The distributions of the first type consider values that jointly exceed the thresholds of all variables. The distributions of the second type consider values that exceed the thresholds for at least one variable and we will use them in this work. The BGPD-II is given by

$$H(x, y) = \frac{-1}{\log G(0, 0)} \log \frac{G(x, y)}{G(\min\{x, 0\}, \min\{y, 0\})} \quad (12)$$

where the marginal distributions of X and Y must be Generalized Pareto Distributions [9], and G so that $0 < G(0, 0) < 1$ is a Bivariate Extreme Value Distribution (BEVD) with non-degenerate margins. As mentioned, this distribution provides a model for observations that are extreme in at least one variable. Following [10], this distribution can be written as

$$G(x, y) = \exp[-V(x, y)]. \quad (13)$$

Note that there exist several families of functions G depending on the parametric function V (see for example [10]). As an example, Figure 2(a) shows 1000 samples generated from a BEVD G in our case-study, and Figure 2(b) shows 1000 samples generated from the corresponding BGPD-II H .

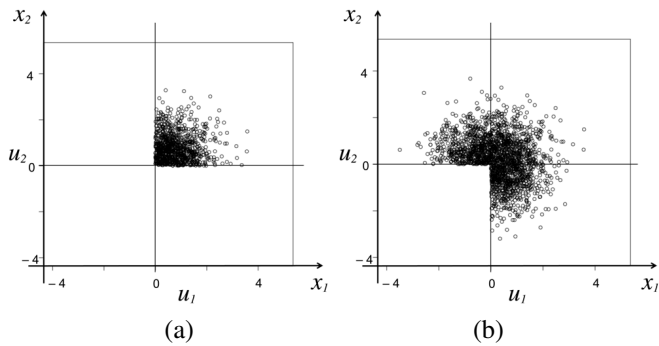


Fig. 2. (a) 1000 samples generated from the BEVD G in our case-study, and (b) 1000 samples generated from the corresponding BGPD-II H .

C. Estimation of the BGPD-II model

1) *The dependence in the extremes:* The estimation of the BGPD-II model requires first to estimate the dependence in the extremes of the multivariate distribution. Figure 3 shows the procedure to follow.

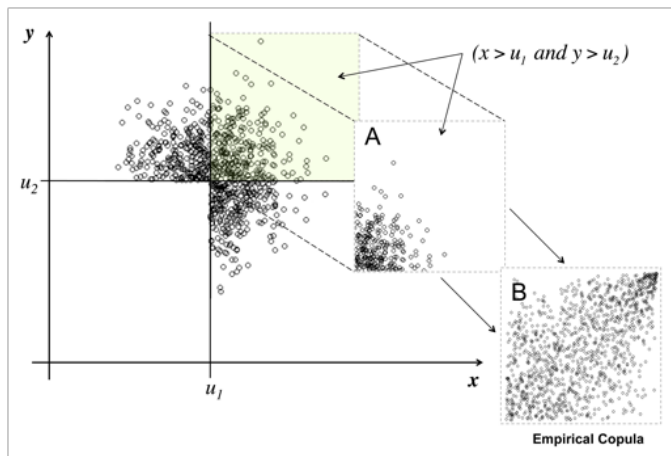


Fig. 3. The dependence in the extremes (α).

First we extract only the extreme samples that verify $x > u_1$ and $y > u_2$ (the samples of the box A in Figure 3). Next, we calculate the empirical copula of these samples (the samples of the box B in Figure 3) [3]. The dependence parameter α of these samples depends on the type of the obtained copula. In the case of the Gumbel copula which belongs to the family of extreme copulas, α is estimated as follows:

$$\hat{\alpha} = \frac{1}{1 - \hat{\tau}} \quad (14)$$

where $\hat{\tau}$ is Kendall's tau factor estimated as follows [12][13]:

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{i < j} \text{sgn}[(x_i - x_j)(y_i - y_j)], \quad i, j = 1, \dots, n \quad (15)$$

where

$$\text{sgn}(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0 \end{cases}$$

and $(x_1, y_1), \dots, (x_n, y_n)$ are n observations from a vector (X, Y) of continuous random variables.

2) *The Negative Logistic model:* The MGPD model is defined by the underlying Multivariate Extreme Value Distribution (MEVD) model, and practical MEVD models are defined by the dependence structure. In order to choose the most adequate model, it is necessary to visualize the dependence in the extremes. After the model is validated visually, it must be statistically validated with a corresponding goodness of fit test. In this paper we describe the Negative Logistic model [10] because it fits our case-study data. We use the EVT package of the R software [11] for validation. This model is given by the function

$$V(x, y) = \frac{1}{x} + \frac{1}{y} - (x^\alpha + y^\alpha)^{-\frac{1}{\alpha}}, \quad \alpha > 0 \quad (16)$$

where α is the dependence on the extremes as estimated by Equation (14). Figure 4 shows the contour plot of the probability density function (PDF) of this distribution with the Negative Logistic model, for different values of the dependence parameter α .

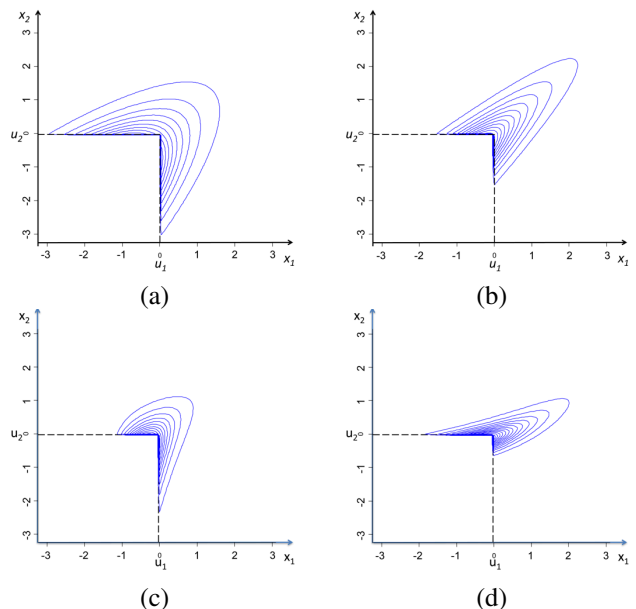


Fig. 4. Contour plot of the BGPD PDF with the Negative Logistic model for dependencies (a) $\alpha = 1.2$, (b) $\alpha = 2$, (c) $\alpha = 2.5$ and (d) $\alpha = 3.5$ and for different marginal distributions.

