



HAL
open science

Development of Built-In Self-Test solutions for RF/mm-wave integrated circuits

Florent Cilici

► **To cite this version:**

Florent Cilici. Development of Built-In Self-Test solutions for RF/mm-wave integrated circuits. Micro and nanotechnologies/Microelectronics. Université Grenoble Alpes, 2019. English. NNT : 2019GREAT072 . tel-03132469

HAL Id: tel-03132469

<https://theses.hal.science/tel-03132469>

Submitted on 5 Feb 2021

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : NANO ELECTRONIQUE ET NANO TECHNOLOGIES

Arrêté ministériel : 25 mai 2016

Présentée par

Florent CILICI

Thèse dirigée par **Salvador MIR**, Directeur de recherche , CNRS
et codirigée par **Sylvain BOURDEL** , Enseignant-chercheur
préparée au sein du **Laboratoire Techniques de l'Informatique
et de la Microélectronique pour l'Architecture des systèmes
intégrés**
dans l'**École Doctorale Electronique, Electrotechnique,
Automatique, Traitement du Signal (EEATS)**

Développement de solutions BIST (Built-In Self-Test) pour circuits intégrés radiofréquences/millimétriques

Development of Built-In Self-Test solutions for RF/mm-wave integrated circuits

Thèse soutenue publiquement le **17 décembre 2019**,
devant le jury composé de :

Monsieur Philippe Ferrari

Professeur, Université Grenoble Alpes, Président

Monsieur Serge Bernard

Directeur de Recherche, LIRMM, Université Montpellier, Rapporteur

Madame Nathalie Deltimple

Maître de Conférences, Bordeaux INP, Laboratoire IMS, Rapporteur

Monsieur Benoît Poussard

Ingénieur, NXP Semiconductors, Toulouse, Examineur

Monsieur Thierry Parra

Professeur, Université de Toulouse III Paul Sabatier, LAAS, Examineur

Monsieur Manuel Barragan

Chargé de Recherche, TIMA, Grenoble, Co-encadrant de thèse

Madame Estelle Lauga-Larroze

Maître de Conférences, Université Grenoble Alpes, RFIC-Lab, Grenoble
Co-encadrant de thèse, Invité

Monsieur Sylvain Bourdel

Professeur, Grenoble INP, RFIC-Lab, Grenoble, Co-directeur de thèse

Monsieur Salvador Mir

Directeur de Recherche, TIMA, Grenoble, Directeur de thèse



Acknowledgements

First, I want to thank my supervisors Salvador Mir, Manuel Barragan, Sylvain Bourdel and Estelle Lauga-Larroze. Their human and technical qualities greatly helped me in making my Ph.D. a successful and most enriching adventure. In this line, I want to sincerely thank Gildas Leger from the Institute of Microelectronics of Seville for his ideas, work and welcome during my stay in Seville. I also want to thank all the researchers, professors, engineers and other colleagues whose support was of immense value: Antonio, Jean-Daniel, Philippe, Florence and Emmanuel from the RFIC Lab, Jean-Michel, Ekta, Walid and Abdelhaziz from IMEP-LAHC and Ayssar from CEA Leti.

I want to especially thank the jury for reviewing my doctoral dissertation and for their involvement my Ph.D. defense: Philippe Ferrari, Serge Bernard, Nathalie Deltimple, Benoît Poussard, Thierry Parra.

I want to especially acknowledge the precious help and efficiency of Alejandro, Mohamed, Frédéric and Ahmed with any computer, design kit or software-related problem.

Circuit fabrication by STMicroelectronics has been funded by the TARANTO project. In this regard, I acknowledge the contribution of P. Ferrari from RFIC-Lab laboratory, D. Gloria, and C. Durand from STMicroelectronics. All the measurements I had to perform would not have been possible without Loïc from CIME Nanotech, Nicolas from IMEP-LAHC and Fabrice from TIMA, I therefore warmly thank them.

I thank Youness, Anne-Laure, Laurence from TIMA, Lorraine from CIME Nanotech and Isabelle from IMEP-LAHC for the administrative support.

Of course thank you to all my close colleagues: Guissepe, Lisinius, Dimitris, Mahdi, Ekta, Jing, Edouard from IMEP-LAHC, Halim, Renato, Marc, Ghislain, Hani, Rshdee from TIMA, Serge, Jordan from RFIC Lab, and everyone who has shared some time and took part in creating a warm work atmosphere.

During the last three years I had the opportunity to teach several classes, especially in Phelma engineering school. This was a most enriching time, so thank you to all my colleagues: Laurent, Fatah, Daniel, Renato, Sylvain, Jean-Michel, Nicolas, Marc and Pierre.

Finally, I want to thank my fiancé, my family and my friends for their invaluable support.

Contents

1	Introduction to RF/mm-wave test	1
1.1	Introduction	1
1.2	Production testing	2
1.2.1	Characterization and production testing	2
1.2.2	Required equipment	2
1.2.3	Wafer probing	2
1.2.4	Calibration	3
1.2.5	Accuracy, repeatability, correlation	3
1.2.6	Test cost	3
1.3	Measures to be performed	4
1.3.1	Small signal measurements	4
1.3.2	Large signal measurements	5
1.3.3	Power efficiency	7
1.4	Alternatives to conventional test	7
1.4.1	Structural test	8
1.4.2	Built-Off Test	8
1.4.3	Built-In Test	8
1.4.4	Machine Learning Indirect Test	9
1.4.5	Conclusion	13
1.5	Calibration against process variations	13
1.5.1	Introduction	13
1.5.2	One-shot statistical calibration	14
1.6	Conclusion	15
2	State of the art	17
2.1	Built-In Self-Test	17
2.1.1	Intrusive methods	17
2.1.2	Non-intrusive methods	21
2.2	Calibration	24
2.2.1	Iterative methods	24
2.2.2	One-shot strategies	25
2.2.3	Towards self-healing	26
2.3	Conclusions	26

3	Assisted test design for non-intrusive MLIT	29
3.1	Introduction	29
3.2	Proposed approach	29
3.2.1	Step 1: Unveiling the root causes of parametric performance variation	30
3.2.2	Step 2: Assisted design of non-intrusive process variation sensors	32
3.3	Case study	33
3.3.1	Basics of linear power amplifier design	33
3.3.2	Design of a mm-wave power amplifier	36
3.3.3	Applying the proposed methodology to the designed PA	43
3.4	Conclusions	53
4	Assisted MLIT design of mm-wave ICs with custom passives	55
4.1	Introduction	55
4.2	Non-intrusive test program generation for generic mm-wave ICs	56
4.2.1	Problem statement	56
4.2.2	Step 1: Assessment of the root causes of the parametric performance variation	56
4.2.3	Step 2: Design of non-intrusive process variation sensors	58
4.3	Case Study	59
4.3.1	Device Under Test	59
4.3.2	Step 1: Assessment of parametric performance variation root causes	63
4.3.3	Step 2: Design of non-intrusive process variation sensors	68
4.4	Experimental results	69
4.4.1	Measurement setup	70
4.4.2	Fabricated set of chips	71
4.4.3	Signature measurement	71
4.4.4	PA measurement results	78
4.4.5	General results	81
4.5	Conclusions	82
5	One-shot calibration against process variation	85
5.1	Introduction	85
5.2	Design of a power amplifier with calibration capabilities	85
5.2.1	Power amplifier design	85
5.2.2	Tuning knobs design	87
5.3	A novel data generation method	92
5.3.1	Previous work	93
5.3.2	Data generation for training calibration models	93
5.3.3	Proposed random sampling results	95
5.3.4	Practical implementation	97
5.3.5	Conclusions	98
5.4	Calibration procedure and results	99
5.4.1	Calibration procedure and implementation	99
5.4.2	Calibration results	102
5.5	Conclusions	104

<i>CONTENTS</i>	vii
Conclusions	107
Publications	109
Bibliography	111
Appendices	117
Résumé	123

List of Figures

1.1	Illustration of incident and reflected waves for a two-port network.	5
1.2	Definition of 1 dB compression point.	6
1.3	Definition of IP3 (for voltage quantities).	7
1.4	Machine learning example: price versus area.	10
1.5	Machine learning example: predicted price versus real price.	11
1.6	Machine learning indirect test principle.	12
1.7	Generic test flow for machine learning indirect test.	12
1.8	a) Iterative calibration using external ATE, b) Iterative calibration using low-cost ATE.	14
1.9	One-shot calibration principle.	14
2.1	Direct loopback setup for integrated transceiver test.	19
2.2	Block diagram showing the required modifications to convert a DUT into an oscillator.	20
2.3	Block diagram of a generic non-intrusive indirect test based on process-aware sensors.	22
2.4	Self-calibration/self-healing using on-chip resources.	26
3.1	Assisted design of non-intrusive process variation sensors - Exploiting the sensors reusability by means of a library.	32
3.2	Basic diagram of a common source NMOS power amplifier.	33
3.3	Selection of the DC operating point using the $I_{ds}(V_{ds})$ curves.	34
3.4	Selection of the DC operating point taking into account R_{ds}	35
3.5	Transistor level schematic of the power amplifier under test.	37
3.6	Layout of the power amplifier under test.	38
3.7	Scattering parameters obtained through post-layout simulations.	39
3.8	Output power versus input power at 60 GHz - Compression point and saturation. . . .	39
3.9	3D view of the MOS transistor connections.	40
3.10	Top and 3D views of a $50\ \Omega$ microstrip line.	41
3.11	T-junction layout and model.	41
3.12	3D view of the decoupling cell.	42
3.13	Top view and model of the decoupling cell.	42
3.14	Exploration of the MC process parameter space for finding the root causes of gain parametric variation for the PA.	44
3.15	Scatterplot of predicted versus actual gain of the PA using the 18 most relevant MC parameters as signatures.	45

3.16	Scatterplot of predicted versus actual gain of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3.	49
3.17	RMS error for gain prediction versus the number of signatures.	49
3.18	Maximum and RMS error for gain prediction versus the number of signatures.	50
3.19	Scatterplot of predicted versus actual specification of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3.	51
3.20	Scatterplot of predicted versus actual specification of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3 (continued).	52
4.1	Conceptual block diagram of the proposed methodology for the generation of non-intrusive indirect test for a generic mm-wave integrated circuit.	57
4.2	Top view and corrected model of the T-junction.	60
4.3	Expected S-parameters with the initial (dotted lines) and corrected (solid lines) models.	61
4.4	Expected S-parameters with the corrected model (dotted lines) and measured S-parameters (solid lines).	62
4.5	Expected output power with the corrected model (dotted lines) and measured output power (solid lines).	62
4.6	Expected PAE with the corrected model (dotted lines) and measured PAE (solid lines).	63
4.7	Transistor level schematic of the PA under test with typical S-parameters boxes for passive custom components.	64
4.8	RMS and maximum error for the prediction of the PA small signal gain as a function of the number of selected MC parameters.	65
4.9	Full-custom microstrip line and relevant BEOL features.	65
4.10	Microphotograph of the PA under test with the proposed non-intrusive process monitors and de-embedding structures.	71
4.11	General diagram of the measurement setup.	72
4.12	Photography of the test bench for large signal measurements.	72
4.13	Photography of the PA being measured.	73
4.14	Photography of the wafer with the circuits position (red dots) and mapping to the wafer reticle.	73
4.15	Gate resistance extraction - Simulation versus measurement.	74
4.16	Measurement results on the thin lines, wide lines and MOM capacitors - Signatures S_1 , S_2 and S_4	76
4.17	Measurement results on the MOS junction capacitance sensors and MOS transistors - Signatures S_5 , S_6 and S_7	77
4.18	Measurement results of the PA performance - I_{DC} , P_{sat}	78
4.19	Measurement results of the PA performance - S_{21} , S_{11} and S_{22}	79
4.20	Measurement results of the PA performance - ICP_{1dB} , OCP_{1dB} , PAE_{max}	80
4.21	Prediction of PA specifications from the set of low-cost signatures versus standard mm-wave functional measurements.	83
5.1	Transistor-level schematic of the power amplifier under study.	87
5.2	Layout of the power amplifier under study.	87
5.3	Scattering parameters of the 2-stage PA.	88

5.4	Output power versus input power - 2-stage PA.	88
5.5	Layout of the proposed variable decoupling cell.	90
5.6	Model of the proposed variable decoupling cell.	90
5.7	Impact of the interconnection model for fixed and variable decoupling cells.	91
5.8	Simplified example for illustrating the effect of the tuning knobs on a node impedance.	92
5.9	Generation of training data sets using a) full factorial DOE and b) random sampling.	94
5.10	Scatterplot of predicted versus actual PA gain for the samples in the validation set.	96
5.11	Figure of merit for the prediction of the PAE as a function of the size of the training data set.	97
5.12	OCEAN script used to simulate the circuit with the calibrated tuning knob values.	98
5.13	Layout of the power amplifier with variable decoupling cells and non-intrusive process monitors.	99
5.14	Predicted versus actual gain using the regression function used for the calibration.	100
5.15	Histograms of the main performances of the PA simulated both before and after calibration.	103
5.16	Histograms of the main performances of the PA simulated both before and after calibration (continued).	104
5.17	Yield before and after calibration.	104

List of Tables

3.1	Performances of the PA under test.	38
3.2	50 Ω transmission line main features.	41
3.3	Proposed library of non-intrusive sensors and their associated signatures for gain prediction.	47
3.4	Identified relevant MC parameters for the prediction of the PA specifications.	51
3.5	Prediction of PA performances using the generated indirect test program.	53
4.1	Identified relevant MC parameters for the prediction of the PA specifications.	66
4.2	Transmission line main performances for the seven selected corners.	68
4.3	Identified relevant BEOL features for the prediction of the PA specifications.	68
4.4	Proposed library of non-intrusive sensors and signatures for gain prediction.	70
4.5	Measured performances of the PA and prediction results.	82
5.1	Performance and design parameters of the PA.	89
5.2	Variable decoupling cell model features.	91
5.3	Comparison of PA performances figures of merit for different training sets.	96
5.4	Performances of the PA before and after calibration.	102

Chapter 1

Introduction to RF/mm-wave test

1.1 Introduction

The market of Systems on Chip (SoC) and Systems in Package (SiP) is continuously growing, pulled by sectors such as automotive, biomedical or 4.0 industry. This is possible especially thanks to advanced nanometric CMOS technologies, which allow the use of high frequencies and the integration of complex functions within tiny SoCs or SiPs.

Nevertheless, such technologies are prone to imperfections which can noticeably alter the devices performances. Process variations during the fabrication of the circuits are the most prominent source of faulty devices. Several kind of parametric process variations can be distinguished: lot-to-lot, wafer-to-wafer, die-to-die, within-die. The first three correspond to global variations whereas the last one corresponds to local random discrepancies, also known as mismatch. Catastrophic defects, also called spot-defects, are another possible source of faulty devices. They are caused by open/short circuits and most of the time, translate into faults.

Today a cleanroom can deliver a circuit for a fraction of the price it cost a few years ago. However, the test of the fabricated circuits still represents a bottleneck in the production line. As a matter of fact, the use of highly integrated and various types of circuits (digital, memory, analog, mixed-signal, RF, . . .) makes it difficult or even impossible to access internal nodes for test purposes. Moreover, the test of analog, mixed-signal, RF and mm-wave circuits usually relies on specification-based test. The time and equipment required to perform these tests still represent an important part of the test cost and, therefore, a major challenge.

In [1], the test of a mobile phone SoC is thoroughly studied. It is estimated that the digital and memory test only represents 13% of the overall test time, while these parts account for more than half of the silicon area. This well illustrates the efficiency of test approaches for digital circuits and the possible progression margin for their analog counterparts.

This Chapter is structured as follows. Section 1.2 gives an insight into production testing while Section 1.3 details the classical measurements that are performed to ascertain device performances.

Section 1.4 reviews the main approaches that are suitable to replace the conventional test procedures and Section 1.5 introduces calibration techniques that can be used to improve the performances of a suboptimal device. A special focus is given on one-shot strategies, that will be further developed in this thesis.

1.2 Production testing

1.2.1 Characterization and production testing

In order to sort out good and bad fabricated samples, a circuit can be tested in several ways. During the early stages of pre-production and production runs, a thorough characterization of the devices is typically performed to ensure their performances. However, several of the measurements that are performed may be redundant and therefore not mandatory to ensure the device behavior. Thus, during the production stage, the number of tests is minimized as far as possible. Nevertheless, even when the product is mature enough, an extensive characterization of each n-th die can be performed [2].

1.2.2 Required equipment

Rack-and-stack configurations are similar to laboratory test equipment and are suited to characterization. They can be specifically optimized for the target Device Under Test (DUT). For optimal test time during the production stage, Automated Test Equipment (ATE) are provided by large test-equipment manufacturers. They are especially suited to large volume production, can test several parts in parallel and are flexible enough to be reconfigured for another DUT in another test campaign. Different kinds of equipment are used depending on the DUT. For high frequency tests, they can consist in Vector Network Analyzer (VNA), spectrum analyzer, RF/mm-wave source, handler, load boards, contactor sockets, wafer prober, probe card, probes, etc. The ATE cost and complexity is highly dependent on the DUT, the tests to perform and the number of devices that are to be tested in parallel.

1.2.3 Wafer probing

In this thesis we only focus on wafer-level test. The interested reader is referred to [2] for testing of packaged circuit. For prototyping and characterization, bench top wafer probing equipment is used. The probes are typically contacted to the wafer using a manual or semi-automatic wafer prober, and measurements are performed using external equipment (VNA, ...). For production testing, a probe card is required in addition to bench top equipment. “A probe card, serving the purpose of the load board and contactor (in an analogy to packaged testing), is a complex printed circuit board that contains a customized arrangement of probe needles or probe tips to allow all of the necessary tester resources to contact all of the bond pads on one or more chips simultaneously” [2].

1.2.4 Calibration

For the measurements to truly reflect the DUT specifications, calibration procedures have to be set up and performed regularly. The losses introduced by the different elements that are used to connect the DUT to the measurement equipment (cable, connector, probes, etc) are especially critical for RF and mm-wave DUTs. A de-embedding procedure, typically a SOLT calibration, is used to remove their effect on the measured data [3]. It is based on 4 measurements: a Short, an Open, a $50\ \Omega$ Load and a Thru, hence the name SOLT. Power sources are also key elements that need to be calibrated, typically using a power-meter.

1.2.5 Accuracy, repeatability, correlation

The accuracy of a test procedure translates how far the test results are from the truth and seems to be the first feature that need to be accounted for. However, repeatability is of paramount importance, often far more than accuracy since the go/no-go decision must be consistent over time. It is also critical to ascertain the correlation, the consistency of the test results with measurements provided by bench top equipment, especially if the test is outsourced to a test house [2].

1.2.6 Test cost

The overall Cost Of Test (COT) is an extremely critical feature of testing. Several parameters are worth considering to compute it. A simple expression is given by the following Equation.

$$COT = \frac{Fixed\ costs \times Recurring\ costs}{Lifetime \times Yield \times Utilization \times Throughput} \quad (1.1)$$

Fixed costs refer mainly to the ATE capital, while recurring costs are implied by the use of the test equipment. The lifetime and utilization (percentage of the time) of the equipment is also taken into account to compute the COT. The yield refers to the percentage of circuits that pass the test and the throughput to the number of circuits that are tested per time unit. The yield is a topic of paramount importance. It is impacted by the process quality of course but also by the ATE accuracy which imposes to take guard-band in the go/no-go decision. Two kinds of error related to yield can imply a revenue loss: false negative and wrong positive. A false negative is a circuit that is thrown away because the ATE failed it, although it was good. These undesirable events are accounted for to compute the yield loss. A wrong positive is a bad circuit that is nevertheless passed by the ATE. This is typically the case with insufficient guard-banding and can be much more costly than a false negative [2]. The metric that accounts for this event is the test escape.

To alleviate the test cost, several possibilities are worth considering [1]: a) increase ATE throughput and b) reduce ATE capital and operational cost (human intervention, energy). The throughput is limited by the test time, the handling time and the number of parallel tests. On the one hand, parallelization and handling time are elements that should not be neglected, but they are greatly dependent on the ATE capabilities and thus must be considered by the ATE suppliers and users. On the other hand,

test time can be reduced by proposing alternate solutions to the classical functional tests. As a matter of fact, they can be particularly long as it is the case with Error Vector Magnitude (EVM) or Bit Error Rate (BER) to name but a few. As the technologies scale down, the available frequencies increase and so does the price of the associated test equipment. In addition to being less time-consuming, test procedures would greatly benefit from using low-cost ATE. This implies lowering the test frequencies and thus, devising new test procedures.

Next section focuses on the classical measurements that are used to ascertain DUT specifications.

1.3 Measures to be performed

Several measurements can be performed depending on the circuit and the target specifications. In this work we have decided to use Power Amplifiers (PA) as case study. Thus, the following subsections review the main kinds of measurement that are typically performed on a power amplifier. They can be classified into three main categories:

- a. Small signal measurements in order to characterize the linear behavior of the circuit.
- b. Large signal measurements in order to characterize the non-linear behavior of the amplifier under large signal stimuli.
- c. Power measurements for quantifying the energy-wise performances of the circuit.

Other types of measurements can be performed on PAs, such as modulated wave measurements to compute the Adjacent Chanel Power Ratio (ACPR) for instance. This will not be discussed in this thesis since we do not target any specific modulation.

We could have used other circuits with other specifications such as Low Noise Amplifier (LNA), Voltage Controlled Oscillator (VCO), mixer or Phased Locked-Loop (PLL) as case study. Since we focus on PAs, their specifications are not discussed in this section.

1.3.1 Small signal measurements

Microwave theory generally uses power quantities instead of voltages and currents [4]. This is due to two main reasons: a) traditionally microwave designs aim at transmitting power between stages and b) high frequency powers are much more easy to measure compared to voltages and currents. To quantify the power transfer, parameters called “scattering parameters” (S-parameters) have been introduced. Let us consider a two-port network with an incident input wave V_1^+ , a reflected input wave V_1^- , and the same for the output V_2^+ and V_2^- , as illustrated in Figure 1.1 [4]. S_{11} and S_{22} respectively represent the accuracy of the input and output matching. S_{21} represents the gain and S_{12} the reverse isolation. Their definitions are given by the expressions 1.2.

$$S_{11} = \left. \frac{V_1^-}{V_1^+} \right|_{V_2^+=0} \quad S_{12} = \left. \frac{V_1^-}{V_2^+} \right|_{V_1^+=0} \quad S_{22} = \left. \frac{V_2^-}{V_2^+} \right|_{V_1^+=0} \quad S_{21} = \left. \frac{V_2^-}{V_1^+} \right|_{V_2^+=0} \quad (1.2)$$

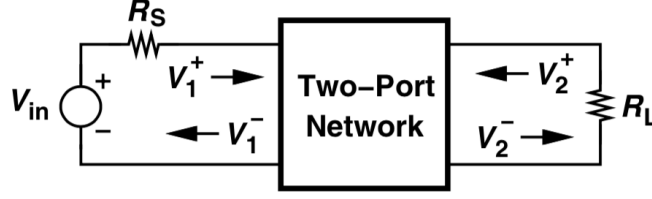


Figure 1.1: Illustration of incident and reflected waves for a two-port network [4].

1.3.2 Large signal measurements

For memoryless linear devices, such as passive components, the output voltage $y(t)$ can be expressed as a function of the input $x(t)$ using the following equation [4]:

$$y(t) = \alpha x(t) \quad (1.3)$$

For a memoryless non-linear device, such as a MOS transistor, the input/output characteristic can be approximated by:

$$y(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) \quad (1.4)$$

For a sinusoidal input, $x(t) = A \cos(\omega t)$, Equation 1.4 can be expressed as,

$$y(t) = \alpha_1 A \cos(\omega t) + \alpha_2 A^2 \cos^2(\omega t) + \alpha_3 A^3 \cos^3(\omega t) \quad (1.5)$$

$$y(t) = \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4} \right) \cos(\omega t) + \frac{\alpha_2 A^2}{2} \cos(2\omega t) + \frac{\alpha_3 A^3}{4} \cos(3\omega t) \quad (1.6)$$

Under large signal stimuli, the circuit behavior can thus be noticeably different from its small-signal operation. A device non-linearity has various impacts such as harmonic distortion, gain compression, cross modulation or intermodulation. For power amplifiers, gain compression and intermodulation are effects that are typically quantified as part of their specifications.

Gain compression

Gain compression phenomenon is highlighted by Equation 1.6, where the gain of the device is $\alpha_1 + (3\alpha_3 A^2)/4$ at the fundamental frequency. With $\alpha_1 \alpha_3 < 0$ (which is typically the case for common source amplifiers), the gain is reduced as A increases. This gain compression phenomenon is quantified by the 1 dB compression point, which corresponds to the signal amplitude at which the gain is 1 dB below its nominal value. The 1 dB compression point can be input or output referred, respectively shortened ICP_{1dB} and OCP_{1dB} . The following expression is used to calculate the ICP_{1dB} .

$$ICP_{1dB} = \sqrt{0.145 \left| \frac{\alpha_1}{\alpha_3} \right|} \quad (1.7)$$

Figure 1.2 shows the output signal A_{out} as a function of the input signal A_{in} in logarithmic scale [4]. This theoretical analysis has been carried out with voltages quantities, but the concept is the same with powers.

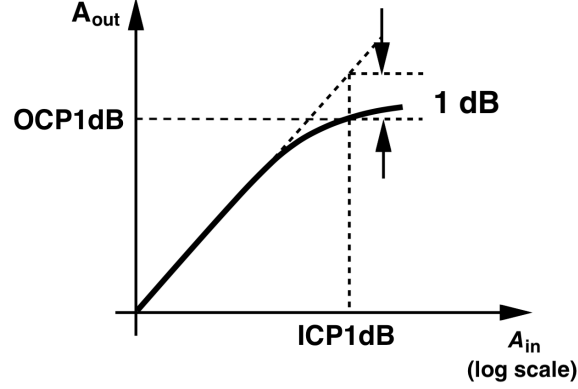


Figure 1.2: Definition of 1 dB compression point.

Third intercept point

Intermodulation (IM) is a phenomenon that must be taken into account for wireless communications. Let us consider a signal of interest at ω_0 and two interferers at ω_1 and ω_2 applied to a non-linear system. According to 1.4, third order intermodulation products (IM3) will appear at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$. If the angular frequency of one of this four IM3 products equals ω_0 , it will corrupt the signal of interest.

A typical way to quantify this phenomenon is the “two-tone” test. Two sinusoidal signals of same amplitude are applied to the device input. The amplitude of the IM3 products are thus

$$A_{IM3} = 20 \log \left(\frac{3}{4} \frac{\alpha_3}{\alpha_1} A^2 \right) \text{ dBc}, \quad (1.8)$$

where the unit dBc denotes that the IM3 product amplitude is normalized and relative to the one of the carrier (A).

Thus, as the input amplitude increases ($\propto A$), the relative IM increases even more ($\propto A^3$). The third intercept point (IP3) is the amplitude at which the IM3 products signal levels are as high as the fundamental. It is based on an extrapolation of the device small signal behavior, without any gain compression. It can be input (IIP3) or output (OIP3) referred and is calculated using the expression 1.9. However, due to gain compression, it is actually impossible to directly measure a device IP3. It is extrapolated using a “two-tone” test as illustrated in Figure 1.3 [4].

$$IIP3 = \sqrt{\frac{4}{3} \left| \frac{\alpha_1}{\alpha_3} \right|} \quad (1.9)$$

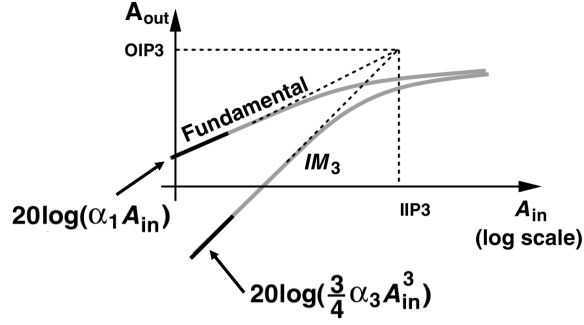


Figure 1.3: Definition of IP3 (for voltage quantities).

Using 1.7 and 1.9, we obtain a most useful formula (1.10). It is only true if the third-order approximation stands for the system since higher-order terms have been neglected.

$$IIP3 \approx ICP_{1dB} + 9.6 \text{ dB}. \quad (1.10)$$

1.3.3 Power efficiency

Power amplifiers are usually the circuits that consume the main part of a transmitter power. Two metrics can be used to quantify the power consumption. The efficiency is defined as

$$\eta = \frac{P_{load}}{P_{supply}}, \quad (1.11)$$

where P_{load} is the power delivered to the amplifier load (the antenna) and P_{supply} is the power drawn from the power supply. The power added efficiency (PAE) takes into account the input power P_{in} that is provided to the PA. It is defined as

$$PAE = \frac{P_{load} - P_{in}}{P_{supply}}. \quad (1.12)$$

1.4 Alternatives to conventional test

As already mentioned, the test equipment cost is a major limitation for RF/mm-wave circuits. Moreover, testing high frequency devices can require long test times and complex setups. Thus, several solutions have been proposed to tackle these issues. They can be classified in four main tendencies: structural test, built-off test, built-in test and machine learning indirect test. Their principle is explained in the next subsections.

1.4.1 Structural test

The structural test approach is similar to digital test procedure and relies on fault modeling. However, such models are difficult to build for analog and RF circuits [5]. This is due to the variety of components that can be used, their complex interactions and the faults that are modeled: catastrophic faults which correspond to short/open circuits and parametric faults which correspond to excessive process variations. The main difficulty for analog circuits is that process variations combinations can result in a specification violation, especially for demanding specifications. In this line, authors in [6] propose a suitable solution for RF circuits.

1.4.2 Built-Off Test

The versatility of ATEs may not be required in high volume production phase as it is the case during the development phase [7]. As a matter of fact, the needs are very different for these two steps. During the development phase, testers must handle wide ranges of powers and frequencies to allow changes in the way the test is performed. On the contrary, during the production phase, only specific combinations of frequencies and powers are needed since the test procedure is well set up. In this line, built-off test proposes to move some of the test equipment features on a custom board. The involvement of the ATE in the test procedure can vary depending on the board abilities, the DUT and the ATE itself. In certain cases a low cost ATE can be used as in [8].

1.4.3 Built-In Test

Built-In Test (BIT) or Built-In Self-Test (BIST) proposes to go even further and embed test features on-chip, with the DUT. This can drastically simplify the test procedure and required equipment. However, several drawbacks of analog BIST techniques still limit their adoption. In [7], the authors especially point out the extra die area required for the BIST circuitry and the extra package cost due the dedicated pins and wire-bondings for test purposes. The need to co-design the DUT with its BIST circuitry is also an important limiting point since this may be time-consuming, costly and may degrade the DUT performances. Moreover, the BIST must be accurate enough to provide relevant information about the DUT.

Thus, for a BIST technique to be really appealing, it must (a) provide a simpler and cheaper way to test the circuit compared to classical functional tests while it should (b) not imply a noticeable area overhead, (c) avoid the need to co-design and (d) be accurate enough.

A large number of BIST techniques have already been proposed. The main methods may use: DC sensing, envelope detector, loopback test, Oscillation Based Test (OBT), embedded VNA, thermal measurement or non-intrusive dummy circuits. These techniques are further discussed in Chapter 2.

1.4.4 Machine Learning Indirect Test

Before presenting machine learning indirect test, the following subsection briefly introduces basics of machine learning knowledge.

Introduction to machine learning

Machine learning is a subfield of computer science and more specifically of artificial intelligence. The term refers to algorithms written in such a way that they can learn how to solve a problem without being explicitly programmed for it. One of the first consequent works in this field was a program designed to learn how to play a game of checkers better than the person who coded it [9]. Nowadays the term “machine learning” is used to describe a wide variety of algorithms with various purposes such as email filtering, web-page ranking in search engines, pattern recognition (face, hand-written characters, etc) or financial market analysis to name but a few. Typically, there is a systematic two-step procedure to use a machine learning algorithm. In the first stage, the algorithm learns from the data it has been given. It is said that the algorithm is learning or is being trained, hence the name learning or training stage. Then, when the algorithm has been trained, it can be used for the purpose it was conceived for. Two main kinds of algorithms can be distinguished depending on the way the learning is performed, in a supervised or unsupervised manner.

Supervised algorithms They are devised to learn on both input and output data. Once the algorithm has been trained, it is able to predict the output based solely on input data.

Let us consider a simple example where we have a data base of houses with their price and area. The problem we wish to solve here is getting the price of a house knowing only its area, which is a typical regression problem. To this end, the data base is fed into the algorithm for it to learn the link that exists between the price and the area. The trained algorithm is basically a regression function/model. It can be used to predict the price of an other house, that is not in the initial data base, knowing only its area. In other words, the model is used to regress the output(s) from the input(s), that is the price from the area in this example.

The problem here has only two dimensions, so it is quite straightforward to represent it visually. In this line, Figure 1.4 displays the price versus the area of a few houses. Now if we train a machine learning algorithm on this data in order to build a regression model, we can plot Figure 1.5.a that shows the predicted value of the houses versus their actual price. Blue dots correspond to the data that are used to train the machine learning algorithm and red circles to the data set apart and used to ascertain its accuracy. This scatter plot is an excellent visual way to represent the model accuracy. It is all the more convenient when the problem dimensionality increases. Suppose for instance that we have other relevant information about the houses such as their age, the garden area and the presence of a swimming pool. In this case, we can leverage four variables to predict the price of the house. This five dimension regression problem would be impossible to represent in a similar way as in Figure 1.4. However, the scatter plot of the predicted versus the real price of the house would be a really good indicator of the prediction accuracy. With this extra *relevant* data we can expect a more accurate

prediction, visually represented by Figure 1.5.b. The relevance of the information is crucial. Imagine for instance that instead of the age, the garden area and the presence of a swimming pool, we provide a non relevant information, let us say the hair color of the previous owner. This data is clearly useless for the price prediction and may disturb the learning, leading to a degraded accuracy. Quantitatively speaking, we can compute the Root Mean Square (RMS) error of the model that corresponds to the average distance between the predicted and the real output target, the price of the house in this example. Mathematically speaking, the RMS error ϵ_{RMS} can be computed using Equation 1.13 where n is the size of the data (the number of houses in this example) and T_{real} and $T_{predicted}$ respectively refer to the real and predicted target values (the price of the house).

$$\epsilon_{RMS} = \sqrt{\frac{\sum_{i=1}^n (T_{real\ i} - T_{predicted\ i})^2}{n}} \quad (1.13)$$

The RMS error of the regression function can be computed on the data that are used for the training to check that the algorithm is learning correctly. However, the model accuracy in itself must be evaluated on independent data i.e. data that was not used to train the algorithm (corresponding to the red circles). Thus, in this work, when we mention the model RMS error, or simply the model error, we refer to the RMS error that is computed on independent data, set apart for verification purposes.

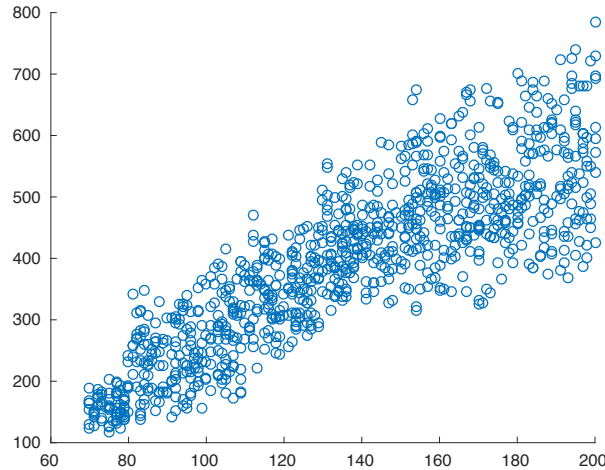


Figure 1.4: Machine learning example: price versus area.

Supervised machine learning algorithms can also be used to solve classification problems, such as spam email filtering. In this case, the algorithm learns on a data base composed of emails with associated boolean values representing whether they are spam or not. Once the algorithm has been trained, it is able to sort out incoming new emails and directly trash the spams. Overall, supervised machine learning algorithms are especially flexible and can be easily implemented to solve a large variety of complex non-linear problems.

In this thesis we will use a classical supervised machine learning algorithm for regression purposes.

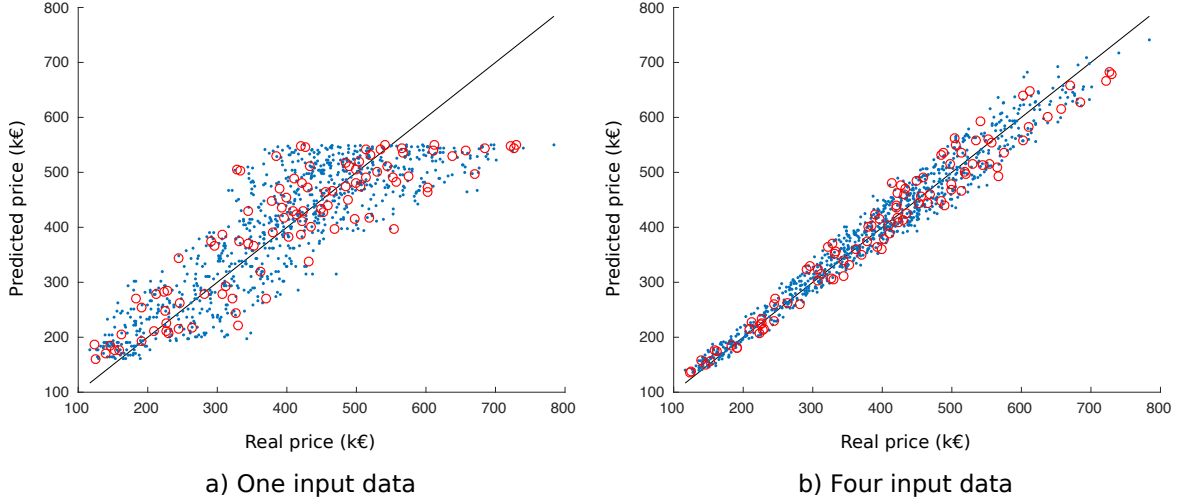


Figure 1.5: Machine learning example: predicted price versus real price.

Unsupervised algorithms Also referred as clustering algorithms, they are aimed at finding categories, patterns in large data bases, without any idea of the output result. This technique is especially leveraged in biology, medicine, genomics or in advertising for market segmentation. We will not use this kind of algorithm in this work.

Machine learning indirect test

Machine learning indirect test (MLIT), first introduced as the so-called alternate test strategy in [10], is a possible solution to alleviate the complexity and cost associated to testing mm-wave circuits. The aim of machine learning indirect test is to replace the measurement of complex functional specifications by a set of simpler measurements, also called signatures. This is a multivariate regression problem, so in mathematical terms, this means we need to build a regression function $f_{S \rightarrow P}$ between the signatures space $\{S_1, S_2, \dots, S_l\}$ and the performances space $\{P_1, P_2, \dots, P_N\}$ as

$$f_{S \rightarrow P} : [S_1, \dots, S_l] \rightarrow [P_1, \dots, P_N]. \quad (1.14)$$

This function can be easily obtained by training a supervised machine learning algorithm to map the signatures to the specifications.

The usual approach to machine learning indirect test is illustrated in Figure 1.6. The procedure is developed in two stages: a learning stage, and a testing stage. During the learning stage, both performances (blue path) and signatures (orange path) are measured from a set of training devices. Then, a machine learning algorithm is trained over the two sets of measurements to build a mapping model. This function, $f_{S \rightarrow P}$ as it is defined in equation 1.14, can be used to predict the performances based on the signatures. In the production/testing stage, signatures are measured for each DUT, and performances are simply inferred using the mapping model obtained in the previous stage. This indirect test

framework has the benefit of a simple interpretation of the test outcome since the estimated specifications are directly compared to their test acceptance windows in a conventional manner. Moreover, indirect measurements are devised to be cheaper than the direct measurement of functional specifications.

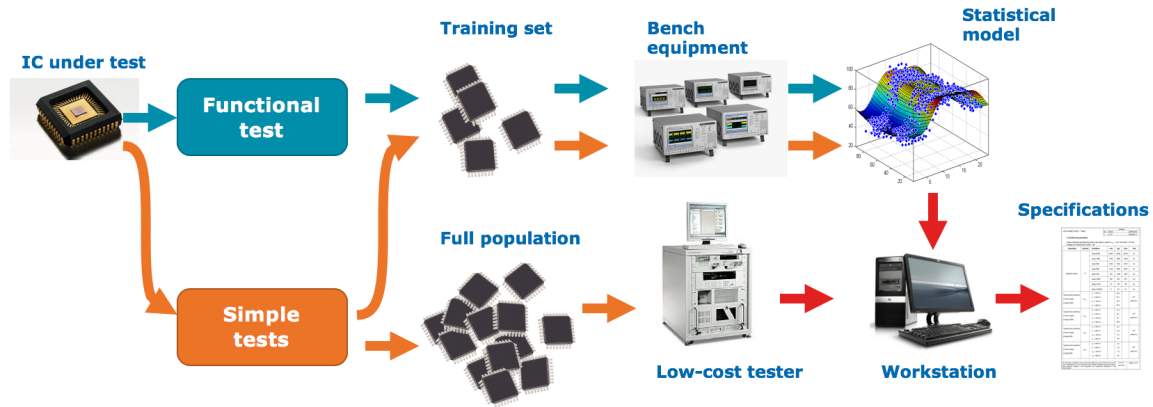


Figure 1.6: Machine learning indirect test principle.

For a practical use of machine learning indirect test, some extra steps should be added to the procedure. Figure 1.7 shows a typical flow for this test paradigm [11]. The first step aims at screening out outliers, circuits that have undergone a catastrophic spot defect. Then, regression models can be trained to learn the correlation between the signatures and the DUT specifications, as explained in the previous paragraph. A third phase would be to propose adaptive algorithms which would reduce even more the test time and cost.

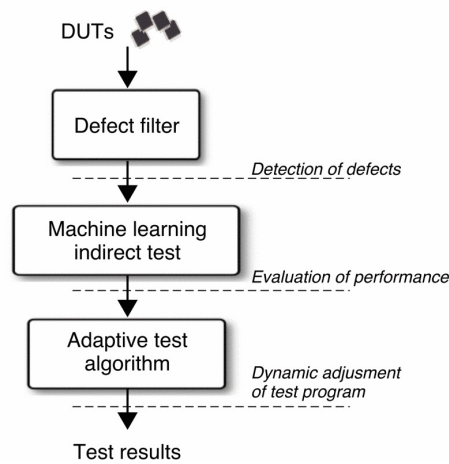


Figure 1.7: Generic test flow for machine learning indirect test [11].

1.4.5 Conclusion

To tackle the complexity and cost associated with testing RF/mm-wave circuits, a great variety of techniques exist. The choice of one above the others depends on the DUT and the conventional test methods that are to be replaced. Overall, there is not yet a generic method that may work for any kind of RF/mm-wave circuit. In our opinion, BIST and machine learning indirect test offer great possibilities that can be combined to automate the implementation of a low-cost test procedure on a variety of circuits.

A complete test method should be able to detect catastrophic faults, that may cause a functional failure, as well as process variations that can induce suboptimal performances. In addition to providing the go/no-go decision, the test outcome can further be used to make up for suboptimal performances. This refers to calibration methods, that are introduced in the next section.

1.5 Calibration against process variations

1.5.1 Introduction

Generally speaking, calibration methods aim at tuning a device after its fabrication to improve its performances. The goal may be to improve the fabrication yield, to obtain the best performance trade-off or to compensate for aging or environmental effects. The calibration relies on dedicated tunable elements, usually called tuning knobs, that allow the trimming of the device.

In [12], authors distinguish three categories of such techniques:

- Calibration that can compensate for static variations: catastrophic or parametric defaults.
- Healing that can compensate for static variations and quasi-static ones: aging, environment.
- Adaptation that can compensate for static, quasi-static and dynamic variation such as fast modifications of the operating conditions, the channel quality, etc.

Such procedures can be implemented off-chip or on-chip with a great variety of approaches which lead to many calibration/healing/adaptation methods. Whatever the selected method, the first step is to quantify the phenomenon that is to be compensated for (process variations, temperature, impedance variation, aging, . . .). Once this has been done, calibration/healing/adaptation can be performed.