IV. TEST METRICS ESTIMATION USING MULTIVARIATE EXTREME DATA ANALYSIS

In this Section, we will first describe the estimation of test metrics using empirical relative frequencies of extreme circuits. Next, we will present the BGPD-II model for their estimation.

A. Estimation of test metrics using relative frequencies of extreme circuits

By using Equations (7) and (8), we can attempt a direct estimation of test metrics using relative frequencies on a sample of

circuits generated using Monte Carlo electrical simulation and statistical learning. A first Monte Carlo simulation of the CUT is required to find the threshold vector \mathbf{u} of the performances and the threshold vector \mathbf{v} of the test measures, to estimate the probability of extreme circuits and to obtain a classifier for speeding up the generation of extreme circuits. A second Monte Carlo simulation takes into account this classifier for the generation of extreme circuits and the estimation of test errors.

Considering first the case of test escapes as given by Equation (7), T_E is defined, within the set of circuits that pass the test, as the probability that the circuit is faulty knowing that it is extreme, times the probability that the circuit is extreme. This can be written as follows:

$$\hat{T}_E = \frac{N_{ed}}{N_e} \times \frac{N_u}{N_p} \quad (17)$$

where N_e is the number of extreme circuits (in the second Monte Carlo circuit simulation), N_{ed} is the number of the faulty extreme circuits (in the second Monte Carlo circuit simulation), N_u is the number of extreme circuits (those exceeding the high threshold \mathbf{u} in the first Monte Carlo circuit simulation), and N_p is the number of circuits that pass the test (in the first Monte Carlo circuit simulation). These parameters are illustrated in Figure 5.

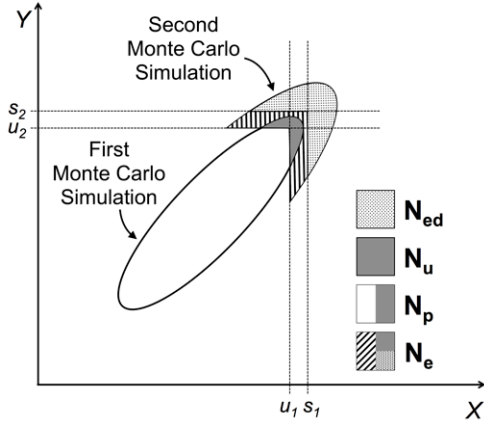


Fig. 5. Parameters used to estimate test metrics using relative frequencies of extreme circuits.

In the same way, for the case of yield loss as given by Equation (8), Y_L is defined within the set of circuits that are functional as the probability that the circuit fails the test knowing that it is extreme, times the probability that the circuit is extreme. This can be written as follows:

$$\hat{Y}_L = \frac{N_{ef}}{N_e} \times \frac{N_v}{N_g} \quad (18)$$

where N_e is the number of extreme circuits (in the second Monte Carlo circuit simulation), N_{ef} is the number of the extreme circuits that fail the test (in the second Monte Carlo circuit simulation), N_v is the number of extreme circuits (those exceeding the high thresholds \mathbf{v} in the first Monte Carlo

circuit simulation), and N_g is the number of circuits that are functional (in the first Monte Carlo circuit simulation).

In summary, the procedure to estimate T_E according to Equation (7) is as follows:

- 1) Run a first Monte Carlo circuit simulation in order to generate N circuits, and build a classifier as in [5].
- 2) Consider $N_p \leq N$ circuits $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{N_p}$ that pass the test with n performances (i.e., $\mathbf{p}_i = (p_i^1, p_i^2, \dots, p_i^n)$ where $i = 1, \dots, N_p$).
- 3) Calculate thresholds u_j for each performance P_j . Each threshold can be calculated in a precise manner using the Mean Excess Function (MEF) as shown in Section III-A if sufficient extreme values have been generated in the first Monte Carlo simulation (Notice that if the threshold is set automatically at a given q-quantile of the distribution like 97%, 3% of the values p_i^j , $i = 1, \dots, N_p$ will be greater than u_j and thus extreme, but it will not be ensured that they will follow a GPD).
- 4) Run a second Monte Carlo circuit simulation in order to generate N_e extreme circuits using the classifier learnt in the first simulation.
- 5) From this new sample, select the N_{ed} circuits that are faulty, that is, those that violate at least one specification.
- 6) Estimate T_E using Equation (17).

A similar procedure is used to estimate Y_L using Equation (18).

B. Estimation of test metrics with a BGPD-II model

We develop next the BGPD-II model for test metrics estimation. From Equation (7) we can write:

$$\begin{aligned} T_E &= \Pr((P_1 > s_1 \vee P_2 > s_2, \dots, P_n > s_n) \mid \\ &\quad (P_1 > u_1 \vee P_2 > u_2, \dots, P_n > u_n)) \times \\ &\quad \Pr(P_1 > u_1 \vee P_2 > u_2, \dots, P_n > u_n) \\ &= [1 - \Pr((P_1 \leq s_1 \vee P_2 \leq s_2, \dots, P_n \leq s_n) \mid \\ &\quad (P_1 > u_1 \vee P_2 > u_2, \dots, P_n > u_n))] \times \\ &\quad [1 - \Pr(P_1 \leq u_1 \vee P_2 \leq u_2, \dots, P_n \leq u_n)] \\ &= [1 - F_A(s_1, s_2, \dots, s_n)] \times \\ &\quad [1 - F_B(u_1, u_2, \dots, u_n)] \end{aligned} \quad (19)$$

where $F_A(s_1, s_2, \dots, s_n)$ is a MGPD-II calculated using Equation (12), so that

$$F_A(s_1, s_2, \dots, s_n) = H(s_1, s_2, \dots, s_n). \quad (20)$$

$F_B(u_1, u_2, \dots, u_n)$ can be expressed in terms of the copula C_P of the performances as

$$F_B(u_1, \dots, u_n) = C_P(F_{P_1}(u_1), \dots, F_{P_n}(u_n)). \quad (21)$$

Figure 6 (a) shows the distributions $F_B(u_1, u_2)$ and $(1 - F_B(u_1, u_2))$ and Figure 6 (b) shows the distributions $F_A(s_1, s_2)$ and $(1 - F_A(s_1, s_2))$ in the bivariate case. The Test escapes T_E is calculated as

$$\begin{aligned} T_E &= [1 - H(s_1, s_2)] \times [1 - C_P(F_{P_1}(u_1), F_{P_2}(u_2))] \\ &= \left[1 - \left(\frac{-1}{\log G(0, 0)} \log \frac{G(s_1, s_2)}{G(\min\{s_1, 0\}, \min\{s_2, 0\})} \right) \right] \times \\ &\quad [1 - C_P(F_{P_1}(u_1), F_{P_2}(u_2))]. \end{aligned} \quad (22)$$

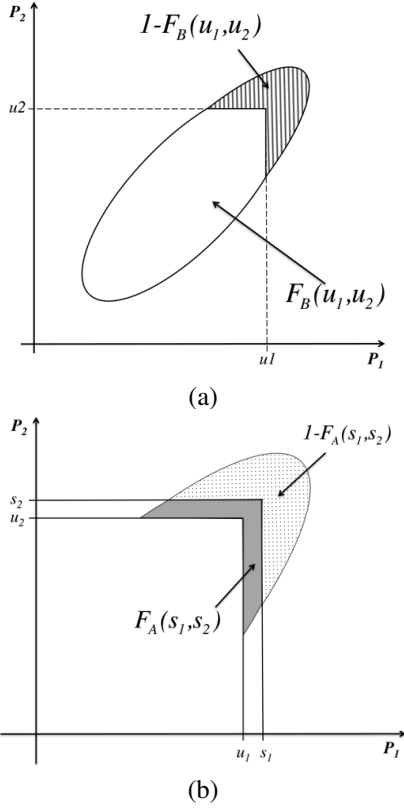


Fig. 6. (a) Distributions $F_B(u_1, u_2)$ and $(1-F_B(u_1, u_2))$, (b) Distributions $F_A(s_1, s_2)$ and $(1-F_A(s_1, s_2))$.

In the same way, we can calculate the Yield Loss Y_L with the extreme value model as

$$Y_L = [1 - F_C(l_1, l_2, \dots, l_m)] \times [1 - F_D(v_1, v_2, \dots, v_m)] \quad (23)$$

where $F_C(l_1, l_2, \dots, l_m)$ is a MGPD-II) calculated using Equation (12). $F_D(v_1, v_2, \dots, v_m)$ can be expressed in terms of the copula C_T of the test measures as

$$F_B(v_1, \dots, v_m) = C_T(F_{P_1}(v_1), \dots, F_{P_n}(v_m)) \quad (24)$$

and the Yield loss Y_L is calculated as

$$Y_L = [1 - H(l_1, l_2)] \times [1 - C_T(F_{T_1}(v_1), F_{T_2}(v_2))] = \left[1 - \left(\frac{-1}{\log G(0, 0)} \log \frac{G(l_1, l_2)}{G(\min\{l_1, 0\}, \min\{l_2, 0\})} \right) \right] \times [1 - C_T(F_{T_1}(v_1), F_{T_2}(v_2))] \quad (25)$$

V. TEST VEHICLE

Our case study is an RF Low Noise Amplifier (LNA) that operates at 2.4 GHz and which is commonly used in narrow band applications like Wifi and Bluetooth. The schematic is shown in Figure 7.

Two sensors have been studied for a low-cost BIT solution. First, an envelope detector (ED) is considered at the output of the LNA. This sensor consists of a half-wave rectifier followed by a low-pass filter that extracts the DC component from the rectified signal [14]. The output of this sensor is directly

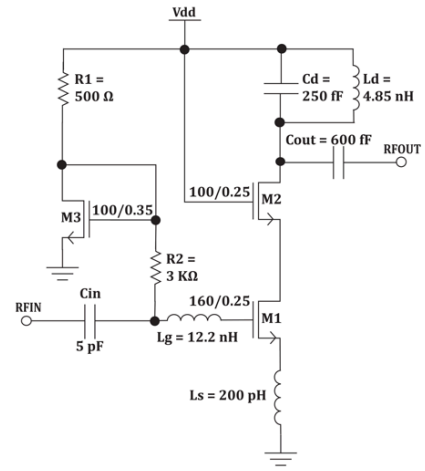


Fig. 7. CMOS inductive degeneration cascode LNA

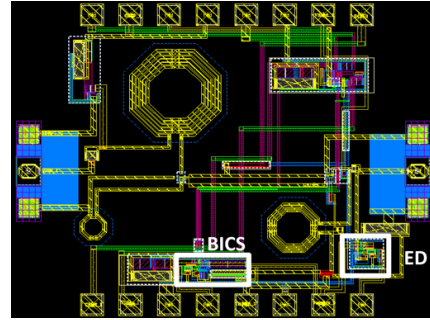


Fig. 8. Layout of the chip.

related to the RF output amplitude of the LNA. Second, a built-in current sensor (BICS) is used. It takes advantage of the small resistance of the power line that connects the core of the LNA to the power supply pad. Whenever a current I_{CUT} flows into the LNA, it generates a small voltage drop across the resistor. This drop is processed by the sensor in order to extract an output voltage proportional to the dynamic power supply current. For measurement, the BICS output is connected to the input of the envelop detector to extract the DC component.