Calibration methods can be divided into iterative test & tune and one-shot approaches [13]. Iterative approaches are more straightforward. The simpler implementation would be a tunable circuit without any test capabilities. The ATE would be used to test & tune the circuit, but obviously this would require prohibitive long test times. Providing BIST capabilities and/or alternate measurements to the device seems to be a more reasonable choice, thus allowing to use low-performances ATE. These two approaches are illustrated in Figures 1.8 a) and b) [14].

In this thesis, we focus on calibration for process variations and more precisely on one-shot methods which are appealing possibilities, easily compatible with machine learning indirect test.

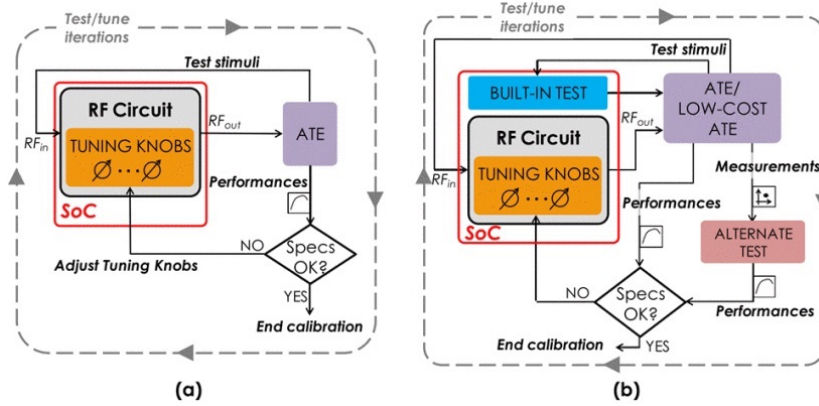


Figure 1.8: a) Iterative calibration using external ATE, b) Iterative calibration using low-cost ATE [14].

1.5.2 One-shot statistical calibration

This subsection provides a brief review of the operation of a generic one-shot statistical calibration strategy which operation principle is illustrated in Figure 1.9 [14].

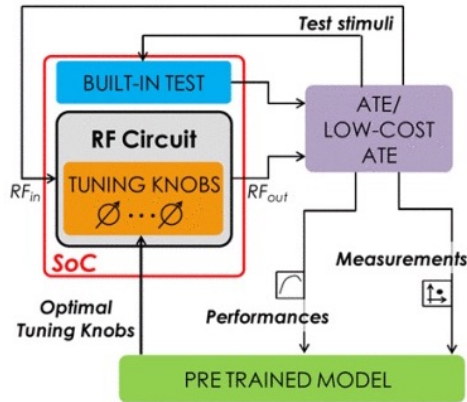


Figure 1.9: One-shot calibration principle [14].

Let us consider a circuit under calibration with a set of N specifications $\{P_1, P_2, \dots, P_N\}$ and k tuning knobs $\{T_1, T_2, \dots, T_k\}$. The main goal of statistical calibration techniques is to find a set of simple measurements (i.e. signatures) $\{S_1, S_2, \dots, S_l\}$ such that a regression function $f_{T,S \rightarrow P}$ can be built as

$$f_{T,S \rightarrow P} : [T_1, \dots, T_k, S_1, \dots, S_l] \rightarrow [P_1, P_2, \dots, P_N]. \quad (1.15)$$

Once such a regression function has been determined, and by measuring vector $\{S_1, \dots, S_l\}$ on a fabricated sample, function $f_{T,S \rightarrow P}$ can be used for exploring the space of $\{T_1, \dots, T_k\}$ and maximize the performances $\{P_1, \dots, P_N\}$. The regression function $f_{T,S \rightarrow P}$ is typically obtained using a supervised machine learning algorithm.

This procedure is divided into two stages: in a first stage –the training stage– for a population of circuits (the training set) both the performances and the signatures under different combinations of tuning knobs are extracted. A machine learning algorithm is then trained to regress the circuit performances from the signatures and tuning knob states. Once the training has been performed, we can take advantage of this regression model (function $f_{T,S \rightarrow P}$ referred to as “pre-trained model” in green in Figure 1.9) to actually perform the calibration to circuits in the production line. During the second stage –the calibration stage– only the set of signatures is measured (typically for the typical tuning knobs combination) and the machine learning model previously obtained is used to determine the best combination of tuning knobs. The process has the advantage of avoiding iterative calibration loops, since the regression function allows finding the optimum calibration point in one shot. Moreover, it is clear that this calibration approach is based on the machine learning indirect test paradigm.

1.6 Conclusion

In this chapter we have introduced RF/mm-wave test. The classical approaches of production testing have been reviewed and the main challenges pointed out. The cost of test equipment and the test time are two major hurdles when trying to reduce the overall test cost. Consequently, alternate solutions such as structural test, built-off test, built-in self-test or machine learning indirect test have been presented. We have also given an insight into calibration strategies to enhance DUT performances after fabrication and more especially into one-shot statistical calibration. Next chapter reviews BIST and calibration approaches that have been proposed in the literature.

Chapter 2

State of the art

Continuous demand for high data-rate transmissions and high-speed systems has led the industry towards mm-wave frequencies. This trend is supported by the continuous scaling of CMOS technologies that enables the fabrication of transistors with transition frequencies higher than 200 GHz. While such advanced fabrication processes allow the integration of mm-wave circuits, they are also especially prone to large process variations that may result in substantial performance degradation and low fabrication yield. Hence, integrated RF and mm-wave circuits require extensive production testing to evaluate the performance of the fabricated parts. However, testing such circuits in the production line is especially demanding since test procedures often involve complex at-speed measurements that require costly dedicated ATE and long test times. Possible solutions to tackle these issues have been proposed and are reviewed in this chapter. In this sense, we have organized the existing literature on the topic in two sections. The first one focuses on BIST while the second one reviews calibration techniques.

2.1 Built-In Self-Test

Built-in self-test strategies for analog and RF circuits are as diverse as the DUTs they target. This section summarizes the main contributions in this prolific research topic. They are presented in two groups whether they are intrusive or non-intrusive methods.

2.1.1 Intrusive methods

Many works have addressed intrusive BIST. In most cases the test circuitry uses one or several of the following elements: envelope detector, DC voltage probe, current sensor, loopback method or oscillation based test. These approaches are reviewed in the followings subsections.

Envelope detector, DC probes and current sensor

These kinds of sensors provide DC or low frequency signals that are easy to understand and process. Envelope detectors can be used to monitor the RF/mm-wave power in critical nodes. In [15], the gain of a 5 GHz LNA is extracted thanks to envelope detectors. The principle is quite simple: one detector senses the input RF signal while another monitors the output signal. The latter can then be variably attenuated and compared to the input envelope. A comparator is thus used to determine whether the gain of the attenuator is low enough to compensate for the LNA gain. The implementation of this approach has a very low overhead of only 7.6% of the overall active area and the gain extraction error is less than 1 dB in measurement. Moreover, the envelope detectors do not imply any design modification for the DUT due their high input impedance at the working frequency. In [16], the authors present a broadband (900 MHz to 2.4 GHz) amplitude detector that can be used for gain and ICP_{1dB} prediction. Measurements performed on a 1.6 GHz LNA show an accuracy better than 1 dBm for ICP_{1dB} extraction. The detector has a low area and it is estimated that 10 detectors would account for less than 5% of the area of a complete transceiver.

Several architectures of envelope detectors exist, depending on the technology, the target performances and the frequency band. Detectors for mm-wave frequencies are difficult to design within a limited surface and may negatively impact the DUT performances. Though, in [17] a small RMS Schottky diode detector is detailed. It allows measurements in the 80-110 GHz band for an area overhead of just $340 \mu m^2$. To avoid the co-design of the DUT alongside with the envelope detector, which may be necessary for mm-wave DUTs, low-coupling coefficient couplers can be leveraged as in [18]. The authors use -20 dB couplers at the input and output nodes of a 28 GHz PA to take off a very small amount of power and quantify it. Aside from removing the need to co-design the detectors with the PA, this approach provides immunity against load impedance mismatch.

Envelope detectors can also be used as part of a more thorough test framework including voltage and current sensors. DC probes or current sensors are used to check the DC operating point of the DUT. Since a good DC operating point is not a sufficient condition for the DUT to work properly, tests at its operating frequency have to be carried out. This is the role of envelope detectors and dynamic current sensors. A dynamic current sensor is used to extract a signature of the power supply dynamic current. These sensors have been used successfully in [5] to predict the conversion gain, compression point, third order intercept point and noise factor of an RF front-end (LNA and VCO). Performance prediction made use of machine learning algorithms to build mapping models between the sensors output and the target performances. In [19], a similar technique is applied to a 2.4 GHz QPSK transceiver to implement a low cost EVM BIST. 3 DC voltage probes, 2 envelope detectors and 1 dynamic current sensor are used to predict the EVM with a 5.8% error. This kind of approach have a very low area overhead and give very good results in simulations. However, they require to co-design the test circuitry with the DUT, which can be a major blocking point for mm-wave devices. Moreover, DC probes or current sensors are not straightforward to implement. DC probes basically consist in a large resistor connected to an internal node, but as the frequency increases it becomes far less ideal and can affect the circuit behavior. Current sensors can also affect the device performances and their design can be a complex task [5, 18].

Another paradigm exploiting an envelope detector is the signature based test. In [20], two high-

frequency close tones are fed into a 2.45 GHz LNA which output is connected to an envelope detector. The signal is processed to extract the main specifications: S-parameters, input and output impedances, noise factor, intermodulation, compression point. A machine learning algorithm is used to learn the correlation between the envelope detector output and the performances. The envelope detector adds a real improvement compared to a traditional signature based approach since the output signal is low frequency, and thus much more easy to process.

An important point for these kind of approaches is the RF/mm-wave signal source. In [15, 16] external sources are used, which is suitable for RF DUTs, but can be much more costly for mm-wave ones. Other works in the same field have led to different solutions. For instance in [21], the authors take advantage of a local oscillator to feed an integrated millimeter-wave radar receiver. In [22], a 60 GHz LNA is tested using the loopback technique which alleviates the need of RF/mm-wave sources as discussed next.

envelope detectors, alone or together with DC probes and DC current sensors, have been successfully leveraged in many works to predict various performances of devices. Overall, these sensors provide DC information that are easy to extract, the results accuracy is good and the area overhead is not too critical, especially for RF DUTs. Nevertheless, although co-design is not mandatory for all RF circuits, it is unlikely to be the same for mm-wave ICs. The use of low-coupling coefficient couplers with envelope detectors can partially alleviate this drawback, but requires extra area.

Loopback

A simple idea for testing transceivers is to connect the transmitter (TX) to the receiver (RX). In this configuration, the need of an RF source is removed and a whole set of functional tests can be easily performed. Recent works still focus on this topic, such as [23], [24] where the concept is used for mm-wave transceivers BIST or to test and calibrate a 2.4 GHz transceiver [25]. Figure 2.1 shows the diagram of a basic direct loopback implementation [26].

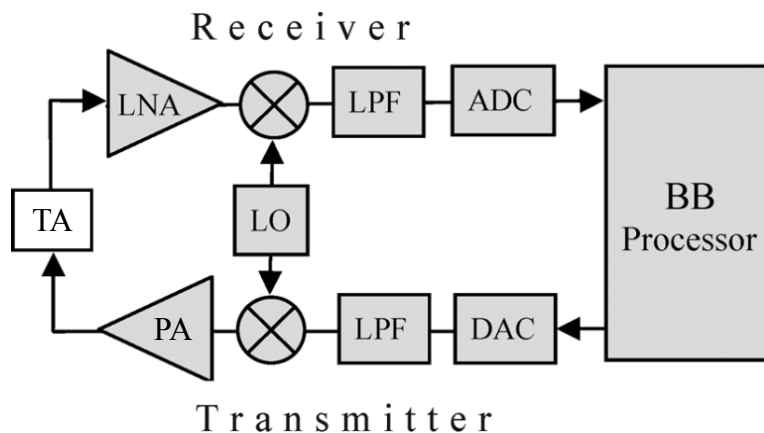


Figure 2.1: Direct loopback setup for integrated transceiver test.

This is a very appealing method and yet it implies several problems. Despite its apparent simplicity,

it may incur a noticeable area overhead, require co-design and may not be able to fully differentiate transmitter and receiver impact in the overall performance. Besides, the absence of radio channel implies to find a way to reduce the signal power at the input of the RX chain to be in the dynamic range of the LNA. A test attenuator (TA in Figure 2.1) can be inserted but its design may be challenging. Another option is to remove the PA in order to reduce the power at the LNA input, but of course the PA is not tested. An extra mixer can also be inserted between TX and RX for certain types of transceivers [26].

The loopback configuration can be setup along side with envelope detectors at critical nodes for a more complete test procedure as in [22]. Performances such as gain, gain compression, intermodulation, I/Q mismatch can be easily extracted.

Oscillation Based Test

Oscillation Based Test (OBT), first introduced in [27] for analog and mixed-signal circuits, proposes to use the DUT to create an oscillating system. Defects will cause drift in the oscillation frequency (and amplitude) making it a convenient way to test a circuit. Figure 2.2 presents a block diagram which illustrates this method [28].

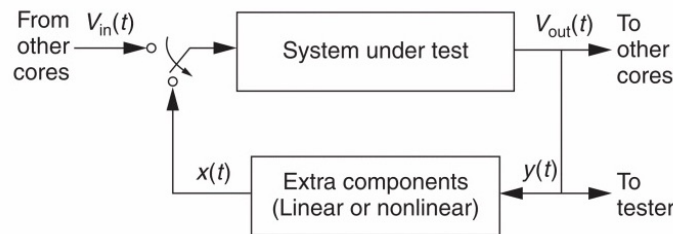


Figure 2.2: Block diagram showing the required modifications to convert a DUT into an oscillator [28].

In [29], a practical implementation of this technique is presented for a 2.4 GHz LNA. The authors take advantage of a frequency-selective on-chip power detector to sense the operating frequency as well as the output power. They achieve a 60% fault coverage in simulation. However, the BIST circuitry implies performance degradations that cannot be neglected of around 0.5 dB for the noise factor and 2 dB for the gain. OBT can also be applied to mm-waves circuits such as in [30] where the concept has been implemented as a BIST solution for a 60 GHz phase shifter. In this work, the phase shifter is used as the resonant element of a 14-17 GHz VCO, whose output frequency is an image of the phase shift.

OBT is an interesting solution for testing RF and millimeter-wave ICs since the test output is sensitive to the performance degradation mechanisms and is simple to exploit. Nevertheless, the extra-circuitry and design modification that may be required represent major challenges that may limit the adoption of OBT for high frequency circuits.

Embedded VNA and spectrum analyzer

Some works propose to design on-chip miniature VNA or spectrum analyzer. The area overhead can be important but they offer a great flexibility and variety of test possibilities. A key advantage for mm-wave circuits is that it alleviates the need of de-embedding procedures. In [31], authors present a 33 MHz - 3 GHz spectrum analyzer which can be used for on-chip PLL characterization. It presents a dynamic range of 60 dB and a resolution of 30 kHz with an area of around 0.27 mm². [32] introduces a dual-port 50-100 GHz VNA that can be used to perform measurements on a wide variety of mm-wave circuits or to provide BIST capabilities. The presented VNA achieves a dynamic range of 55 - 72.5 dB at an IF RBW of 100 kHz and a 3σ phase error of 0.3°. The work in [33] proposes a dedicated BIST circuitry for evaluating the S-parameters of phased-arrays with an accuracy better than 1 dB in amplitude and 7° in phase. Although promising silicon results have been reported, the complexity and area overhead of the on-chip instrumentation is high and careful co-design with the DUT is necessary. Overall, this kind of approaches have a great potential but suffer from their complexity and area. They are not suited for all kind of integrated devices though they can be relevant for specific cases.

2.1.2 Non-intrusive methods

Though really relevant, most intrusive BIST techniques suffers from a major drawback: they require to co-design the test circuitry with the DUT and are architecture-dependent. An exception would be embedded VNAs, but the extra area they require may be a stumbling block for most DUTs. Several works thus target non-intrusive test approaches.

Thermal measurement

Non-intrusive temperature sensors can be used to detect defects within a DUT. A major advantage of temperature sensors is that they are able to detect catastrophic defects in the components they monitor in a non-intrusive way. When the DUT is powered on, part of the electric power is dissipated thermally by Joule effect and conducted through the substrate. Sensors can thus monitor the thermal evolution of several locations of the DUT. In [34], thermal sensing is applied to a 2.4 GHz LNA to monitor its DC operating point and RF behavior. In [35], temperature sensors are used to evaluate the thermal noise level of a 60 GHz LNA. In [18], they are used to measure the DC power dissipation of a 28 GHz PA. This approach is especially interesting but it needs to be used together with another test technique to provide a thorough test outcome.

Dummy circuits

This approach is sensitive to parametric process variations (lot-to-lot, wafer-to-wafer, die-to-die). The idea is to monitor these variations by means of dummy analog elements, similar to the ones used in the DUT. By placing them next to their kind, they would undergo the same process quality. The choice of the duplicate analog elements is the crucial point in this approach. A monitor can be a

complete analog stage or simple components called Process Control Monitors (PCM). Then, DC and low frequency measurements on these sensors, that are not electrically connected to the DUT, are used as test signatures for training machine learning regression models. This method is thus part of the so-called “non-intrusive indirect test” and relies on process-aware sensors. The concept was first proposed in [36] for mixed-signal test applications, and later extended to the RF and mm-wave domain in [37, 38]. Figure 2.3 illustrates this test paradigm [37].

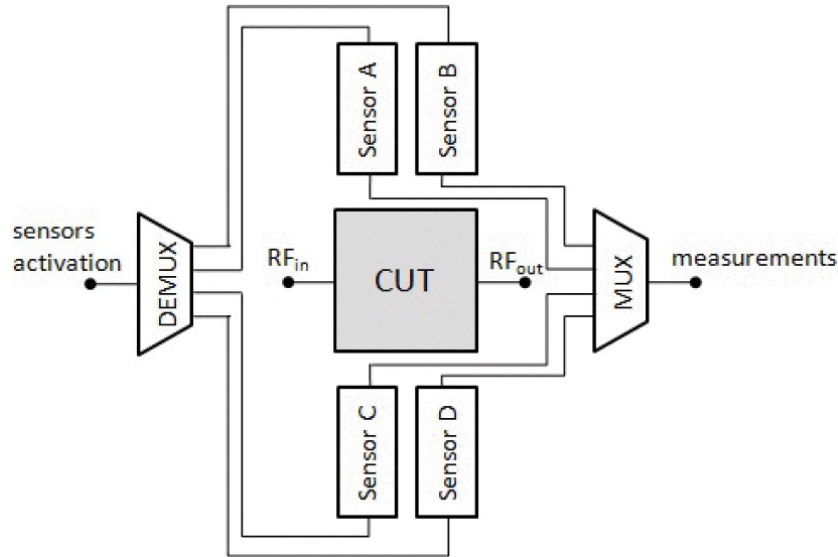


Figure 2.3: Block diagram of a generic non-intrusive indirect test based on process-aware sensors [37].

In [37], several analog stages and PCMs are used successfully to predict the gain, NF, 1 dB compression point and IIP3 of 2.4 GHz LNA. The average prediction error is excellent: less than 0.1 dB for gain and NF, less than 0.2 dB for the compression point and IIP3. In [38], the approach is applied to a 60 GHz LNA. Considering the different sensitive elements in the DUT, several dummies are chosen: common source stage, cascode stage, N+ polysilicium resistor, MOM capacitor, transistor and micro-strip transmission line. They allow the prediction of S-parameters, noise factor and third order intercept point. In simulation the results are good for most specifications with a RMS error of 0.5 dB for S_{21} and S_{11} , 0.1 dB for NF and 0.3 dB for IIP3. Results for S_{22} are not as good as for the other specifications with a RMS error of 1.7 dB and a maximum error of 4.1 dB.

Another approach, based on the same paradigm, is presented in [39] where a 80 GHz LNA is tested by the intermediate of a ring oscillator. Since the oscillator undergoes the same process variations as the transistors it is designed next to, any drift in the size of the transistors results in a drift in the oscillating frequency. As the frequency varies almost linearly with the width of the MOS transistors, it is an excellent process variation monitor which output is easy to measure. However, in a fabrication process many of the transistor features can vary and not only its width, so this study is a little too simplistic.

Non-intrusive test approaches are especially relevant for high frequency circuits since they do not load any internal node. However, it is highly dependent on design knowledge: sensors are proposed

based on an expert analysis on the circuit topology and no systematic approach for defining the set of signatures is proposed. Moreover, it implies a few challenges due to the use of machine learning tools. Next section thus focuses on machine learning indirect test and tries to review the main contributions in this domain.

Non-intrusive machine learning indirect test limitations

Although especially appealing, non-intrusive indirect test still have to face many challenges. Some of the elements of the following state-of-the-art of the machine learning indirect test framework are based on the review work in [11]. In our opinion its main limitation is that there is no systematic approach for generating an appropriate set of input signatures: the process has to be hand-repeated for every different DUT. From the point of view of Electronic Design Automation (EDA) tools, this is similar to developing an Automatic Test Pattern Generator (ATPG) for a given test campaign. Such a systematic approach would provide the set of input measurements (signatures) that maximizes the quality of the machine learning regression model. Previous efforts in this line include the optimization strategy in [40, 41], where a piece-wise linear stimulus is optimized to minimize the prediction error of the regression model. While this is a sound strategy, the choice and the parametrization of the input stimulus are still *ad hoc* and the computational effort is huge since Monte Carlo simulations have to be performed for each iteration of the optimization loop. Finding the appropriate set of input signatures to build a reliable regression model is a classical problem in the field of statistics and a variety of feature selection strategies have been proposed to deal with this issue. In this line, different feature selection techniques have been adapted to machine learning indirect test approaches to select the best subset of features within an initial set of potential features [42–46].

A direct approach to feature selection consists in preselecting a subset of features based on some statistical observations, before training any regression models. This approach, widely used when the number of initial features is high, is known as filtering. Some filtering techniques recently proposed include Principal Component Analysis (PCA) [42], multivariate distance correlation [43], Kendall's tau correlation metric [44], etc. However, these filtering techniques only capture the most significant variation components. In order to further improve the selection algorithm, it is necessary to resort to wrapper approaches [47]. Basically, a wrapper algorithm uses a machine learning prediction model as a black box within an optimization loop, with the objective of finding the smallest subset of signatures that minimizes the prediction error in an independent validation set [45, 48]. Finally, hybrid filter-wrapper methodologies have been proposed to reduce the computational burden of the wrapper approach. Hybrid techniques guide the search in the input signature space based on the dynamic evaluation of a correlation metric [46].

Another prominent issue to implement a machine learning indirect test strategy lies in the data required to build the regression models. As a matter of fact, lengthy Monte Carlo simulations are typically used to generate this data. For the regression functions to be accurate enough, generated data must cover the complete set of possible variations since models cannot be extrapolated outside the boundaries of the training samples. To cover a 6σ variation range, a huge number of simulations would thus be mandatory. Some works have focused on alleviating this computational burden. We can cite for instance the statistical blockade method [49–52]. This technique alters the selection likelihood

of Monte Carlo samples to generate extreme instances of the circuit. This ensures that faulty devices will be generated and therefore, that regression models will include both passing and failing circuits.

Besides the computational cost, another limitation is that this simulation-based approach assumes that Monte Carlo models are correct and accurate, which can be reasonably questioned. In addition to these models inaccuracies, some elements may not have any Monte Carlo models. Any error in the modeling of the circuit process variations can negatively and noticeably impact the test outcome as detailed in [53]. However, the authors highlight that even if the test validation is not so good, the test might still be relevant since the test accuracy cannot be fully ascertained by simulations. Bayesian model fusion [54] has been proposed as a possible solution to this issue. Regression models are still trained on simulated data but they are then adapted using real measurements. It assumes that modeling errors only *slightly* alter the regression models but not their general structure. In the case that Monte Carlo models are not available or too inaccurate, the training must be performed on a representative set of fabricated samples, from different wafers and lots over an important period of time.

Thus, the computational effort and the lack of a truly generic method for the signature proposal seems to be the main roadblocks in the adoption of machine learning indirect test.

2.2 Calibration

In the last few years, embedded calibration strategies have been proposed as promising solutions to alleviate the performances drift induced by process variation, especially for integrated circuits operating in the RF and mm-wave frequency bands. As introduced in section 1.5, the basic idea consists in including some measuring capabilities and/or dedicated tuning knobs that allow a) the assessment of the circuit performance after fabrication and b) the correction of an excessive deviation from the target design goals. Calibration methods can be divided into iterative test & tune and one-shot approaches [13] and the main contributions are reviewed in the next subsections.

2.2.1 Iterative methods

In [55], authors propose a tapped gate inductor to calibrate the input matching of a 1.9 GHz LNA based on the measurement of the LNA bias current. Another calibration method based on alternate measurements is applied in [56] to a 25 GHz differential Colpitts VCO. One bias voltage is tuned thanks to a brute force search.

The work in [25] introduces a 2.4 GHz transceiver calibration targeting yield recovery. Authors take advantage of an envelope detector together with a loopback framework to tune the system using bias and supply voltages. Calibration is based on a gradient search algorithm guided by a machine learning model that maps the output of an envelope detector to the circuit performances. This approach is a real improvement compared to a lengthy brute force search. A similar method is presented in [57] where two kinds of tunable elements are leveraged: software tuning-knobs (distortion coefficients for baseband signal generation) and hardware ones (bias and supply voltages).

The work in [18] proposes a tunable transmission line stub which can be shorted at different discrete lengths in order to change the loads of a 28 GHz PA. The proposed calibration strategy iteratively searches the best possible performance by varying the tuning knobs state. Current, power and temperature sensors are used for monitoring the performance of the PA during the calibration process.

2.2.2 One-shot strategies

Authors in [58] present a calibration strategy for a 1.9 GHz LNA, based on a machine learning model for the prediction of the calibrated tuning knob values. In this case, an embedded envelope detector is used for predicting the performance of the LNA from the response to an optimized external stimulus. Authors propose two tuning knobs for the LNA: a bias voltage and an architecture-specific PMOS IMD sinker. A limitation of this approach is that it assumes tuning knobs are orthogonal. Although desirable, this feature can easily be untrue for other knobs than the ones they have used.

The work in [59–61] proposes a one-shot calibration strategy for a 1.57 GHz RF LNA. The performance of the fabricated LNA is estimated by using an embedded VCO and an output peak detector for exciting the LNA and acquiring its response, respectively. The bias voltages and power supply voltage are used as tuning knobs. The calibration strategy makes use of a machine learning model to predict the best values of the tuning knobs based on the output of the peak detector. Compared to [58] this approach does not assume anymore the tuning-knobs orthogonality but make nonetheless an assumption: tuning-knobs and process variations have an orthogonal impact of the device performances. Thanks to this hypothesis, during the calibration phase of the circuit, signatures are extracted for the typical knob setting only, hence the name “mid-point alternate test-based calibration” of this method. This assumption is much more likely to be true than the previous one and should have little impact in the final results.

On the other hand, the work in [14, 62] proposes a one-shot calibration strategy for a 2.4 GHz PA relying on non-intrusive process variation sensors that do not tap into the nodes of the PA. This approach does not imply any orthogonality assumption. Bias voltages and power supplies are used as tuning knobs for calibration. This is actually a common disadvantage of the previously presented calibration techniques: most of the calibration capability of these techniques comes from varying the power supply of the device under calibration. From a system level point of view, providing such a control over the power supply may require a complete re-design of the power management circuitry that may degrade the performance and increase the complexity of the system.

A comparative study of these three one-shot calibration techniques is presented in [13]. Authors conclude that the method in [58] is faster when dealing with important numbers of tuning-knobs, although calibration results are not as good as with the two others techniques. The method based on non-intrusive sensors [14] is slightly less good than the one in [59] but the difference is not relevant due to the limited number of fabricated samples (eighty). Moreover, the method in [14] is completely non-intrusive which is a key advantage. They also point out that Bayesian model fusion can be used to leverage simulation data as it is the case with non-intrusive test (see 2.1.2).

Others works have been carried out on this topic, such as the one in [63] where the authors propose the one-shot calibration of an RF amplitude detector to increase its dynamic range. They leverage the

sensing of its DC offset used to monitor the process variations to digitally correct its output. A simple lookup table built on MC simulation data is used to propose a suited digital correction.

2.2.3 Towards self-healing

Fully implementing a calibration strategy on-chip paves the way to self-healing architectures. The calibration can thus be performed during the lifetime of the circuit, during the idle times for instance, to make up for aging or environmental modifications. This self-calibration/self-healing principle is illustrated in Figure 2.4 [14].

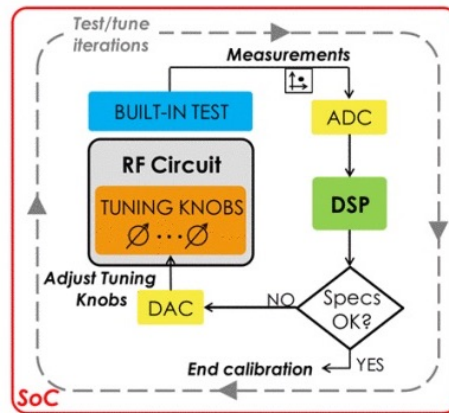


Figure 2.4: Self-calibration/self-healing using on-chip resources [14].

For instance in [58], a Digital Signal Processor (DSP) was leveraged to tune the device. Only little modifications would be necessary to avoid the need of an external ATE and thus implement a self-healing system. In this line, the work in [18] introduces a complete autonomous self-healing PA with astonishing capabilities. Figure 2.4 presents a generic implementation of such techniques. Note that self-healing systems typically use test & tune iterations as in [18] but it is not mandatory since pre-trained regression models can be implemented within the DSP.

Many works have been carried out in the self-healing and self-adaptivity fields. Since this is not the scope of this thesis we will not deal with the subject here. The interested reader is referred to [12] for an overview of the topic and bibliographic references.

2.3 Conclusions

Several BIST implementations targeted at different RF and mm-wave circuit families have been presented in the last few years. Thus, a classic RF/mm-wave test method that has been implemented on chip is the loop-back test, that gives a convenient way to test a complete transceiver by connecting the transmitter output to the receiver input. Despite its apparent simplicity, it may imply a noticeable area overhead, a careful co-design of DUT and test circuitry and may not be able to fully differentiate

transmitter and receiver impact in the overall performance. Moreover it obviously requires a complete transceiver so standalone circuits cannot be tested this way.

Other test approaches have been presented such as integrated VNAs for on-chip testing or Oscillation Based Test. However interesting, the complexity of the on-chip implementation and the area overhead are elements that would greatly limit the adoption of these methods.

On the other hand, simple on-chip DC-sensors, such as power detectors, envelope detectors, temperature sensors, etc. have been also proposed for simplifying the read out of high-frequency signals. These sensors are often used in combination to machine learning algorithms with the goal of simplifying (or avoiding the need of) high-frequency measurements. These works replace the measurement of complex RF specifications by simpler low-frequency signatures obtained from the sensors. Then, a linear model or a regression function obtained using machine learning algorithms is used to map the signatures to the DUT performance. However, since these sensors usually tap into critical circuit nodes of the DUT, they usually require a careful co-design to avoid the degradation of the DUT performances. Nevertheless, the impact of envelope detectors for mm-wave circuits can be mitigated using low-coupling coefficient couplers.

The recently proposed non-intrusive indirect test paradigm may be a solution to the aforementioned issues. This test strategy proposes to integrate a set of non-intrusive process variation sensors in the close proximity of the DUT but not electrically connected to it, in such a way that both the DUT and the monitors will undergo the same process variations. Then, a machine learning regression model is used to infer the performance of the DUT from the outputs of the process monitors. Thus, this non-intrusive test is aimed at detecting excessive process variations that may degrade the functionality of the DUT. Moreover, it has the additional advantage of not loading the DUT nodes, so no complex co-design of DUT and sensors is needed. Furthermore, the output of the process monitor circuits is usually DC or low-frequency which eliminates the need of costly high-frequency testers.

However, recent attempts to extend the non-intrusive test strategy to the mm-wave domain show serious practical limitations and designing appropriate non-intrusive sensors to estimate accurately all the specifications is still a key open problem. Thus, the lack of systematic approaches to propose sensors and signatures may greatly limit the adoption of the non-intrusive test strategy. Methods to tackle this issue are presented in chapters 3 and 4.

While testing a device is obviously a necessity, calibration methods to mitigate the effect of excessive process variation are worth considering. In this sense, one-shot calibration methods that leverage machine learning regression tools are an obvious choice to complement a non-intrusive machine learning indirect test strategy. If calibration procedures seem well defined, there is still room for innovation on the tunable elements that are leveraged to perform the calibration. As a matter of fact, the power supply is often used as a tuning knob but this can greatly affect the power management unit design. Architecture-specific tunable elements are also worth mentioning but overall there is a lack of generic tuning knobs that can be easily reused in other designs. Chapter 5 thus focuses on on-shot calibration and proposes a novel tunable element for mm-wave ICs.

Chapter 3

Assisted test design for non-intrusive machine learning indirect test

3.1 Introduction

This chapter introduces a generic method to implement a non-intrusive machine learning indirect test based on process variation sensors. As mentioned in the previous chapter, several works have already proven the feasibility of this test approach for RF and mm-wave circuits [37–39]. However, non-intrusive sensors are proposed based on an expert knowledge of the circuit design and no systematic approach for defining them have been presented. The lack of automation of this test paradigm is an important limitation that we try to tackle in this chapter. Thus, we present a method aimed at being as automated and systematic as possible to ease the implementation of a non-intrusive machine learning indirect test on a given circuit. In this chapter, we focus on mm-wave circuits designed exclusively with components that have Monte Carlo models. Chapter 4 details an extension of this method to target circuits that use custom passive structures.

3.2 Proposed approach

Let us consider a generic non-intrusive machine learning indirect test scenario in which we intend to regress a set of performances $\{P_1, P_2, \dots, P_m\}$ of a given DUT from a set of signatures $\{S_1, S_2, \dots, S_n\}$ measured from a set of non-intrusive process variation sensors. In other words, we want to build a function $f_{S \rightarrow P}$ defined as follows

$$f_{S \rightarrow P} : [S_1, \dots, S_n] \rightarrow [P_1, P_2, \dots, P_m]. \quad (3.1)$$

The problem we address in this work is, given the netlist of a particular DUT in a given technology and a set of target performances, how to systematize the proposal of the set of signatures and the design of the associated sensors.

The methodology that we propose is divided into two interrelated steps. The first step is aimed at unveiling the root causes of parametric performance variation in the simulation environment. The second step is devised to find process variation sensors that are sensitive to the identified set of root causes. Let us describe these steps separately.

3.2.1 Step 1: Unveiling the root causes of parametric performance variation

The key idea for finding the root causes of performance variation at the design stage is to take advantage of the available information in the MC models included in the Process Design Kit (PDK). Indeed, in a MC simulation environment, we have direct access to the complete set of root causes behind the DUT performance variation, that is, the set of statistical model parameters generated for each MC iteration. There are two sets of statistical parameters, describing global parametric process variations and local random mismatch. The non-intrusive machine learning indirect test approach works under the assumption that the effect of mismatch can be neglected with respect to process variations. We have checked in simulation that this assumption was correct for the selected technology. Thus, our methodology explores the space of MC global process variation parameters $\{MC_1, MC_2, \dots, MC_l\}$ to find the minimum subset that explains the observed parametric performance variation within a given accuracy threshold.

Exploring the space of process variation parameters can be seen as a feature selection problem in which we consider the l MC parameters as candidate features to regress a given target performance using a machine learning regression model. Thus, we propose to adapt the Brownian distance correlation-directed search, previously presented in [46], according to Algorithm 1. The goal of this algorithm is to find a small subset of process parameters $\{MC_1, MC_2, \dots, MC_l\}$ that explains most of the performance degradation. It has to be applied for each target performance of the DUT.

This is a hybrid filter-wrapper algorithm that leverages the advantages of both wrappers and filters. A typical wrapper algorithm would iteratively train different models on different subsets of MC parameters in order to keep the best one. This technique gives accurate results but is extremely lengthy, since machine learning algorithms are trained for each iteration of the wrapper algorithm. On the other hand, filters are fast but inaccurate. They do not train any regression model, hence their speed, but can only detect first order correlation. To summarize, a wrapper algorithm would find the best subset of MC process parameters, whereas a filter would only find one of most important ones, but very quickly. The hybrid filter-wrapper algorithm that is presented here leverages the fast computing of a filter—the Brownian distance correlation in this case—to propose only relevant features with regard to the missing information. By filtering all the possible process parameters to propose a relevant one, the procedure is extremely fast. A key point of the algorithm is the concept of correlation, that relies on the evaluation of the Brownian Distance correlation [43, 64]. This non-linear correlation metric is sensitive to multivariate non-linear dependencies. The interested reader is referred to [46] for a complete explanation of the algorithm.

Given a target performance, the proposed algorithm first ranks the MC process parameters with respect to their distance correlation to the target performance and trains an initial regression model with the most correlated parameter. Then, we compute the residues of the fit (i.e. the individual errors

Algorithm 1 Exploring the space of MC process parameters using a Brownian Distance Correlation-Directed Search.

- 1: Filtering: Compute the Brownian distance correlation of all the MC process parameters with the target performance
 - 2: Add the associated most correlated process parameter to the input space
 - 3: **while** *continue* **do**
 - 4: Train a machine learning algorithm to build a model mapping the input space onto the performance
 - 5: Compute the residues of the model (the difference between the real and the predicted values)
 - 6: Compute the mean square of the residues (i.e. the RMS error of the model)
 - 7: **if** The RMS error is less than or equal to a given threshold **then**
 - 8: *continue* $\leftarrow 0$
 - 9: **else**
 - 10: *continue* $\leftarrow 1$
 - 11: Filtering: Compute the distance correlation of each of the remaining MC process parameters with the residue
 - 12: Add the most correlated MC process parameter to the input space
 - 13: **end if**
 - 14: **end while**
 - 15: **return** The input space (the most significant MC process parameters) and the mean square error of the fit (the RMS error of the model)
-

for all the training samples). After that, the algorithm ranks the remaining MC process parameters with respect to their distance correlation to the residues of the fit and selects the best candidate to be added to the set of the most significant MC parameters. By iterating this procedure, the search is guided so that each iteration of the algorithm should identify a MC parameter that adds *relevant* information that was previously missing. The output of the algorithm is the set of the p MC process parameters that best explains the variation of the selected target performance. In the end we also have built a regression function $f_{MC \rightarrow P}$ from the subset of p relevant MC parameters to the target performance P defined as follows

$$f_{MC \rightarrow P} : [MC_1, \dots, MC_p] \rightarrow P. \quad (3.2)$$

The process can then be iterated to cover the rest of the target performances of the considered DUT and thus build a more complex and complete regression model defined as

$$f_{MC \rightarrow P} : [MC_1, \dots, MC_{p'}] \rightarrow [P_1, \dots, P_m]. \quad (3.3)$$

It is important to notice that the identified MC parameters may be different for each target specification. Consequently, a set of relevant p' MC parameters may be used to regress the complete set of performances using Equation 3.3.

3.2.2 Step 2: Assisted design of non-intrusive process variation sensors

The following step is to devise simple process variation sensors that provide a signature sensitive to the set of the p' most significant MC process parameters selected by the previous algorithm. Again, we take advantage of the information contained in the PDK of the technology to identify the underlying physical meaning of the selected process parameters. This way we can guide the design of appropriate non-intrusive sensors to target the desired information. Ideally, a dedicated sensor may be designed for each one of the selected process parameters. However, signatures, in practice, may contain contributions from several process parameters. The devised sensors, and their associated signatures, are then stored in a library for further reuse. Indeed, as the methodology is applied to different DUTs in the same technology, sensors may be reused if the same MC process parameters are deemed significant for new DUTs. In fact, once this library covers the complete set of MC parameters, the design of the test protocol becomes a simple selection problem within the tests in the library. The methodology only calls for human intervention if there is no sensor in the library associated to a given process parameter. This simple strategy leads the proposed methodology to capture the expert knowledge of the designer for the generation of new tests and sensors. It is worth noticing that, since the methodology guides the design of the sensors towards a target MC process parameter, sensors are actually independent of the topology of the DUT. In fact, no expert design knowledge on the DUT is required for any of the stages of the proposed methodology. An experienced designer is of course required to conceive new sensors with their associated signature. However, he does not have to know the DUT he is designing the sensors for. Thus, two different teams can work separately on the DUT and the test design. Figure 3.1 illustrates this assisted test design approach based on a sensors library (at first $i = 1$).

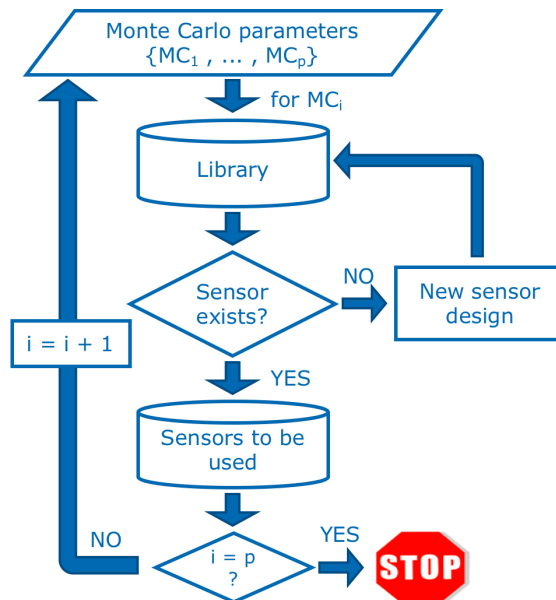


Figure 3.1: Assisted design of non-intrusive process variation sensors - Exploiting the sensors reusability by means of a library.

As a sanity check, we implement a validation for each developed signature S_i . It consists in applying a Brownian distance correlation directed-search in the complete space of MC process parameters to train a regression model towards the signature S_i . In other words, we select the most relevant MC parameters to train a regression function $f_{MC \rightarrow S_i}$ in order to ensure that the proposed signature is actually sensitive to the desired MC parameters.

The final output of the proposed test design algorithm is the set of non-intrusive sensors and the actual set of measurements (i.e. the signatures associated to the sensors) $\{S_1, \dots, S_n\}$ for the regression of the DUT performances $\{P_1, \dots, P_m\}$.

3.3 Case study

The selected DUT for our case study is a class A one-stage mm-wave power amplifier designed in a 55 nm CMOS technology from ST Microelectronics. Despite the apparent simplicity of the circuit, it is difficult to identify *a priori* the dominant source of performance degradation of the PA, since it is affected by variations of both the active and passive components. This makes it an interesting case study for the proposed test generation methodology. We have thus designed a 60 GHz power amplifier whose design and performances are detailed in this section.

3.3.1 Basics of linear power amplifier design

This subsection briefly details how to design a class A PA using a common source NMOS transistor, according to Figure 3.2. It is assumed that the reader is familiar with RF/mm-wave design, especially with the behavior of MOS transistor, passive components, the conjugate match concept, etc. For further knowledge about linear PA design please refer to [65] (chapters 1 and 2).

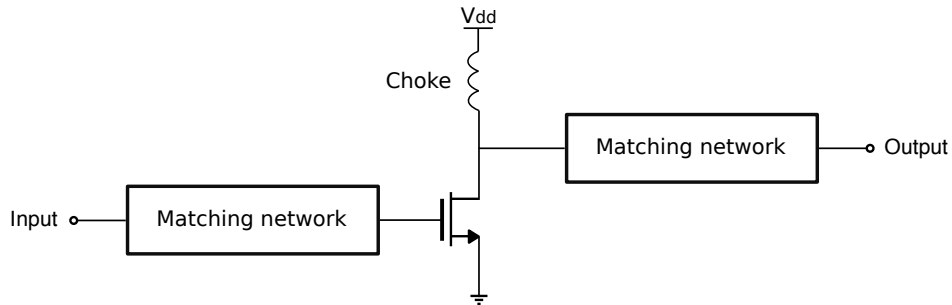


Figure 3.2: Basic diagram of a common source NMOS power amplifier.