The chip design and layout are realised with a $0.25\mu m$ BiCMOS process technology provided by NXP semiconductors. The layout of the LNA and the BIT sensors is shown in Figure 8. Special care is taken to account for layout-induced parasitics, including capacitive, resistive, inductive, as well as mutual inductance effects. The extraction of parasitics is carried out using the Assura tool and the circuits are re-sized to meet the performance requirements. Post-layout simulations using Spectre RF show that the nominal main performances of the LNA are: $NF = 2.72$ dB, $S_{11} = -22$ dB, $Gain = 16.4$ dB, $IIP1 = -7.5$ dBm and $IIP3 = 3$ dBm.

We have a data set of a Monte Carlo circuit simulation that contains 10^6 samples with the LNA performances and the BIT sensor outputs [15]. For each sample, we have as output parameters $\{NF, S_{11}, Gain\}$ and the test measures

$\{T_{ED}, T_{CS}\}$. IIP1 and IIP3 are not available since their simulation takes a very long time and it is not feasible for 10^6 circuits. This simulation took about three months. In order to use the bivariate extreme value model described in this work for the estimation of T_E and Y_L , we select two out of the three performances. We have also the two test measures. We notice that the three-dimensional problem with three performances needs to consider all pairs of performances and a more complex model that is beyond the scope of this paper. The performances that we have considered for illustration in this work are *Gain* and *NF*, but it could have been any other pair. The specifications of the two performances are set at $k_1\sigma$, i.e.,

$$\begin{aligned} Gain &\geq s_{Gain} = \mu_{Gain} - k_1 \cdot \sigma_{Gain} \\ NF &\leq s_{NF} = \mu_{NF} + k_1 \cdot \sigma_{NF}. \end{aligned}$$

The test limits on T_{ED} and T_{CS} are set at k_2 , i.e.,

$$\begin{aligned} T_{ED} &\geq l_{T_{ED}} = \mu_{T_{ED}} - k_2 \cdot \sigma_{T_{ED}} \\ T_{CS} &\geq l_{T_{CS}} = \mu_{T_{CS}} - k_2 \cdot \sigma_{T_{CS}}. \end{aligned}$$

VI. TEST METRICS ESTIMATION

In this Section, we will use the BGPD-II model to evaluate the test metrics of the RF LNA BIT technique presented above using a reduced number of extreme samples and we will compare the results with those obtained using the full data set.

A. Direct estimation using the full data set

Since we have a set of 10^6 samples, we can directly estimate test metrics with ppm precision using the estimators given by Equations (1) and (2). Figure 9 shows a scatterplot of 10^5 LNA samples from the full data set.

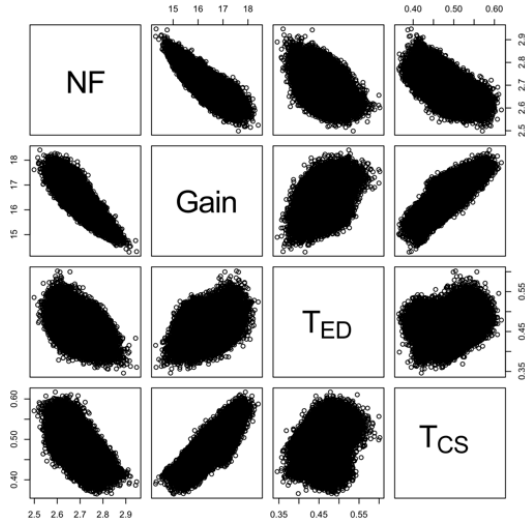


Fig. 9. 10^5 samples of LNAs generated from the Monte Carlo circuit level simulation.

We have fixed the specifications at $k_1 = 4$, whereas the test limits are set at k_2 ranging from 2 to 6 with a step of 0.1. The obtained results for the T_E and the Y_L are illustrated by the curves represented by the circular markers in Figure 10.

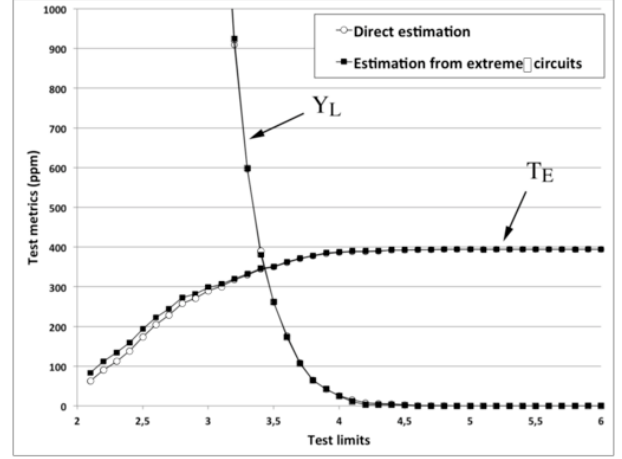


Fig. 10. Comparison of the test metrics (T_E and Y_L) obtained by the direct estimation and by the estimation over the set of $38 \cdot 10^3$ extreme circuits.

B. Estimation using relative frequencies of extreme circuits

We estimate here the test metrics considering the procedure presented in Section IV-A based on relative frequencies of extreme circuits, and we will compare them with those obtained using the direct estimation with 10^6 circuits.

We first consider $N = 5000$ Monte Carlo circuit simulations (in practice, we pick up the first 5000 instances from the set of 10^6). For different values of the test limits, from $k_2 = 2$ to $k_2 = 6$ with a step of 0.1, we select in each iteration only circuits that pass the test. For example, for $k_2 = 4$ the number of the circuits that pass the test is equal to $N_p = 5000$. Then, this set is used to calculate the thresholds $u_{NF} = 2.78\text{dB}$ and $u_{Gain} = 15.47\text{dB}$ using the *MEF* function presented in III-A. With these thresholds, we consider the second Monte Carlo circuit simulation that requires a classifier in order to generate a set of $N_e = 38 \cdot 10^3$ extreme circuits. In practice, we pick up the first $38 \cdot 10^3$ extreme circuits from the set of 10^6 . Among these, we select N_{ed} circuits that are faulty, that is, those that violate at least one specification. In the case where $k_1 = 4$ and $k_2 = 4$ the value of N_{ed} equals 147. The test metrics can then be estimated for each value of k_2 as mentioned above using Equations (17) and (18). To estimate the Y_L we start with the circuits that are functional instead of those that pass the test. The obtained results of the T_E and the Y_L are illustrated by the curves represented by the squared markers in Figure 10.

We can see that the results obtained using $N_e = 38 \cdot 10^3$ extreme circuits are very close to those obtained for the full set of 10^6 circuits, which represents a reduction of 96%. In summary, an accurate estimation of test metrics is obtained with a fast Monte Carlo simulation. Figure 11 shows the tolerance intervals $[\hat{m} - 3\hat{\sigma}, \hat{m} + 3\hat{\sigma}]$ that are obtained by repeating the computation 100 times for different data sets and by estimating the mean \hat{m} and the standard deviation $\hat{\sigma}$ of each test metric.

Note that these results are obtained with extreme circuits obtained from the full data set. This corresponds to the case of having an ideal classifier. In general, we will not have

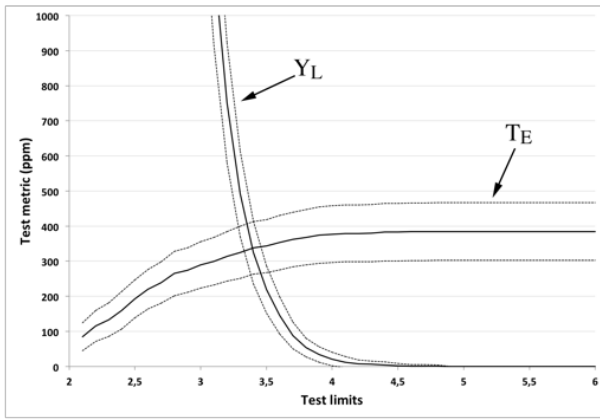


Fig. 11. Tolerance intervals for the test metrics (T_E and Y_L) as a function of $k_1 = 4$ and $k_2 = 2$ to 6 with a step of 0.1 (estimated from a set of $38 \cdot 10^3$ extreme circuits).

a full data set and the required classifier will make errors, biasing the distributions for the calculation of both terms in Equations (17) and (18). It must be noticed that it was not possible for us to build an actual classifier since we only have the data set. Alternatively, we have considered the injection of errors by randomly selecting a set of extreme and non extreme circuits and changing the status of the selected extreme circuits to non extreme circuits, and the status of the selected non extreme circuits to extreme circuits. We consider the case of 1000, 2000, 3000, 4000, 5000 and 10000 misclassified extreme circuits and we analyze the effect for the case of the Test Escapes T_E only, for simplicity, as shown in Figure 12. For the case of the test limits placed at 4σ , Figure 13 shows the mean error in the estimation of the test metrics for each level of misclassified circuits.

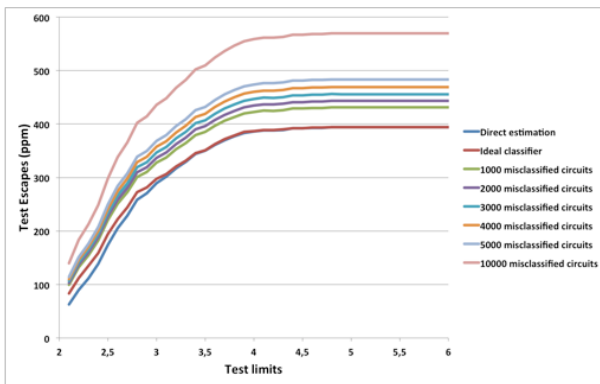


Fig. 12. Estimation of the mean Test Escapes for different error levels of misclassified extreme circuits using relative frequencies.

C. Estimation using the BGPD-II model

We estimate now the test metrics by constructing the model presented in Section IV-B and will compare the results of the different approaches. As before, we first run $N = 5000$ Monte Carlo circuit simulations for extracting the threshold of exceedances and the copula functions required in Equations

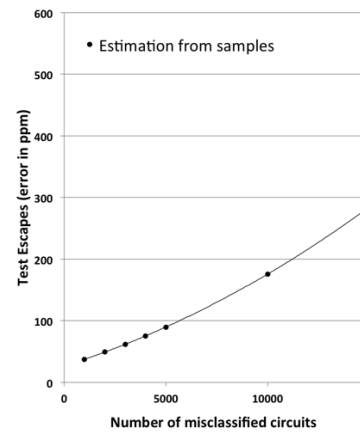


Fig. 13. Mean estimation error of the Test Escapes as a function of the level of misclassified extreme circuits using relative frequencies for the test limits placed at 4σ .

(22) and (25). In practice, we pick up the first 5000 instances from the set of 10^6 and we select only circuits that pass the test for the case of T_E . For example, for $k_2 = 4$ the number of the circuits that pass the test is equal to $N_p = 5000$. The obtained values of the thresholds are already calculated before $u_{NF} = 2.78\text{dB}$ and $u_{Gain} = 15.47\text{dB}$.