Output power Once a technology and a supply voltage have been selected, the starting point of any PA design is the maximum power to be delivered to the load, typically a $50\ \Omega$ antenna. This maximum power, called saturation power P_{sat} , can be computed using Equation 3.4 where V_{dd} and $V_{ds_{sat}}$ respectively refer to the power supply voltage and the minimum drain-source voltage that is

required for the transistor to be in current saturation region. R_{opt} is the optimum load that needs to be presented to the transistor drain in order to deliver the desired P_{sat} . It is computed using Equation 3.5. It is important to highlight the fact that the conjugate match theorem is not applied here because of the low voltage swing the transistor drain can undergo. This concept is called loadline match and is “a real-world compromise which is necessary to extract the maximum power” from a transistor, or at least the target power, within the power supply boundaries [65].

$$P_{sat} = \frac{V_{max}}{\sqrt{2}} \times \frac{I_{max}}{\sqrt{2}} = \frac{V_{max} \times I_{max}}{2} = \frac{V_{max}^2}{2 \times R_{opt}} = \frac{(V_{dd} - V_{ds_{sat}})^2}{2 \times R_{opt}} \quad (3.4)$$

$$R_{opt} = \frac{(V_{dd} - V_{ds_{sat}})^2}{2 \times P_{sat}} \quad (3.5)$$

DC operating point Once the optimum load has been chosen, the DC bias current is obtained using Equation 3.6. The DC operating point selection is visually represented with the loadline on the $I_{ds}(V_{ds})$ curves in Figure 3.3.

$$I_{DC} \approx \frac{V_{dd}}{R_{opt}} \quad (3.6)$$

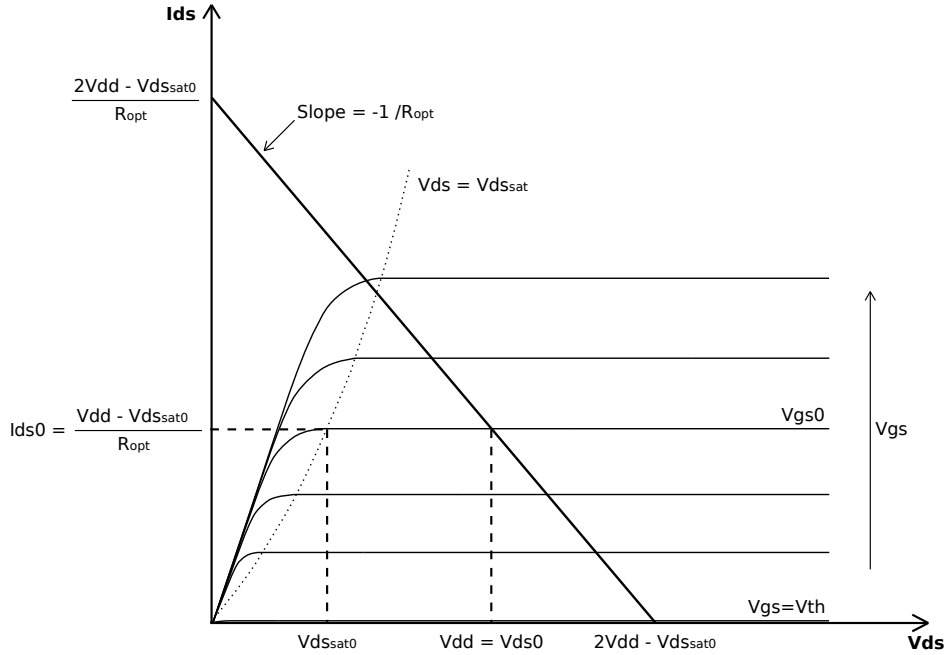


Figure 3.3: Selection of the DC operating point using the $I_{ds}(V_{ds})$ curves.

Additionally, if the quadratic behavior is assumed by neglecting the channel length modulation and the short channel effects, the DC current can also be expressed as in Equation 3.7. We can take

advantage of this expression to choose the W/L ratio and the V_{gs} voltage.

$$I_{DC} \approx \frac{\mu_n \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot (V_{gs} - V_{th})^2 \quad (3.7)$$

As this expression shows, the required DC current can thus be achieved by infinite combinations of W/L and V_{gs} values. However, two other features of the PA need to be taken into account: its power gain G_p and its compression point CP_{1dB} . In order to achieve a high compression point, the gate bias voltage V_{gs} needs to be high enough for the AC signal not to get under V_{th} , thus remaining in class A operation even under large signal operation. Nevertheless, it should not be too high to preserve the gate integrity over time. Typically, V_{gs} should have a value a little bit higher than $\frac{V_{th} + V_{dd}}{2}$. Once V_{gs} has been set up, the W/L ratio is chosen in order to have the desired I_{DC} according to Equation 3.7 and the target DC operating point is achieved.

Actually, if we consider the drain-source resistance of the MOS transistor R_{ds} , the DC current is in fact given by Equation 3.8 and the DC operating point selection is slightly different as displayed in Figure 3.4.

$$I_{DC} = \frac{V_{dd} - V_{ds_{sat}}}{R_{opt} // R_{ds}} \quad (3.8)$$

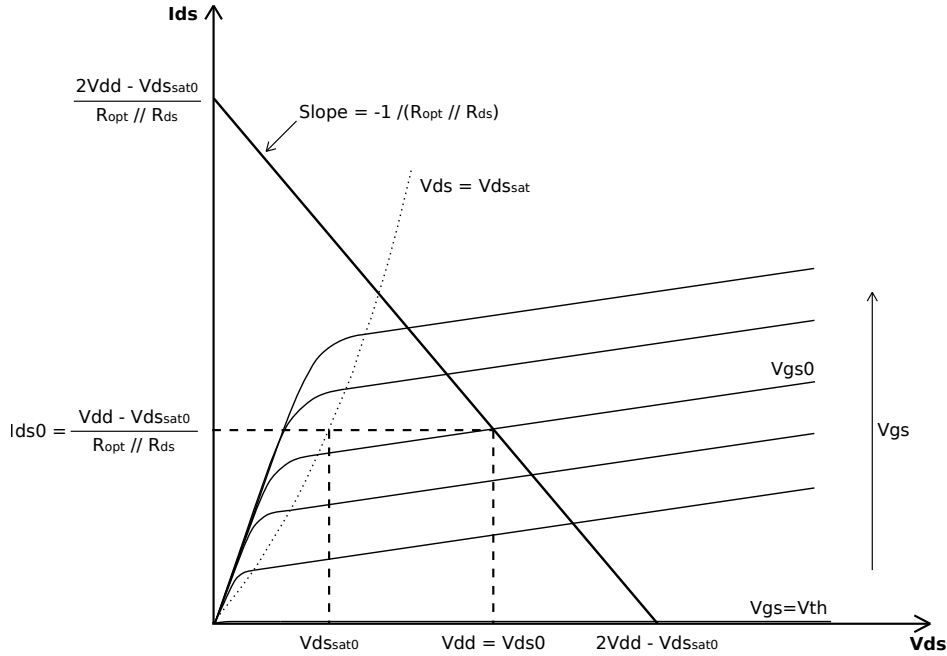


Figure 3.4: Selection of the DC operating point taking into account R_{ds} .

Since the MOS dimensions have been fixed by the previous approximation, we can obtain the R_{ds} value using an electrical simulator and thus get a more accurate value of I_{DC} using Equation 3.8. To obtain such a current, the dimensions of the transistor must be tuned, thus slightly changing the R_{ds}

value. We should keep on using alternatively Equations 3.8 and 3.7 to finally find the optimum DC operating point, after a few iterations.

Gain and input matching Once the operating point has been chosen, we can focus on other design aspects such as the gain and the matching. The power gain G_p of the PA can be computed using Equation 3.9 [66], where Y_{in} is the input admittance of the MOS transistor and α_{in} and α_{out} respectively represent the input and output matching network losses. In this expression, the transconductance g_m^* , is the transistor transconductance degenerated by the source inductor L_s (see Figure 3.5). g_m^* can be computed using Equation 3.10 [66]. The parasitic inductor L_s is due to the layout of the transistor and represents the inductive path between the real source (n^+ silicon) and the external ground. Due to its strong impact on the circuit behavior, its value needs to be carefully evaluated. The consequences of this parasitic inductor are twofold: it lowers the g_m value and it increases the real part of the input admittance $Real(Y_{in})$. Consequently, the power gain is even more lowered, but the stability and the input matching are easier to achieve.

$$G_p = \frac{(g_m^* \cdot R_{opt} // R_{ds})^2}{R_{opt} \cdot Real(Y_{in})} - \alpha_{in} - \alpha_{out} \text{ dB} \quad (3.9)$$

$$g_m^* = \frac{g_m}{\sqrt{1 + (g_m(2\pi f_0 L_s))^2}} \quad (3.10)$$

Output matching Theoretically, the output network should synthesize a real impedance R_{opt} but should also provide a positive imaginary part in order to cancel, at the working frequency, the negative imaginary part due to the drain-source capacitance C_{ds} . In order to optimize the behavior of the amplifier, load-pull simulations can be performed to tune the impedance presented by the output network. These simulations leverage the large signal models of the transistor to optimize the output load for a given input power in order to slightly increase the P_{sat} , the gain or the PAE, depending on the desired target.

3.3.2 Design of a mm-wave power amplifier

PA general presentation and main performances

For this design we have used a 55 nm BiCMOS technology from ST Microelectronics. In this work we have only used MOS transistors, so from now on, we only refer to this technology as a CMOS one. The Back-End-Of-Line (BEOL) of this technology is composed of 8 metal layers: metals 1 to 5 are thin, metals 6 and 7 are thick and metal 8 is ultra-thick. Such a BEOL is especially suited to conceive passive devices with low loss, even at mm-wave frequencies. The typical supply voltage for the selected MOS transistors is 1.2 V.

Based on the theoretical equations and guidelines introduced in the previous subsection, a class A one-stage 60 GHz PA has been designed to target a saturation power P_{sat} of 10 dBm. This power

corresponds to an optimum output load $R_{opt} \approx 40 \Omega$. Figure 3.5 shows the transistor-level schematic of the circuit. The PA input and output nodes are matched to a 50Ω impedance at 60 GHz using microstrip line stubs [67]. The matching network structure is based on shunt stubs with MOM capacitors to synthesize the shunt (i.e., a low impedance at the working frequency). These cells also play the role of decoupling cells for the DC power supplies. Figure 3.6 shows the layout of the PA. The transistor is delimited by the red box, the DC blocking capacitors by the yellow ones and the RF decoupling cells by the green ones. The distance between the DC pads and the PA core is imposed by our measurement setup.

The 9 main performances of the designed PA, obtained after post-layout simulation, under typical corner operation conditions are listed in Table 3.1, alongside with their simulated standard deviation. The 9 performances $\{P_1, \dots, P_9\}$ are the target we intend to regress from a set of non-intrusive measurements. The expected standard deviation was computed using 2,000 Monte Carlo instances, sampled using Latin Hypercube engine with SpectreRF simulator in Cadence ADE XL. It is important to keep in mind that we have taken advantage of the Monte Carlo models of the transmission lines in this study. Therefore, the simulated standard deviation takes into account process variation on all the components that were used in the design. Figure 3.7 displays the simulated scattering parameters, obtained after post-layout simulations. Based on the S-parameters, the Rollet stability factor (K-factor) has been computed to ensure the unconditional stability of the amplifier. Figure 3.8 displays the output power versus the input power at 60 GHz in order to spotlight the compression and saturation of the amplifier.

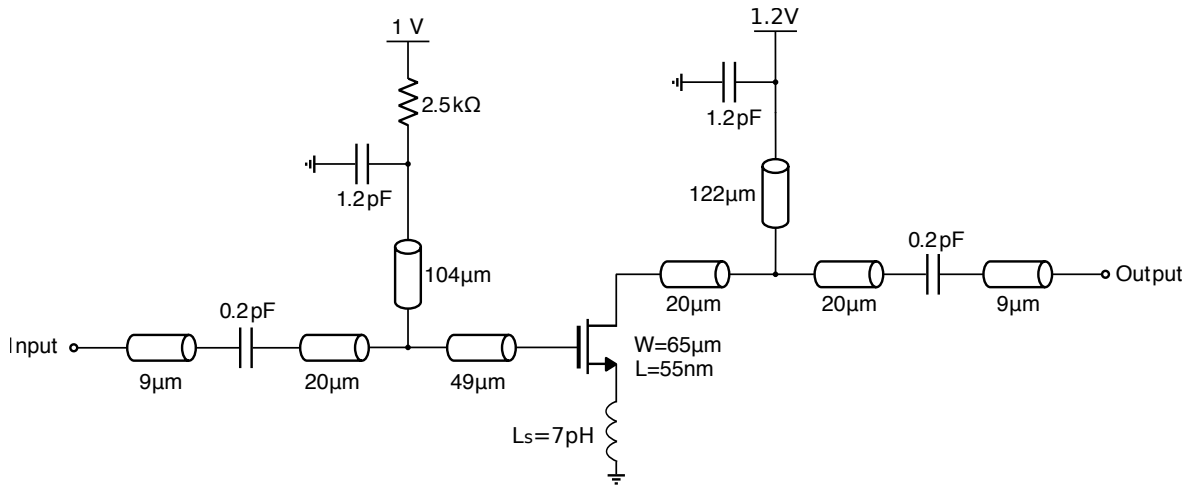


Figure 3.5: Transistor level schematic of the power amplifier under test.

PA design specifics

Some design features need to be carefully taken into account in order to build a consistent simulation model of the PA. The most critical elements are reviewed here after.

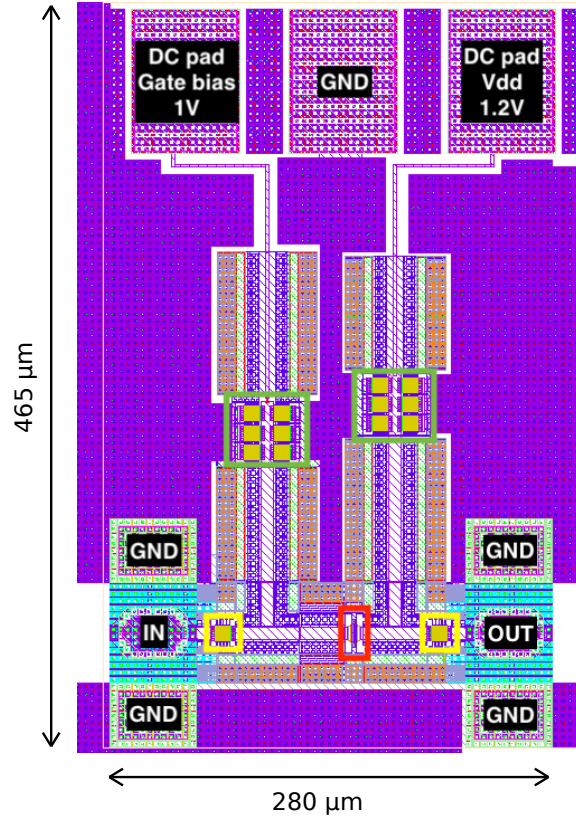


Figure 3.6: Layout of the power amplifier under test.

Table 3.1: Performances of the PA under test.

Performances	Unit	Value	Standard deviation
DC current, I_{DC}	mA	28.2	1.6
Gain (S_{21})	dB	5.6	0.7
Power Added Efficiency, PAE	%	11	2.0
S_{11}	dB	< -20	4.5
S_{22}	dB	< -20	4.9
S_{12}	dB	< -13	0.8
Input referred 1dB Compression Point, ICP_{1dB}	dBm	0.7	0.6
Saturation output power, P_{sat}	dBm	10.5	0.2
Input referred third-order intercept point, $IIP3$	dBm	11	0.7

MOS transistor DC operating point and dimensions Following the method presented in the previous subsection, we have chosen the DC operating point of the transistor: $I_{ds}=28$ mA and $V_{gs}=1$ V. The corresponding dimensions for the transistor are $W=65$ μm and $L=55$ nm. In order to reduce the parasitic capacitances, and thus increase the transition frequency, we have of course used the minimum channel length of the technology i.e. 55 nm. The transistor is divided into 48 fingers of 1.35 μm which

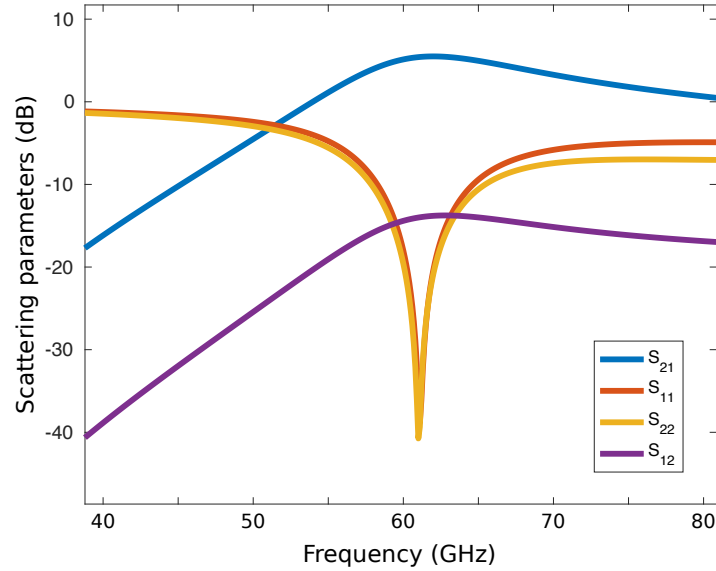


Figure 3.7: Scattering parameters obtained through post-layout simulations.

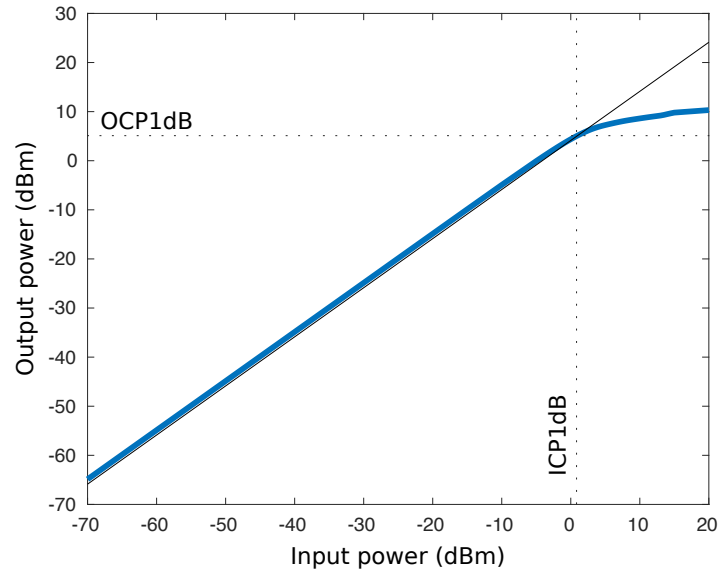


Figure 3.8: Output power versus input power at 60 GHz - Compression point and saturation.

is a tradeoff between the gate resistance, the parasitic capacitances and the shape of the transistor. These dimensions led to an output resistance of $R_{ds} = 160 \Omega$. As mentioned in the previous section, we have taken this resistance into account to choose the DC operating point since it is not negligible compared to $R_{opt} = 40 \Omega$.

MOS transistor connection The transistor connection with the matching networks and the ground is crucial. In order to reduce the parasitics, and more especially C_{dg} and L_s , we have used the layout displayed in Figure 3.9.a. Gate and drain access are in metal 8 (purple) which corresponds to the metal used for the microstrip transmission lines. The source connection makes use of 48 metal fingers, stacked from metal 1 to metal 4 and then shorted together using metals 5 to 8. Inductor L_s corresponds to the inductive path of all the fingers in parallel. For this design we have used an empirical value of 7 pF. According to the measurements we performed on fabricated samples, a more realistic value for this design would have been 3 pH (see subsection 4.3.1).

To avoid antenna problems during the fabrication, a diode was inserted between the bulk and the gate. It is spotlighted in Figure 3.9.b, which displays a few fingers of the MOS transistor, where metals 5 to 8 have been removed for more relevance.

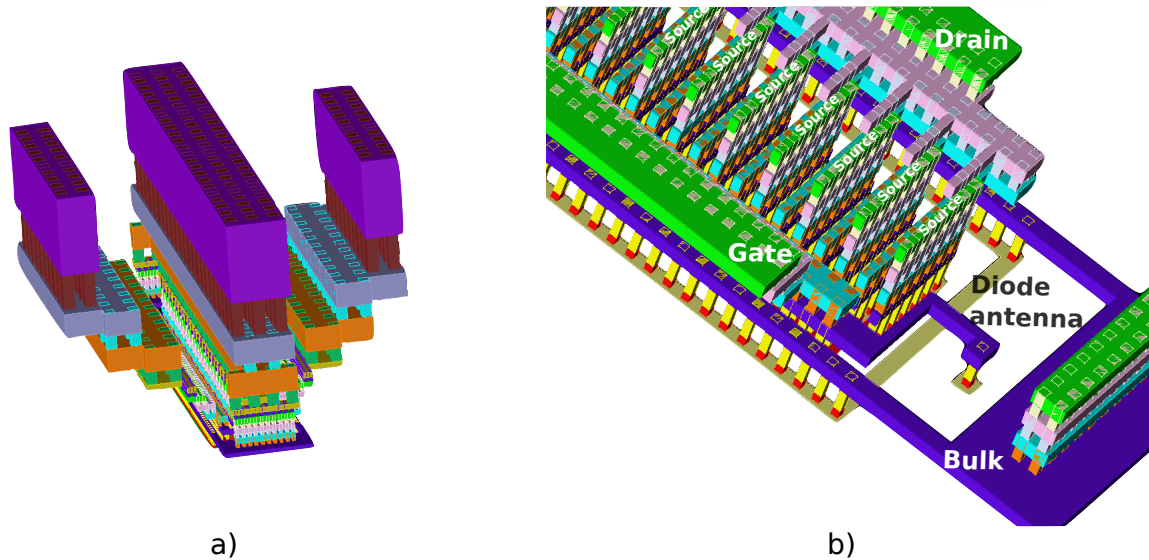


Figure 3.9: 3D view of the MOS transistor connections.

Transmission lines We have taken advantage of the accurate and flexible models that are provided in the PDK. These models have the advantage of having Monte Carlo models that are required for this study. We have used $50\ \Omega$ microstrip lines since they provide several assets, especially the fact that they can be easily connected to the transistor we have designed since its size after layout is close to the $50\ \Omega$ width. Moreover they have quite low losses and the size of the matching networks are reasonable for the impedance transformations we aim. We might have been able to slightly reduce their size by combining lines with different characteristic impedances but we expected more accurate and reliable results using a classical $50\ \Omega$ impedance. Figure 3.10 displays simplified top and 3D views of a $50\ \Omega$ microstrip line. Metal 1 (M1) is the ground floor, metals 2 to 7 (M2 to M7) are used for complying with the local density rules and are connected to the ground while metal 8 (M8) is the signal strip.

Small holes have been inserted in the ground floor for complying with the maximum local density rule for M1 as shown in Figure 3.10.a, though they are not displayed in Figure 3.10.b for a better readability. The main features of this transmission line are summarized in Table 3.2.

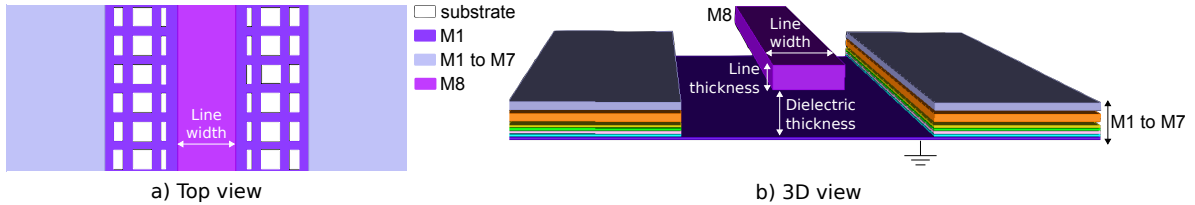


Figure 3.10: Top and 3D views of a 50 Ω microstrip line.

Table 3.2: 50 Ω transmission line main features.

Feature	Unit	Value
line width	μm	7.7
line thickness	μm	2.9
M8 conductivity σ	S/m	47.0
dielectric thickness	μm	5.68
dielectric constant ϵ	–	3.91

The selected matching network architecture is based on parallel stubs and T-junctions are not modeled in the PDK. Having considered the relative small width of the lines, we decided not to perform any electromagnetic simulations. We simply modeled the T-junction with three small pieces of transmission line whose dimensions accurately matched the layout as displayed in Figure 3.11. As observed in measurements and then confirmed by electromagnetic simulations, this was a too simple model (see subsection 4.3.1).

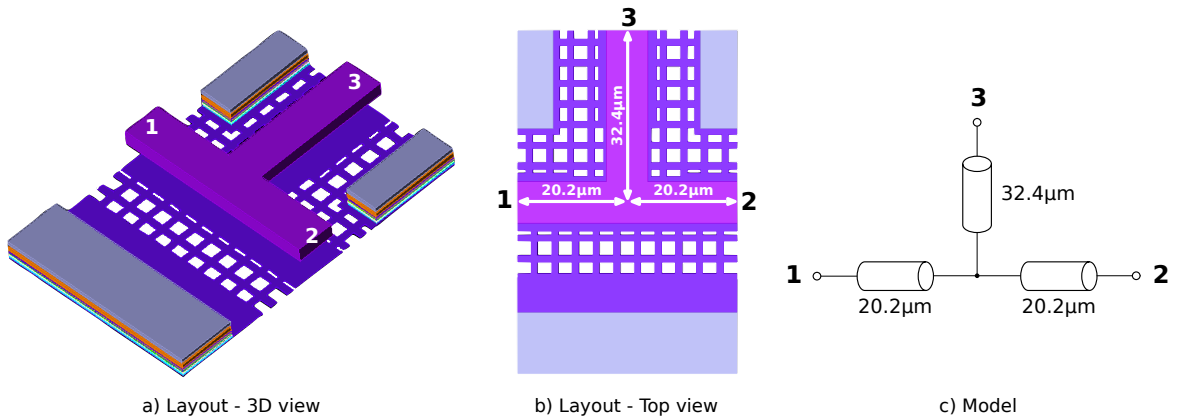


Figure 3.11: T-junction layout and model.

Decoupling cells The role of this cell is twofold: it serves both as a decoupling cell and as an RF shunt at 60 GHz. It was conceived by connecting together 6 MOM capacitors of 0.2 pF in order to achieve a low impedance at the working frequency ($0.5 - j1.8 \Omega$). The capacitors were connected together using a M8 strip in order to ease the connection with the microstrip stubs. Figure 3.12 shows a 3D view of this cell and Figure 3.13 shows a top view along side with the model that was used for simulations. Note that the physical length of the M8 strip (purple) does not match the length of the microstrip lines that were used to model it. This model was constructed using Keysight Momentum simulations.

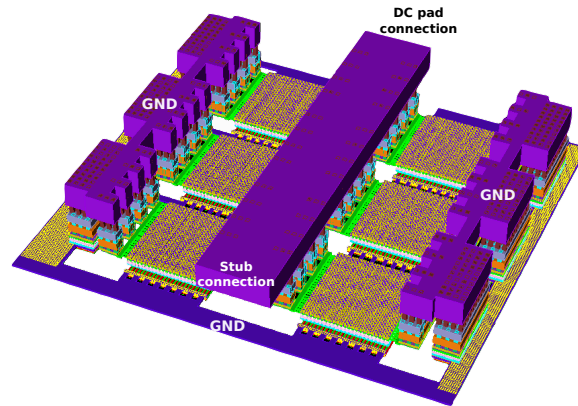


Figure 3.12: 3D view of the decoupling cell.

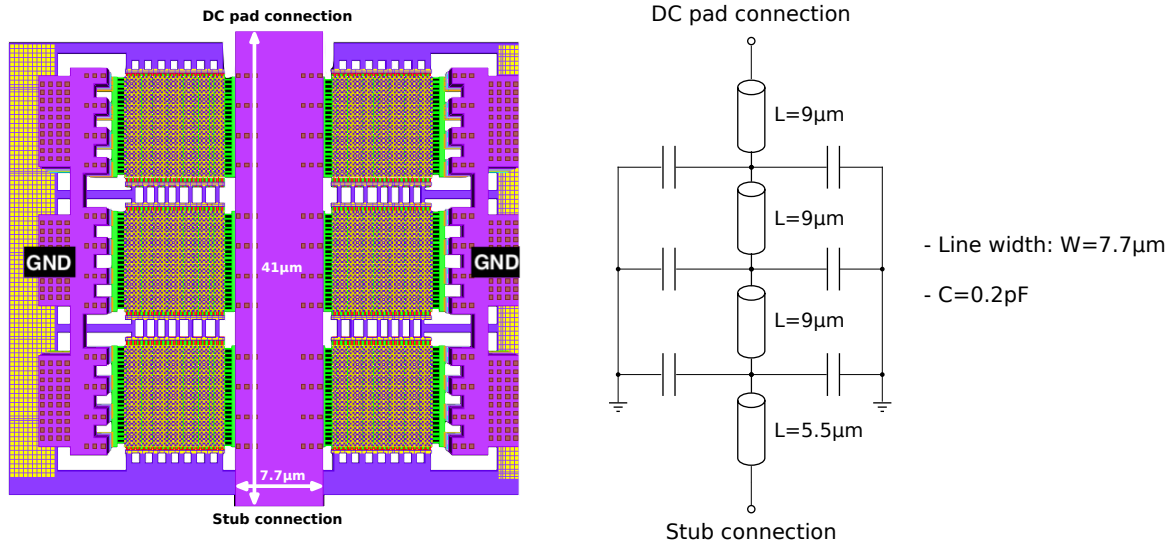


Figure 3.13: Top view and model of the decoupling cell.

RF Ground-Signal-Ground (GSG) pads They need to be accounted for since they are part of the circuit and will not be de-embedded. Based on the team experience and measurements that had been performed using the same RF pads in the the past years, the pads are modeled using a simple parallel capacitor. Its extracted value is between 25 fF and 35 fF depending on the measurements and the RF probes that were used. We chose the worse value of 35 fF for designing the amplifier. It is worth noticing that this parallel capacitor not only models the pad capacitance itself but also the additional coupling between the probe and the ground. For the measurements, we have finally used different probes than the ones available at the time the circuit was designed. Consequently, the model [pad+probe] needs to be modified (see subsection 4.3.1).

3.3.3 Applying the proposed methodology to the designed PA

The proposed test design methodology has been applied to the PA as a practical case study to illustrate the technique. The goal of this case study is to generate a set of appropriate non-intrusive sensors that provide meaningful signatures for testing the PA using a machine learning indirect test strategy. In the end, we should have a set of n signatures $\{S_1, \dots, S_n\}$ that allow to regress the 9 main performances of the PA $\{P_1, \dots, P_9\}$ using a model $f_{S \rightarrow P}$ defined as

$$f_{S \rightarrow P} : [S_1, \dots, S_n] \rightarrow [P_1, \dots, P_9]. \quad (3.11)$$

Step 1: Unveiling the root causes of parametric performance variation

A set of 2,000 instances of the selected PA has been generated using latin hypercube sampling and the MC models provided in the PDK of the technology. A total of 200 instances of the PA are set apart to be used as an independent verification set and the remaining 1,800 instances are used for training. The first step of the proposed methodology, as previously described, is to find the root-causes of parametric variation within the set of MC parameters defined in the technology. The MC simulation of the PA in the selected 55 nm CMOS technology includes 154 independent process parameters $\{MC_1, \dots, MC_{154}\}$ that define the search space. The exploration of the MC parameter space uses the proposed Brownian distance correlation-directed search for regressing the PA specifications. A perceptron Neural Network is used as machine learning regression algorithm. It has been implemented in Matlab, using the Neural Network Toolbox. The complete procedure is illustrated for only one of the specifications, namely the small signal gain S_{21} of the PA at 60 GHz. The process is then iterated for the rest of the specifications.

Figure 3.14 presents the model RMS error as a function of the number of MC parameters selected by the search algorithm, for the complete set of 1800 training instances (marked with blue diamonds), and the 200 samples in the independent verification set (marked with red circles). For illustration purposes, instead of setting a target prediction error as stopping condition for the algorithm, the search has been set to find and rank the 20 most relevant MC parameters, as shown in Figure 3.14. It is interesting to notice that the proposed algorithm guides the selection of MC parameters in such a way that it yields a monotonic improvement of the prediction error. Moreover, it is clear to see that this improvement of the prediction error tends to slow down between the fifteenth and the twentieth

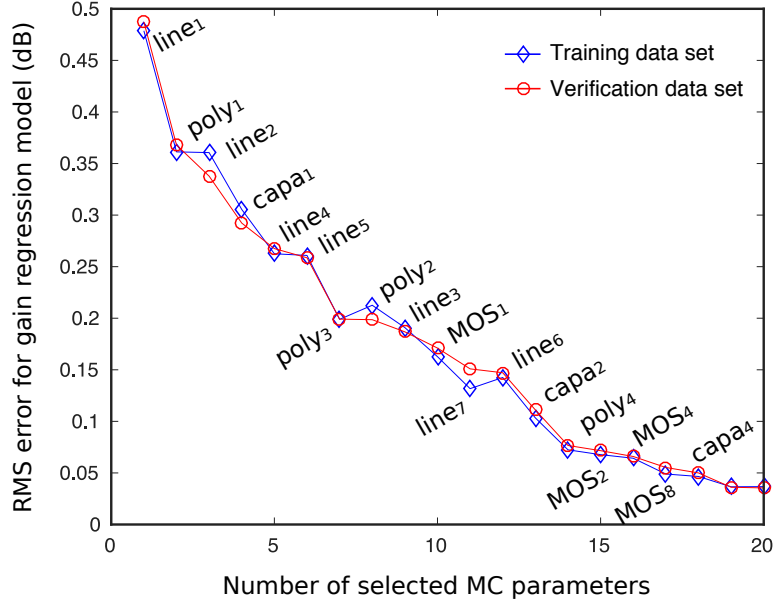


Figure 3.14: Exploration of the MC process parameter space for finding the root causes of gain parametric variation for the PA.

iteration of the search. This means additional MC parameters do not bring significant improvements to the prediction error. In this line, the 18 most significant parameters have been selected and are labelled next to the corresponding data points in the figure. For confidentiality reasons we cannot disclose the real name and meaning of the MC parameters. $poly_i$ parameters refer to the MC process parameters considered for polysilicium transistor gate, MOS_j to other MOS transistors features, $capa_k$ to MOM capacitors features and $line_l$ to transmission line features.

Figure 3.15 shows the scatterplot of the predicted versus actual gain of the PA when we train a neural network regression model using these 18 selected MC parameters as input features. Blue dots represent the training data set, while red circles mark the samples in the independent verification set. It is clear that the selected MC parameters have a very strong correlation to the target performance. The gain was predicted using a regression function $f_{MC \rightarrow S_{21}}$ defined as follows

$$f_{MC \rightarrow S_{21}} : [line_1, poly_1, line_2, capa_1, line_4, line_5, poly_3, poly_2, line_3, MOS_1, line_7, line_6, capa_2, poly_4, MOS_2, MOS_8, MOS_4, capa_4] \rightarrow S_{21}. \quad (3.12)$$

Step 2: Assisted design of non-intrusive process variation sensors

Once the most relevant MC parameters for the prediction of the targeted performance have been identified, the second step in the presented test design methodology is to propose appropriate non-intrusive sensors that yield signatures strongly correlated to these parameters. The design of these

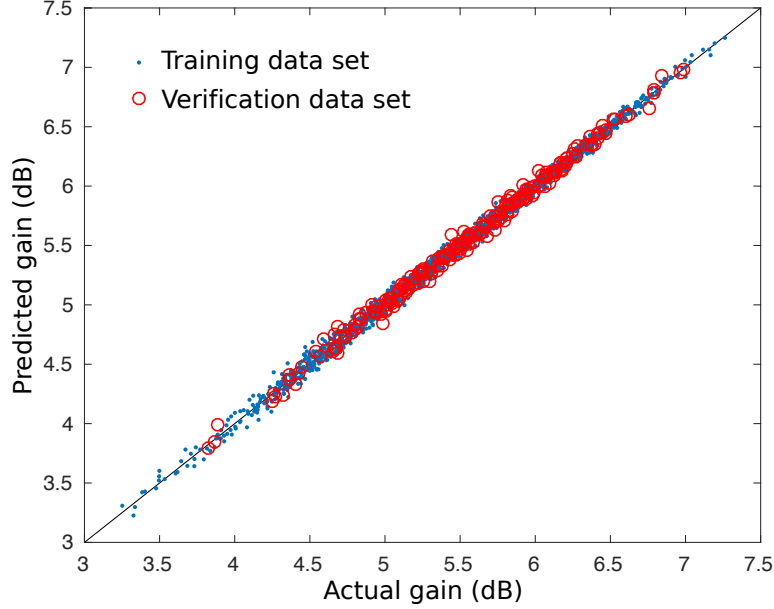


Figure 3.15: Scatterplot of predicted versus actual gain of the PA using the 18 most relevant MC parameters as signatures.

sensors and their associated measurements is guided by the physical meaning of the selected MC parameters. In order to verify that the 18 selected MC parameters are covered by the proposed signatures, we perform our Brownian distance correlation-directed search in the space of process parameters for regressing each one of the proposed signatures. In other words, for each signature S_i we try to build a function $f_{MC \rightarrow S_i}$ to verify that the signature is relevant, that it is sensitive to the target MC parameters

$$f_{MC \rightarrow S_i} : [MC_1, \dots, MC_p] \rightarrow S_i. \quad (3.13)$$

Let us consider the 18 selected MC parameters in the order they are ranked in Figure 3.14 and find a relevant signature for each of them:

- 1) $line_1$ describes the inductive behavior of a microstrip line. Thus, we propose to design a high-impedance line and measure two signatures $S_1: Imag(S_{11})$ and $S_2: Imag(S_{21})$ at 100 MHz. They should be sensitive to this MC parameter in particular and complement each other. To check to which MC parameter the signatures are sensitive to, we take advantage of the Brownian distance correlation-directed search algorithm. It happens that S_1 and S_2 are sensitive to $line_1$ as they were devised for, but also $line_2, line_3, line_4$ and $line_5$ for S_1 and $line_2, line_4, line_5$ and $line_6$ for S_2 . Hence, these MC parameters should not require specific signatures.
- 2) $poly_1$ corresponds to an important feature of the gate polysilicon of MOS transistors. Thus, we simply propose to measure the gate resistance of a dummy transistor. This signature S_3 is sensitive to $poly_1, poly_2, poly_3, poly_4, MOS_7$. Parameters $poly_2, poly_3, poly_4$ are covered by this signature so they should not require specific signatures.
- 3) $line_2$ describes the resistance of a microstrip line. Although it has already been addressed by

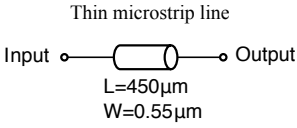

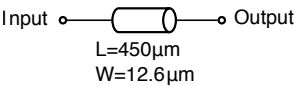

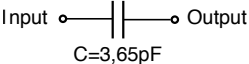

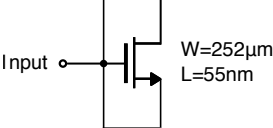
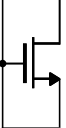
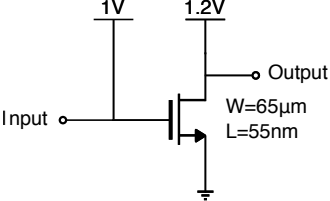
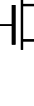
signatures S_1 and S_2 , a resistance measurement S_4 on the high impedance line seems to be a more relevant and reliable sensor.

- 4) $capa_1$ describes MOM capacitor features related to their losses. We thus propose to implement a large MOM capacitor and measure $S_5: Real(Y_{11})$ at 100 MHz. This signature is correlated to $capa_1, capa_4, capa_2, capa_5$ and $capa_3$.
- 5) $line_4$ describes a dielectric feature of the lines that is already addressed by signatures S_1 and S_2 .
- 6) $line_5$ describes a dielectric feature of the lines that is already addressed by signatures S_1 and S_2 .
- 7) $poly_3$ is used to describe the current of a transistor. Although it is already addressed by signature S_3 , we decide to add another signature S_6 that we deem more relevant. It corresponds to a DC current measurement on the existing MOS transistor sensor.
- 8) $poly_2$ describes a feature of the gate polysilicon of MOS transistors. It is already addressed by signature S_1 .
- 9) $line_3$ describes a dielectric feature of the lines that is already addressed by signature S_1 .
- 10) MOS_1 describes a feature of NMOS transistors related to their junction capacitance. We propose to add a junction capacitance sensor by shorting the source, gate and drain access of a NMOS transistor and measure signature $S_7: Imag(S_{11})$ at 100 MHz.
- 11) $line_7$ describes a dielectric feature of the lines. It is not addressed by signatures S_1 and S_2 so we propose to implement a low impedance microstrip line and measure $S_8: Imag(S_{11})$ at 100 MHz.
- 12) $line_6$ describes a dielectric feature of the lines and is already addressed by signature S_2 and S_8 .
- 13) $capa_2$ describes MOM capacitor features related to their capacitance value. Although this parameter has already been addressed by signature S_5 , we propose to reuse the large MOM capacitor and measure $S_9: Imag(S_{21})$ at 100 MHz. As a matter of fact, this measurement is more straightforward to sense a change in the capacitance value.
- 14) $poly_4$ describes a feature of the gate polysilicon of MOS transistors. It is already addressed by signature S_3 .
- 15) MOS_2 describes a feature of NMOS transistors related to their junction capacitance and is already addressed by signature S_7 .
- 16) MOS_4 describes a feature of NMOS transistors and is already addressed by signature S_6 .
- 17) MOS_8 describes a feature of NMOS transistors and is already addressed by signatures S_6 and S_7 .
- 18) $capa_4$ describes a feature of MOM capacitors and is already addressed by signatures S_5 and S_9 .

Table 3.3 details the schematics of the developed sensors and the associated signatures. It also includes the list of the most significant MC parameters that explain the variation of each proposed signature for validation. As it can be seen, all 18 relevant MC parameters are covered by the devised signatures. It is worth noticing that several sensors are sensitive to other process parameters ($capa_3, capa_5, MOS_3, MOS_5, MOS_6, MOS_7$). Their importance is of second order for the concerned sensors.

In total, for covering the identified 18 most significant MC parameters we have developed 5 non-intrusive sensors and 9 signatures labelled as $\{S_1, \dots, S_9\}$. Notice that signatures are sensitive to multiple MC parameters, which reduces the number of necessary measurements. The proposed set

Table 3.3: Proposed library of non-intrusive sensors and their associated signatures for gain prediction.

Process variations sensor	Signature description	MC parameters covered
<p>Thin microstrip line</p>  <p>Input  Output L=450µm W=0.55µm</p>	<ul style="list-style-type: none"> • S_1: Imaginary part of S_{11} @ 100 MHz • S_2: Imaginary part of S_{21} @ 100 MHz • S_4: Resistance 	<ul style="list-style-type: none"> • $line_1, line_2, line_3, line_4, line_5$ • $line_1, line_2, line_4, line_5, line_6$ • $line_2$
<p>Wide microstrip line</p>  <p>Input  Output L=450µm W=12.6µm</p>	<ul style="list-style-type: none"> • S_8: Imaginary part of S_{11} @ 100 MHz 	<ul style="list-style-type: none"> • $line_4, line_5, line_6, line_7, line_1$
<p>MOM Capacitor</p>  <p>Input  Output C=3,65pF</p>	<ul style="list-style-type: none"> • S_5: Real part of Y_{11} @ 100 MHz • S_9: Imaginary part of S_{21} @ 100 MHz 	<ul style="list-style-type: none"> • $capa_1, capa_4, capa_2, capa_5, capa_3$ • $capa_2, capa_4, capa_5$
<p>Junction capacitance sensor</p>  <p>Input  Output W=252µm L=55nm</p>	<ul style="list-style-type: none"> • S_7: Imaginary part of S_{11} @ 100 MHz 	<ul style="list-style-type: none"> • $MOS_1, MOS_2, poly_3, MOS_8, MOS_6, MOS_7$
<p>NMOS transistor</p>  <p>Input  Output 1V 1.2V W=65µm L=55nm</p>	<ul style="list-style-type: none"> • S_3: Gate resistance • S_6: Bias current 	<ul style="list-style-type: none"> • $poly_1, poly_2, poly_3, poly_4, MOS_7$ • $poly_3, MOS_8, MOS_4, MOS_5, MOS_7, MOS_3$

of signatures contains DC measurements and low frequency measurements performed at 100 MHz. Notice that the 100 MHz frequency is a subjective choice, far lower than the operating frequency of the device. These low frequency measurements are simpler and cheaper to extract than 60 GHz measurements on the PA under test. From an electrical point of view, the proposed signatures contain information about phase-shifts in the elements of the PA (S_1, S_2, S_8), capacitive behavior (S_7, S_9), electrical losses (S_3, S_4, S_5) and the operation point of the NMOS transistor (S_6).

The developed non-intrusive sensors and their corresponding signatures are then stored in a library for reuse in future indirect test programs. As it was previously mentioned, the developed sensors are independent of the DUT topology, since they are devised to target specific root causes of parametric variation in a given technology. The reuse of the library of non-intrusive sensors for other DUTs offers an additional layer of automation to the proposed test design strategy.

As a final validation of the performance of the proposed assisted test design strategy, we applied the resulting non-intrusive machine learning indirect test to the PA. A perceptron Neural Network is trained to regress the gain of the PA under test from the identified 9 input signatures $\{S_1, \dots, S_9\}$ extracted from the generated non-intrusive sensors, that is we train the function $f_{S \rightarrow S_{21}}$ defined as follows

$$f_{S \rightarrow S_{21}} : [S_1, \dots, S_9] \rightarrow S_{21}. \quad (3.14)$$

Figure 3.16 shows the scatterplot of the predicted versus actual gain over the 2,000 generated instances of the PA under test. Blue dots mark the training data set, while red circles represent the samples in the independent verification set. As it can be seen, the proposed test program offers a good estimation of the gain performance across the complete variation range. The RMS error of the prediction for samples in the independent verification set is of only 0.07 dB, which is very comparable to the regression in Figure 3.15 that employed the 18 most significant MC process parameters as input signatures and yields an RMS prediction error of 0.05 dB. This result is a good indicator of the relevance of the proposed set of signatures to capture the variation of the target performance.

It is also interesting to train regression models with subsets of the set of 9 signatures. As a matter of fact, if less accuracy is required, all the signatures may not be required. In this line, the Brownian distance correlation directed-search algorithm was used on the signatures to rank them and spotlight relevant subsets. Figure 3.17 displays the RMS error for gain prediction as a function of the number of signatures that are used, in the order they are selected by the algorithm. Blue diamonds correspond to the training samples while red circles represent independent verification samples. Signature names are labelled next to their respective points. For instance, the 5 most relevant signatures are able to provide a gain prediction accuracy around 0.25 dB in average. Figure 3.18 displays again the RMS prediction error in the independent verification set and compares it with the maximum error of the regression model. With 5 signatures, the maximum prediction error is less than 0.6 dB and with 9 signatures it is around 0.2 dB. The signatures are selected in the same order as they were in Figure 3.17.

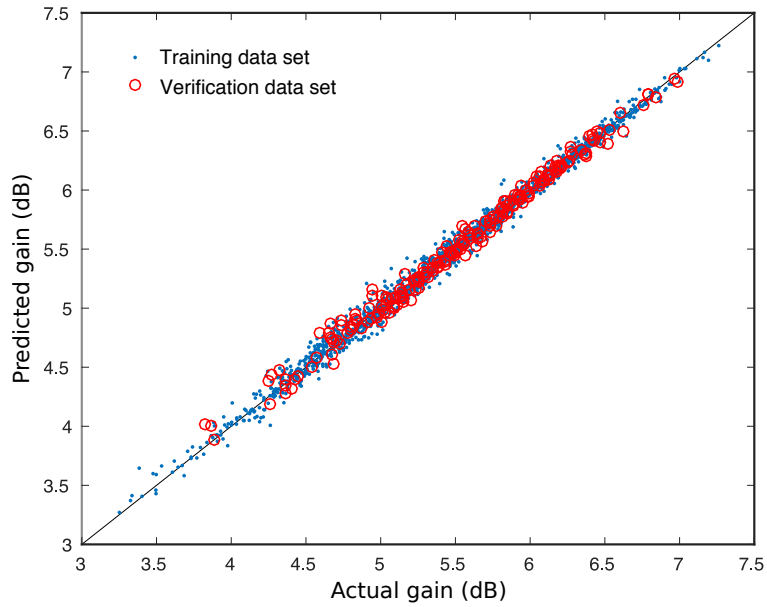


Figure 3.16: Scatterplot of predicted versus actual gain of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3.

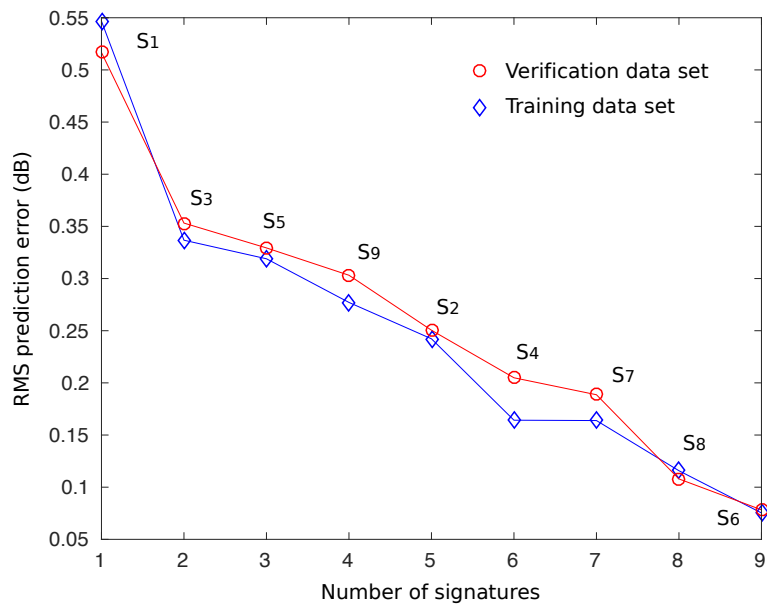


Figure 3.17: RMS error for gain prediction versus the number of signatures.

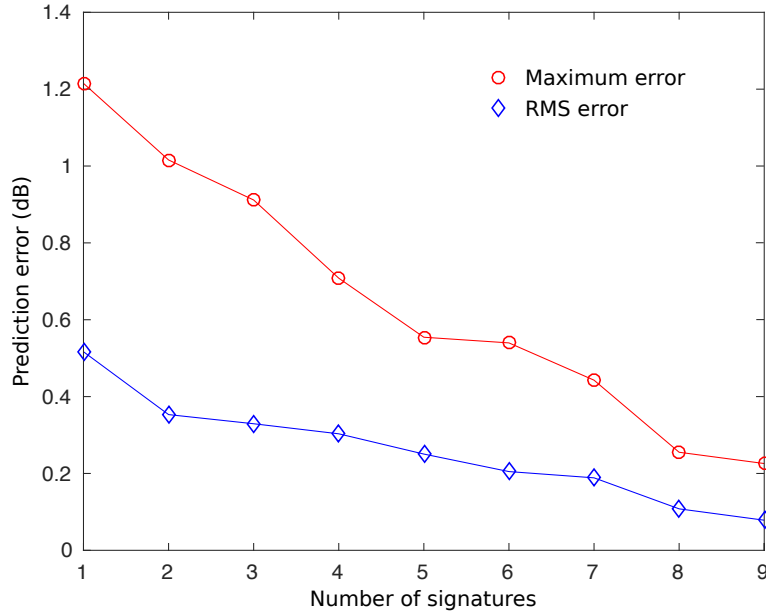


Figure 3.18: Maximum and RMS error for gain prediction versus the number of signatures.

Applying the methodology to target the other PA specifications

Finally, the complete test design process can be repeated for the rest of the specifications of the PA under test. The results are listed in Table 3.4. This table shows the sets of identified MC parameters that are relevant for the prediction of each PA specification, ranked by relevance. For each specification, the selection of MC parameters has been stopped when the prediction error does not improve significantly by incorporating additional parameters. As it can be seen, the degradation of the selected specifications due to process variations is dominated by a subset of only 22 different MC parameters. Interestingly, if we consider the complete set of PA target performances, the described assisted test design algorithm do not lead to any additional sensor or signature.

In order to validate the resulting non-intrusive indirect tests, perceptron neural network models $f_{S \rightarrow P}$ have been trained for regressing the complete set of PA specifications using the identified sets of signatures $\{S_1, \dots, S_9\}$. Figures 3.19 and 3.20 display the scatter plot of the predicted versus actual specification for the other target specifications of the PA. Visually, the prediction results are excellent, although they do not seem as good for S_{11} and S_{22} as they are for the other specifications. However, for these two plots the scale is linear ($0.1 = -10$ dB and $0.05 = -13$ dB) so the accuracy is in fact fairly good. Table 3.5 lists the RMS prediction error for samples in the independent verification set for each specification. The RMS prediction error was averaged over 100 trained neural networks to be as representative of the models performance as possible. In order to evaluate the quality of the regression models, Table 3.5 compares the obtained RMS errors with the minimum, maximum and standard variation of the specification in MC simulations. It is clear that the models trained

with the automatically identified set of signatures offer accurate predictions for all the considered specifications.

Table 3.4: Identified relevant MC parameters for the prediction of the PA specifications.

Specification	Identified Monte Carlo parameters (ranked by relevance)
Gain (S_{21})	$line_1, poly_1, line_2, capa_1, line_4, line_5, poly_3, poly_2, line_3, MOS_1, line_7, line_6, capa_2, poly_4, MOS_2, MOS_4, MOS_8, capa_4$
S_{11}	$line_1, poly_1, poly_3, line_2, line_4, poly_2, line_6, line_5, line_3, capa_1, line_7, capa_2, MOS_1, poly_4$
S_{22}	$line_1, poly_3, line_4, poly_1, line_7, line_6, line_3, capa_2, MOS_4, capa_4, line_2, poly_2, MOS_1, MOS_2, capa_5$
S_{12}	$line_1, poly_1, poly_3, line_2, capa_1, line_5, line_4, poly_2, line_3, MOS_1, line_7, line_6, capa_2, poly_4, MOS_2, MOS_4, capa_4$
Power Added Efficiency, PAE	$line_1, poly_1, line_2, line_4, capa_1, poly_3, line_5, line_3, capa_2, line_6, line_7, MOS_1, poly_2, MOS_2, poly_1, capa_4, MOS_4$
DC current, I_{DC}	$poly_3, MOS_8, MOS_4, MOS_5, MOS_7, MOS_3, line_2$
Input referred 1 dB Compression point, ICP_{1dB}	$poly_1, line_1, line_2, poly_3, poly_2, capa_1, line_5, MOS_1, line_4, line_3, poly_4, line_6, line_7, MOS_2, capa_3, capa_2, MOS_8, capa_4$
Saturation output power, P_{sat}	$line_1, capa_2, capa_1, line_2, line_4, line_5, line_3, capa_4, MOS_1, line_6, line_7, poly_1, MOS_5, MOS_8, MOS_2, MOS_3, capa_5, poly_3, MOS_4$
Input referred third-order intercept point, $IIP3$	$line_1, poly_1, line_2, capa_1, line_4, poly_3, line_5, poly_2, MOS_1, line_3, line_6, line_7, poly_4, MOS_2, MOS_8, MOS_4, capa_2$

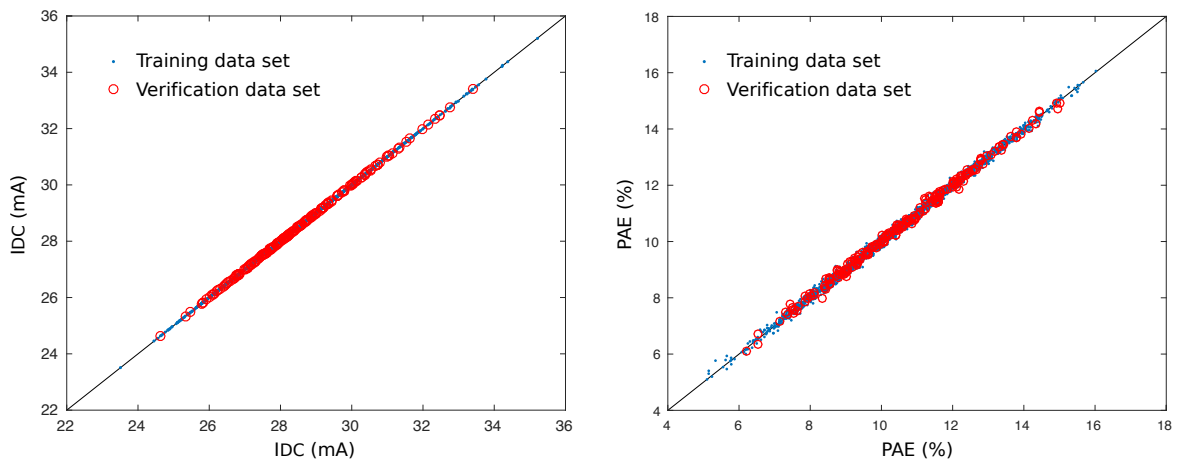


Figure 3.19: Scatterplot of predicted versus actual specification of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3.

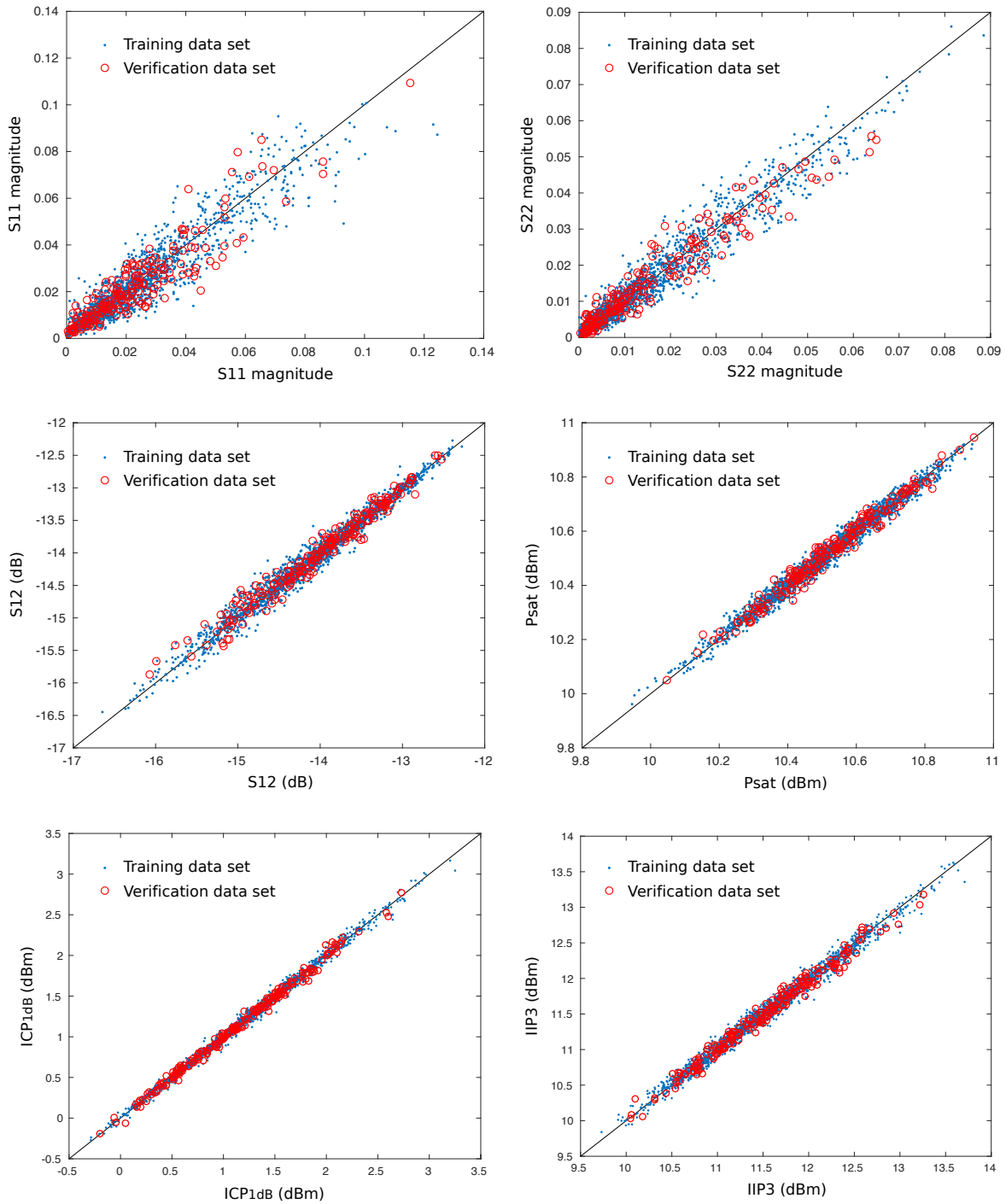


Figure 3.20: Scatterplot of predicted versus actual specification of the PA using the generated indirect test program with the 9 non-intrusive signatures of Table 3.3 (continued).

Table 3.5: Prediction of PA performances using the generated indirect test program.

Specification	Unit	min_{spec}	max_{spec}	σ_{spec}	RMS prediction error
PAE	%	5.10	16.02	1.98	0.16
Gain (S_{21})	dB	3.25	7.26	0.69	0.07
S_{11} (magnitude)	–	0.0001	0.1244	0.0196	0.007
S_{22} (magnitude)	–	0.0001	0.0885	0.0147	0.0037
S_{12}	dB	–16.65	–12.28	0.75	0.12
ICP_{1dB}	dBm	–0.29	3.25	0.59	0.07
P_{sat}	dBm	9.95	10.94	0.17	0.02
$IIP3$	dBm	9.73	13.71	0.67	0.10
DC current, I_{DC}	mA	23.51	35.21	1.59	<0.01 using S_4 and S_6 only

3.4 Conclusions

We have presented a systematic methodology for generating a non-intrusive indirect test program for mm-wave circuits. The proposed strategy represents a promising step towards the full automation of test generation for non-intrusive machine learning test. The proposed methodology has been illustrated on a 60 GHz PA designed in a 55 nm CMOS technology from ST Microelectronics. We have shown that the presented assisted test design algorithm enables us to find the root causes of parametric performance degradation and guides the design of appropriate non-intrusive sensors and signatures for accurate performance prediction. This study has been carried out using simulation data only by taking advantage of the MC models.

In the next chapter we present an extension of this generic method that targets circuits with custom passive components. To assert the feasibility of such a technique, measurements have been performed on fabricated samples.

Chapter 4

Assisted test design for machine learning indirect test of mm-wave circuits using custom passive components

4.1 Introduction

In the last chapter we have presented a first attempt at systematizing the design of an indirect test program for mm-wave integrated circuits. However, our methodology assumes that Monte Carlo process variation models are available for all the elements in the DUT. This is not usually the case in mm-wave integrated circuit designs, where passive components such as transmission lines, couplers, baluns, etc. are often custom-built in the BEOL of the technology. In this chapter, we extend our work to cover also custom-built passive devices. This way, we define a truly systematic methodology for designing an indirect test program for a generic mm-wave IC.

Moreover, our methodology is supported by experimental results in a set of 65 GHz power amplifiers fabricated in ST Microelectronics 55 nm CMOS technology. To our knowledge, this is the first time that the non-intrusive test strategy of mm-wave integrated circuits is validated with silicon results.

The chapter is organized as follows. Section 4.2 details our generic non-intrusive BIST strategy and the proposed methodology for non-intrusive sensors design. In Section 4.3 the presented methodology is illustrated with a step-by-step application on a 65 GHz PA. Section 4.4 validates the proposed test strategy based on experimental results from 21 fabricated samples of the 65 GHz PA. Finally, Section 4.5 summarizes the main contributions of this work.

4.2 Non-intrusive test program generation for generic mm-wave integrated circuits

4.2.1 Problem statement

Let us consider a generic mm-wave IC that may contain both active and passive components modeled in the PDK and full-custom passive devices built in the BEOL of the technology. The problem we address is basically the same as in the last chapter. The only, and major, difference is that the circuit can use custom passive components. Consequently, we still consider a mm-wave IC that has m performances $\{P_1, \dots, P_m\}$ that have to be characterized during production test. The goal of this methodology is to propose a non-intrusive test program containing: a) a set of non-intrusive process monitors, and b) a set of n low-frequency signatures $\{S_1, \dots, S_n\}$, extracted from these non-intrusive process sensors. The key idea in the methodology is that the designed signatures will capture the main causes of performance degradation due to process variations, in such a way that we can build a machine learning regression model $f_{S \rightarrow P}$ in order to accurately infer the circuit performances from the measurement of the signatures as,

$$f : [S_1, \dots, S_n] \rightarrow [P_1, \dots, P_m]. \quad (4.1)$$

As a starting point, the proposed methodology assumes that we have access to: a) the netlist and layout of the mm-wave IC, and b) the technological information available in the PDK of the technology, including the Monte Carlo process variation models and basic information about the BEOL layers. No particular assumptions are made about the nature of the mm-wave device itself. It is important to highlight that the starting point is different than the one we have used in the last chapter where we only needed a) the netlist of the DUT, and b) the MC process variation models.

The proposed methodology is schematically illustrated in Figure 4.1. Similarly to the method we have presented in the last chapter, it is divided into two main steps. The first step aims at finding the root causes of performance degradation in the simulation environment. The second step is aimed at designing process-aware sensors that are sensitive to the identified set of degradation root causes. Let us describe these two steps separately.

4.2.2 Step 1: Assessment of the root causes of the parametric performance variation

This first step in our methodology intends to find which of the components in the DUT have the most relevant contributions to the overall parametric performance variation. As mentioned above, in a generic mm-wave DUT, we can distinguish two families of components: standard components that are modeled in the PDK of the technology and full-custom passive components built by the designer in the BEOL of the technology. From now on, as it is usually the case in mm-wave designs, we assume that we have a netlist of the DUT where all custom-built components are replaced by a box of typical S-parameters obtained by electromagnetic simulation. The contributions of these two families of components to the DUT performance degradation are considered separately.

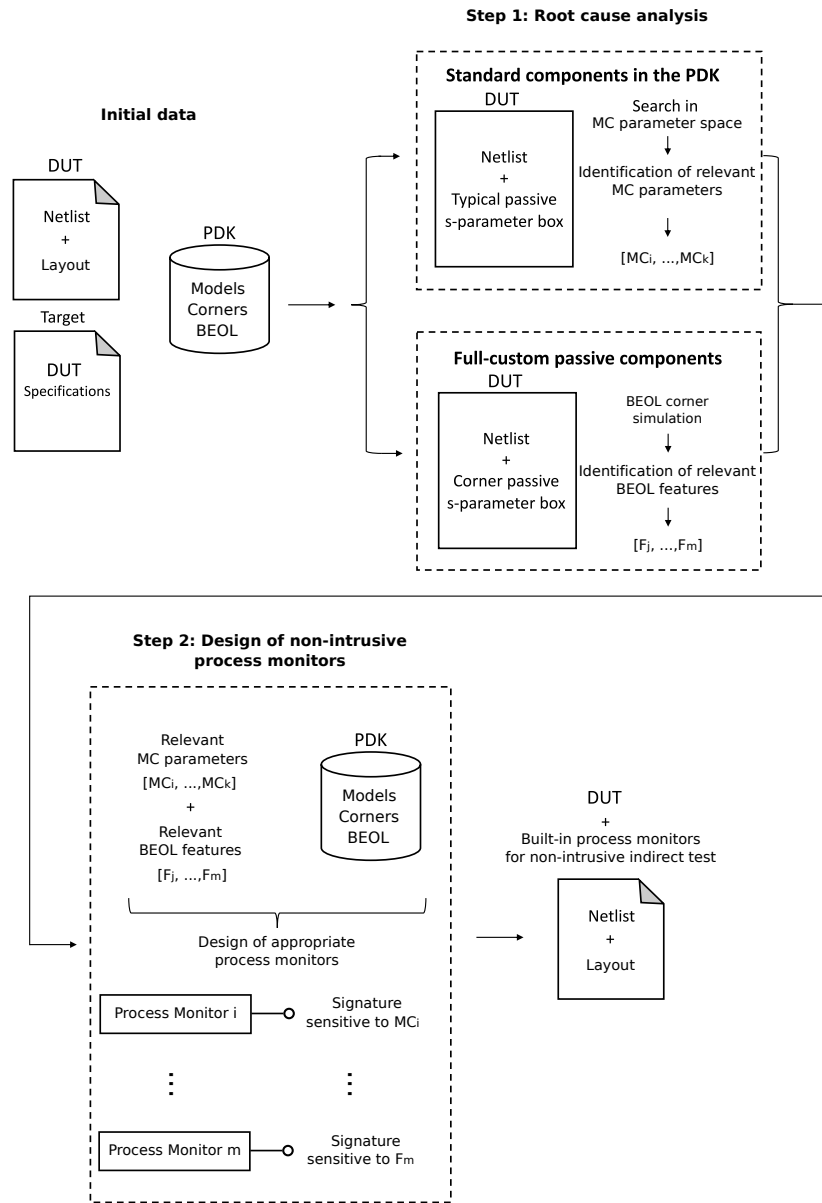


Figure 4.1: Conceptual block diagram of the proposed methodology for the generation of non-intrusive indirect test for a generic mm-wave integrated circuit.

Performance degradation due to standard components process variations

The performance degradation that is due to standard components process variations is typically simulated using MC sampling. Therefore, we propose to follow the methodology we have detailed in the last chapter (subsection 3.2.1) for finding the root-causes of the performance degradation, that is the list of MC process parameters that explain the main part of the parametric performance variation.

Performance degradation due to full-custom component parametric variations

Full-custom passive components are usually designed using electromagnetic simulation tools in the typical fabrication corner of the technology. In order to evaluate the degradation of the DUT performance due to process variations of full-custom components, we rely on the fabrication corner information available in the PDK. In this line, we assume that we have access to the information about fabrication corners for the different layers of the BEOL. This includes BEOL features such as thickness and dielectric constant variations for dielectric corners, and thickness and conductivity variations for metal layers corners. With this information, corner passive components can be generated using an electromagnetic simulation tool. The generated corner passive components can be then introduced in the DUT netlist and simulated in the electrical domain.

In order to determine which are the most relevant BEOL variability features for predicting the performance of the DUT, we propose to analyze the sensitivity of the DUT performance to each feature. In this line, we will introduce a scalar metric, $\Delta P(F)$, that will be used as a proxy of the sensitivity. The proposed $\Delta P(F)$ metric is defined as the normalized variation of the target performance P when we vary a given BEOL feature F between its minimum, typical and maximum technological corners while all other BEOL features are kept at their typical values, that is,

$$\Delta P(F) = \frac{|P(F_{typ}) - P(F_{min})| + |P(F_{typ}) - P(F_{max})|}{P(F_{typ})} \quad (4.2)$$

where $P(F_{typ})$, $P(F_{min})$, and $P(F_{max})$ correspond to the DUT performance obtained by electrical simulation when the BEOL feature F is set to its typical, minimum and maximum fabrication corner, respectively.

In order to further reduce the number of necessary simulations, it has to be noticed that this evaluation should be guided using expert design knowledge on the particular custom passive component. Typically, only a limited subset of BEOL features will have an impact on its performance, which may significantly reduce the simulation effort for this evaluation. Thus, analytical design equations that link the geometry of the passive component to its electromagnetic behavior may be used to quickly identify the subset of candidate BEOL features for which the ΔP metric should be computed. In this line, a variety of analytical models can be found in the literature for typical passive components in millimeter-wave circuits such as spiral inductors [68], slow-wave coplanar waveguides [69], MIM capacitors [70], microstrip lines [71, 72], interdigitated capacitors [73], etc.

4.2.3 Step 2: Design of non-intrusive process variation sensors

Once the previous step has been completed, we have identified a set of MC parameters and BEOL features that are relevant for explaining the degradation of the DUT performance due to process variations. The next step in the proposed methodology is to design simple tests that target the measurement of these parameters so that they can be used as signatures for predicting the DUT performance. In this line, we will rely on non-intrusive process monitor circuits sensitive to the identified parameters.

For the identified MC parameters, we propose to use the same approach as in the last chapter which is to make use of the information in the PDK of the technology to determine their physical

meaning. Together with the identified BEOL features, this information can be used to guide the design of appropriate non-intrusive process monitors. As in the last chapter, the designed process monitors and their associated signatures may be stored in a library for further reuse. This way, if we were to design a test program for a different DUT in the same technology, it would be possible to reuse many of the same process monitors and signatures already designed.

The final result of the proposed methodology is the set of non-intrusive sensors and their associated signatures that can be used for regressing the DUT performances.

4.3 Case Study

This section details an experimental case study in order to illustrate the proposed test generation methodology. We have taken advantage of the PA that was designed as a case study in the last chapter to extend our test methodology to circuits with custom passive components. In this sense, the major difference is that we did not use the Monte Carlo models of the transmission lines. However, since the circuit had already been sent to fabrication at the time we focused on this extension, the layout of the PA and the non-intrusive sensors is exactly the same. After measurements, the PA actually works at 65 GHz instead of 60 GHz. Thus, we tried to enhance the modeling of the circuit to obtain simulation results that were in better accordance with the measurements. The PA model modifications are discussed in the following paragraphs. Then, in subsections 4.3.2 and 4.3.3, the proposed methodology is applied to the PA. Simulations were performed using the corrected model.

4.3.1 Device Under Test

RF pad

As mentioned in subsection 3.3.2, we have modeled the RF pad and the probes by a parallel capacitor of 35 fF. We have performed many measurements using the same RF pads and probes, on different structures (PA, standalone passive components), between a few MHz and 145 GHz. Based on these measurements, we have settled on a model composed of a parallel capacitor of 7 fF. The model is slightly different with frequencies higher than 100 GHz for which a parallel resistor can be inserted to model the losses. The value of this extra resistor is a few $k\Omega$, depending on the measurement and the accuracy of the measured device model. For the PA that we have designed, the losses of the pad can be neglected. According to our analysis, this dramatic change for the pad model from 35 fF down to 7 fF could be explained by the use of high performance probes (50 μm GSG 145 GHz Infinity Probes form Cascade) that exhibit low parasitic coupling to allow high frequency measurements.

MOS source connection

The MOS source connection is of paramount importance and has a dramatic effect on the gain and input matching. For this design we have used an empirical value of 7 pF whereas according to the

measurements we should have used 3 pH. As a result, the input matching is degraded while the gain is noticeably increased.

Power supply voltage

The typical supply voltage for the transistor we have used is 1.2 V. Nevertheless, in order to account for the resistance between the external voltage source and the actual V_{dd} pad, we have decided to use $V_{dd} = 1.3$ V.

Transmission line T-junction

As mentioned in subsection 3.3.2, the T-junction model was too simple and required to be modified. Consequently, we performed Momentum simulations to create a more accurate model, whose schematic is displayed in Figure 4.2.b. As it can be seen, the microstrip lines in the T-junction model (schematic) do not match the real size of the implementation (layout). The line width, both for the layout and the model, corresponds to the $50\ \Omega$ width that is $7.7\ \mu\text{m}$. The length difference between the layout and the model is not relevant between ports 1 and 2, but it is much more important between ports 1 and 3 (or ports 2 and 3). Thus, the electrical lengths of the stubs slightly differ from the ones they were initially designed for. This model based on Momentum simulations provide results that are noticeably different from the ones obtained with the first model.

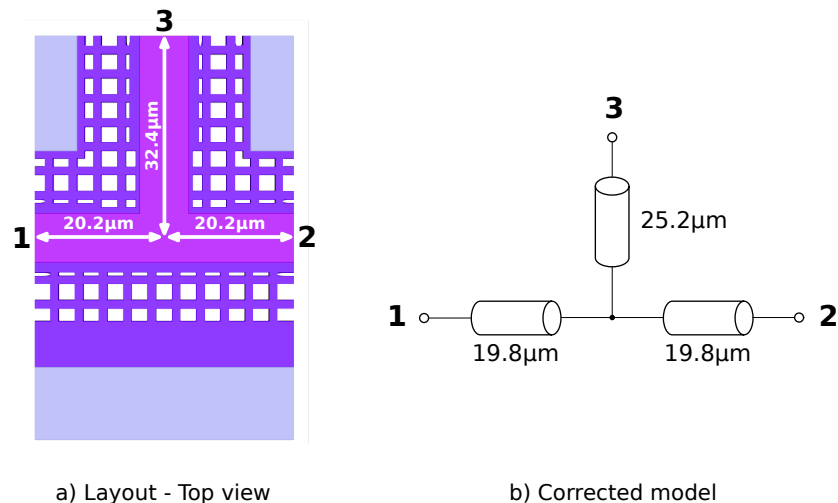


Figure 4.2: Top view and corrected model of the T-junction.

PA performances

Having taken into account the errors we had made during the initial design, we can obtain the performances of the amplifier with the corrected model. Figure 4.3 displays the S-parameters of the

PA with the initial model and compares it with the ones obtained with the corrected model, in dotted and solid lines respectively. We especially notice the shift in the working frequency, the higher gain value and the slightly degraded matching.

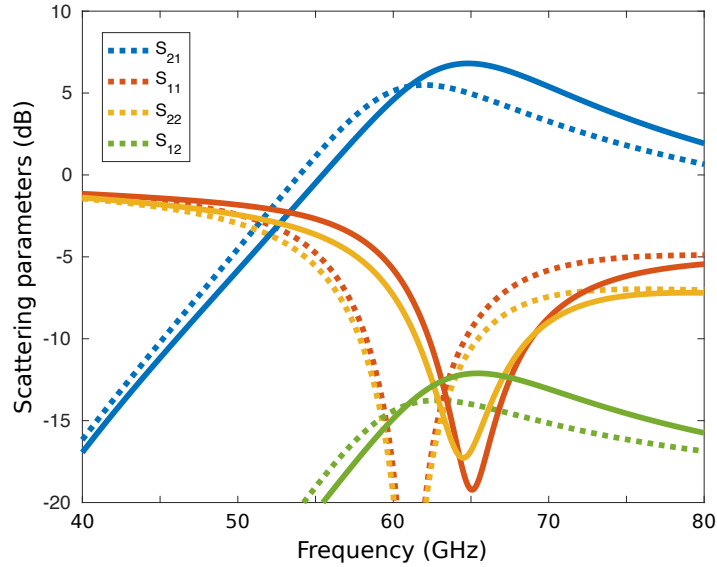


Figure 4.3: Expected S-parameters with the initial (dotted lines) and corrected (solid lines) models.

Finally, we can compare the corrected model to real measured data. In this line, Figure 4.4 displays the S-parameters obtained in measurements (solid lines) and with the new PA model (dotted line). Figure 4.5 displays the output power versus the input power, in simulation (dotted line) and measurement (solid line) while 4.6 presents the PAE curves in simulation (dotted line) and measurement (solid line).

Although the operating frequency and the general shape of the simulated S-parameters comply with the measurements, there is still a noticeable difference, especially with S_{11} . According to our analysis, this difference is due to an important process variation on the MOS transistor and a slightly inaccurate model provided by the PDK. This assumption is supported by two observations:

- a. The measured DC current is typically 22.5 mA whereas the simulated value is around 28 mA.
- b. We have fabricated a standalone MOS transistor and compared its performances with the model provided by the PDK. S-parameters measurements after TRL calibration do not comply with the model, even with an adequate hand-picked process corner selected to match the current consumption.

In the rest of this chapter, simulations will be performed using the corrected model of the PA and measured data will be provided to prove the feasibility of our method.

The goal of this case study is to design a complete non-intrusive indirect test protocol to assert the PA main specifications. That is, we target the design of non-intrusive process monitors and a set of associated simple measurements that allows the prediction of the PA main performances using a

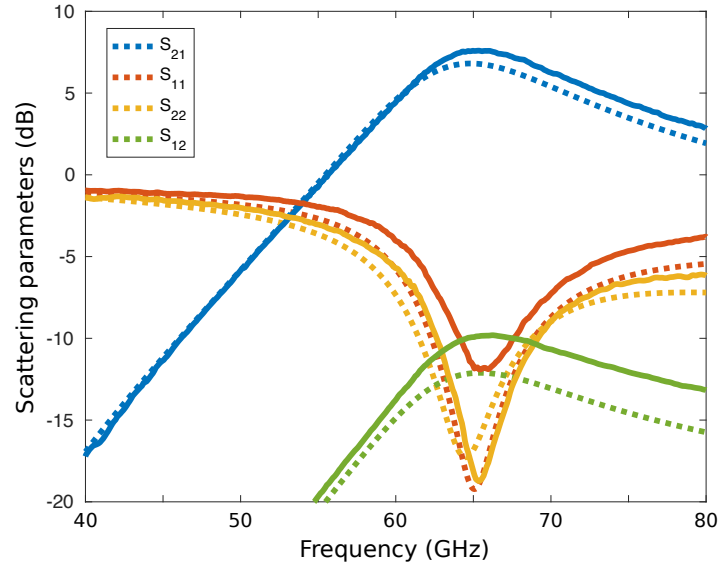


Figure 4.4: Expected S-parameters with the corrected model (dotted lines) and measured S-parameters (solid lines).

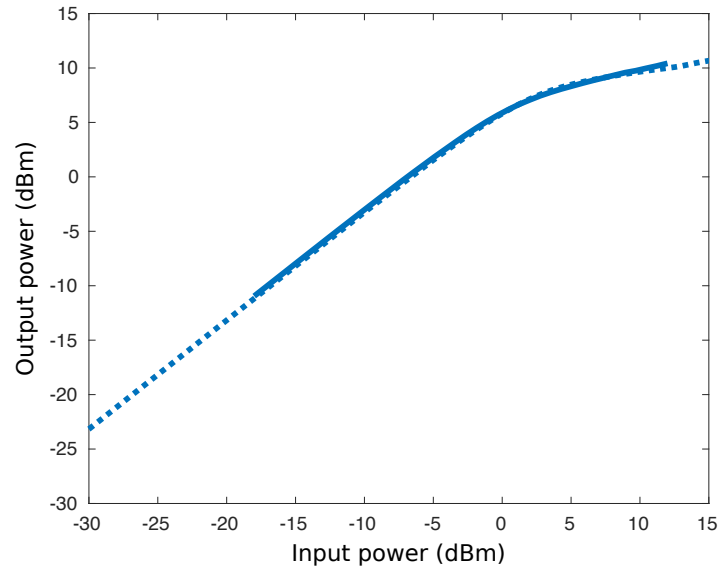


Figure 4.5: Expected output power with the corrected model (dotted lines) and measured output power (solid lines).

machine learning regression model. In the following subsections, the methodology is applied step by step.

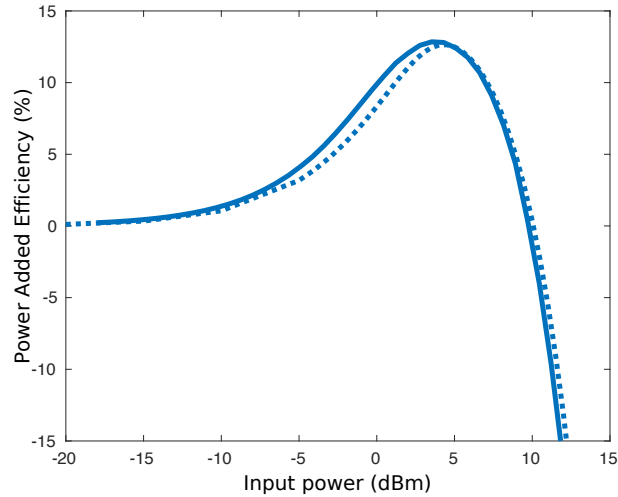


Figure 4.6: Expected PAE with the corrected model (dotted lines) and measured PAE (solid lines).

4.3.2 Step 1: Assessment of parametric performance variation root causes

As previously explained, at the design stage the methodology distinguishes degradation causes due to process variations of standard components modeled in the PDK of the technology and BEOL variations of full-custom passive components. As an application example, the method will be fully detailed using the small signal gain S_{21} as the target specification of the PA under test. The procedure can be then iterated for the rest of the performances.

Performance degradation due to standard components process variations

Firstly, all full-custom passive devices in the DUT netlist are replaced by S-parameter boxes, obtained by electromagnetic simulation in Keysight Momentum, describing their performance in the typical corner, as in Figure 4.7. A data set of 1000 instances of the resulting PA was then generated using latin hypercube sampling and the Monte Carlo models in the PDK of the technology. These instances were simulated using the SpectreRF simulator to extract the PA performances.

The MC simulation of the DUT in the 55 nm CMOS technology includes more than 500 independent process parameters. It is worth noticing that in the previous chapter we only used 154 MC process parameters, because we had first removed the parameters targeting specifically other components (bipolar transistors, MIM capacitors, etc.). In this study we chose not to perform this pre-filtering step and let it to the wrapper algorithm (Algorithm 1 in subsection 3.2.1). As described in the previous section, we explore the space of MC parameters to find the root causes of performance degradation. For performing this search, we rely on the proposed Brownian distance correlation-directed search as we did in Chapter 3.

As an application example, the procedure is fully detailed for the small signal gain S_{21} as the target

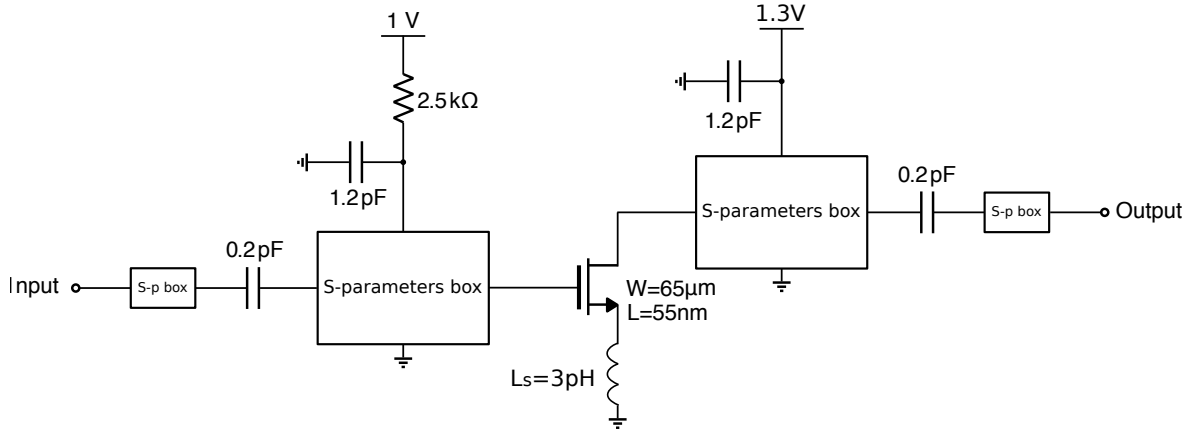


Figure 4.7: Transistor level schematic of the PA under test with typical S-parameters boxes for passive custom components.

specification of the PA under test. The procedure can be then iterated for the rest of the specifications. Thus, Figure 4.8 shows the RMS error and the maximum absolute error of the PA gain prediction in an independent set of PA instances, as a function of the number of MC process parameters selected by the search algorithm. As it can be seen, the proposed search identifies a set of 17 MC process parameters that allow the prediction of the PA gain with an error below 0.01 dB. For confidentiality reasons we cannot disclose the actual names of the MC parameters in the PDK. Instead, we provide a label that describes its physical meaning. Thus, parameters labelled as $poly_i$ refer to variations affecting to the polysilicon transistor gate, MOS_i refers to other MOS transistor features, and $capa_i$ refers to variations of MOM capacitor features, as in the previous chapter. Obviously, there is not anymore line parameters $line_i$ since in this study we did not use the MC model of the lines. In this sense, they can be regarded as full custom passive components and their process variations are tackled in the next subsection.

This process is then iterated for the rest of the PA specifications. The results are listed in Table 4.1. This table shows the sets of identified MC parameters that are relevant for the prediction of each PA specification, ranked by relevance. For each specification, the selection of MC parameters has been stopped when the prediction error does not improve significantly by incorporating additional parameters. As it can be seen, the degradation of the selected specifications due to process variations is dominated by a subset of only 17 different parameters.