Figure 14 shows the marginal distributions $F_{NF}(x)$ and $F_{Gain}(x)$ that fit a GPD. The parameters of these distributions are given in Table I and they have been validated by the Kolmogorov-Smirnov test. Figure 15 shows the empirical copula $C_P(F_{NF}(x), F_{Gain}(x))$ which is Gaussian with a correlation factor $\rho = -0.84$. With all these parameters we calculate the second term of Equation (22) that gives

$$1 - C_{Gauss}(F_{NF}(u_{NF}), F_{Gain}(u_{Gain})) = 0.04008141.$$

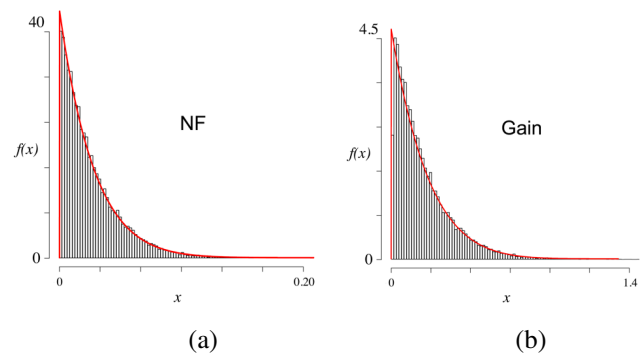


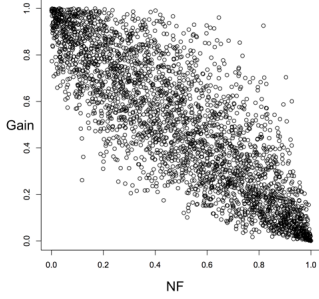
Fig. 14. The marginal distributions for NF (a) and $Gain$ (b) for the circuits that pass the test (exceedances beyond threshold u).

We calculate next the first term of the T_E . Considering the thresholds found in the previous step, we run the second Monte Carlo circuit simulation that requires a classifier in order to generate a set of $38 \cdot 10^3$ extreme circuits. In practice, we pick up the first $38 \cdot 10^3$ extreme circuits from the set of 10^6 that

TABLE I

THE PARAMETERS OF THE MARGINAL DISTRIBUTIONS OF NF AND $Gain$

Marginal Distribution	location μ	scale σ	shape ξ
NF	0	0.0249576	-0.0904312
$Gain$	0	0.011563367	0.0249576

Fig. 15. The empirical copula $C_P(Gain(x), NF(x))$.

verify $Gain < u_{Gain}$ or $NF > u_{NF}$. The obtained circuits are presented in Figure 16.

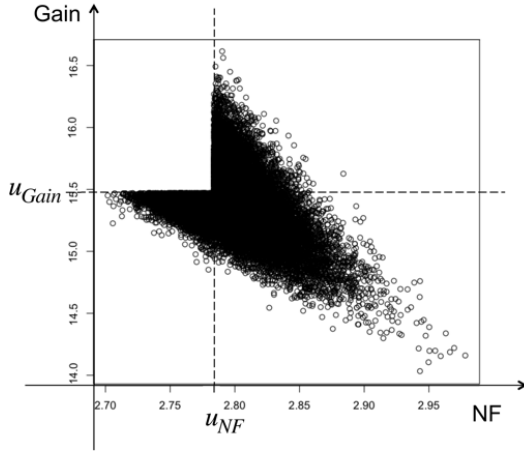
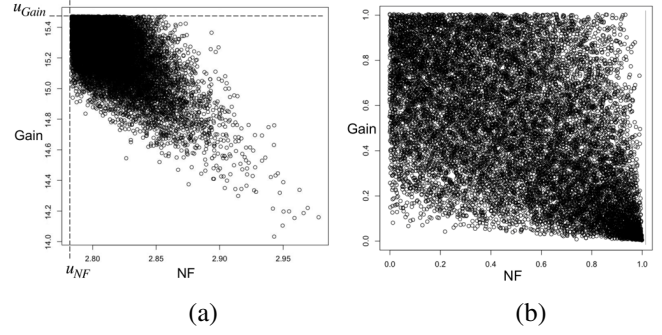


Fig. 16. Extreme circuits that pass the test.

By comparing Figures 16 and 4(b), we see that the model is likely a BGPD, the marginal distributions have already proven to be GPDs, and the extreme dependence model $V(x, y)$ is the Negative Logistic model given by Equation (16). This has been validated using the test proposed in [10]. This model depends on the dependence parameter α which is estimated from the copula of the extreme samples that verify $NF > u_{NF}$ and $Gain < u_{Gain}$ as illustrated in Section III-C.1. Figure 17(a) shows these samples and Figure 17(b) shows their empirical copula. This copula is a Gumbel copula and has been validated with the Cramér-Von Mises test [16]. Therefore, α is estimated using Equation (14).

We can now calculate the BEVD $G(x, y)$ given by Equation (13) and with this the first term of Equation (22). For example,

Fig. 17. (a) Extreme samples that verify $NF > u_{NF}$ and $Gain < u_{Gain}$, (b) their empirical copula.

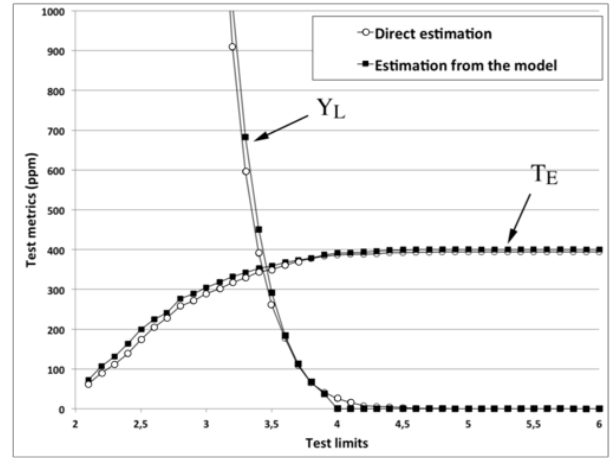
for the test limit $k_2 = 3.5$, the obtained value is

$$1 - \left(\frac{-1}{\log G(0, 0)} \log \frac{G(s_1, s_2)}{G(\min\{s_1, 0\}, \min\{s_2, 0\})} \right) = 0.00883411$$

and the value of the Test escapes T_E for this test limit is

$$T_E = 0.04008141 \times 0.00883411 = 0.000354 \text{ (354ppm)}.$$

We have calculated the T_E for different values of the test limits, from $k_2 = 2$ to $k_2 = 6$ with a step of 0.1 as shown with circular markers in Figure 18.