Performance degradation due to full-custom component parametric variations

This step requires knowledge of the BEOL feature variations described in the PDK of the technology. In the selected technology, we remind that the BEOL is composed of 8 metal layers. Metal layers 1 to 5 are thin, metal layers 6 to 7 are thick and layer 8 is an ultra-thick metal layer. The full-custom transmission lines in the PA under test employ the last metal layer for the signal while the metal 1 layer provides the ground plane, as represented in Figure 4.9. Metal layers 2 to 7 are dummy metal

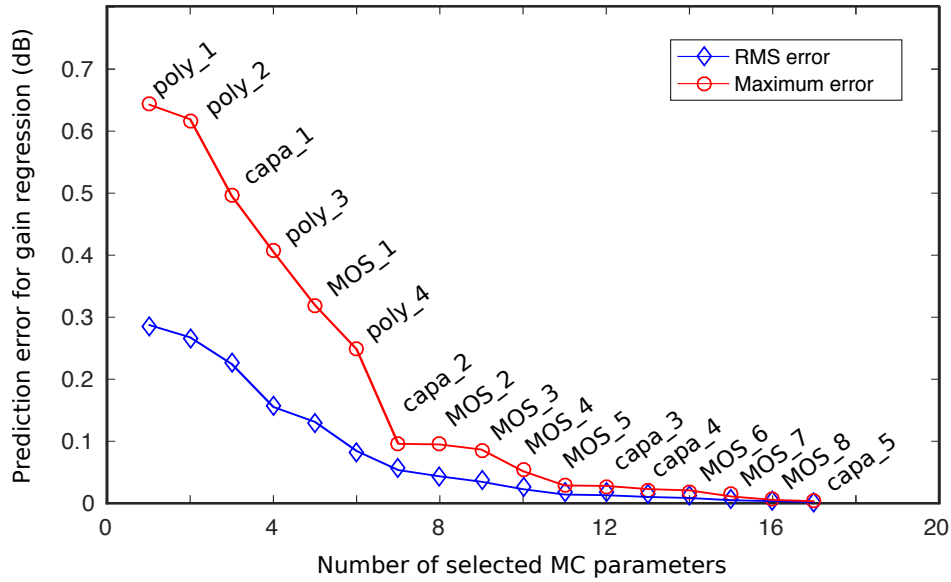


Figure 4.8: RMS and maximum error for the prediction of the PA small signal gain as a function of the number of selected MC parameters.

structures connected to ground that are needed for respecting the local metal density rules of the technology. As we have already mentioned, these lines are the same as the ones that we have used in the last chapter (provided in the PDK), except that they are simulated in Momentum.

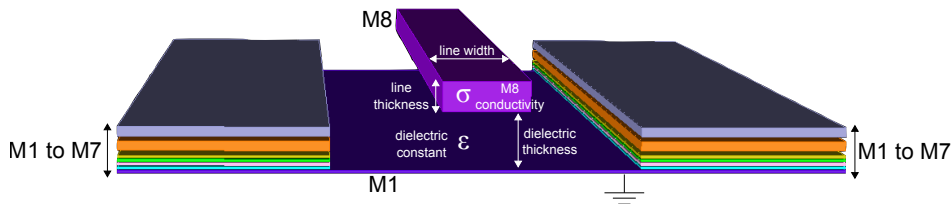


Figure 4.9: Full-custom microstrip line and relevant BEOL features.

Following the proposed methodology for analyzing the root causes of performance degradation related to full-custom components, we should vary the BEOL features (e.g., metal thickness, dielectric constant, metal conductivity, etc.) according to the PDK of the technology. Then we build S-parameter boxes for each corner of the passive components using Momentum, introduce these variations into the PA netlist and compute the associated ΔP metrics using the electrical simulator SpectreRF. However, the procedure can be greatly simplified by using some design knowledge on the custom components, as detailed in the following paragraph.

Table 4.1: Identified relevant MC parameters for the prediction of the PA specifications.

Specification	Identified Monte Carlo parameters (ranked by relevance)
Gain (S_{21})	$poly_1, poly_2, capa_1, poly_3, MOS_1, poly_4, capa_2, MOS_2, MOS_3, MOS_4, MOS_5, capa_3, capa_4, MOS_6, MOS_7, MOS_8$
S_{11}	$poly_1, poly_2, poly_3, capa_1, poly_4, MOS_1, capa_2, MOS_4, MOS_2, capa_4, MOS_7$
S_{22}	$poly_1, poly_3, poly_2, MOS_1, capa_1, MOS_2, poly_4, capa_2, MOS_4, capa_4, MOS_7, MOS_3, MOS_8, MOS_5, capa_3, MOS_6$
Power Added Efficiency, PAE	$poly_1, capa_1, poly_3, MOS_1, poly_2, capa_2, MOS_8, capa_3, MOS_5, MOS_6, capa_5, MOS_4, MOS_7$
DC current, I_{DC}	$poly_3, MOS_8, MOS_4, MOS_5, MOS_3$
Input referred 1 dB Compression point, ICP_{1dB}	$poly_1, poly_2, MOS_1, capa_1, poly_1, poly_3, MOS_2, MOS_4, capa_4, MOS_7, capa_2, capa_3, MOS_6, MOS_8, MOS_3$
Output referred 1 dB Compression point, OCP_{1dB}	$poly_1, capa_1, poly_3, capa_2, poly_2, MOS_3, MOS_4, capa_4, MOS_5, poly_4, MOS_7, capa_5, MOS_1, capa_3, MOS_8$
Saturation output power, P_{sat}	$capa_1, poly_1, poly_3, MOS_1, capa_2, poly_2, MOS_2, MOS_4, MOS_8, MOS_3, poly_4, MOS_5, capa_3, capa_4, MOS_6, MOS_7$

Simplifying the procedure for microstrip lines In our case, we can take advantage of the theoretical equations proposed in [71, 72, 74, 75] for microstrip lines design for simplifying the number of BEOL features that will be considered in the study. It should be noted that although these equations require some corrections for advanced silicon technologies, they still provide a consistent theoretical basis for predicting first-order behavior.

In [71] the author introduces an equation to compute the effective permittivity of the line dielectric ϵ_{eff} . This equation was then corrected in [72]. It takes into account that most of the electric field goes down into the dielectric, which relative permittivity is ϵ_r , but a small part part goes up into the air first, which relative permittivity is 1. The effective permittivity can be computed using the following equation [72], where w is the line width and h is the dielectric thickness:

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \times \left(1 + 12 \frac{h}{w}\right)^{-\frac{1}{2}} \quad (+0.04 \left(1 - \frac{w}{h}\right)^2 \text{ if } w/h \leq 1). \quad (4.3)$$

In [72] the author also introduces equations to compute the characteristic impedance Z_c of the line:

$$Z_c(w, h, \epsilon_{eff}) = \frac{\eta_0}{\sqrt{\epsilon_{eff}}} \times f(w, h), \quad (4.4)$$

where $\eta_0 = 120\pi$ is the free-space characteristic impedance and

$$f(w, h) = \frac{2\pi}{\ln\left(\frac{8h}{w} + \frac{w}{4h}\right)} \quad \text{if } w/h \leq 1 \quad (4.5)$$

$$f(w, h) = \frac{w}{h} + 1.393 + 0.667 \ln\left(\frac{w}{h} + 1.444\right) \quad \text{if } w/h > 1 \quad (4.6)$$

In order to take into account the strip thickness t , which is especially large compared to the dielectric thickness h in silicon technologies, the work in [71] proposes to replace the line width w by an equivalent width w_{eq} that accounts for the line thickness t . It can be computed using the following equations.

$$w_{eq} = w + \frac{t}{\pi h} \left[1 + \ln \left(\frac{4\pi w}{t} \right) \right] \quad \text{if } w/h \leq \frac{1}{2\pi} \quad (4.7)$$

$$w_{eq} = w + \frac{t}{\pi h} \left[1 + \ln \left(\frac{2h}{t} \right) \right] \quad \text{if } w/h > \frac{1}{2\pi} \quad (4.8)$$

Other equations to compute the effective dielectric constant, characteristic impedance of a microstrip line and take into account its thickness have been proposed is [74, 75]. The study of the transmission line losses have also lead to several works [71, 74, 75]. The accuracy of these different line models for silicon technologies are discussed in [76]. The equations in [77] propose a first order approximation of the conduction and dielectric losses, α_c and α_d respectively:

$$\alpha_c = \frac{\sqrt{\omega \mu_0 / (2\sigma)}}{Z_c w} (Np/m), \quad (4.9)$$

$$\alpha_d = \frac{\pi \epsilon_r (\epsilon_{eff} - 1) \tan \delta}{\lambda \epsilon_{eff} (\epsilon_r - 1)} (Np/m), \quad (4.10)$$

where ω is the angular frequency, μ_0 is the permeability of free space, λ is the wavelength in the line and δ is the skin depth, which can be computed using the following equation [77].

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \sigma}} \quad (4.11)$$

Note that by definition, $\lambda = \frac{\lambda_0}{\sqrt{\epsilon_{eff}}}$, where λ_0 is the free-space wavelength.

These equations show that microstrip line performances are highly dependent on the line width w , thickness t and conductivity σ , dielectric thickness h and dielectric constant ϵ_r . These parameters are schematically represented in Figure 4.9. Their variation ranges cannot be disclosed for confidentiality reasons. Nevertheless, we can mention that the dielectric constant and the line width are parameters with a small process variation that would not change noticeably the performance of this passive structure. Hence, they are excluded from the analysis and we will only consider three parameters: dielectric thickness, line thickness and M8 conductivity.

Compute the ΔP metric for each selected variation of the BEOL Consequently, in order to compute the ΔP metric for the selected BEOL features and the set of target specifications, we have built seven corner models of the microstrip lines using the Momentum electromagnetic simulator: typical corner, thin dielectric, thick dielectric, thin line, thick line, low- σ metal, high- σ metal. The main performances of these models are presented in Table 4.2. The PA under test was simulated using these corner models to infer the ΔP metrics associated to the BEOL features. The results are listed in Table 4.3. This table shows the sets of identified BEOL features that are relevant for the prediction of each PA specification, ranked by relevance, together with the obtained ΔP values, computed using Equation 4.2. According to the obtained results, the variations of the dielectric thickness, metal 8 thickness and metal 8 conductivity have a significant impact on the S-parameters and the PAE of the PA and a moderate impact on the linearity and saturation power. Concerning the current consumption, no significant impact is observed for any of the BEOL features.

Table 4.2: Transmission line main performances for the seven selected corners.

	Typical	Dielectric thickness		M8 thickness		M8 σ	
		min	max	min	max	min	max
Characteristic impedance Z_c (Ω)	45.9	43.8	47.7	46.8	45.2	46.0	45.8
Loss constant α (dB/mm)	0.67	0.75	0.64	0.64	0.69	0.71	0.64
Propagation constant β (deg/mm)	131.3	132.2	130.8	131.1	131.5	131.5	131.0

Table 4.3: Identified relevant BEOL features for the prediction of the PA specifications.

Specification	Identified BEOL features ($\Delta P(F)$)
Gain (S_{21})	dielectric thickness (1.09), metal 8 thickness (0.25), metal 8 σ (0.25)
S_{11}	dielectric thickness (3.25), metal 8 σ (1.25), metal 8 thickness (1.02)
S_{22}	dielectric thickness (5.29), metal 8 thickness (1.29), metal 8 σ (0.36)
Power Added Efficiency, PAE	dielectric thickness (2.62), metal 8 thickness (0.79), metal 8 σ (0.54)
DC current, I_{DC}	–
Input referred 1 dB Compression point, ICP_{1dB}	dielectric thickness (0.54), metal 8 σ (0.33), metal 8 thickness (0.07)
Output referred 1 dB Compression point, $OC P_{1dB}$	dielectric thickness (0.54), metal 8 thickness (0.24), metal 8 σ (0.08)
Saturation output power, P_{sat}	dielectric thickness (0.81), metal 8 thickness (0.27), metal 8 σ (0.04)

4.3.3 Step 2: Design of non-intrusive process variation sensors

As a result of the previous step, we have identified two subsets of parameters related to process variations of standard and full-custom components, respectively, that are deemed relevant for explaining the degradation of the DUT performance. The following step in our methodology consists in designing a set of non-intrusive process monitors that generate simple signatures that are strongly correlated to the identified parameters. The design of the non-intrusive process monitors and their associated signatures is guided by the physical meaning of the identified parameters. As mentioned in the introduction of this chapter, the PA and the sensors had already been sent to fabrication when we focused on this study. Since the model that has been used in simulation for this study is close to the one we

have used in Chapter 3, we expect to obtain the same sensors and signatures for components from the PDK. In Chapter 3, two microstrip sensors had been designed and implemented. Following this extended method for circuits with custom passives components, we wish to find signatures on the very same sensors.

Table 4.4 lists the set of non-intrusive sensors together with their associated signatures and the process parameters they targeted. It must be noticed that sensors and signatures that cover MC parameters are exactly the same as in Chapter 3. Following the identification of relevant BEOL features (Table 4.3) we have proposed two signatures on the existing microstrip line sensors. Signature S_1 is a resistance measurement on the thin line, which is the same as signature S_4 in Chapter 3. Signature S_2 is a phase measurement on the wide line and should be especially sensitive to the dielectric thickness. It is worth noticing that this second signature is different from the signature S_8 proposed in Chapter 3 (obtained by leveraging MC models) but its role is basically the same, i.e. sensing the dielectric thickness. Based on the BEOL feature study, we chose not to perform any S-parameters measurements on the thin line. Moreover it is noteworthy that we chose to perform many measurements at 1 GHz instead of 100 MHz to achieve a better resolution.

4.4 Experimental results

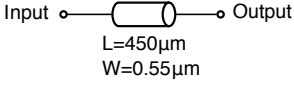
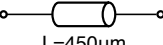
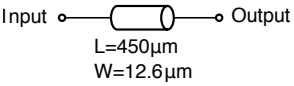
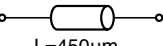
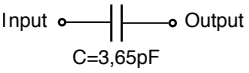

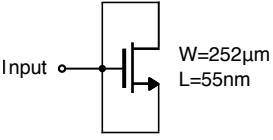
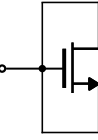
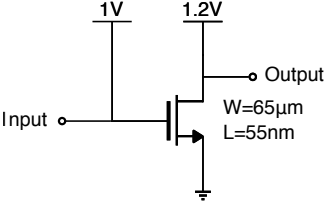
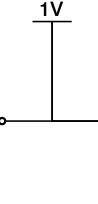
In order to verify the feasibility and performance of the developed non-intrusive indirect test program, we have performed measurements on fabricated samples of the PA and its associated sensors. The goal of our validation will be to show the equivalence between the direct measurement of the PA specifications using mm-wave equipment, and the proposed indirect measurement based on predicting the PA specifications from the set of developed signatures using a machine learning model.

Figure 4.10 shows a microphotograph of the fabricated prototype, including the PA under test, the non-intrusive process monitors and a set of de-embedding structures. The PA under test is labelled as PA, a corresponds to the thin transmission line, b to the wide transmission line, c to the MOM capacitor, d to the NMOS junction capacitance sensor and e to the NMOS transistor. Structure e_2 is another NMOS transistor, designed to extract the gate resistance, but in the end not used in this study. Devices T, R, L and S are de-embedding structures, thru, reflect (open), line and short respectively. They were used to de-embed the sensors access for verification and design debugging and they are not actually needed for this study.

Although it may seem that the area overhead of the proposed test structures makes this approach unfeasible, it is convenient to remark that the developed prototype has been conceived as a proof-of-concept case study and as such, dedicated pads for probe test have been added to each non-intrusive sensor. In an actual system-level implementation, the non-intrusive sensors will be placed in the unused space within a transceiver and the access pads can be easily multiplexed to save area. Indeed, the active area of the sensors (highlighted in blue in Figure 4.10) represents only 0.0185 mm^2 , dominated mainly by the wide transmission line process monitor structure, while the PA under test occupies an area of 0.1430 mm^2 excluding pads.

In total, we have received 21 fabricated samples of the prototype PA that were characterized in the

Table 4.4: Proposed library of non-intrusive sensors and signatures for gain prediction.

Process variations sensor	Signature description	Process parameters covered
<p>Thin microstrip line</p>  <p>Input  Output L=450μm W=0.55μm</p>	<ul style="list-style-type: none"> • S_1: Resistance 	<ul style="list-style-type: none"> • Metal 8 thickness, Metal 8 σ
<p>Wide microstrip line</p>  <p>Input  Output L=450μm W=12.6μm</p>	<ul style="list-style-type: none"> • S_2: Phase of S_{21} @ 1 GHz 	<ul style="list-style-type: none"> • Dielectric thickness, Metal 8 thickness, Metal 8 σ
<p>MOM Capacitor</p>  <p>Input  Output C=3.65pF</p>	<ul style="list-style-type: none"> • S_3: Real part of Y_{11} @ 1 GHz • S_4: Imaginary part of S_{21} @ 1 GHz 	<ul style="list-style-type: none"> • $capa_1, capa_4, capa_2, capa_5, capa_3$ • $capa_2, capa_4, capa_5$
<p>Junction capacitance sensor</p>  <p>Input  Output W=252μm L=55nm</p>	<ul style="list-style-type: none"> • S_5: Imaginary part of S_{11} @ 1 GHz 	<ul style="list-style-type: none"> • $MOS_1, MOS_2, poly_3, MOS_8, MOS_6, MOS_7$
<p>NMOS transistor</p>  <p>Input  Output 1V 1.2V W=65μm L=55nm</p>	<ul style="list-style-type: none"> • S_6: Gate resistance • S_7: Bias current 	<ul style="list-style-type: none"> • $poly_1, poly_2, poly_3, poly_4, MOS_7$ • $poly_3, MOS_8, MOS_4, MOS_5, MOS_7, MOS_3$

laboratory. The test setup is detailed in the next subsection.

4.4.1 Measurement setup

Measurements were performed using Anritsu ME7838D VNA, Cascade S300 semi-automatic probe station and 50 μ m GSG 145 GHz Infinity Probes from Cascade. Calibration was performed using the LRRM method on a Cascade impedance standard substrate P/N 138-356. DC supply was provided by a Keysight B2902A unit, through the VNA bias tee or using 100 μ m SGS multi-contact Wedge Picoprobes from GGB. For large signal measurements, a 55-65 GHz PA (SP6010-30-20W from Spacek Labs) was used to drive the input of the PA under test. Power calibration was performed using Agilent

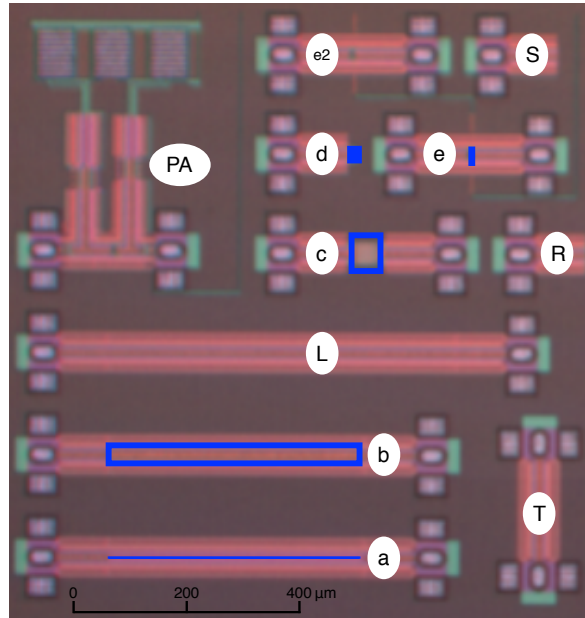


Figure 4.10: Microphotograph of the PA under test with the proposed non-intrusive process monitors and de-embedding structures.

V8486A power sensor, connected to a Helwett-Packard 437B power-meter.

The signatures associated to each non-intrusive process monitor were extracted using $50\ \mu\text{m}$ GSG 145 GHz Infinity Probes from Cascade on the same VNA and probe station for AC signatures. For DC signatures, the operating point of the monitors was provided through the VNA bias tee.

Figure 4.11 displays the general diagram of the measurement setup while Figure 4.12 shows a photography of the setup for performing large signal measurements. The VNA is in the background of the orange dashed rectangle and the PA driver is in the green rectangle. For small signal measurements, the PA driver was removed. Figure 4.13 shows a photography of the PA being measured, seen through the *eVue* camera (in pink in Figure 4.12).

4.4.2 Fabricated set of chips

22 samples of the PA and its associated sensors were fabricated. Figure 4.14 shows a photography of the piece of wafer on which we performed the measurements, together with the circuit mapping to the wafer reticle. Red dots correspond to the circuits position.

4.4.3 Signature measurement

The signatures associated to each non-intrusive process monitor were extracted using $50\ \mu\text{m}$ GSG 145 GHz Infinity Probes from Cascade on the same VNA and probe station for AC signatures. For DC

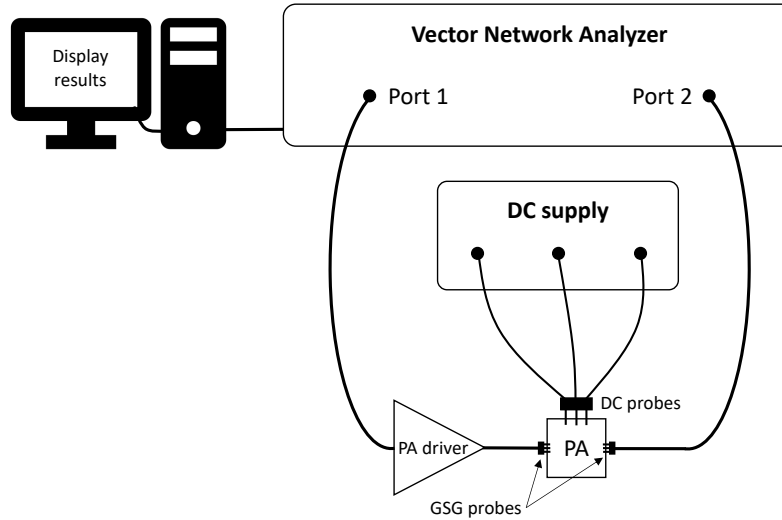


Figure 4.11: General diagram of the measurement setup.

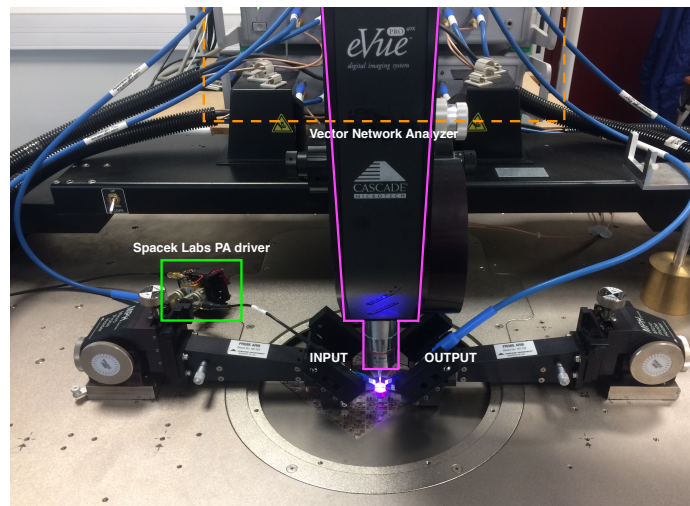


Figure 4.12: Photography of the test bench for large signal measurements.

signatures, the operating point of the monitor circuits was provided through the VNA bias tee.

For the thin microstrip line sensor, resistance (S_1) was calculated using Ohm's law with a 1 mV voltage difference across the line. Small signal measurements were performed at 1 GHz on the wide microstrip line, the MOS junction capacitor and the MOM capacitor (S_2 , S_3 , S_4 and S_5). It should be noted that the measurement of S_3 shows a very poor repeatability due to a high dependency on the probe placement in our test setup. Due to this experimental limitation we will not use signature S_3 in the proposed indirect test. Fortunately, most of the information present in S_3 is redundant with the

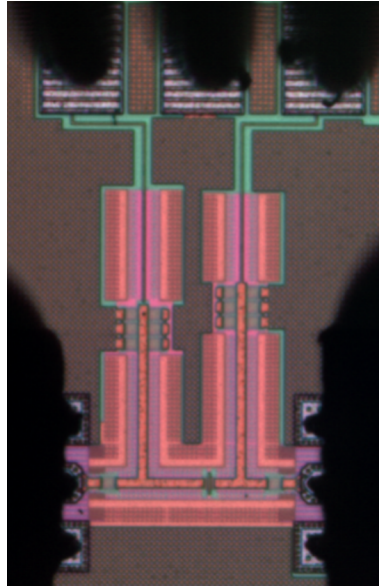


Figure 4.13: Photography of the PA being measured.

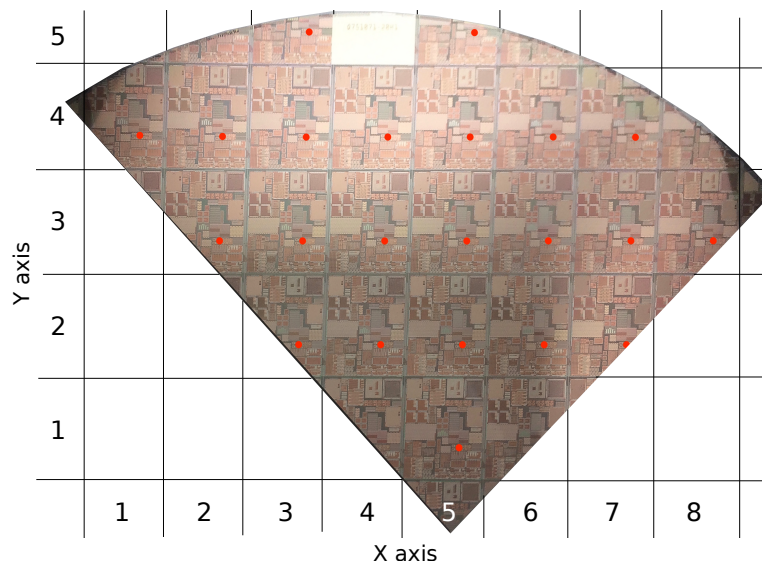


Figure 4.14: Photography of the wafer with the circuits position (red dots) and mapping to the wafer reticle.

information in S_4 , which should mitigate, to some extent, the effect of removing this signature.

Gate resistance measurement

Gate resistance process variations (signature S_6 in Table 4.4) were inferred from small signal measurements using the following expression [78],

$$S_6 = \frac{\text{Real}(Y_{11})}{\text{Imag}(Y_{11})^2}. \quad (4.12)$$

where admittance parameter Y_{11} refers to the input node of the NMOS transistor process monitor in Table 4.4. Note that it is possible to extract the gate resistance at much lower frequencies, 1 GHz for instance, in a simulation environment. In this sense, Figure 4.15 compares the simulation and the real measurement (on chip (5 ; 3) according to the wafer mapping introduced in Figure 4.14). The gate resistance is of course independent of the frequency and is directly obtained in simulation by looking into the device parameter (yellow dashed line). The value extracted thanks to Equation 4.12 (in red) is very close to the real value, even in low frequencies. However, this expected result was not obtained in measurements, as clearly shown by the blue curve. As a matter of fact, due to the smallness of $\text{Real}(Y_{11})$ and $\text{Imag}(Y_{11})$ in low frequencies and a little measurement noise, it is far more difficult in practical measurements. Taking in account this phenomenon, we should have designed differently the gate resistance sensor.

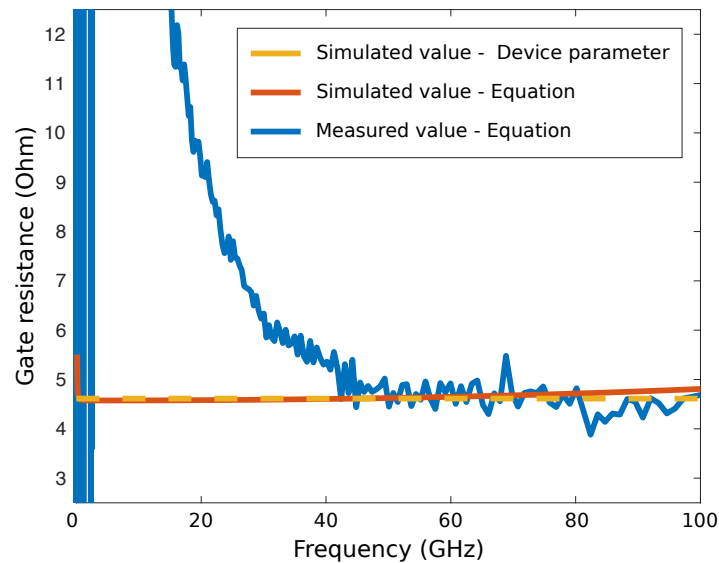


Figure 4.15: Gate resistance extraction - Simulation versus measurement.

The actual gate resistance value can be obtained in measurements between 40 GHz and 80 GHz. However, the information we care about is the process variation, that is the relative difference between the different chips. For this study, we have decided to extract the gate resistance using the proposed equation by averaging the measured data between 14 GHz and 17 GHz. This way, we do not rigorously get the gate resistance values for the different chips, but instead values with a strong correlation to the real gate resistances. This means that a gate resistance of 4.5Ω , measured between 14 GHz and 17 GHz, will give a value around 13Ω . Although the absolute value is false, the information about

the relative difference between the different samples is still available. This is the lowest frequency range in which we manage to find the information about gate resistance process variation. Although the measurement frequency is higher than expected, the relevant information from this sensor has yet been extracted.

Signature measurement results

The 7 signatures presented in Table 4.4 have been measured for the 22 available chips. As we have already mentioned, signature S_3 will not be used in this study due to a poor repeatability of the measure over the available samples. Figures 4.16 and 4.17 display the measured signatures S_1 , S_2 , S_4 , S_5 , S_6 and S_7 . For each signature, it is relevant to notice the performance gradients and the outliers on the wafer edge. These observations are the direct consequences of the process parameter gradients over the wafer and the side effects. It is also worth noticing that on chip (5 ; 5) both the PA and the MOS transistor sensor were non functional. Consequently, for the rest of this study we only consider the 21 functional samples.

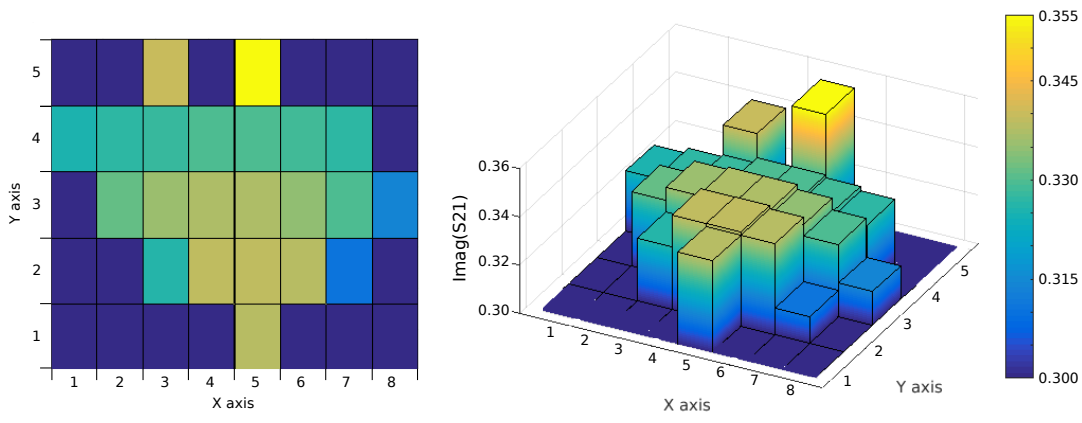
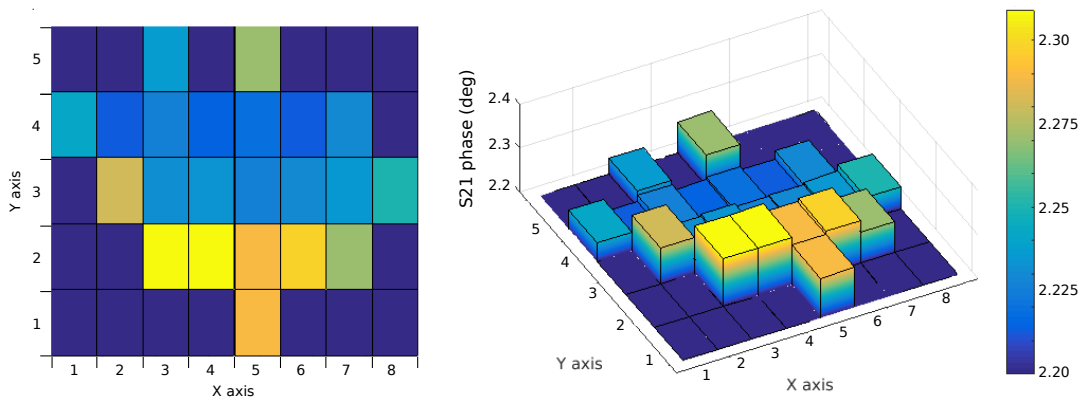
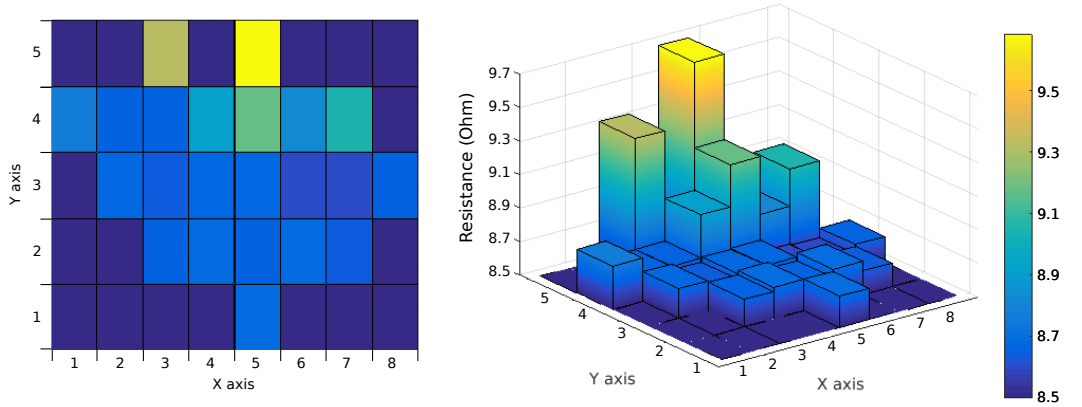
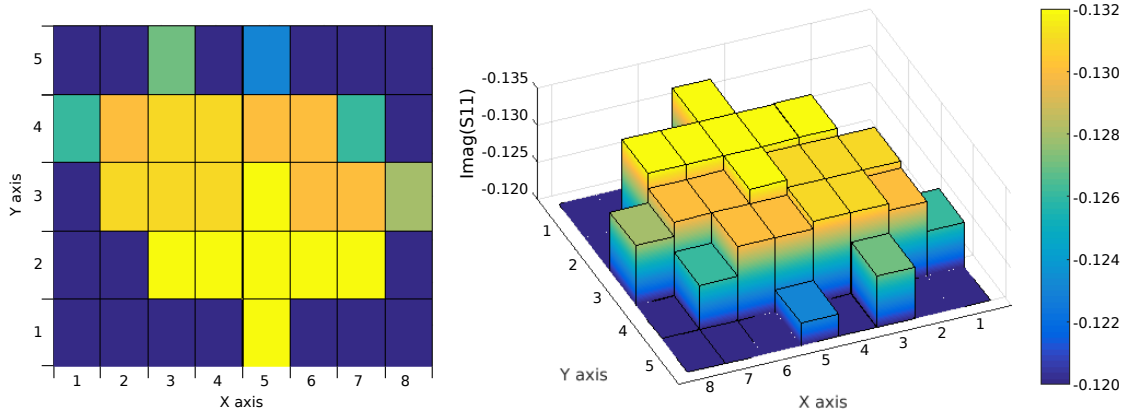
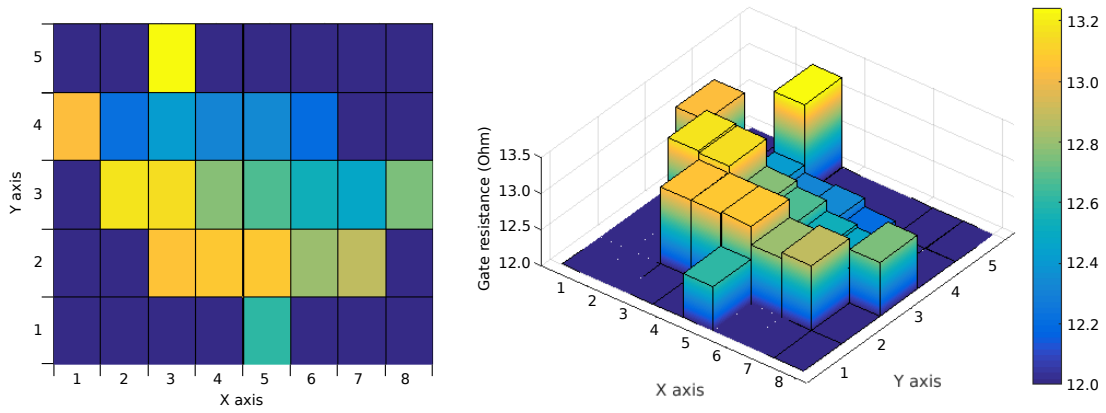


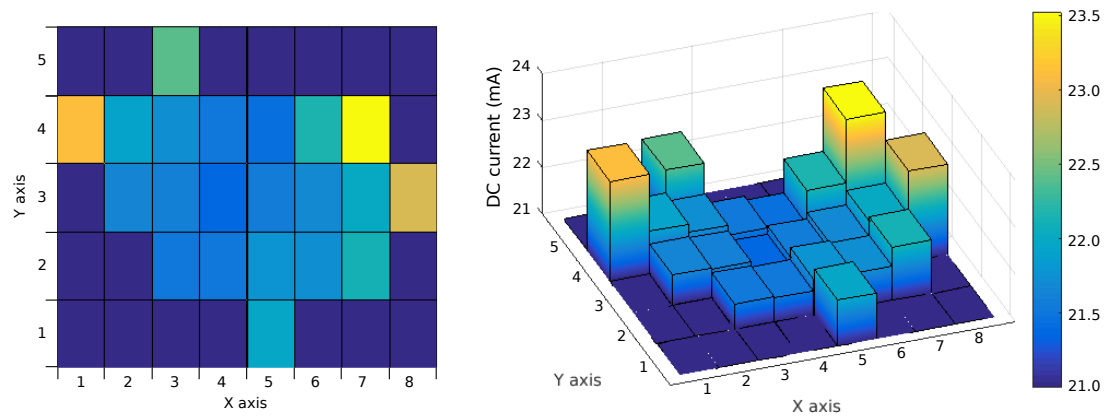
Figure 4.16: Measurement results on the thin lines, wide lines and MOM capacitors - Signatures S_1 , S_2 and S_4 .



Signature S5 - MOS junction capacitance sensor - S11 imaginary part @ 1 GHz



Signature S6 - MOS transistor - Gate resistance @ 14-17 GHz (Ohm)



Signature S7 - MOS transistor - DC current (mA)

Figure 4.17: Measurement results on the MOS junction capacitance sensors and MOS transistors - Signatures S_5 , S_6 and S_7 .

4.4.4 PA measurement results

Out of the 22 available chips, 21 were fully functional (both PA and sensors). Figures 4.18, 4.19 and 4.20 display the performance variation across the wafer for each specification of the PA. It is worth noticing that for the PAE, we only consider 19 samples out of the 21 available ones since the measurement failed on 2 of them. We could have performed the measurements another time but the results would not have been accurate due the important degradation of the RF pads after several characterizations.

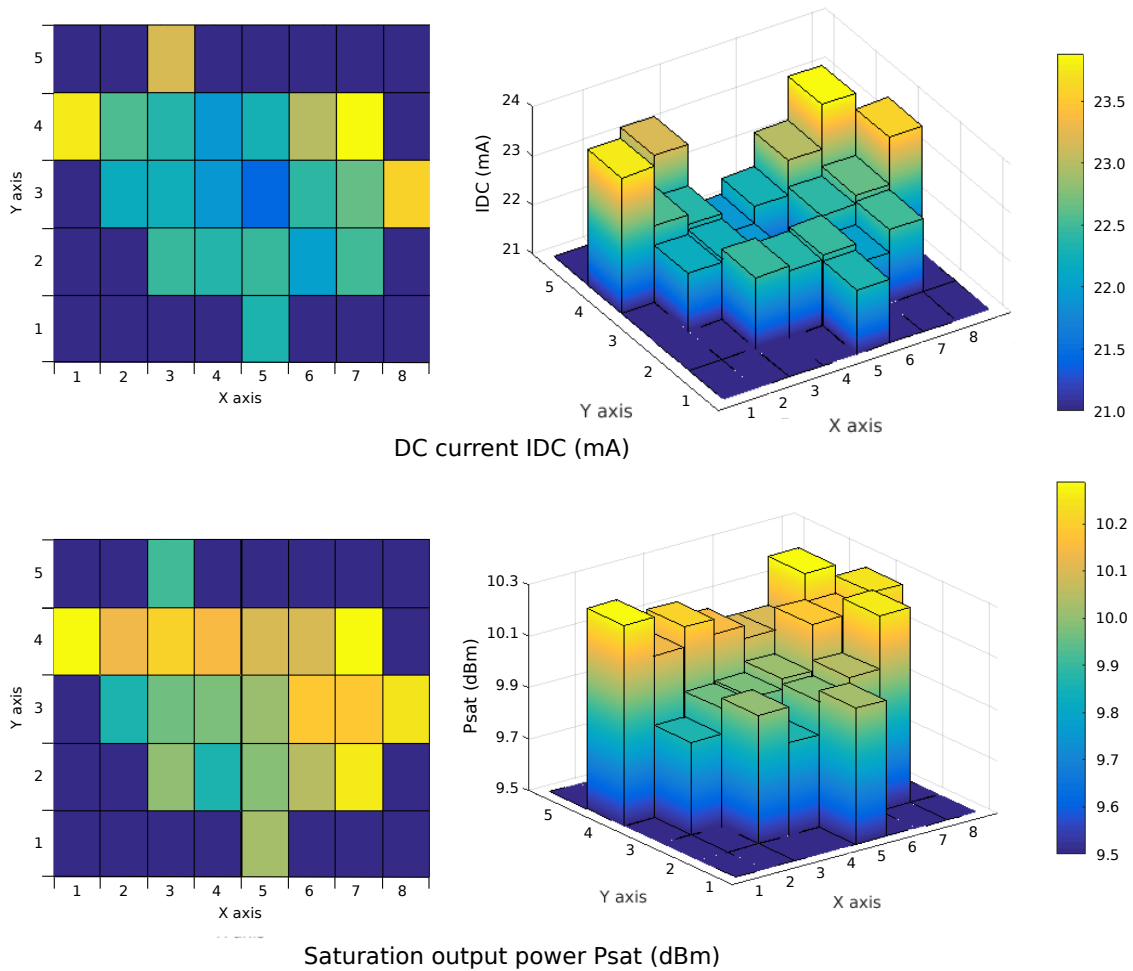
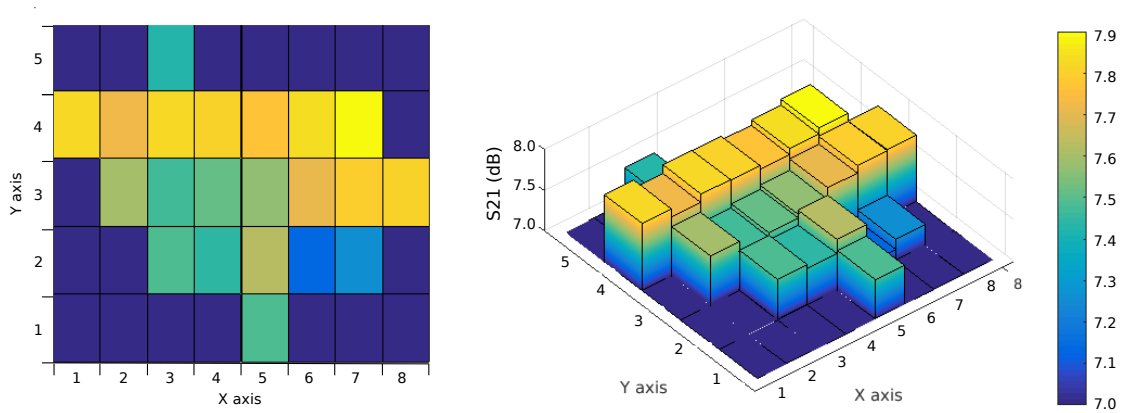
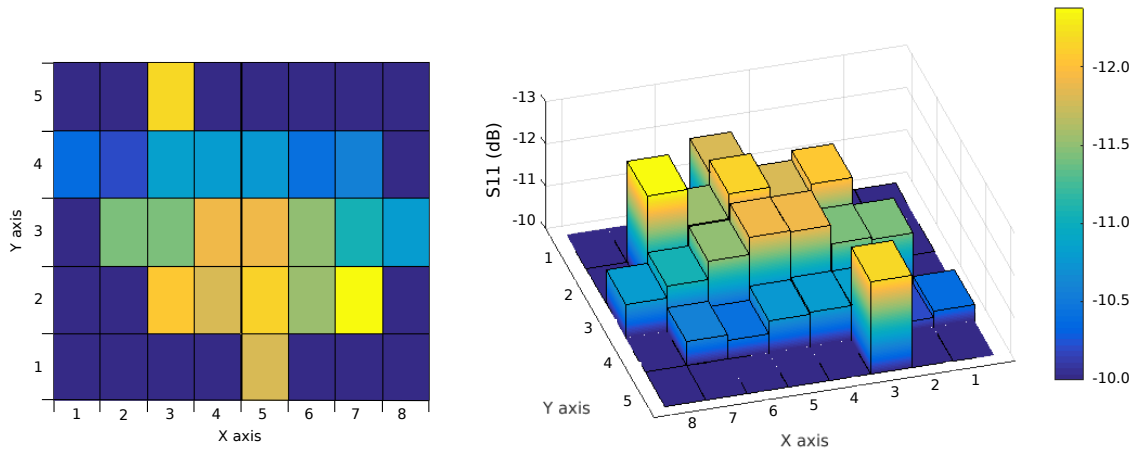


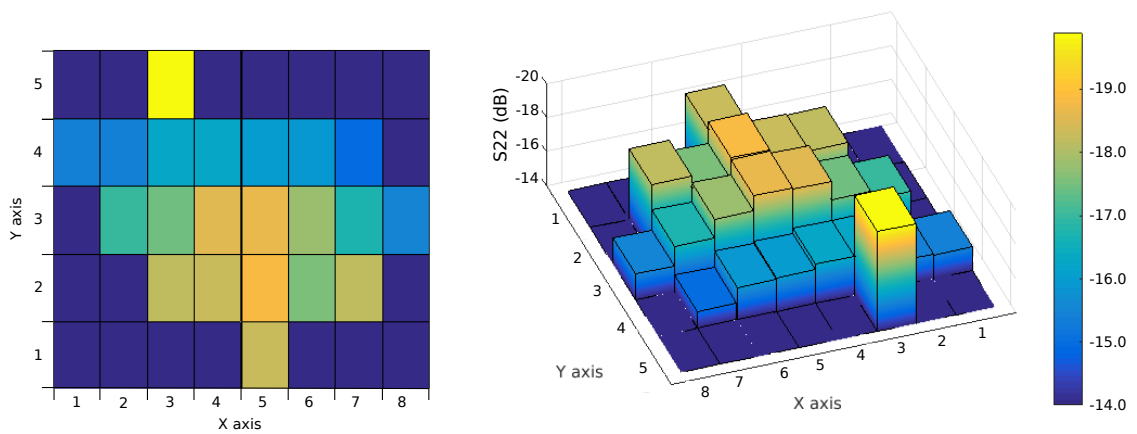
Figure 4.18: Measurement results of the PA performance - I_{DC} , P_{sat} .



Small signal gain S_{21} (dB)



Input small signal matching S_{11} (dB)



Output small signal matching S_{22} (dB)

Figure 4.19: Measurement results of the PA performance - S_{21} , S_{11} and S_{22} .

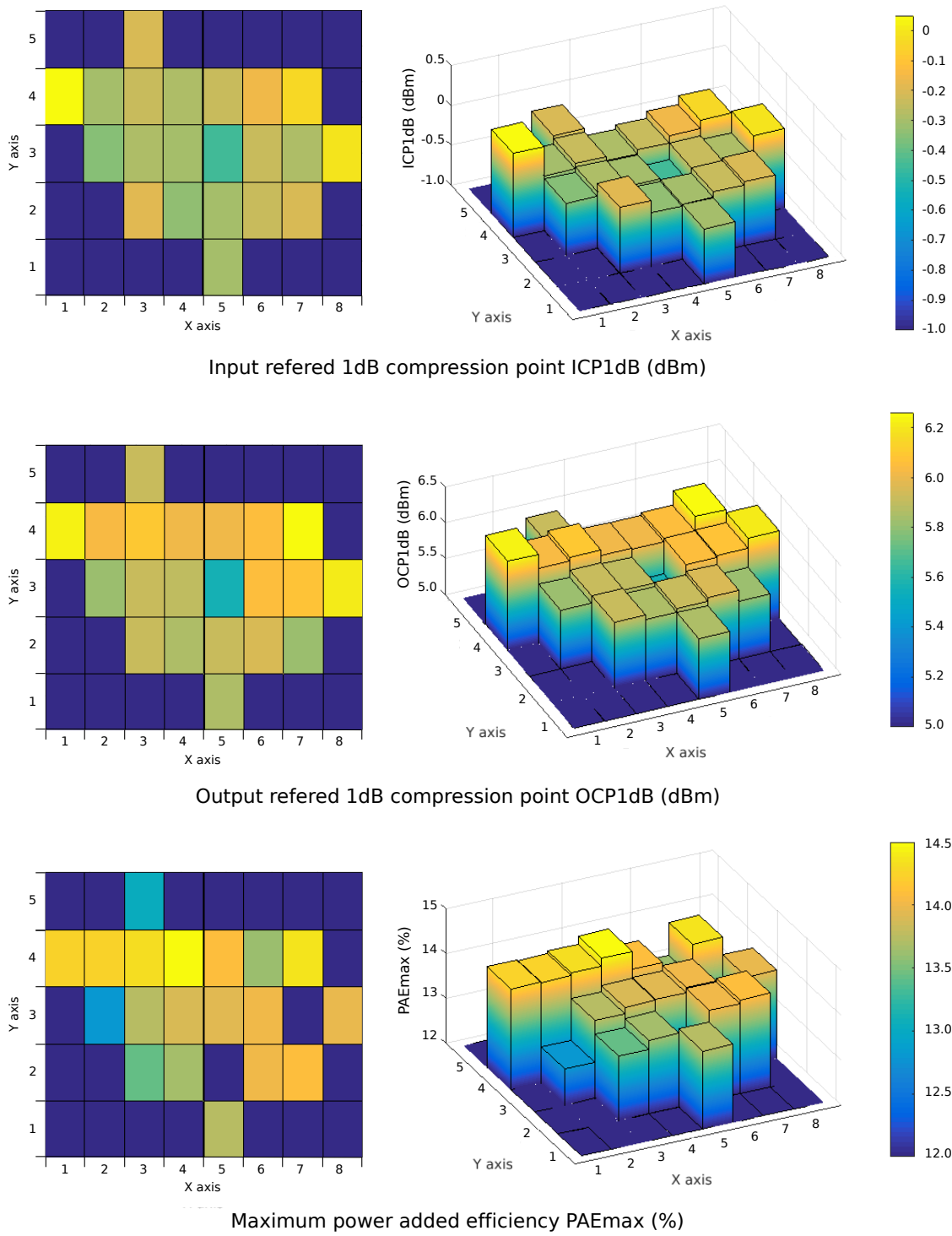


Figure 4.20: Measurement results of the PA performance - ICP_{1dB} , OCP_{1dB} , PAE_{max} .

4.4.5 General results

The goal of the proposed indirect test is to infer the PA specifications from the extracted signatures. In other words it means we need to build a function $f_{S \rightarrow P}$ as in

$$f_{S \rightarrow P} : [S_1, S_2, S_4, S_5, S_6, S_7] \rightarrow [S_{11}, S_{22}, S_{21}, I_{DC}, ICP_{1dB}, OCP_{1dB}, PAE_{max}, P_{sat}]. \quad (4.13)$$

As it was presented in the introduction, mapping the signatures to the target specifications using a machine learning algorithm requires the training of the regression model $f_{S \rightarrow P}$. Typically, this training will be performed during the pre-production phase of the DUT on a data sample containing a few hundred (or thousand) devices. In our case study, we do not have access to such fabrication volume. Hence, special techniques are used to train the machine learning algorithm with our limited set of 21 circuit instances. In this line, we use the well-known leave-one-out cross-validation technique to partition the data set. One instance of the prototype circuit is left apart as independent verification set, while the 20 remaining circuits are used to train the machine learning algorithm. The process is then repeated changing the left-out circuit.

Figure 4.21 shows scatterplots of the obtained results for each of the considered specifications of the PA under test. The vertical axis represents the prediction of the PA under test specifications obtained from the proposed set of low-cost signatures using a perceptron Neural Network implemented in Matlab. The horizontal axis represents the actual specification of the PA under test, obtained by standard mm-wave measurements. The dashed red line represents the perfect 1:1 regression line for visual reference. Each point in the scatterplot corresponds to an independent verification point in the leave-one-out cross-validation scheme. Points highlighted in red corresponds to outlier devices in the distribution of specifications and signatures. Given that machine learning models, as it is well-known, do not have good extrapolation capabilities, the predictions for these outlier devices are expected to be worse than for the rest of centered devices. As it can be seen, even with our limited number of training devices, we observe a significant correlation between the predictions and the actual specifications for all the considered PA performances. Even for the extreme devices highlighted in red, the prediction error remains contained. In order to evaluate the quality of the overall regression, Figure 4.21 includes the RMS prediction error, ϵ_{RMS} , for each considered specification. The obtained RMS prediction error values show a very good accuracy for the prediction of all the considered PA specifications. It has to be noted that the error ϵ_{RMS} was computed on the samples corresponding to the blue circles only since the red ones are outlier devices.

Table 4.5 offers additional insight into the obtained results. This table shows a direct comparison between the standard deviation (σ) in the functional measurement of all the considered specifications of the PA under test, and the RMS prediction error obtained in the prediction of each specification from the proposed set of low-cost signatures. In this line, we define a Figure of Merit (FoM) as the ratio between the standard deviation of the functional measurement $\sigma_{performance}$ and the RMS prediction error μ_{error} , for each specification, as in

$$FoM = \frac{\sigma_{performance}}{\mu_{error}}. \quad (4.14)$$

Thus, a FoM larger than one indicates that the regression model predicts the specification with a better accuracy than the standard deviation of the specification in the considered set of devices. As it can be observed, the RMS prediction error is significantly smaller than the specification standard deviation for all the considered specifications which leads to high values of FoM. This is a strong indication of the accuracy of the proposed indirect test, even more noteworthy if we consider the very limited number of samples at our disposal for this case study.

Table 4.5: Measured performances of the PA and prediction results.

PA performance	Unit	Functional measurement			RMS prediction error	FoM
		Minimum	Maximum	Std. dev. σ		
DC current, I_{DC}	mA	21.1	24.0	0.64	0.18	3.58
Gain, S_{21}	dB	7.13	7.91	0.21	0.13	1.62
S_{11}	dB	-12.4	-10.2	0.67	0.27	2.47
S_{22}	dB	-19.9	-15.0	1.38	0.61	2.27
Saturation output power, P_{sat}	dBm	9.86	10.29	0.14	0.08	1.71
Input referred 1 dB Compression point, ICP_{1dB}	dBm	-0.45	0.05	0.12	0.09	1.38
Output referred 1 dB Compression point, $OC P_{1dB}$	dBm	5.57	6.26	0.16	0.14	1.18
Maximum Power Added Efficiency, PAE_{max}	%	12.82	14.50	0.44	0.37	1.19

4.5 Conclusions

We have presented a systematic methodology for designing non-intrusive built-in indirect test strategies for mm-wave circuits. The proposed methodology targets the test of generic mm-wave ICs containing both standard devices modeled in the PDK of the technology and full-custom passive components built in the BEOL of the technology. For a given DUT, the proposed test generation strategy is based on analyzing the root-causes of parametric performance degradation. This analysis guides the design of a set of non-intrusive process monitors with associated low cost signatures sensitive to the identified degradation root-causes. The generated indirect test is completed with a machine learning model that allows inferring the DUT specifications from the set of low-cost signatures.

The feasibility of the proposed technique has been experimentally validated on a 65 GHz PA fabricated in ST Microelectronics 55 nm CMOS technology. A set of non-intrusive process monitors has been designed according to the proposed methodology and integrated together with the PA under test. Experimental measurements on a set of 21 fabricated PAs show a very good agreement between the direct functional measurement of the PA specifications using dedicated mm-wave test equipment and the proposed indirect test based on low-cost low-frequency non-intrusive signatures.

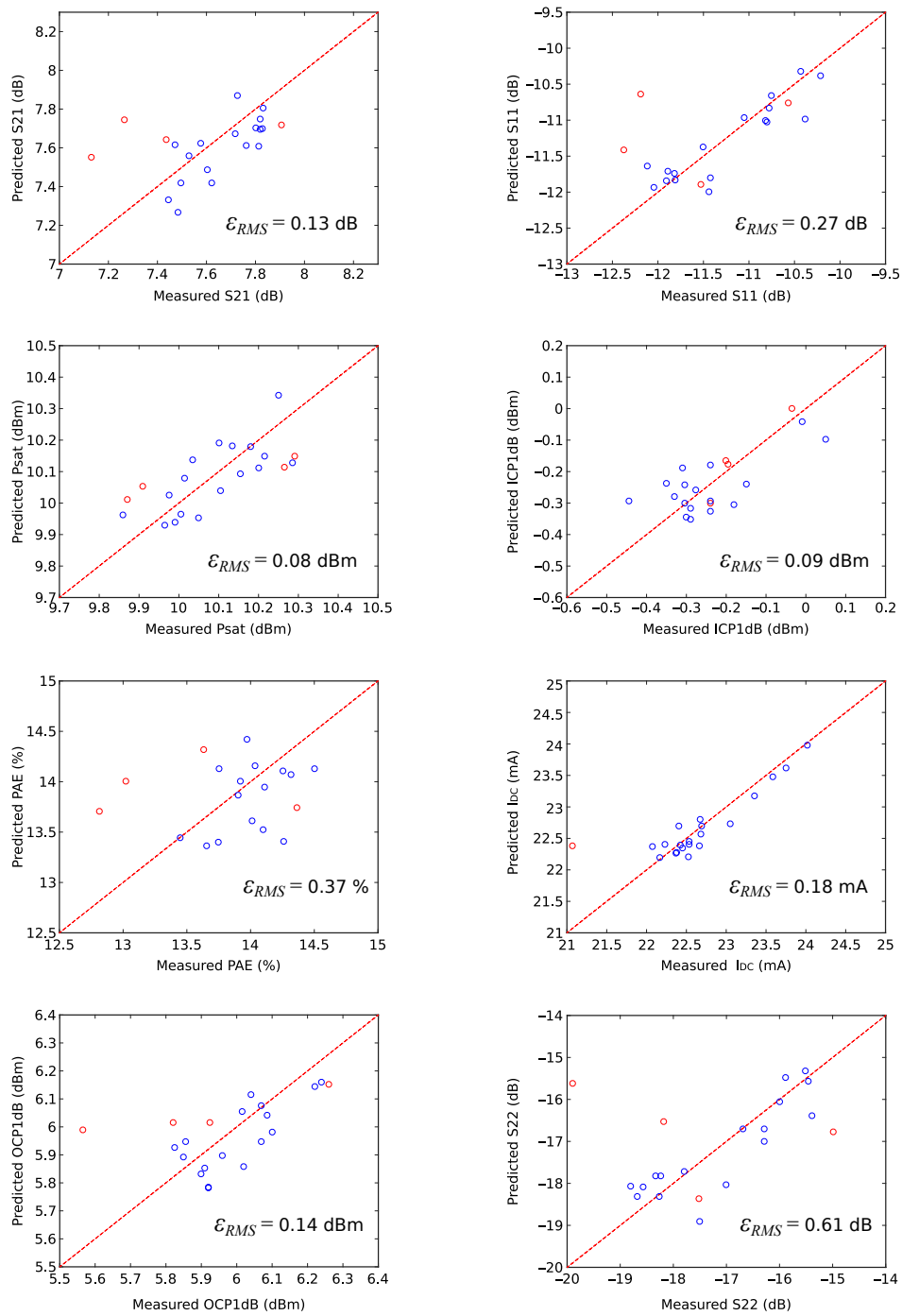


Figure 4.21: Prediction of PA specifications from the set of low-cost signatures versus standard mm-wave functional measurements.

Chapter 5

One-shot calibration against process variation

5.1 Introduction

The practical integration of mm-wave systems using current nanometric fabrication processes is still a challenging task. Process variations in modern nanometric fabrication processes may have a significant impact on the performance of mm-wave circuitry, to the point where production yield can be severely degraded. In order to counter these effects, calibration methods have been introduced, as presented in Section 1.5. In the last few years, embedded calibration strategies have been proposed as a promising solution to alleviate the aforementioned issues for integrated circuits operating in the RF and mm-wave frequency bands, as the ones reviewed in Section 2.2.

In this chapter, we present a calibration strategy for increasing the fabrication yield of a 60 GHz PA. We propose a programmable decoupling cell as a tunable element that can be easily integrated into existing mm-wave circuits for calibration purposes. Calibration is performed in one-shot, which avoids running multiple iterations of the calibration loop, based on an indirect estimation of the PA performance derived from a set of non-intrusive process monitors. These process monitors are not electrically connected to the PA under calibration and are optimized to be sensitive to the main degradation mechanisms of the PA performance, following the method developed in Chapter 3.

5.2 Design of a power amplifier with calibration capabilities

5.2.1 Power amplifier design

The demonstrator vehicle for this work is a 2-stage class A 60 GHz power amplifier designed in a 55 nm CMOS technology provided by STMicroelectronics. The choice of a 2-stage structure provides a compelling validation for the proposed calibration strategy since the interactions between the

different circuit elements make calibration a complex task.

Theoretical basis

A two-stage power amplifier is basically composed of an input driver stage and an output power stage. The output stage can be designed following the guidelines introduced in subsection 3.3.1. The main goal of the input stage is to increase the gain of the PA. In mm-wave frequencies, an interstage matching network is used to provide a conjugate match at both the output of the driver stage and the input of the power stage, in order to maximize the overall gain. The design of the input stage differs from the one of the output, only by the output matching. For the power stage, the output impedance targets a specific real value R_{opt} in order to achieve the target saturation power, whereas for the driver stage, we wish to achieve a conjugate match.

The use of two non-linear stages have to be accounted for when computing the overall $IIP3$ and compression point [4]. If we use the same formalism as in subsection 1.3.2, the compression point can be approximated using Equation 5.1, where $ICP_{1dB,1}$ and $ICP_{1dB,2}$ respectively refer to the driver and power stage ICP_{1dB} . α_1 refers to the small signal gain of the input stage. The same expression can be used to compute the $IIP3$ by replacing $ICP_{1dB,i}$ by $IIP3,i$.

$$\frac{1}{ICP_{1dB}^2} \approx \frac{1}{ICP_{1dB,1}^2} + \frac{\alpha_1^2}{ICP_{1dB,2}^2} \quad (5.1)$$

Practical design implementation

Following the guidelines presented in the previous subsection we have designed a PA which is composed of two stages of common source MOS transistors in class A operation. The power supply voltage is 1.2 V and the gate bias voltage for the two transistors is set to 0.85 V. The matching networks use a structure based on microstrip transmission line stubs. Figure 5.1 shows the transistor-level schematic of the circuit and Figure 5.2 displays its layout. The transistors are in red rectangles, the MOM capacitors used for DC blocking in yellow ones and the decoupling cells in green ones. Figure 5.3 displays the scattering parameters of the PA while Figure 5.4 displays the output power versus the input power. Unlike the first PA we have designed, the output network is not matched to a $50\ \Omega$ impedance. We have performed load-pull simulations in order to maximize the output compression point. Table 5.1 summarizes the main performances obtained by electrical post-layout simulations under typical corner conditions and nominal tuning knob configuration.

As it is shown in Figure 5.1, programmable decoupling cells have been implemented in the PA. The decoupling capacitor banks have been modified by embedding varactors to provide variable termination loads for the stubs in the matching networks. From a designer point of view they act as both a standard decoupling cell and an impedance tuner. The functionality of these structures is further explored in the next subsection.

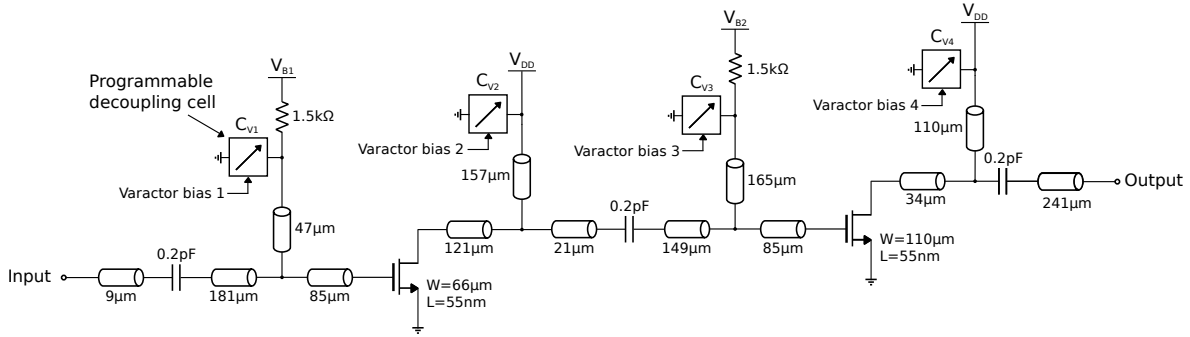


Figure 5.1: Transistor-level schematic of the power amplifier under study.

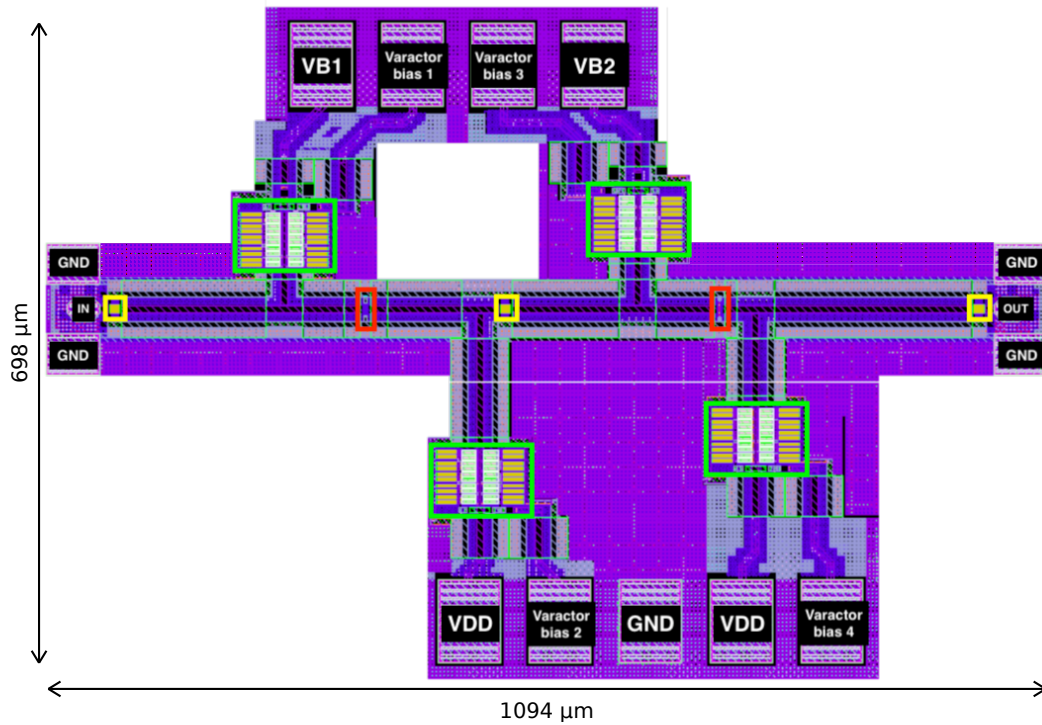


Figure 5.2: Layout of the power amplifier under study.

5.2.2 Tuning knobs design

General presentation

We have implemented two different tunable elements within the PA for calibration purposes: two DC gate biases, labeled V_{B1} and V_{B2} in Figure 5.1, and four variable decoupling cells, labeled C_{V1} to C_{V4} in Figure 5.1. These six tuning knobs $\{T_1, \dots, T_6\}$ are externally controlled with DC voltages since this is a proof-of-concept prototype. The impact of these tuning knobs on the PA under calibra-

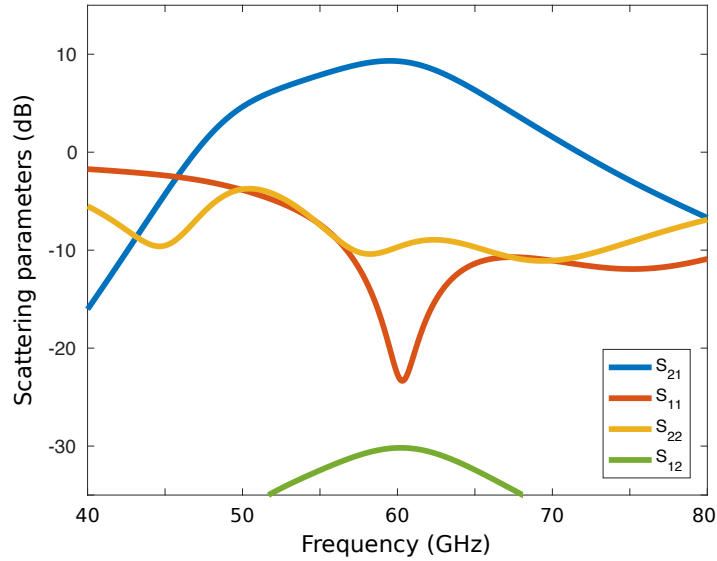


Figure 5.3: Scattering parameters of the 2-stage PA.

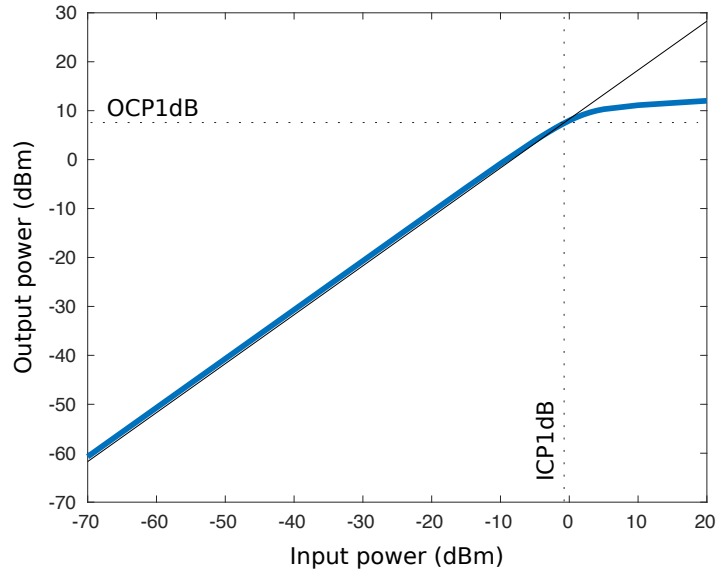


Figure 5.4: Output power versus input power - 2-stage PA.

tion is twofold: by programming the tuning knobs, it is possible to modify the DC power consumption of the circuit and the impedance at different internal nodes. Thus, the DC gate biases V_{B1} and V_{B2} control the current consumption of the two stages of the amplifier, but they also have an impact on the gate and drain impedances of the transistors since MOS transistor parasitic capacitances are bias-dependent. The variable decoupling cells are used as termination loads for the stubs in the matching networks. Their impedance value is low enough to work as a decoupling cell at 60 GHz, and since their impedance can be programmed, they provide the ability to modify the impedance matching between the different elements of the PA. The proposed tuning knobs avoid complex modifications of

Table 5.1: Performance and design parameters of the PA.

	DC current, I_{DC}	51.1 mA
	Gain (S_{21})	9.3 dB @ 60 GHz
	Power Added Efficiency, PAE	12.9 %
Output referred 1dB-Compression Point, OCP_{1dB}		7.1 dBm
	Saturation output power, P_{sat}	11.6 dBm
	S_{11}	< -20 dB
	S_{12}	< -30 dB

the circuit basic structure, as it was the case in [18, 55, 58], and it does not require to program the power supply of the system, as it was needed in [25, 58, 60–62].

The operation principle of the proposed variable decoupling cells can be detailed as follows. Basically, a decoupling cell is a capacitor whose value is high enough at the working frequency to provide an AC short circuit at the operating frequency which implies a constant DC voltage in a given node. In our case, the fixed capacitor has been replaced with a variable one, implemented with a varactor and a MOM capacitor in series. Due to quality factors and layout constraints, 20 small variable cells are connected together to achieve a low enough impedance value at the working frequency: 1.25 pF, i.e. $0.5 - j1.7 \Omega$ at 60 GHz. The decoupling cell features an imaginary impedance value that can be programmed between -2.3Ω and -1.1Ω at 60 GHz. The quality factor changes between 2 and 4.8. Although narrow, it will be shown that this variation range is sufficient for the purpose of this work. For comparison, notice that the impedance of a standard decoupling cell built with six 0.2 pF MOM capacitors in parallel displays a similar impedance ($0.5 - j1.8 \Omega$) and quality factor $Q = 4.2$. According to electrical simulations, the quality factor of the decoupling cell has a negligible impact on the performances of the amplifier, therefore it is not a critical design element.

Model of the variable decoupling cells

The model of the interconnection between the 20 small capacitive cells is a key design element to consider. Figure 5.5 shows the layout of the proposed variable decoupling cell, Figure 5.6 shows the model used to simulate it and Table 5.2 details the model features. The varactors are spotlighted by the green dashed rectangles and the MOM capacitors are in yellow. The dimensions of the microstrip lines that are used to model the interconnections do not match the physical size of the interconnections, as it was already the case with the classic decoupling cell (Figure 3.13). Same as with the classic decoupling cell, the modeling of the interconnection between the stub connection and the DC pad connection was achieved through Keysight Momentum simulations. RLCK extractions using QRC parasitic extraction tool were leveraged to check that the other interconnections (between the MOM capacitors, the varactors and the polysilicon resistors) could be neglected at the operating frequency.

Figure 5.7 displays the impedance of a basic decoupling cell (six 0.2 pF MOM capacitors in parallel) with and without interconnection model, in yellow and red respectively. The interconnection model is the one we have used for the first PA design, as displayed in Figure 3.13. With the variable decoupling

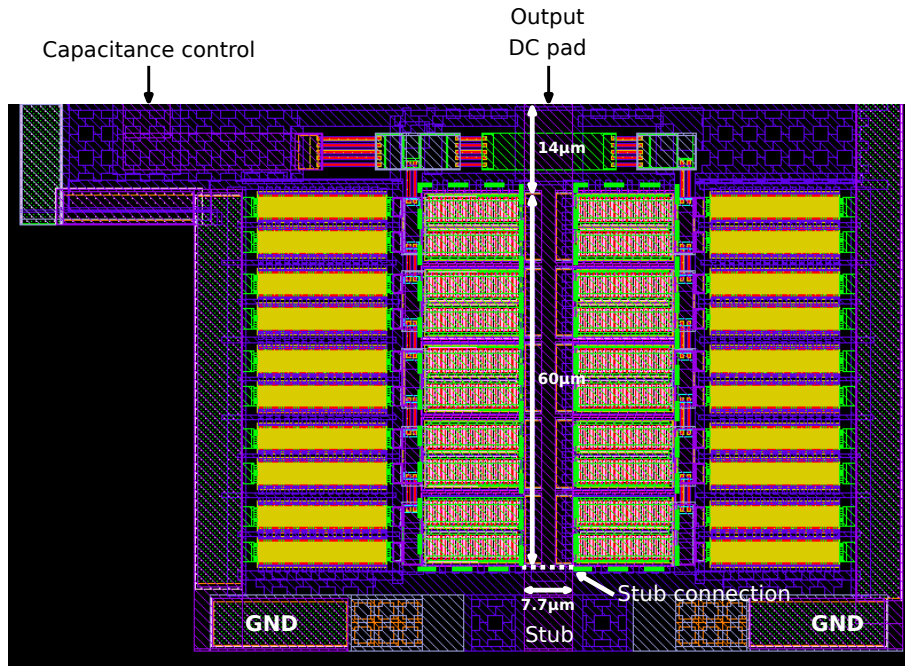


Figure 5.5: Layout of the proposed variable decoupling cell.

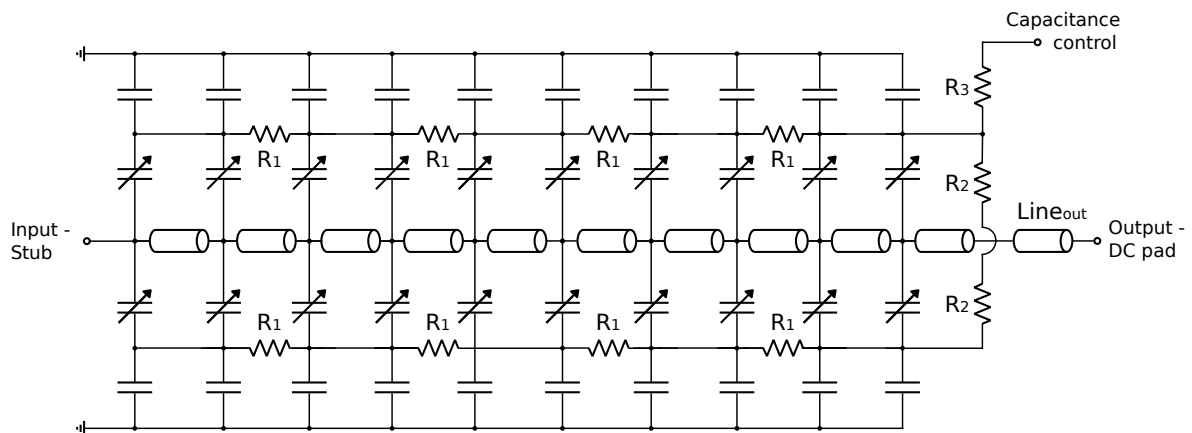


Figure 5.6: Model of the proposed variable decoupling cell.

cell, the importance of modeling the interconnections is even more crucial, as displayed with the blue and green lines.

Table 5.2: Variable decoupling cell model features.

Component feature	Unit	Value
Line width	μm	7.0
Line length	μm	4.5
Output line $Line_{out}$ width	μm	7.7
Output line $Line_{out}$ length	μm	14
MOM capacitance	fF	148
Typical varactor capacitance	fF	108
Resistance R_1	$\text{k}\Omega$	2.6
Resistance R_2	$\text{k}\Omega$	0.8
Resistance R_3	$\text{k}\Omega$	2.1

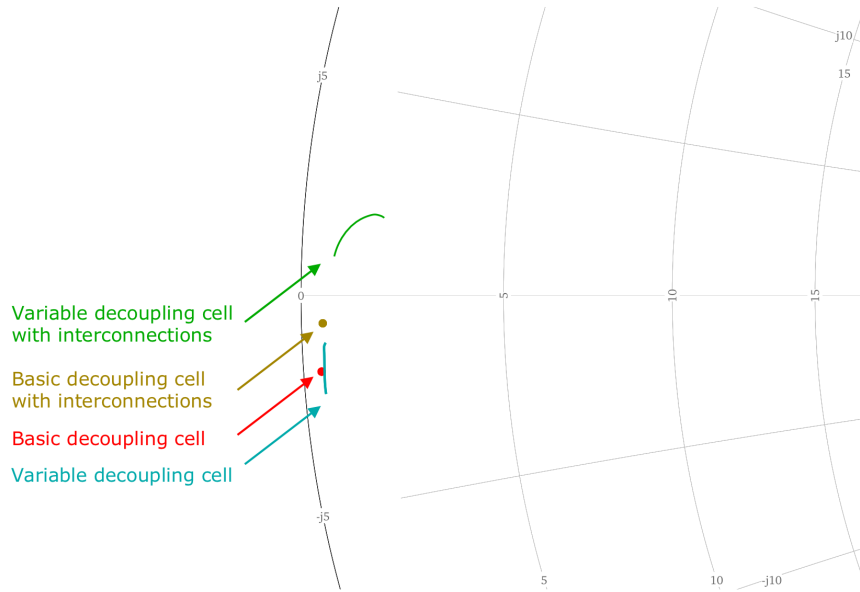


Figure 5.7: Impact of the interconnection model for fixed and variable decoupling cells.

Example using the proposed variable decoupling cell

In order to better understand how both process variations and tuning knobs impact the impedance of the PA internal nodes (and, by extension, its performance), their functionality is illustrated with the simplified example in Figure 5.8. Let us consider an output matching network whose purpose is to match the optimal impedance Z_{opt} presented to the drain of a transistor in order to maximize the PAE for a given input power. In an ideal case, the network is able to present an impedance $Z_{network} = Z_{opt}$. However, Z_{opt} and $Z_{network}$ actually depend on the process variations undergone by the PA, which results in $Z_{network} \neq Z_{opt}$. A good combination of tuning knob values for the transistor gate bias and the output network decoupling cell may recover the ideal matching. As a matter of fact,

by changing the gate bias the required Z_{opt} is modified and by changing the decoupling cell bias the actual $Z_{network}$ may be also modified to make $Z_{network} \approx Z_{opt}$ after calibration. For illustration purposes, Figure 5.8 shows a Smith chart on which the aforementioned impedances are plotted. The required Z_{opt} for several process corners combinations obtained by electrical simulations are plotted in red dots. For a PA in the typical corner, the Z_{opt} excursion obtained by changing the transistor gate bias is plotted in pink dashed line and the $Z_{network}$ excursion obtained by changing the decoupling cell bias are plotted in blue line. It is clear to see that the complete corner variation area is reachable by varying these tuning knobs.

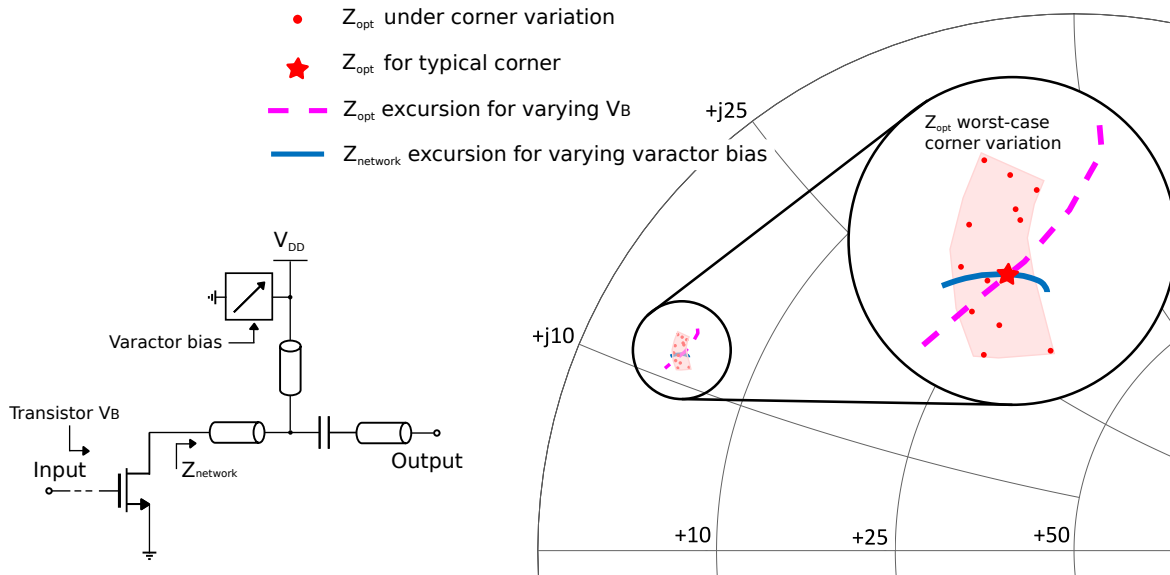


Figure 5.8: Simplified example for illustrating the effect of the tuning knobs on a node impedance.

5.3 A novel data generation method

We have presented the design of a PA with calibration capabilities. As we have already mentioned, we wish to take advantage of this capability to implement a one-shot calibration procedure. The technique relies on a pre-trained regression model $f_{T,S \rightarrow P}$ as presented in 1.5.2. Building such a model requires an important amount of data, including the low-cost measurements and Deviate Under Calibration (DUC) performances, varying tuning-knob values, for a significant number of different devices. The computational burden of generating these data in a simulation environment may be prohibitive, and this issue is accentuated when data comes from real measurements in the production line. In this section we address this problem and propose a method to dramatically reduce the amount of data needed for building the regression models.

5.3.1 Previous work

In the last few years, different one-shot calibration strategies have been proposed for RF and mm-wave circuits. In this subsection we review the main ones with a special focus on the building of the regression model.

One-shot techniques that avoid iterative optimizations were explored in [59, 61] to calibrate a 1.57 GHz RF LNA. The performance of the fabricated LNA is estimated by using an embedded VCO and an output peak detector for exciting the LNA and acquiring its response, respectively. The bias voltages and power supply voltage are used as tuning knobs. The calibration strategy makes use of a machine learning model to predict the best values of the tuning knobs based on the output of the peak detector. This machine learning model is trained using a classic full factorial Design of Experiments (DOE) approach. Thus, over 27,000 simulations were performed in [59] (3 tuning knobs, 3 knob positions, 1,000 Monte Carlo instances), and over 49,392 actual measurements were acquired in [61] (3 tuning knobs, 7 knob positions, 144 fabricated LNA samples) for data generation. This work has recently been extended in [79] where an on-chip neural network implementation is proposed. The data sets that are used for the training are the same as in [59].

A similar approach was employed in [14, 62], where the one-shot calibration paradigm was applied to a 2.4 GHz PA. The main difference lies in the use of non-intrusive dummy circuits that are sensitive to process variations. Four tuning knobs were proposed: 2 bias voltages and 2 supply voltages. Machine learning models map the output of the sensors to the optimum combination of tuning knobs. Data generation for training the calibration model is also based on a full factorial DOE combined with Monte Carlo simulation. The training data set in [14] is composed of 100 Monte Carlo circuit samples and 305 tuning knobs combinations, that is, a total of 30,500 instances.

As previous works show, building calibration models based on full factorial DOE techniques is a time-consuming task that requires a huge number of simulations and/or measurements. The size of the required data set increases with the complexity of the DUC and the number of tuning knobs. Consequently, we propose an efficient alternative to full factorial DOE for the generation of this data set.

5.3.2 Data generation for training calibration models

In order to build an accurate calibration model, the training set has to cover a significant variation range for both process variations and tuning knob positions. In fact, machine learning models do not extrapolate, so ideally the training data set should cover the full possible variation range. At the same time, the size of the training set has to be increased exponentially with the dimensionality of the problem in order to avoid sparse sets that would lead to inaccurate regressions.

The classical approach for generating the training data set relies on DOE techniques for defining an appropriate set of positions for different tuning knobs that covers the calibration space. Then, the effect of these tuning knob positions in the circuit performance is measured over a wide population of DUC samples to account for process variations. In these applications, the standard DOE approach is the so-called full factorial design, that would consider all possible combinations among a pre-defined set

of positions for all the tuning knobs in the DUC. Thus, for a device with k tuning knobs, if we consider that each tuning knob has p_i pre-defined positions, where $i = 1, \dots, k$, then a total of $M = \prod_{i=1}^k p_i$ combinations have to be considered for each circuit sample. In a simulation environment, these circuit samples are generated by Monte Carlo sampling and, ideally, should cover the full process variation range of the DUC.

Figure 5.9a shows a visual representation of the training data set obtained by DOE for a generic one-shot statistical calibration scenario. The data set contains the set of performances and signatures for all tuning knob combinations, in the considered Monte Carlo process variation range.