Fig. 18. Comparison of the test metrics (T_E and Y_L) obtained by the direct estimation with the full data set and by the estimation using the BGPD-II model with $38 \cdot 10^3$ extreme circuits.

The same procedure is used to estimate the Y_L with the following changes:

- Replace the set of circuits that pass the test by circuits that are functional.
- The Y_L is calculated with Equation (25), where the threshold u_1 must be replaced by the threshold $v_1 = -0.4109976$ and the threshold u_2 by $v_2 = -0.4176371$.
- Replace the copula C_P by C_T . We found that the copula C_T is Gaussian with a correlation factor $\rho = 0.188215$.
- The marginal distributions $F_{T_{ED}(x)}$ and $F_{T_{CS}(x)}$ are the Generalized Pareto Distributions. Their histograms fit a

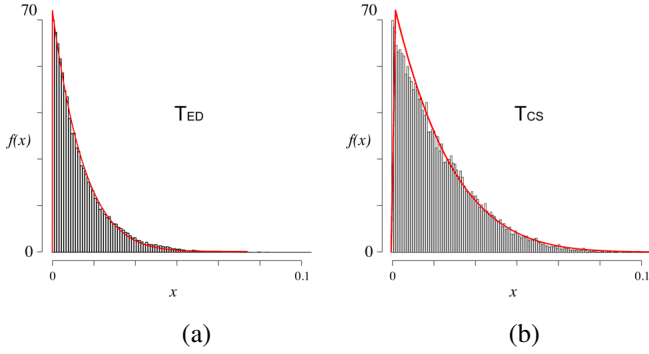


Fig. 19. The marginal distributions of the exceedances of the test measures T_{ED} (a) and T_{CS} (b).

TABLE II

THE PARAMETERS OF THE MARGINAL DISTRIBUTIONS OF T_{ED} AND T_{CS}

Marginal Distribution	location μ	scale σ	shape ξ
T_{ED}	0	0.0117069	-0.02050995
T_{CS}	0	0.0130899	-0.182502

GPD as shown in Figure 19 with the parameters given in Table II validated by the Kolmogorov-Smirnov test.

As an example, for a test limit of $k_2 = 3.5$, the first term of Equation (25) is equal to 0.006193789, the second term is equal to 0.05097219, and $Y_L = 0.006193789 \times 0.05097219 = 0.000315711$ (316ppm). We have calculated Y_L for different values of the test limits, from $k_2 = 2$ to $k_2 = 6$ with a step of 0.1 as shown in Figure 18. In summary, we can see that the results obtained with the BGPD-II are very close to those obtained with the full set of 10^6 circuits. The tolerance intervals are shown in Figure 20.

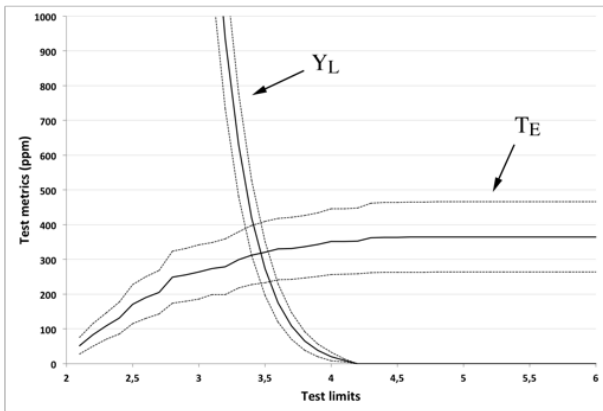


Fig. 20. Tolerance intervals for T_E and Y_L as a function of limits $k_1 = 4$ and $k_2 = 2$ to 6 with a step of 0.1 (estimated using the BGPD-II model with $38 \cdot 10^3$ extreme circuits).

Here again, we note that these results are obtained with extreme circuits from the full data set which corresponds to the case of an ideal classifier. As for the previous Section, we consider next the case of 1000, 2000, 3000, 4000, 5000 and

10000 misclassified extreme circuits as shown in Figure 21 for the case of T_E . For the case of the test limits placed at 4σ , Figure 22 shows the model mean error in the estimation for each level of misclassified circuits with square markers that we can compare with the mean error obtained with direct relative frequencies. The mean error in the estimation using the relative frequencies of extreme circuits is higher when the level of misclassified circuits is below 2500. In other words, the BGPD-II model is more accurate when the level of misclassified circuits is below 6.6% which is a very reasonable assumption.

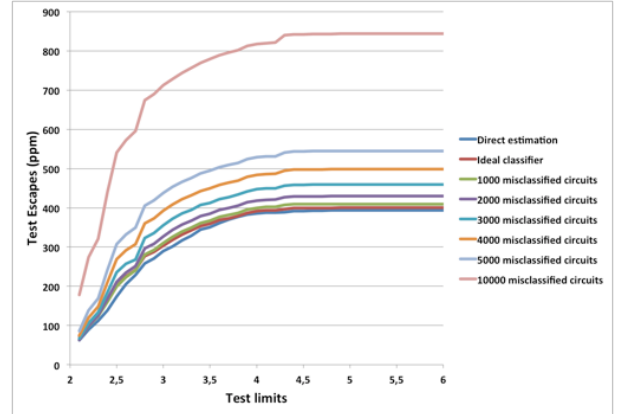


Fig. 21. Estimation of the mean Test Escapes for different error levels of misclassified extreme circuits using the BGPD-II model.