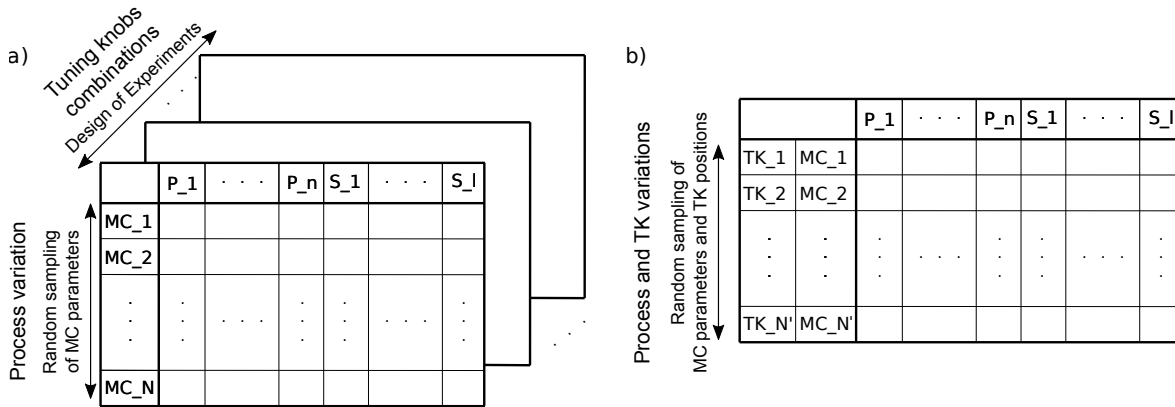


Figure 5.9: Generation of training data sets using a) full factorial DOE and b) random sampling.

Standard DOE techniques are feasible for a reduced number of knobs and knob positions. For example, an analog tuning knob may be usually represented with only three values: *min*, *max*, and *typ*. However, accurately capturing non-linear knob behavior and complex knob interactions require a higher granularity. To tackle this issue, we propose an alternative to DOE techniques for statistical calibration that allows to capture accurately the behavior of the tuning knobs and their interactions, while reducing the size of the training data set. Instead of sampling a limited set of pre-defined values for the tuning knob positions, we propose to randomly sample each knob position. In this way, in a simulation environment, both process and tuning knob variations may be similarly generated within a Monte Carlo simulation, as visually represented in Figure 5.9b. A simple implementation of this technique consists in associating a probability density function to each tuning knob position that covers its full variation range. A uniform probability density function would produce a uniform variation of the tuning knobs for the samples in the training data set. Depending on the application scenario, the probability density function can be also tailored to each particular tuning knob to enrich the information in the data set. In this work we demonstrate that this random sampling strategy reduces significantly the size of the training data set necessary for capturing process and tuning knob variations.

5.3.3 Proposed random sampling results

In this subsection, we show a direct comparison between the classical DOE sampling and the proposed random sampling of knob positions.

The presented calibration results were obtained through post-layout Monte Carlo simulations using the spectreRF simulator. Four different calibration scenarios were considered for comparison purposes. The training data sets in each scenario contain:

- a) 5,000 instances of the DUC: process variations and tuning-knobs values are generated using the proposed random sampling strategy. Uniform probability density functions were associated to each tuning knob covering its full variation range.
- b) 5,040 instances of the DUC: 35 Monte Carlo iterations were performed for a set of $2^4 \times 3^2 = 144$ selected knobs combinations (i.e. *min* and *max* for each varactor and *min*, *typ* and *max* for the two gate bias voltages).
- c) 50,400 instances of the DUC: 350 Monte Carlo iterations were performed for a set of 144 selected knobs combinations (same as in scenario b)).
- d) 50,301 instances of the DUC: 69 Monte Carlo iterations were performed for a set of $3^6 = 729$ selected knobs combinations (i.e. *min*, *typ* and *max* for the 6 tuning knobs).

Data sets a) and b) allows a direct comparison between the proposed data generation methodology and the DOE technique used in previous work [14, 59, 61, 62] using a similar number of training samples. Data sets c) and d) rely as well on DOE, but we have multiplied by 10 the number of training instances of the PA compared to b). Data set c) should offer a better coverage of process variations, since the number of Monte Carlo iterations has been increased, while data set d) should improve the coverage of the tuning knob impact on the PA performance.

Following the statistical one-shot calibration methodology described in subsection 1.5.2, we have trained regression functions $f_{T,S \rightarrow P}$ for each performance of the PA, for the four training data sets. This regression function is used for calibrating an independent set of 500 PA instances generated by Monte Carlo simulation. The obtained results are compared based on the figure of merit (FoM) already defined in Equation 4.14 as

$$FOM = \frac{\sigma_{performance}}{\mu_{error}}, \quad (5.2)$$

where $\sigma_{performance}$ is the standard deviation of the performance and μ_{error} is the RMS prediction error. For instance, if $FoM = 3$ that means that in average, the performance is predicted with an error of $\sigma/3$, which is a reasonably fair accuracy.

Table 5.3 displays the FOM results for the four considered scenarios. For scenarios a) and b) (i.e., 5 000 samples) and for all the performances, the obtained FoMs are significantly better using random sampling than with the DOE technique. For scenarios c) and d), even with 10 times more data in the training set, results provided by the DOE method still do not match scenario a) that uses random sampling. Regarding scenarios c) and d), it appears that data set c) lacks information about the effect of the tuning knobs for large signal performances and that data set d) lacks information about process variations, especially for small-signal performances.

Table 5.3: Comparison of PA performances figures of merit for different training sets.

Training set \ Performance		Figure of Merit			
		a)	b)	c)	d)
I_{DC}		90	59	68.5	68.3
CP_{1dB}		8.3	2.8	3.8	5.45
PAE		7.2	2.7	3.45	4.4
P_{sat}		6.4	2.0	3.4	4.0
S_{21}		6.4	2.0	2.6	2.3
S_{11}		4.6	1.2	2.5	2.45

In order to show a direct comparison of the regression capabilities of the trained regression function f , Figure 5.10 represents the scatterplot of the predicted versus actual values of the PA small signal gain for the 500 instances in the independent validation set. It is clear that training the regression model with data set a) (in green) results in a better accuracy.

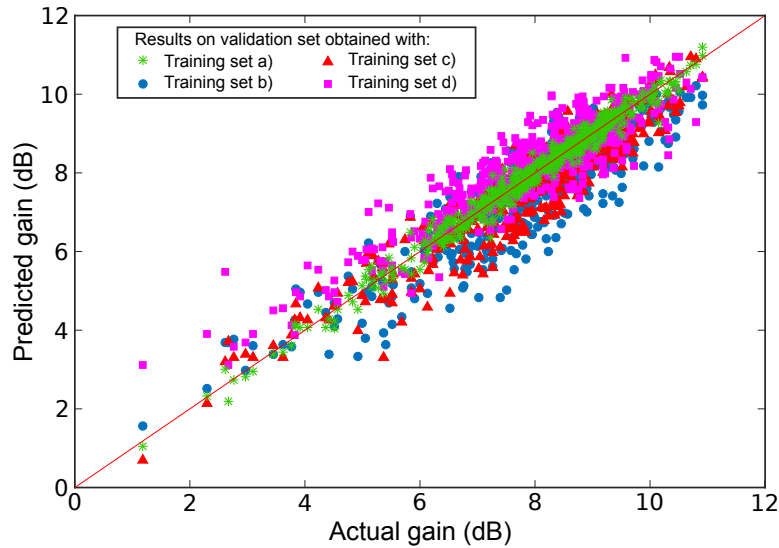


Figure 5.10: Scatterplot of predicted versus actual PA gain for the samples in the validation set.

For further insight, Figure 5.11 shows the FoM corresponding to the PAE calibration for scenarios a), c) and d) as a function of the number of samples in the training set (obtained by random subsampling of the data sets). It is interesting to notice the linear behavior of the FoM calculated on set a) as the size increases. In comparison, the FoM calculated on set c) tends to saturate as the number of training instances increases, probably due to the lack of information about the effect of the tuning knobs. Similarly, data set d) shows the same behavior, but this time probably due to a lack of process

variation information. It seems clear that the proposed random sampling methodology offers a better coverage of the training space while reducing the number of required simulations.

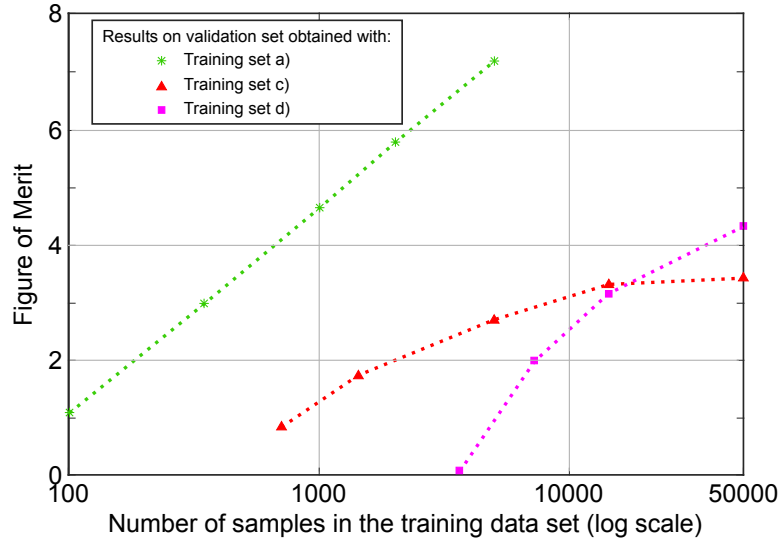


Figure 5.11: Figure of merit for the prediction of the PAE as a function of the size of the training data set.

5.3.4 Practical implementation

The practical implementation of the classical full factorial DOE and the proposed random sampling implied a few challenges. We took advantage of Cadence SKILL programming language and more especially on the Open Command Environment for Analysis (OCEAN) which is based on SKILL.

Full factorial DOE An OCEAN script was written to generate the data following the full factorial DOE paradigm. The script used for generating data set b) is displayed in Annex 1.

Random sampling To implement the random sampling of both process parameters and tuning knob values, we took advantage of the Monte Carlo engine of Cadence ADE XL. To this end, we modified the MC process parameters generation file by adding 6 fake process parameters, one for each of the 6 devised tuning knobs. Using the same syntax, we were able to generate uniform distributions of random values, between their respected boundaries, for the 6 tuning knob values.

Simulating the circuit with the calibrated tuning knob values In order to obtain the true performances of the PA after its calibration, the calibrated values of the 6 tuning knobs needed to be changed from their typical values for each Monte Carlo iteration. To this end, we developed the OCEAN script presented in Figure 5.12 that was used as a pre-run script in ADE XL. This kind of script can be

especially leveraged to run pre-simulations and then modify the variables or the other tests that should be performed in the main simulation, depending on the results of this pre-run. Therefore, it was quite straightforward to implement the change of the tuning knob values before each iteration of Monte Carlo sampling.

```

mc_num=ocnxIMCIterNum()
data_file= infile("List_calibrated_TK_values.txt")

for(i=1 mc_num
fscanf(data_file"%f %f %f %f %f %f" a b c d e f ))
close(data_file)

ocnxIAddOrUpdateOutput("Calibrated_TK1" a)
...
ocnxIAddOrUpdateOutput("Calibrated_TK6" f)

ocnxIUpdatePointVariable("TK1" sprintf( nil "%L" a))
...
ocnxIUpdatePointVariable("TK6" sprintf( nil "%L" f))

```

Acquire the MC iteration number.

Define the file containing the calibrated tuning knob values.

Read the file and set variables a, ..., f to the desired calibrated tuning knob values, depending on the MC iteration.

Actually set the tuning knob values to those that has been read in the file (stored in variables a, ..., f).

Update the values that are displayed in the ADE XL window.

Figure 5.12: OCEAN script used to simulate the circuit with the calibrated tuning knob values.

5.3.5 Conclusions

We have introduced an efficient methodology for the generation of appropriate training data sets for statistical calibration applications. The proposed technique is based on the joint random sampling of both process and tuning knob position variations. Compared to previous works based on full factorial DOE, this method dramatically reduces the size of the required training data set for an accurate training of statistical calibration models. Moreover, this method is especially appealing for complex systems with an important number of tuning knobs, since sweeping all the possible combinations would be unrealistic.

For the one-shot calibration detailed in this chapter we have used this random sampling method to build the regression models. Next section details the implementation of the calibration procedure and presents the results of this work.

5.4 Calibration procedure and results

5.4.1 Calibration procedure and implementation

The presented one-shot calibration strategy has been validated by post-layout electrical simulations of the two-stage 60 GHz PA. Figure 5.13 displays the layout of the PA including the variable decoupling cells and the set of non-intrusive process monitors for calibration, implemented in the selected 55 nm CMOS technology. Following the assisted design procedure presented in Chapter 3, we have designed six process monitors, labelled 1 to 6 in Figure 5.13: a high-impedance transmission line (1), a low impedance transmission line (2), a MOM capacitor (3), a varactor (4), a MOS junction capacitance monitor (5), and a MOS transistor (6). 10 signatures S_1, \dots, S_{10} were extracted from these sensors. Basically, signatures S_1, \dots, S_9 are the same as for the one-stage PA (Table 3.3) and S_{10} is the imaginary part of scattering parameter S_{21} measured at 100 MHz on the standalone varactor.

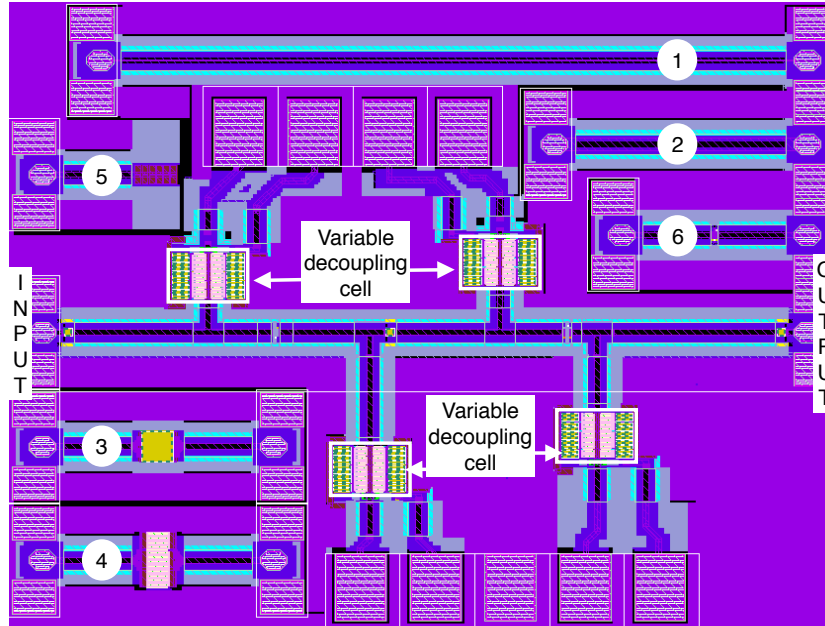


Figure 5.13: Layout of the power amplifier with variable decoupling cells and non-intrusive process monitors.

Based on the previously presented data generation method, Monte Carlo sampling was used for generating a training set of 5,000 PA instances. These training samples are then used to build the regression function $f_{T,S \rightarrow P}$ used for this one-shot statistical calibration, as detailed in Equation 5.3. Figure 5.14 illustrates the regression function accuracy with the prediction of the small signal gain S_{21} of the amplifier. The RMS prediction error is 0.21 dB and the maximum regression error is 0.62 dB. Each blue dot corresponds to one of the 5,000 PA instances. It is clear that this prediction error imposes a fundamental limit to the accuracy of the calibration. However, in a practical scenario this prediction error can be mitigated by improving the training phase (for instance by increasing the size

of the training set).

$$f_{T,S \rightarrow P} : [T_1, \dots, T_6, S_1, \dots, S_{10}] \rightarrow [P_1, P_2, \dots, P_6]. \quad (5.3)$$

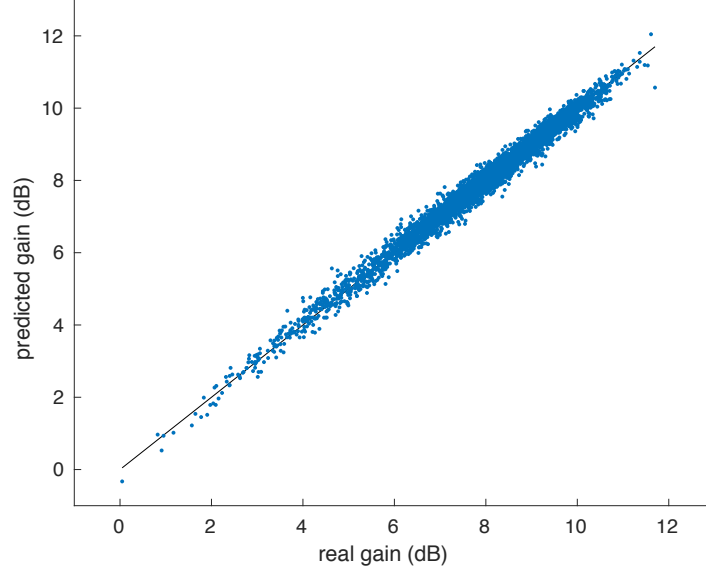


Figure 5.14: Predicted versus actual gain using the regression function used for the calibration.

With regard to the goals for the calibration of our PA under study, initially we target the combination of tuning knobs that allows the circuit to pass all of its specifications while maximizing gain, saturation power, 1dB output compression point and PAE and minimizing power consumption. If no solution is found, the calibration algorithm looks for the tuning knob combination that gives the closest result to the typical PA performance. The input matching is not targeted by this calibration, but the algorithm checks that it remains good enough (below -11 dB in the worst case). The output matching is not targeted either because the relevant performance for our PA is the saturation power P_{sat} . It is worth noticing that another advantage of one-shot statistical calibration techniques is that the calibration goals can be changed *a posteriori* without the need for additional processing. Algorithm 2 details the algorithm that has been implemented to explore the space of reachable performances by varying the tuning knobs values. It leverages the regression model $f_{T,S \rightarrow P}$ and two cost functions CostFunc1 and CostFunc2 for deciding which tuning knob combination is the best. The second one is only used when no calibration solution that passes all the specifications is found. Typical performances P_{typ} correspond to the ones obtained with a typical corner simulation. The target performances P_{target} can be chosen depending on the calibration goal: minimizing or maximizing the different specifications, with different degrees of importance since they can also be seen as weights in the cost function. $TK1_0$ to $TK6_0$ correspond to the 6 tuning knobs initial values: 0 V for the variable decoupling cells and 0.85 V for the gate bias voltages. We consider 9 possible values for the 4 decoupling cell biases and 7 for the gate biases, resulting in $9^4 \times 7^2 = 321,489$ tuning knob combinations, named N in Algorithm 2. Although large, this number is not a limiting point due to the very fast computing of $f_{T,S \rightarrow P}$.

Algorithm 2 Exploring the space of reachable performances to find the best tuning knob combination.

- 1: Specifications before calibration: $P_0 = [S21_0, P_{sat_0}, OCP1dB_0, PAE_0, IDC_0, S11_0] = f_{T,S \rightarrow P}(TK1_0, \dots, TK6_0, S_1, \dots, S_{10})$
 - 2: $best_{combination} \leftarrow 1$ (Initially, the best calibration solution is the default one.)
 - 3: $calibrated \leftarrow 0$ (Initially, the circuits has not been calibrated.)
 - 4: Compute the first cost function CostFunc1 on the default tuning knob position.
 $CostFunc1_{best} = |S21_0 - S21_{target}| / S21_{target} + \dots + |IDC_0 - IDC_{target}| / IDC_{target}$
 - 5: **for** $i=1$ **to** $N-1$ (Number of tuning knob combinations) **do**
 - 6: Specification for the i^{th} tuning knob combination:
 $P_i = f_{T,S \rightarrow P}(TK1_i, \dots, TK6_i, S_1, \dots, S_{10})$
 - 7: $CostFunc1_i = |S21_i - S21_{target}| / S21_{target} + \dots + |IDC_i - IDC_{target}| / IDC_{target}$
 - 8: **if** P_i passes all specifications **and** $CostFunc1_i < CostFunc1_{best}$ **and** $S11 < -11$ dB **then**
 - 9: $best_{combination} \leftarrow i$
 - 10: $calibrated \leftarrow 1$
 - 11: **else**
 - 12: $best_{combination} \leftarrow 1$
 - 13: **end if**
 - 14: **end for**
 - 15: **if not** calibrated **then**
 - 16: Compute the second cost function CostFunc2 on the default tuning knob position.
 $CostFunc2_{best} = |S21_0 - S21_{typ}| / S21_{typ} + \dots + |IDC_0 - IDC_{typ}| / IDC_{typ}$
 - 17: **for** $i=1$ **to** $N-1$ **do**
 - 18: $P_i = f_{T,S \rightarrow P}(TK1_i, \dots, TK6_i, S_1, \dots, S_{10})$
 - 19: $CostFunc2_i = |S21_i - S21_{typ}| / S21_{typ} + \dots + |IDC_i - IDC_{typ}| / IDC_{typ}$
 - 20: **if** $CostFunc2_i < CostFunc2_{best}$ **and** $S11 < -11$ dB **then**
 - 21: $best_{combination} \leftarrow i$
 - 22: **else**
 - 23: $best_{combination} \leftarrow 1$
 - 24: **end if**
 - 25: **end for**
 - 26: **end if**
 - 27: **return** The best tuning knob combination: $best_{combination}$
-

5.4.2 Calibration results

The proposed calibration strategy was then applied to an independent set of 500 PA instances for validation. For each sample in the validation set, the set of monitor outputs is acquired, and the regression function is used in combination to the pre-defined calibration goals to find the optimum value of the tuning knobs using Algorithm 2.

The obtained results are summarized in Table 5.4. This table shows the target specification limits for the PA performance, together with the mean value and standard deviation of each specification, before and after calibration, obtained by post-layout simulation. Regarding S_{11} , before calibration its mean value was around -19 dB, while after calibration it is slightly degraded to -18.2 dB. The worst case S_{11} before calibration is -12 dB and after calibration -11 dB. It is interesting that the calibration has tightened the standard deviation of P_{sat} much more than OCP_{1dB} . The reason for this is that tuning knob C_{V4} has a significant effect on P_{sat} that allows for fine tuning this performance.

Figures 5.15 and 5.16 shows the histograms of the main performances of the PA, before and after calibration, for the 500 samples in the validation set. The red lines represent the specification limits. It is clear to see the yield improvement for all the specifications and the re-centering of the gain and the saturation power around the upper performance limit. In this line, Figure 5.17 shows the total yield improvement together with the yield improvement for the different PA specifications. Overall, the yield has been improved from 59% to 83% and PAE is increased by 0.7% in average thanks to the proposed one-shot calibration.

Table 5.4: Performances of the PA before and after calibration.

Performance	Target specification	mean	standard deviation
S_{21} (dB)	$\in [8.5, 10]$	9.16 \rightarrow 9.36	0.84 \rightarrow 0.56
S_{11} (dB)	–	-18.96 \rightarrow -18.22	4.22 \rightarrow 3.98
P_{sat} (dBm)	< 12	11.59 \rightarrow 11.67	0.22 \rightarrow 0.14
OCP_{1dB} (dBm)	> 6.5	7.15 \rightarrow 7.20	0.34 \rightarrow 0.31
PAE (%)	maximize	12.68 \rightarrow 13.38	1.18 \rightarrow 1.09
I_{DC} (mA)	< 60	51.54 \rightarrow 51.49	4.02 \rightarrow 3.91

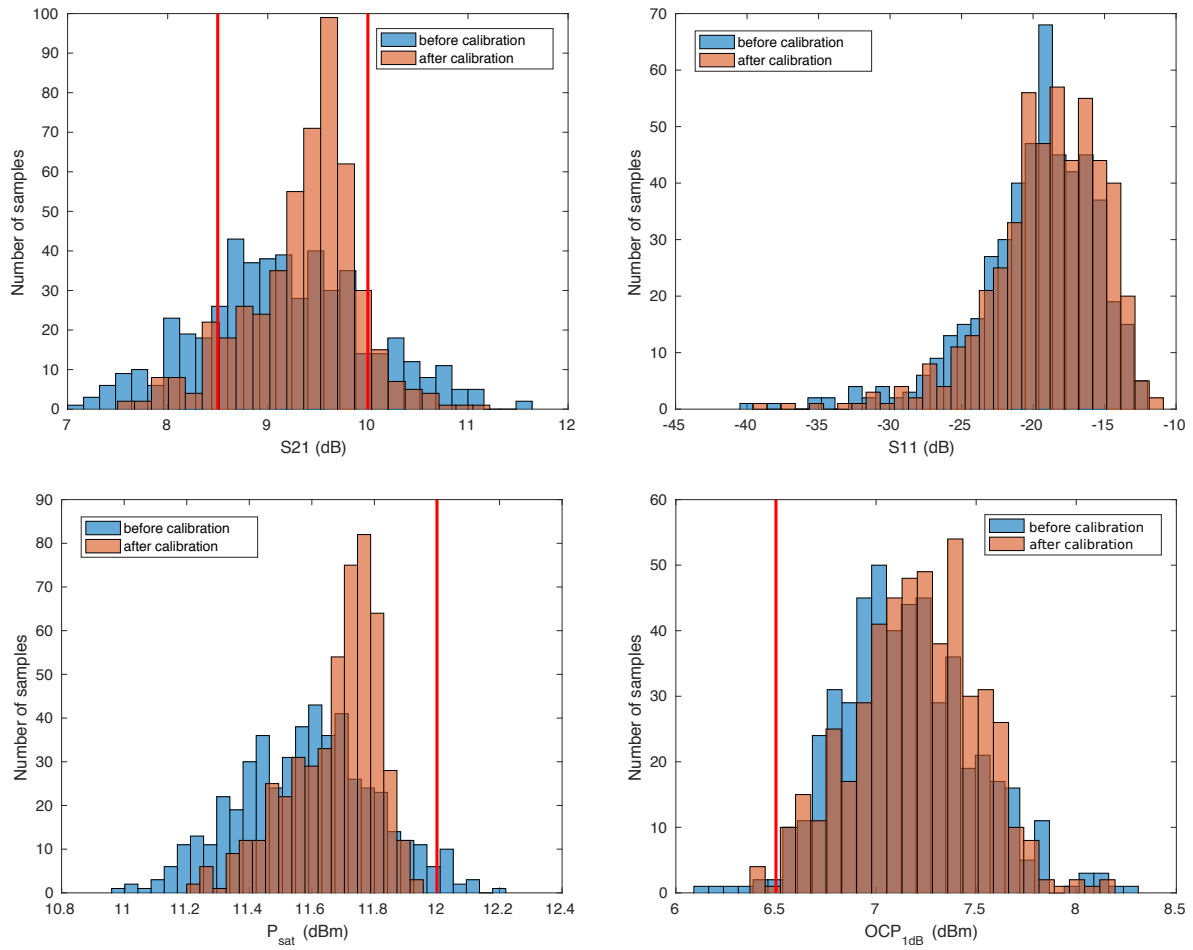


Figure 5.15: Histograms of the main performances of the PA simulated both before and after calibration.

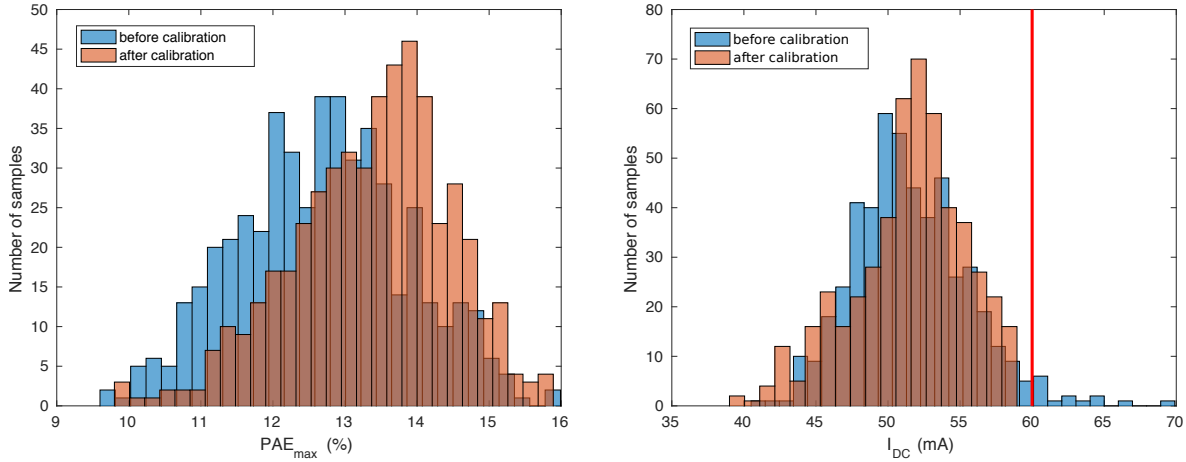


Figure 5.16: Histograms of the main performances of the PA simulated both before and after calibration (continued).

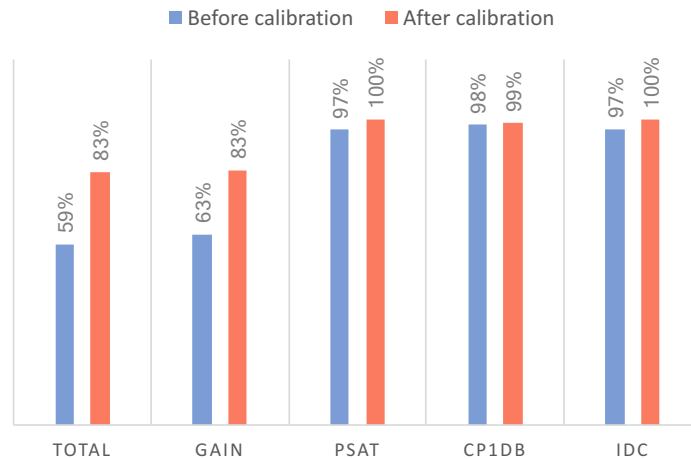


Figure 5.17: Yield before and after calibration.

5.5 Conclusions

In this chapter, we focused on the implementation of a one-shot statistical calibration against process variation.

In this line, we have presented a two-stage 60 GHz power amplifier with one-shot calibration capability. The proposed calibration takes advantage of a tunable element implemented as a variable decoupling cell to change the circuit behavior. Non-intrusive process monitors, placed within the empty spaces of the circuit and not electrically connected to it, are used for predicting the best tuning

knob configuration based on a machine learning regression model.

A novel data generation methodology has also been presented. It greatly relaxes the computational work load necessary to build the regression model. This technique is based on the joint random sampling of both process and tuning knob position variations. The amount of data that is used to train the neural network is 50 times less important compared to the classical method based on full factorial DOE.

The proposed circuitry has been implemented in STMicroelectronics 55 nm CMOS technology. The feasibility and performance of the proposed calibration strategy have been validated by electrical post-layout simulations. Obtained results show that the proposed calibration strategy improves the overall design yield from 59 % to 83 % and the PAE by 0.7 % in average.

Conclusions

In this thesis we have presented several contributions to tackle the cost and complexity associated with testing mm-wave ICs. In this sense, we have focused on two main topics: a) non-intrusive machine learning indirect test and b) one-shot statistical calibration. Our approaches have been validated on circuits designed in ST Microelectronics 55 nm CMOS technology, either in simulations or in measurements.

First, we have presented a systematic methodology for generating a non-intrusive indirect test program for mm-wave ICs. The method leverages the Monte Carlo models of the IC components so it can only be applied to circuits designed with components that have such models. This approach takes advantage of machine learning algorithms to find the root causes of parametric performance degradation during the fabrication of the circuit. Then, our methodology guides the design of appropriate non-intrusive sensors and signatures for accurate performance prediction. This technique is a promising step towards the full automation of test generation for non-intrusive machine learning test. This study has been illustrated on a 60 GHz PA and has been carried out using simulations data only by taking advantage of the MC models provided in the PDK.

We have then extended this work to target generic mm-wave circuits that may use custom passive components. This second method can be applied to mm-wave ICs containing both standard devices modeled in the PDK of the technology and full-custom passive components build in the BEOL of the technology. Similarly to the first method we have introduced, the proposed test generation strategy is based on analyzing the root-causes of parametric performance degradation. This analysis leads to design of a set of non-intrusive process monitors with associated low cost signatures sensitive to the identified process-related performance degradation causes. Finally, a machine learning algorithm is trained to infer the specifications of the DUT from the set of low-cost signatures. The feasibility of the proposed technique has been experimentally validated on a 65 GHz PA. A set of non-intrusive process monitors has been designed according to the proposed methodology and integrated together with the PA under test. Experimental measurements on a set of 21 fabricated PAs have been performed. Overall, there is a very good agreement between the classical specification measurements and the proposed indirect test based on low-cost low-frequency non-intrusive signatures.

Finally, in order to tackle the performance degradation induced by process variations, we focused on the implementation of a one-shot statistical calibration against process variation. In this line, we have presented a two-stage 60 GHz power amplifier with one-shot calibration capability. The proposed calibration takes especially advantage of a tuning knob, implemented as a variable decoupling

cell. Non-intrusive process monitors, placed within the empty spaces of the circuit and not electrically connected to it, are used for predicting the best tuning knob configuration based on a machine learning regression model. The design of non-intrusive sensors was guided by the methodology previously presented. A novel data generation technique has also been introduced. It greatly relaxes the computational burden necessary to build the regression model. The approach is based on the joint random sampling of both process and tuning knob position variations. The feasibility and performance of the proposed calibration strategy have been validated by electrical post-layout simulations. Obtained results show that the proposed calibration strategy improves the overall design yield from 59 % to 83 % and the PAE by 0.7 % in average. The two-stage PA and its associated process sensors have been fabricated and measurements will soon be performed to further prove the feasibility of our approach.

Our future works will try to tackle several challenges. Firstly, novel architectures for the process sensors will be studied. For instance, a novel gate resistance sensor would be interesting in order to ease the measurement at lower frequencies. As a matter of fact, sensing the gate resistance is a key element in the implementation of the proposed methodology since it can dramatically alter the PA performances and is very likely to negatively impact the behavior of any kind of amplifier or active device that uses MOS transistors. Secondly, it could be relevant to apply the non-intrusive indirect test paradigm to devices working at higher frequencies and that make use of complex hand-designed passive structures. As a matter of fact the proposed technique is all the more appealing for high frequency devices since the test equipment is especially costly. Applying the proposed non-intrusive indirect test to a complete mm-wave transceiver could also be an interesting perspective. It would allow the prediction of many specifications on the different circuits and thus would provide a compelling validation of the method. Finally, we will study test techniques sensitive to spot defects that would allow to complement the proposed non-intrusive test based on process variation sensors.

Publications

Posters:

[1] Florent Cilici, Manuel J. Barragan, Salvador Mir, Estelle Lauga-Larroze, Sylvain Bourdel, “Conception en vue du test d’un amplificateur de puissance à 60 GHz”, in *Journées Nationales du Réseau Doctoral en Micro-nanoélectronique 2017 (JNRDM)*, November 2017, Strasbourg, France. **[Best Poster Award]**.

[2] Manuel J. Barragan, Gildas Leger, Florent Cilici, Estelle Lauga-Larroze, Sylvain Bourdel, Salvador Mir, “On the use of causal feature selection in the context of machine-learning indirect test”, in *2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, March 2019, Florence, Italy, pp. 276-279.

Conference papers:

[3] Florent Cilici, Manuel J. Barragan, Salvador Mir, Estelle Lauga-Larroze, Sylvain Bourdel, “Assisted test design for non-intrusive machine learning indirect test of millimeter-wave circuits” in *2018 IEEE 23rd European Test Symposium (ETS)*, May 2018, Bremen, Germany, pp. 1-6. **[Best Paper Award]**.

[4] Florent Cilici, Manuel J. Barragan, Salvador Mir, Estelle Lauga-Larroze, Sylvain Bourdel, Gildas Leger, “Yield recovery of mm-wave power amplifiers using variable decoupling cells and one-shot statistical calibration” in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2019, Sapporo, Japan, pp. 1-5.

[5] Florent Cilici, Gildas Leger, Manuel J. Barragan, Salvador Mir, Estelle Lauga-Larroze, Sylvain Bourdel, “Efficient generation of data sets for one-shot statistical calibration of RF/mm-wave circuits” in *2019 International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, July 2019, Lausanne, Switzerland, pp. 1-4.

Journals:

[6] **Submitted** Florent Cilici, Manuel J. Barragan, Estelle Lauga-Larroze, Sylvain Bourdel, Gildas Leger, Loïc Vincent, and Salvador Mir, “A non-intrusive machine learning-based test methodology for millimeter-wave integrated circuits” in *IEEE Transactions on Microwave Theory and Techniques*, pp. 1-14.

Bibliography

- [1] F. Poehl, F. Demmerle, J. Alt, and H. Obermeir, "Production test challenges for highly integrated mobile phone SOCs —A case study," in *2010 15th IEEE European Test Symposium*, 2010, pp. 17–22.
- [2] K. B. Schaub and J. Kelly, *Production Testing of RF and System-on-a-Chip Devices for Wireless Communications*. Artech House, 2004.
- [3] Agilent Technologies, Inc, *Advanced Calibration Techniques for Vector Network Analyzers*, 2006.
- [4] B. Razavi, *RF microelectronics, second edition*. Pearson, 2011.
- [5] L. Abdallah, H. Stratigopoulos, S. Mir, and C. Kelma, "RF Front-End Test Using Built-in Sensors," *IEEE Design & Test of Computers*, vol. 28, no. 6, pp. 76–84, 2011.
- [6] H. Stratigopoulos and S. Sunter, "Efficient monte carlo-based analog parametric fault modelling," in *2014 IEEE 32nd VLSI Test Symposium (VTS)*, 2014, pp. 1–6.
- [7] J. Ferrario, R. Wolf, and S. Moss, "Architecting millisecond test solutions for wireless phone RFICs," in *Proceedings. International Test Conference*, 2002, pp. 1151–1158.
- [8] J. Park, J. W. Lee, J. Chung, K. Han, J. A. Abraham, E. Byun, C. Woo, and S. Oh, "At-speed Test of High-Speed DUT Using Built-Off Test Interface," in *2010 19th IEEE Asian Test Symposium*, 2010, pp. 269–274.
- [9] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210–229, 1959.
- [10] P. N. Variyam and A. Chatterjee, "Enhancing test effectiveness for analog circuits using synthesized measurements," in *VLSI Test Symposium*, 1998, pp. 132–137.
- [11] G. Leger and M. J. Barragan, "Mixed-signal test automation: Are we there yet?" in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1–5.
- [12] M. Andraud and M. Verhelst, "From on-chip self-healing to self-adaptivity in analog/RF ICs: challenges and opportunities," in *2018 IEEE 24th International Symposium on On-Line Testing And Robust System Design (IOLTS)*, July 2018, pp. 131–134.
- [13] Y. Lu, K. S. Subramani, H. Huang, N. Kupp, K. Huang, and Y. Makris, "A comparative study of one-shot statistical calibration methods for analog / RF ICs," in *2015 IEEE International Test Conference (ITC)*, October 2015, pp. 1–10.

- [14] M. Andraud, H. Stratigopoulos, and E. Simeu, "One-shot calibration of RF circuits based on non-intrusive sensors," in *2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2014, pp. 1–6.
- [15] Y. Huang, H. Hsieh, and L. Lu, "A build-in self-test technique for RF low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 5, pp. 1035–1042, 2008.
- [16] A. Valdes-Garcia, R. Venkatasubramanian, J. Silva-Martinez, and E. Sanchez-Sinencio, "A broadband CMOS amplitude detector for on-chip RF measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, pp. 1470–1477, 2008.
- [17] C. Lee, W. Choi, R. Han, H. Shichijo, and K. K. O, "Broadband Root-Mean-Square Detector in CMOS for On-Chip Measurements of Millimeter-Wave Voltages," *IEEE Electron Device Letters*, vol. 33, no. 6, pp. 752–754, 2012.
- [18] S. M. Bowers, K. Sengupta, K. Dasgupta, B. D. Parker, and A. Hajimiri, "Integrated self-healing for mm-wave power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 3, pp. 1301–1315, March 2013.
- [19] A. Serhan, L. Abdallah, H. Stratigopoulos, and S. Mir, "Low-cost EVM built-in test of RF transceivers," in *2014 9th International Design and Test Symposium (IDT)*, 2014, pp. 51–54.
- [20] 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: Theory and Applications*, vol. 27, no. 3, pp. 277–288, June 2011.
- [21] D. Kissinger, R. Agethen, and R. Weigel, "A versatile built-in test architecture for integrated millimeter-wave radar receiver front-ends," in *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, 2012, pp. 254–258.
- [22] S. B. Sleiman, A. Akour, W. Khalil, and M. Ismail, "Millimeter-wave BiST and BiSC using a high-definition sub-ranged detector in 90nm CMOS," in *2010 53rd IEEE International Midwest Symposium on Circuits and Systems*, 2010, pp. 477–480.
- [23] R. L. Schmid, P. Song, C. T. Coen, A. C. Ulusoy, and J. D. Cressler, "A W-band integrated silicon-germanium loop-back and front-end transmit-receive switch for Built-in-self-test," in *2015 IEEE MTT-S International Microwave Symposium*, May 2015, pp. 1–4.
- [24] S. Zeinolabedinzadeh, A. C. Ulusoy, R. L. Schmid, F. Inanlou, I. Song, T. Chi, J. S. Park, H. Wang, and J. D. Cressler, "A W-Band SiGe Transceiver with Built-in Self-Test," in *2019 IEEE 19th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Jan 2019, pp. 1–4.
- [25] V. Natarajan, A. Banerjee, S. Sen, S. Devarakond, and A. Chatterjee, "Yield Recovery of RF Transceiver Systems Using Iterative Tuning-Driven Power-Conscious Performance Optimization," *IEEE Design & Test*, vol. 32, no. 1, pp. 61–69, Feb 2015.
- [26] J. J. Dabrowski and R. M. Ramzan, "Built-in loopback test for IC RF transceivers," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 6, pp. 933–946, 2010.
- [27] K. Arabi and B. Kaminska, "Oscillation-test strategy for analog and mixed-signal integrated circuits," in *Proceedings of 14th VLSI Test Symposium*, 1996, pp. 476–482.

- [28] G. Huertas, D. Vazquez, E. J. Peralias, A. Rueda, and J. L. Huertas, "Testing mixed-signal cores: a practical oscillation-based test in an analog macrocell," *IEEE Design & Test of Computers*, vol. 19, no. 6, pp. 73–82, 2002.
- [29] H. P. Nel, T. Stander, and F. C. Dualibe, "Built-In Oscillation-Based Self-Testing of a 2.4 GHz LNA in 0.35 μ m CMOS," in *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2018, pp. 837–840.
- [30] M. Margalef-Rovira, M. J. Barragan, E. Sharma, P. Ferrari, E. Pistono, and S. Bourdel, "An oscillation-based test technique for on-chip testing of mm-wave phase shifters," in *2018 IEEE 36th VLSI Test Symposium (VTS)*, 2018, pp. 1–6.
- [31] A. P. Jose, K. A. Jenkins, and S. K. Reynolds, "On-chip spectrum analyzer for analog built-in self test," in *23rd IEEE VLSI Test Symposium (VTS'05)*, 2005, pp. 131–136.
- [32] I. Nasr, J. Nehring, K. Aufinger, G. Fischer, R. Weigel, and D. Kissinger, "Single and dual-port 50-100 GHz integrated vector network analyzers with on-chip dielectric sensors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 9, pp. 2168–2179, 2014.
- [33] O. Inac, F. Golcuk, T. Kanar, and G. M. Rebeiz, "A 90–100-GHz Phased-Array Transmit/Receive Silicon RFIC Module With Built-In Self-Test," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 10, pp. 3774–3782, Oct 2013.
- [34] L. Abdallah, H. Stratigopoulos, S. Mir, and J. Altet, "Defect-oriented non-intrusive RF test using on-chip temperature sensors," in *2013 IEEE 31st VLSI Test Symposium (VTS)*, 2013, pp. 1–6.
- [35] A. Tang, F. Hsiao, D. Murphy, I. Ku, J. Liu, S. D'Souza, N. Wang, H. Wu, Y. Wang, M. Tang, G. Virbila, M. Pham, D. Yang, Q. J. Gu, Y. Wu, Y. Kuan, C. Chien, and M. F. Chang, "A low-overhead self-healing embedded system for ensuring high yield and long-term sustainability of 60GHz 4Gb/s radio-on-a-chip," in *2012 IEEE International Solid-State Circuits Conference*, 2012, pp. 316–318.
- [36] A. Zjajo, M. J. Barragan, and J. P. de Gyvez, "BIST Method for Die-Level Process Parameter Variation Monitoring in Analog/Mixed-Signal Integrated Circuits," in *2007 Design, Automation & Test in Europe Conference & Exhibition*, 2007, pp. 1–6.
- [37] L. Abdallah, H. Stratigopoulos, S. Mir, and C. Kelma, "Experiences with non-intrusive sensors for RF built-in test," in *2012 IEEE International Test Conference*, 2012, pp. 1–8.
- [38] A. Dimakos, H. Stratigopoulos, A. Siligaris, S. Mir, and E. De Foucauld, "Built-in test of millimeter-wave circuits based on non-intrusive sensors," in *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2016, pp. 505–510.
- [39] T. Mahzabeen, R. M. Henderson, and B. Banerjee, "Built-in self-test (BIST) algorithm to mitigate process variation in millimeter wave circuits," in *WAMICON 2014*, 2014, pp. 1–3.
- [40] P. N. Variyam and A. Chatterjee, "Specification-driven test design for analog circuits," in *Proceedings 1998 IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems (Cat. No.98EX223)*, 1998, pp. 335–340.
- [41] ———, "Specification-driven test generation for analog circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 19, no. 10, pp. 1189–1201, 2000.