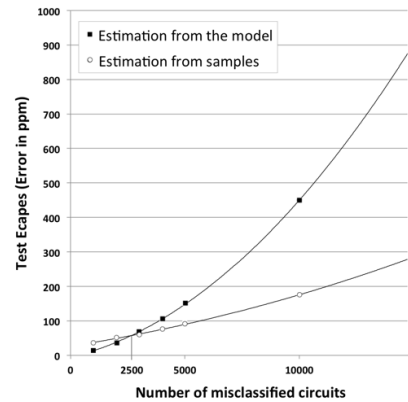


Fig. 22. Comparison of the T_E mean estimation error as a function of the level of misclassified extreme circuits for the two approaches (relative frequencies and the BGPD-II model) and the test limits placed at 4σ .

As future work, we expect to develop the MGPD-II model beyond two dimensions which will allow a study of the effect of the number of dimensions in the estimation error. Also, we will be able to compare the use of the MGPD-II approach described here with the technique proposed in [8], where the multidimensional problem has been transformed into a one-dimensional problem with the use of a univariate EVT model. This comparison could be done in terms of the number of extreme random variables, the test metrics estimation error and the number of extreme circuit samples required. We note,

however, that in our approach we do not require all variables to exceed the threshold at the same time as in the case of [8], and since we take advantage of the estimation of tail dependencies, it is likely that our approach will require a lower number of extreme samples with the corresponding reduction in electric circuit simulation time.

VII. CONCLUSION

A multivariate EVT method to estimate analog test metrics such as T_E and Y_L has been presented in this paper. To our knowledge, this is the first time that test metrics are estimated using a multivariate extreme value model. An RF LNA BIT technique is used as a case-study, showing very accurate results for the estimation of test metrics. With the specifications set at 4σ , a set of only $38 \cdot 10^3$ extreme circuits is considered instead of the full set of 10^6 circuits required to obtain ppm precision, which represents a reduction of 96%. This set of $38 \cdot 10^3$ extreme circuits can easily be generated by Monte Carlo circuit simulation, in combination with a statistical learning technique. We have presented two ways to estimate the test metrics using the extreme circuits. The first one is based on the direct estimation of relative frequencies over the set of extreme circuits. The second method is based on estimating a BGPD-II model from the set of extreme circuits. The estimation error is mostly dependent on the level of errors that the statistical learning classifier will produce. We have been able to show the advantage of the EVT model over a direct estimation under the presence of a reasonable number of misclassified extreme circuits. Further work will be directed to develop a multivariate model of higher dimensionality and to evaluate the quality of the estimation in terms of the number of extreme random variables and the number of extreme circuit samples required.

REFERENCES

- [1] K. Beznia, A. Bounceur, R. Euler, S. Mir, A Tool For Analog/RF BIST Evaluation Using Statistical Models of Circuit Parameters, *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, 2015, 20(2), Article No. 31.
- [2] A. Bounceur, S. Mir, E. Simeu, and L. Rolíndez. Estimation of test metrics for the optimisation of analogue circuit testing. *Journal of Electronic Testing: Theory and Applications*, 23(6), 2007, pp. 471-484.
- [3] A. Bounceur, S. Mir and H-G. Stratigopoulos. Estimation of analog parametric test metrics using copulas. *IEEE Trans. on CAD of ICs*, Sep. 2011, 30(9), pp. 1400-1410.
- [4] H. Stratigopoulos, S. Mir, and A. Bounceur. Evaluation of analog/RF test measurements at the design stage. In IEEE International Conference on CAD (ICCAD), 28(4), April 2009, pp. 582-590.
- [5] A. Singhee and R. A. Rutenbar, Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design. *IEEE Trans. on CAD of ICs*, 28(8), 2009, pp. 1176-1189.
- [6] H. Stratigopoulos. Test Metrics Model for Analog Test Development. *IEEE Trans. on CAD of ICs*, 31(7), 2012, pp. 1116-1128.
- [7] A. Spyronasios, L. Abdallah, H. Stratigopoulos and S. Mir. On Replacing an RF Test with an Alternative Measurement: Theory and a Case Study. In IEEE Asian Test Symposium, New Delhi, 2011, pp. 365-370.
- [8] H. Stratigopoulos, P. Faubet, Y. Courant and F. Mohamed. Multidimensional analog test metrics estimation using extreme value theory and statistical blockade. In 50th Design Automation Conference (DAC), June, 2013.
- [9] H. Rootzén and N. Tajvidi, Multivariate generalized Pareto distributions. *Bernoulli*, 12(10), 2006, pp. 917-930.
- [10] P. Rakonczai and A. Zempléni. Bivariate generalized Pareto distribution in practice: models and estimation. *Environmetrics*, 23(3), pp. 219-227, May 2012.
- [11] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2006. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- [12] H. Abdi. Kendall rank correlation. *Encyclopedia of Measurement and Statistics*. N. Salkind (Ed.). Sage Publications, 2007.
- [13] K. Beznia, A. Bounceur, S. Mir, R. Euler. Parametric test metrics estimation using non-Gaussian copulas. In IEEE International Mixed-Signals, Sensors, and Systems Test Workshop (IMS3TW'11), Santa Barbara, USA, 2011.
- [14] L. Abdallah, H. Stratigopoulos, C. Kelma and S. Mir. Sensors for built-in alternate RF test. In 15th IEEE European Test Symposium (ETS), 2010, 49-54.
- [15] K. Huang, H. Stratigopoulos, L. Abdallah, S. Mir, A. Bounceur. Multivariate Statistical Techniques for Analog Parametric Test Metrics Estimation. In 8th IEEE International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS'13), Abu Dhabi, United Arab Emirates, March 2013, pp. 6-11.
- [16] I. Kojadinovic and M. Holmes. Tests of independence among continuous random vectors based on Cramér-von Mises functionals of the empirical copula process. *Journal of Multivariate Analysis*, 100(6), 2009, pp. 1137-1154.