- [42] H. Ayari, F. Azais, S. Bernard, M. Comte, M. Renovell, V. Kerzerho, O. Potin, and C. Kelma, "Smart selection of indirect parameters for DC-based alternate RF IC testing," in *VLSI Test Symposium*, 2012, pp. 19–24.
- [43] M. J. Barragan and G. Leger, "Efficient selection of signatures for analog/RF alternate test," in *European Test Symposium*, 2013.
- [44] A. Gomez-Pau, L. Balado, and J. Figueras, "Efficient Production Binning Using Octree Tessellation in the Alternate Measurements Space," *Trans. on Computer-Aided Design of Int. Circ. and Syst.*, vol. 35, no. 8, pp. 1386–1395, 2016.
- [45] M. Barragan and G. Leger, "A Procedure for Alternate Test Feature Design and Selection," *IEEE Design & Test*, vol. 32, no. 1, pp. 18–25, Feb. 2015.
- [46] G. Leger and M. J. Barragan, "Brownian distance correlation-directed search: A fast feature selection technique for alternate test," *Integration, the VLSI Journal*, vol. 55, pp. 401–414, Sep. 2016.
- [47] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, March 2003.
- [48] H. Stratigopoulos, P. Drineas, M. Slamani, and Y. Makris, "RF Specification Test Compaction Using Learning Machines," *Transactions on VLSI Systems*, 2009.
- [49] M. J. Barragan, H. Stratigopoulos, S. Mir, H. Le Gall, N. Bhargava, and A. Bal, "Practical Simulation Flow for Evaluating Analog and Mixed-Signal Test Techniques," *IEEE Design & Test*, vol. PP, no. 99, pp. 1–1, 2016.
- [50] A. Singhee and R. Rutenbar, "Statistical Blockade: Very Fast Statistical Simulation and Modeling of Rare Circuit Events and Its Application to Memory Design," *Trans. on Computer-Aided Design of Int. Circ. and Syst.*, vol. 28, no. 8, pp. 1176–1189, Aug. 2009.
- [51] H. Stratigopoulos, M. J. Barragan, S. Mir, H. Le Gall, N. Bhargava, and A. Bal, "Evaluation of low-cost mixed-signal test techniques for circuits with long simulation times," in *International Test Conference*, Oct. 2015, pp. 1–7.
- [52] H. Stratigopoulos, "Test Metrics Model for Analog Test Development," *Trans. on Computer-Aided Design of Int. Circ. and Syst.*, vol. 31, no. 7, pp. 1116–1128, 2012.
- [53] G. Leger and M. J. Barragan, "Questioning the reliability of Monte Carlo simulation for machine learning test validation," in *European Test Symposium*, May 2016, pp. 1–6.
- [54] J. Liaperdos, H. Stratigopoulos, L. Abdallah, Y. Tsiatouhas, A. Arapoyanni, and X. Li, "Fast deployment of alternate analog test using Bayesian model fusion," in *Design, Automation & Test in Europe*, 2015, pp. 1030–1035.
- [55] T. Das, A. Gopalan, C. Washburn, and P. R. Mukund, "Self-calibration of input-match in RF front-end circuitry," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 12, pp. 821–825, 2005.
- [56] S. Sun, F. Wang, S. Yaldiz, X. Li, L. Pileggi, A. Natarajan, M. Ferriss, J. Plouchart, B. Sadhu, B. Parker, A. Valdes-Garcia, M. A. T. Sanduleanu, J. Tierno, and D. Friedman, "Indirect Performance Sensing for On-Chip Self-Healing of Analog and RF Circuits," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 8, pp. 2243–2252, 2014.

- [57] V. Natarajan, R. Senguttuvan, S. Sen, and A. Chatterjee, "ACT: Adaptive Calibration Test for Performance Enhancement and Increased Testability of Wireless RF Front-Ends," in *26th IEEE VLSI Test Symposium (vts 2008)*, 2008, pp. 215–220.
- [58] D. Han, B. S. Kim, and A. Chatterjee, "DSP-Driven Self-Tuning of RF Circuits for Process-Induced Performance Variability," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 2, pp. 305–314, 2010.
- [59] N. Kupp, H. Huang, P. Drineas, and Y. Makris, "Post-production performance calibration in analog/RF devices," in *2010 IEEE International Test Conference*, 2010, pp. 1–10.
- [60] N. Kupp, H. Huang, Y. Makris, and P. Drineas, "Improving Analog and RF Device Yield through Performance Calibration," *IEEE Design & Test of Computers*, vol. 28, no. 3, pp. 64–75, May 2011.
- [61] Y. Lu, K. Subramani, H. Huang, N. Kupp, and Y. Makris, "Silicon Demonstration of Statistical Post-Production Tuning," in *2015 IEEE Computer Society Annual Symposium on VLSI*, July 2015, pp. 628–633.
- [62] M. Andraud, H. Stratigopoulos, and E. Simeu, "One-Shot Non-Intrusive Calibration Against Process Variations for Analog/RF Circuits," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 11, pp. 2022–2035, Nov 2016.
- [63] N. Barabino and F. Silveira, "Digitally Assisted CMOS RF Detectors With Self-Calibration for Variability Compensation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 5, pp. 1676–1682, 2015.
- [64] G. J. Székely and M. L. Rizzo, "Brownian distance covariance," *The Annals of Applied Statistics*, vol. 3, no. 4, pp. 1236–1265, 2009.
- [65] S. C. Cripps, *RF Power Amplifiers for Wireless Communications, Second edition*. Artech House, 2006.
- [66] A. Serhan, "Conception et réalisation de fonctions millimétriques en technologie BiCMOS 55nm," Ph.D. dissertation, Université Grenoble Alpes, 2015.
- [67] D. Pozar, *Microwave Engineering, Third edition*. Wiley, 2005, pp. 229 – 232.
- [68] A. Scuderi, T. Biondi, E. Ragonese, and G. Palmisano, "A lumped scalable model for silicon-integrated spiral inductors," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 6, pp. 1203–1209, 2004.
- [69] A. Bautista, A. Franc, and P. Ferrari, "Accurate parametric electrical model for slow-wave cpw and application to circuits design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 12, pp. 4225–4235, 2015.
- [70] J. P. Mondal, "An experimental verification of a simple distributed model of mim capacitors for mmic applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 35, no. 4, pp. 403–408, 1987.
- [71] M. V. Schneider, "Dielectric loss in integrated microwave circuits," *The Bell System Technical Journal*, vol. 48, no. 7, pp. 2325–2332, 1969.
- [72] E. O. Hammerstad, "Equations for microstrip circuit design," in *1975 5th European Microwave Conference*, September 1975, pp. 268–272.

- [73] R. Igreja and C. J. Dias, "Analytical evaluation of the interdigital electrodes capacitance for a multi-layered structure," *Sensors and Actuators A: Physical*, vol. 112, no. 2, pp. 291–301, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924424704000779>
- [74] H. A. Wheeler, "Transmission-Line Properties of a Strip on a Dielectric Sheet on a Plane," *IEEE Transactions on Microwave Theory and Techniques*, vol. 25, no. 8, pp. 631–647, 1977.
- [75] E. Hammerstad and O. Jensen, "Accurate Models for Microstrip Computer-Aided Design," in *1980 IEEE MTT-S International Microwave symposium Digest*, 1980, pp. 407–409.
- [76] T. Quemerais, "Conception et Etude de la Fiabilité des Amplificateurs de Puissance Fonctionnant aux Fréquences Millimétriques en Technologies CMOS Avancées," Ph.D. dissertation, Université de Grenoble, 2010.
- [77] D. Pozar, *Microwave Engineering, Third edition*. Wiley, 2005.
- [78] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1, pp. 342–359, 2002.
- [79] A. Antonopoulos, G. Volanis, Y. Lu, and Y. Makris, "Post-Production Calibration of Analog/RF ICs: Recent Developments and A Fully Integrated Solution," in *IEEE Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, July 2019, pp. 1–4.

Appendices

Annex 1 - OCEAN script used to generate the data following the full factorial DOE paradigm.

```

;Loop to simulate 144 tuning knobs combinations.
for(i 1 144

;===== Set to XL mode
ocnSetXLMode()
ocnxlProjectDir( "simulations directory path" )
ocnxlTargetCellView( "Library_name" "Cell_name" "adexl" )
ocnxlResultsLocation( "" )
ocnxlSimResultsLocation( "" )
ocnxlMaxJobFail( 20 )

ocnxlSession = ocnxlGetSession()

;===== Tests setup

;Test, results and expressions to be saved.
;Automatically generated by exporting the ADE setup as an OCEAN XL
script.
ocnxlBeginTest("Library_name:Cell_name:1")
simulator( 'spectre )
design( "Library_name" "Cell_name" "config")
modelFile(
    '("Monte Carlo model file path (*.scs file)" "" )

analysis('hb ?HBTranMode "No" ?tstab "0" ?saveinit ""
        ?autoharms "" ?autotstab "" ?oversample
list("1") ?funds list("RF1")
        ?maxharms list("3") ?errpreset "liberal" ?
xdbccompression "yes" ?xdbrapid "no"
        ?xdbgain "power" ?xdbref "linear" ?xdblevel
list("1") ?xdbharm list("1")
        ?xdbsource "/PORT0" ?xdbload "/PORT1" )
analysis('pss ?engine "Harmonic Balance" ?flexbalance "yes" ?
oversamplefactor ""
        ?fund "60G" ?harms "3" ?errpreset "liberal" ?
param "prf1"
        ?start "4" ?stop "8" ?sweepStep "1" ?values
list("15")
        ?step "" ?write "" ?writefinal "" ?swapfile ""
        ?checkpss "" ?backtracking "" )
analysis('dc ?save0ppoint t )
analysis('sp ?ports list("/PORT0" "/PORT1") ?freq "60G" ?start
"" ?stop "" )

desVar("Variable_name" value)
...
desVar("Variable_name" value)

```

```

envOption(
    'analysisOrder list("dc" "sp" "pss" "psp" "pxf" "pnoise"
    "pstb" "pac" "hb" "hbxf" "hbnoise" "hbstb" "hbac" "pz" "dcmatch"
    "stb" "tran" "envlp" "ac" "lf" "noise" "xf" "qpss" "qpac" "qpnoise"
    "qpxf" "qpssp" "sens" "acmatch")
)
save( 'i "current_to_save_1" "current_to_save_2" ...
"current_to_save_n" )
temp( 27 )

ocnxlOutputTerminal("signal_to_save" ?save t)
    ...
ocnxlOutputTerminal("signal_to_save" ?save t)

ocnxlOutputExpr("expr_to_save" ?name "Expression_name" ?plot t)
    ...
ocnxlOutputExpr("expr_to_save" ?name "Expression_name" ?plot t)

ocnxlEndTest() ; "Library_name:Cell_name:1"

;===== Sweeps setup

;Reading the file that contains the tuning knobs combinations.
file_to_read = infile("sweep_TK_144.txt")
for(j 1 i
fscanf(file_to_read "%s %s %s %s %s %s" a b c d e f ))
close(file_to_read)

;Changing the values of the variables corresponding to the tuning
knobs
ocnxlSweepVar("TK1" a)
ocnxlSweepVar("TK2" b)
ocnxlSweepVar("TK3" c)
ocnxlSweepVar("TK4" d)
ocnxlSweepVar("TK5" e)
ocnxlSweepVar("TK6" f)

;Even variables that will not change must be declared the same way.
ocnxlSweepVar("variable_name" "value")
    ...
ocnxlSweepVar("variable_name" "value")

;===== Checks and Asserts setup
ocnxlPutChecksAsserts(?netlist nil)

;===== Job setup
ocnxlJobSetup( '(
    "blockemail" "1"
    "configuretimeout" "300"
    "distributionmethod" "Local"
    "lingertimeout" "300"
    "maxjobs" "8"
    "name" "ADE XL Default"
    "preemptivestart" "1"

```

```

        "reconfigureimmediately" "1"
        "runtimeout" "-1"
        "showerrorwhenretrying" "1"
        "showoutputlogerror" "0"
        "startmaxjobsimmed" "1"
        "starttimeout" "300"
        "usesameprocess" "1"
    ) )

;===== Run Mode Options
;Setting up the complete Monte Carlo simulation: 35 is the number of
Monte Carlo instances in this example
ocnxlMonteCarloOptions( ?mcMethod "global" ?mcNumPoints "35" ?
mcNumBins "" ?mcStopEarly "0" ?mcStopMethod "Significance Test" ?
samplingMode "random" ?saveProcess "1" ?saveMismatch "0" ?
useReference "0" ?donominal "1" ?saveAllPlots "0" ?monteCarloSeed
"" ?mcStartingRunNumber "" ?dumpParamMode "yes" )

;===== Run command

;Actually running the simulation
ocnxlRun( ?mode 'monteCarlo ?nominalCornerEnabled t ?
allCornersEnabled t ?allSweepsEnabled t)
ocnxlOutputSummary(?yieldSummary t ?exprSummary nil ?specSummary
nil ?detailed nil)
ocnxlOpenResults()

historyName_id = ocnxlGetCurrentHistory()

;For debug purposes
;printf("\nThe current session name is: %s\n" ocnxlSession)
;printf("\nThe current history name is: %s\n" historyName_id)

;Saving the results
axlWriteMonteCarloResultsCSV(ocnxlSession, historyName_id)

;===== End XL Mode command
ocnxlEndXLMode()

;Closing the FOR loop
)

```


Résumé

Le marché des systèmes sur puces (SoC) ou en package (SiP) jouit d'une croissance continue, tiré par des secteurs comme l'automobile, le biomédical ou l'industrie 4.0. Le développement de ces domaines est notamment rendu possible par des technologies CMOS avancées qui permettent l'intégration de fonctions complexes et à hautes fréquences.

Les technologies silicium récentes sont particulièrement enclines aux imperfections durant la fabrication des circuits. La variation des procédés est la source principale de circuits non-fonctionnels. Des défauts catastrophiques localisés (circuits ouverts, courts-circuits) peuvent aussi survenir et se traduisent la plupart du temps par un circuit défaillant.

Les coûts de fabrication des circuits intégrés ont fortement baissé au cours des dernières décennies. Cependant, le test des circuits fabriqués dans la chaîne de production reste un point bloquant particulièrement notable. En effet, l'intégration de différents types de fonctions (numériques, mixtes, analogiques, mémoires, RF, . . .) complexifie l'accès aux nœuds internes du circuit. De plus, le test des circuits analogiques, RF et millimétriques est typiquement basé sur l'évaluation de leurs spécifications. Le coût des appareils de mesure ainsi que les temps de test représentent ainsi une grande part du coût total et de nombreux défis restent à relever dans le domaine du test.

Afin de réduire le coût du test, différentes approches ont été proposées dans la littérature. Ces méthodes proposent en particulier de remplacer les approches conventionnelles pour alléger les contraintes sur les équipements de test. Le Built-in test ou Built-in Self Test (BIST) est un type d'approche qui propose d'embarquer des éléments de test au sein même de la puce. Parmi les méthodes de BIST les plus pertinentes, on peut notamment citer des approches intrusives faisant appel à des capteurs de tension DC, de courant, à des détecteurs de tension/puissance ou encore à des boucles de type loopback. Bien que pertinentes, ces méthodes ne sont pas génériques et nécessitent de co-concevoir le circuit et les éléments de test, ce qui peut être particulièrement limitant pour les circuits millimétriques.

Le test indirect basé sur de l'apprentissage automatique propose de remplacer la mesure directe et onéreuse des performances par l'évaluation simple de signatures à basse fréquence. Illustrée par la Figure A, la méthode se déroule en deux temps : une phase d'apprentissage et une phase de test. Lors de la première étape, les signatures ainsi que les performances directes sont mesurées sur un ensemble réduit d'échantillons de façon à construire un modèle de régression qui effectue la correspondance entre les signatures et les performances du Dispositif Sous Test (DST). Dans un second temps, pour les circuits sur la chaîne de production, les signatures sont mesurées et le modèle est utilisé afin de déduire les performances des circuits sans avoir à les mesurer directement.

Un autre paradigme de test particulièrement approprié aux circuits millimétriques repose sur l'utilisation de circuits dummy. Il s'agit de structures simples qui copient certains éléments du circuit à tester et qui sont positionnées proches de ce dernier. Ainsi, une variation des procédés de fabrication impacte de façon similaire les éléments dummy et le circuit lui-même. Les structures dummy utilisées pour le test sont ainsi des capteurs de variation des procédés de fabrication. En effectuant des mesures de signatures simples sur ces structures, on peut alors obtenir des informations pertinentes sur la qualité des procédés de fabrication et donc par conséquent sur les performances du circuit. On parle de test indirect non-intrusif et la Figure B en montre le schéma de principe. Cette approche non-intrusive peut ainsi compléter efficacement une méthode de test indirect basée sur de l'apprentissage automatique.

L'utilisation du test indirect non-intrusif est particulièrement adaptée aux circuits millimétriques du fait que les noeuds internes du circuit ne sont pas chargés. Néanmoins, la mise en œuvre de la méthode dépend fortement de l'expertise du concepteur et il n'existe pas de méthode systématique visant à définir le jeu de signatures. Dans cette thèse nous proposons une solution à cette absence d'automatisation.

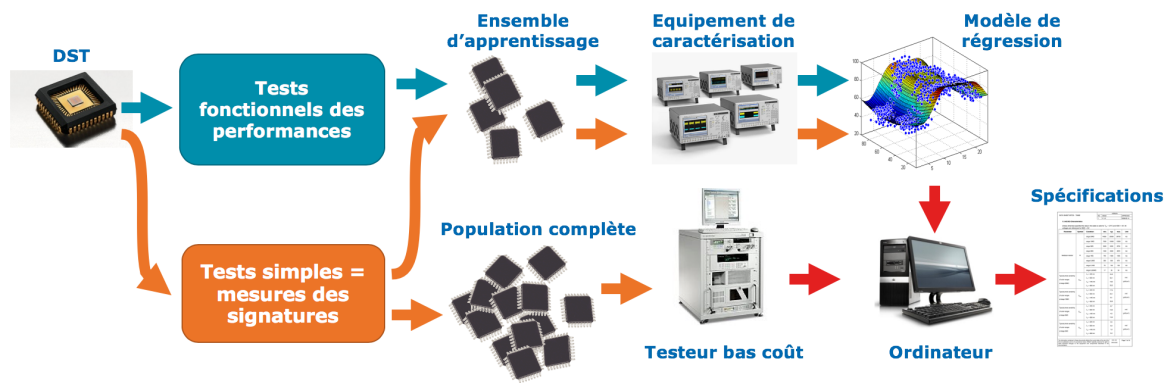


FIGURE A: Principe du test indirect basé sur de l'apprentissage automatique.

En plus de tester les circuits, des méthodes de calibrage peuvent également être implémentées de façon à améliorer les performances des dispositifs sous-optimaux. Ces méthodes proposent d'intégrer des éléments variables au sein du circuit afin de modifier son comportement après fabrication. La particularité des méthodes dites *one-shot* est de proposer une combinaison de valeurs pour les éléments variables à partir d'une seule série de mesures. Ainsi ces méthodes ne requièrent pas d'itération mesures/paramétrisation du circuit. On peut ainsi envisager d'utiliser des signatures mesurées sur des capteurs de variation des procédés de fabrication pour prédire le meilleur réglage du circuit. Dans ce sens, les méthodes de calibrage *one-shot* sont un choix évident pour compléter une stratégie de test indirect non-intrusive utilisant de l'apprentissage automatique.

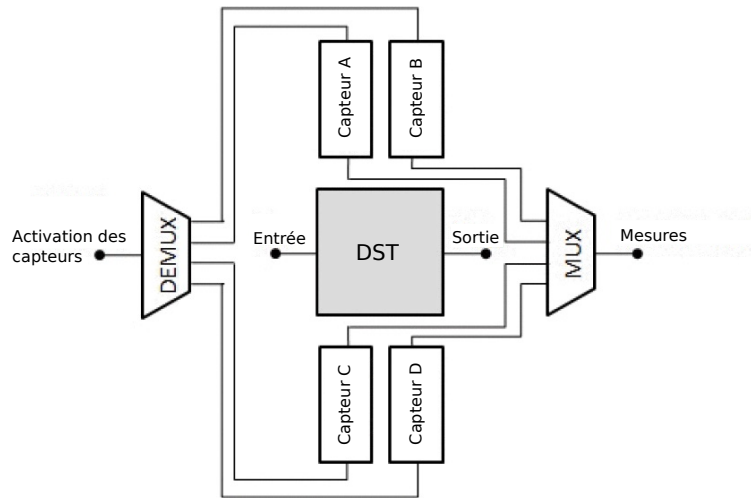


FIGURE B: Schéma de principe du test indirect non-intrusif utilisant des capteurs sensibles aux procédés de fabrication [37].

Méthode générique de test indirect non-intrusif basé sur l'apprentissage automatique

Dans le cadre de ces travaux de thèse, nous avons développé une méthode générique permettant d'implémenter un test indirect non-intrusif basé sur l'apprentissage automatique. Comme nous l'avons déjà souligné, la proposition des capteurs et des signatures est un élément limitant lors de l'utilisation de cette technique. Les signatures sont typiquement proposées par le concepteur du circuit parmi l'infini des possibilités qui s'offrent à lui. Seule sa connaissance pointue du fonctionnement du circuit guide son choix. Nous nous proposons donc de répondre au problème suivant : connaissant la netlist d'un DST et ses spécifications cibles, comment systématiser la proposition d'un jeu de signatures et la conception des capteurs non-intrusifs associés ?

La méthode que nous proposons de suivre se déroule en deux étapes :

- Etape 1 : Mise en évidence des causes des variations des performances.
- Etape 2 : Conception assistée de capteurs non-intrusifs de la qualité des procédés de fabrication.

En explorant l'espace des paramètres MC, on peut trouver les causes des variations des performances. Cette exploration pouvant être vue comme un problème de sélection, on propose d'adapter l'algorithme de sélection proposé dans [46]. Pour une performance cible donnée, cet algorithme est utilisé afin d'obtenir une liste triée des paramètres MC qui vont le plus impacter la performance du DST. A chaque itération de l'algorithme, un modèle de regression est entraîné afin de prédire la performance à partir de la seule connaissance des paramètres MC pertinents. L'étape suivante consiste à s'intéresser au sens physique de chaque paramètre MC sélectionné afin de proposer un capteur et une signature associée pertinente qui soit sensible à ce paramètre en particulier.

Afin de valider cette méthode sur un cas d'étude, un amplificateur de puissance (PA) à 60 GHz a été conçu en technologie CMOS 55 nm de ST Microelectronics. Le résultat de la première étape est affiché Figure C pour le gain S_{21} de l'amplificateur. On peut voir l'évolution de l'erreur de la prediction du

gain diminuer au-fur-et-à mesure que l’algorithme sélectionne des paramètres MC pertinents. Au final, 18 paramètres MC ont été sélectionnés pour construire un modèle $f_{MC \rightarrow S_{21}} : [MC_1, \dots, MC_{18}] \rightarrow S_{21}$ avec une erreur de prédiction de 0,05 dB.

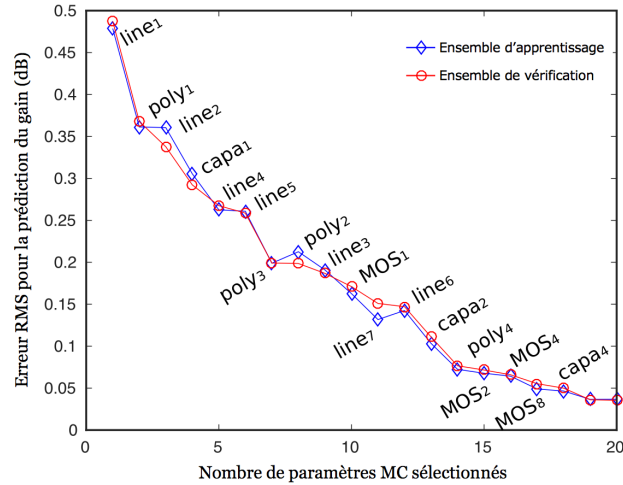


FIGURE C: Evolution de l’erreur RMS de prédiction du gain en fonction du nombre de paramètres MC sélectionnés par l’algorithme.

L’étape suivante consiste à proposer un ensemble pertinent de capteurs et de signatures. Cette étape aboutit à un jeu de 5 capteurs (lignes de transmission basse et haute impédance, capacité MOM, capteur de capacité jonction, transistor MOS) et de 9 signatures DC ou basse fréquence. Finalement, en utilisant ces 9 signatures S_1 à S_9 , on parvient à construire un modèle de prédiction du gain de l’amplificateur : $f_{S \rightarrow S_{21}} : [S_1, \dots, S_9] \rightarrow S_{21}$. La Figure D affiche le gain prédit en fonction du gain réel. L’erreur RMS en utilisant les signatures est 0,07 dB, ce qui est très comparable à l’erreur RMS en utilisant les 18 paramètres MC (0,05 dB) et très inférieur à l’écart-type pour le gain qui est de 0,69 dB. La méthode complète peut ensuite être appliquée pour les autres performances cibles du PA. De façon intéressante, l’étude complète n’apporte pas de nouveaux capteurs ou signatures et le jeu de 9 signatures suffit à prédire toutes les performances avec une très bonne précision.

Nous avons ainsi développé une méthode systématique pour concevoir un programme de test non-intrusif pour circuits millimétriques. La méthode exploite les modèles Monte Carlo et ne requiert aucune connaissance spécifique du DST. La méthode assiste le concepteur dans la proposition des capteurs de variation des procédés de fabrication et des signatures associées. La limitation principale de la méthode est qu’elle nécessite des composants avec des modèles Monte Carlo. La partie suivante de ces travaux s’attache alors à étendre l’approche pour la rendre applicable aux circuits utilisant des composants passifs *full custom*, comme c’est usuellement le cas en conception millimétrique.

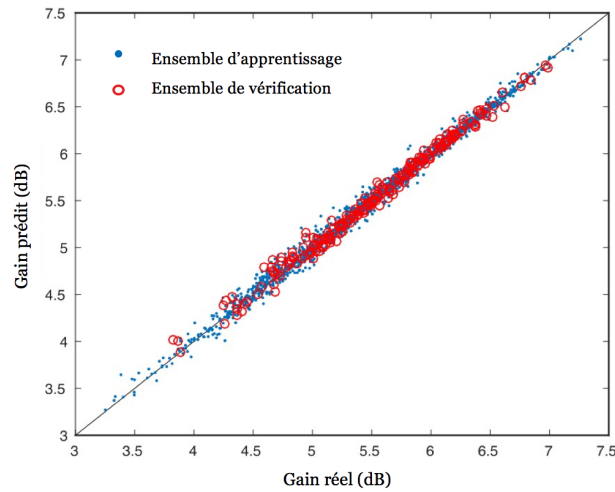


FIGURE D: Gain prédit par le modèle versus gain réel de l'amplificateur.

Méthode générique étendue

Afin d'étendre la méthode de test indirect non-intrusif proposée, nous nous intéressons aux circuits utilisant des composants passifs *full custom* (ligne, balun, coupleur, ...). Nous proposons de décomposer cette méthode selon les deux mêmes étapes que précédemment, avec une spécificité pour les composants *full custom*.

- Etape 1 : Mise en évidence des causes des variations des performances.
 - Pour les composants standards avec modèles Monte Carlo, on utilise l'algorithme de sélection.
 - Pour les composants passifs *full custom* on effectue une étude de sensibilité.
- Etape 2 : Conception assistée de capteurs non-intrusifs de la qualité des procédés de fabrication.

Afin de mettre en évidence les causes des variations des performances, distinguons deux cas selon le type de composants. Pour les composants standards ayant des modèles Monte Carlo, on propose de remplacer les composants passifs *full custom* par des boîtes de paramètres S typiques. On peut alors utiliser l'algorithme de sélection et tirer ainsi parti de la méthode précédente. On obtient alors une liste triée des paramètres MC les plus pertinents. Concernant les composants *full custom*, on va chercher à identifier à quelles variations de caractéristique du Back End Of Line (BEOL) le DST est sensible. Pour ce faire, nous proposons d'évaluer les performances du DST en faisant varier une à une les caractéristiques du BEOL et en remplaçant les boîtes de paramètres S typiques par celles du corner considéré. De cette analyse de sensibilité on tire une liste triée des caractéristiques du BEOL qui peuvent dégrader le plus significativement les performances du DST. On peut ensuite procéder à l'étape de conception des capteurs. Dans cette optique, on s'appuie sur le sens physique des paramètres MC et sur les caractéristiques clés du BEOL pour proposer un jeu de capteurs et de signatures adéquat. Cette méthode est résumée par le schéma Figure E.

Cette méthode générique étendue est appliquée au PA précédemment étudié. La différence majeure étant que les lignes microruban ont été simulées avec le simulateur électromagnétique Keysight

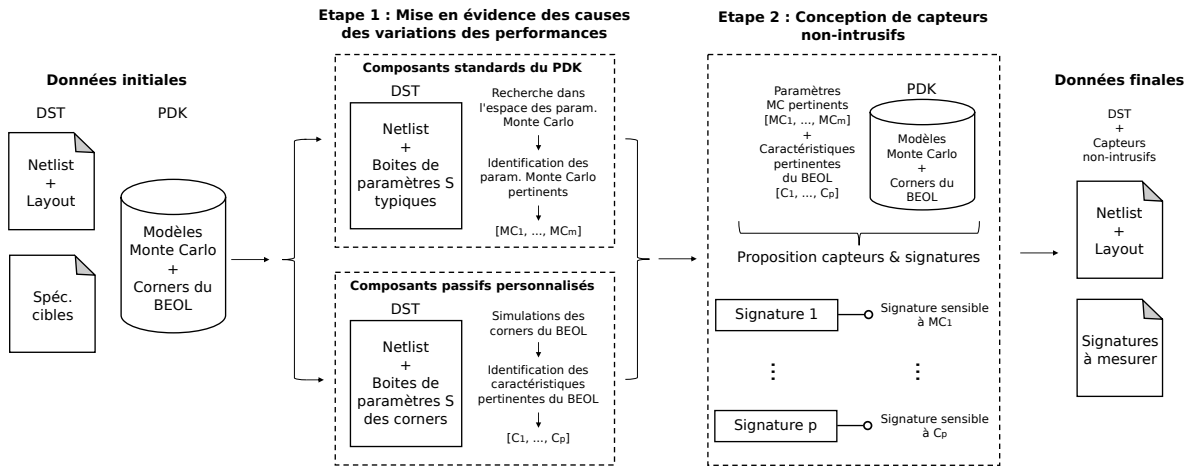


FIGURE E: Principe de la méthode de test générique étendue proposée.

Momentum. Afin de réduire le nombre de corners à simuler, nous avons mis à profit les informations fournies par le fondeur ainsi que les équations théoriques des lignes de transmission proposées dans [71, 72, 74, 75]. Au final, nous avons identifié 3 caractéristiques du BEOL pertinentes pour les performances du PA : l'épaisseur de diélectrique sous les lignes, l'épaisseur et la conductivité du ruban signal. Cela résulte en seulement 7 corners : le typique, le minimum et maximum des 3 caractéristiques identifiées. La liste des paramètres MC pertinents a également été obtenue en utilisant l'algorithme de sélection. Comparée au cas d'étude précédent, l'étape de conception des capteurs a abouti à la même proposition concernant les composants standards et à des signatures similaires sur les mêmes structures de lignes, pour un total de 5 capteurs et de 7 signatures S_1 à S_7 .

Le PA ainsi que les capteurs ont été fabriqués et 22 échantillons ont été mesurés. La Figure F montre le schéma du banc de test ainsi qu'une photographie d'un échantillon fabriqué. Les capteurs sont marqués par les lettres A à E. Grâce aux mesures réalisées, nous avons construit un modèle de regression permettant de prédire les différentes performances, en utilisant seulement la connaissance des signatures. La Figure G montre les résultats correspondants à la prédiction du gain par la fonction $f_{S \rightarrow S_{21}} : [S_1, \dots, S_7] \rightarrow S_{21}$.

Nous avons ainsi proposé une méthode systématique pour concevoir un programme de test non-intrusif pour circuits millimétriques. La technique exploite les modèles Monte Carlo ainsi que les informations de fabrication du BEOL et ne requiert aucune connaissance spécifique du DST. Malgré un nombre très réduit d'échantillons fabriqués, les résultats expérimentaux obtenus sur un PA millimétrique sont très encourageants et mettent en lumière une très forte corrélation entre le test bas coût proposé et la mesure directe des spécifications.

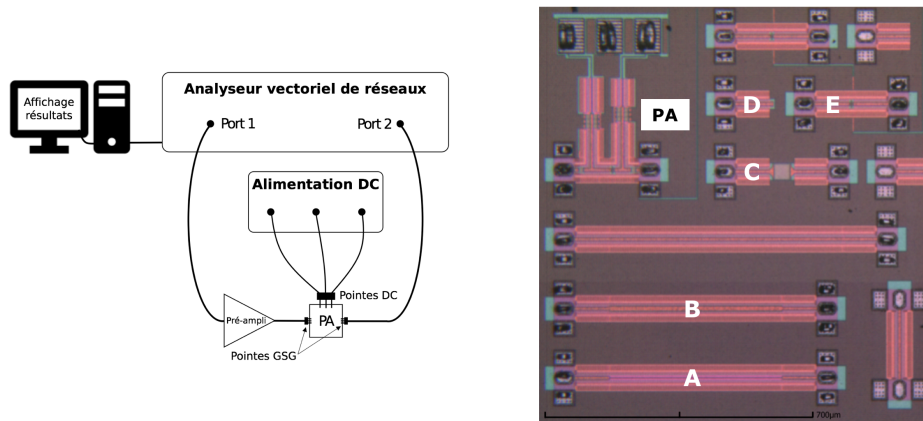


FIGURE F: Schéma du banc de test et photographie d'un échantillon fabriqué.

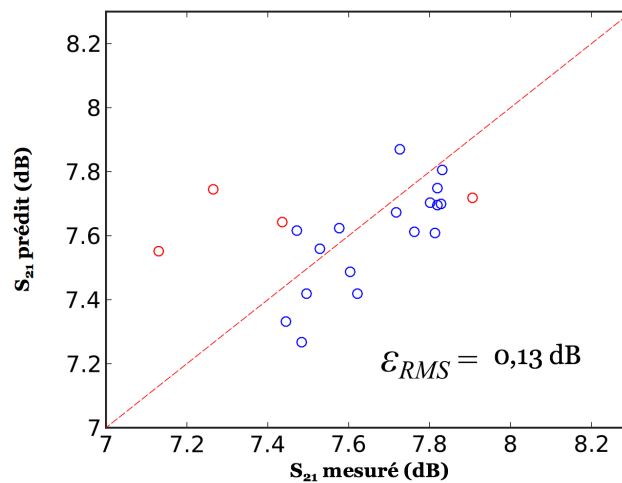
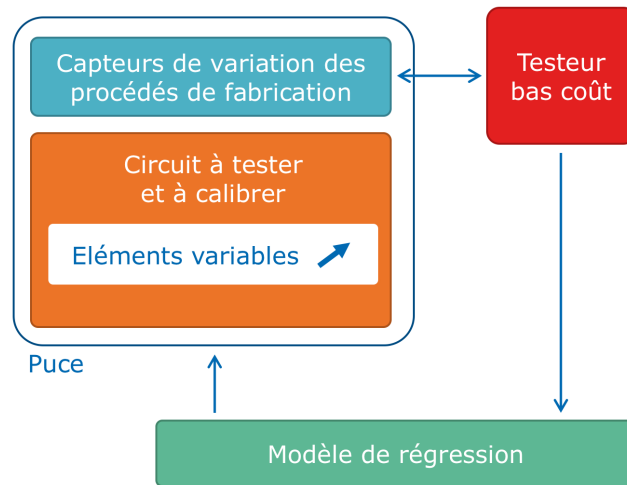


FIGURE G: Gain prédit versus gain mesuré.

Méthode de calibrage *one-shot*

Les informations délivrées par les capteurs de variation des procédés de fabrication ont permis la mise en place d'un test bas-coût. Cependant, elles peuvent également être exploitées afin d'implémenter une technique de calibrage visant à compenser une qualité de fabrication sub-optimale. Dans le cadre de cette thèse, nous nous sommes intéressé à une méthode non-itérative dite *one-shot* et qui re-exploite les signatures afin de calibrer le DST. La Figure H présente le principe de cette méthode.

Un PA à deux étages à 60 GHz intégrant des éléments variables a été conçu dans cette optique. Cet amplificateur peut être calibré après fabrication en utilisant 6 éléments variables : 4 cellules de découplage conçues à base de varactors ainsi que 2 tensions de polarisation de grille. L'approche détaillée dans la première partie a été utilisée afin d'obtenir un jeu de capteurs non-intrusifs et de signatures à partir desquelles les performances du DST seront prédites. Au total, 6 capteurs ont été conçus avec 10 signatures associées. Il est particulièrement intéressant de remarquer que les capteurs

FIGURE H: Schéma de principe d'un calibrage *one-shot*.

sont les mêmes que ceux proposés pour le PA à un étage auquel est venu s'ajouter un varactor en standalone. La Figure I montre le layout de ce circuit avec ses structures de tests non-intrusives.

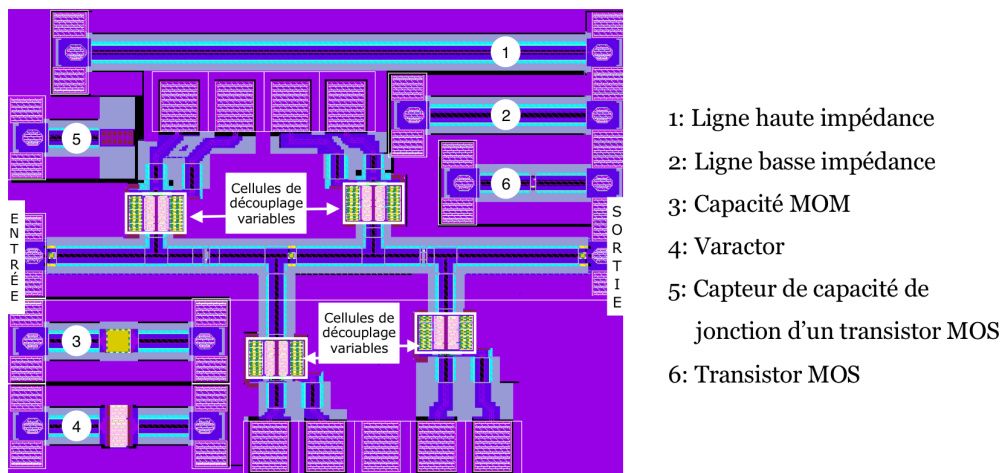


FIGURE I: Layout du PA avec possibilité de calibration ainsi que son jeu de capteurs non-intrusifs de variation des procédés de fabrication.

Afin de construire le modèle de régression permettant de prédire la meilleure combinaison d'éléments variables (Tuning Knobs en anglais, abrégé TK) pour calibrer le PA, nous avons effectué 5000 simulations MC en faisant varier de façon aléatoire à la fois les procédés de fabrication et les commandes des éléments variables. De cette façon nous avons entraîné la fonction de regression suivante : $f_{S,TK \rightarrow S_{21}} : [S_1, \dots, S_{10}, TK_1 \dots, TK_6] \rightarrow S_{21}$. Une fois ce modèle construit, il a été mis à profit pour prédire le gain d'un nouvel échantillon, et ce quel que soit les contrôles des éléments variables. On peut ainsi explorer l'espace des performances atteignables et choisir le meilleur réglage pour un échantillon donné. La Figure J met en lumière l'amélioration du rendement de conception (design yield) pour le

gain, sur un ensemble de 500 échantillons MC indépendants.

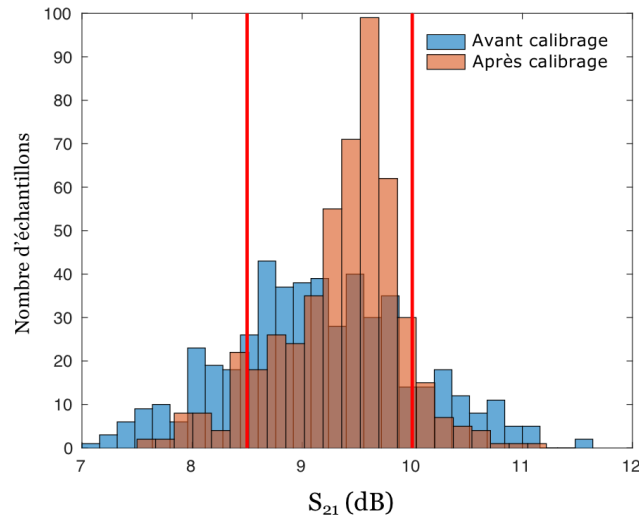


FIGURE J: Distribution du gain avant (bleu) et après (rouge) calibration. Les lignes verticales rouges marquent les spécifications cibles.

En re-utilisant la méthode de test non-intrusive que nous avons développée, nous avons implémenté avec succès un calibrage *one-shot* afin d'augmenter le rendement en conception d'un PA à deux étages à 60 GHz. Pour l'ensemble des spécifications fixées, ce dernier a été amélioré de 59 % à 83 %. Ce circuit a été fabriqué et les résultats de mesures seront investigués prochainement.

Conclusion

Dans ces travaux de thèse, nous avons proposé une méthode générique de test indirect non-intrusif. Cette méthode exploite les modèles Monte Carlo fournis par le fondeur afin de trouver les causes fondamentales des variations des performances d'un dispositif. La méthode propose également de concevoir de façon assistée un jeu de capteurs de variation des procédés de fabrication ainsi que les signatures associées. Cette approche a été validée sur un PA à un étage à 60 GHz. Sa principale limitation est qu'elle nécessite de disposer de modèles Monte Carlo, ce qui n'est typiquement pas le cas pour des composants passifs *full custom*. Nous avons alors travaillé à une extension de cette méthode afin de la rendre utilisable sur n'importe quel circuit intégré. Des mesures sur une vingtaine d'échantillons sont venues valider cette approche. En plus de quantifier les performances, nous sommes intéressés à des méthodes de calibrage. Dans ce sens, nous avons implémenté un calibrage *one-shot* sur un PA à deux étages à 60 GHz. Pour calibrer le dispositif, nous avons tiré parti d'une cellule de découplage variable.

Afin de poursuivre dans cette thématique de recherche, il serait intéressant de proposer de nouvelles architectures de capteurs et de mettre en place une méthode de test indirect non-intrusif pour d'autres circuits, à plus hautes fréquences et utilisant des composants *full custom* complexes. L'étude d'un émetteur-récepteur complet pourrait également être une perspective intéressante afin de se focaliser sur de nombreuses et diverses performances et ainsi valider d'autant mieux cette méthode.

Titre : Développement de solutions BIST (Built-In Self-Test) pour circuits intégrés radiofréquences/millimétriques

Résumé : Les technologies silicium récentes sont particulièrement enclines aux imperfections durant la fabrication des circuits. La variation des procédés peut entraîner une dégradation des performances, notamment aux hautes fréquences. Dans cette thèse, plusieurs contributions visant la réduction des coûts et de la complexité du test des circuits millimétriques sont présentées. Dans ce sens, deux sujets principaux ont fait l'objet de notre attention : a) le test indirect non-intrusif basé sur l'apprentissage automatique et b) la calibration non-itérative "one-shot". Nous avons en particulier développé une méthode générique pour implémenter un test indirect non-intrusif basé sur l'apprentissage automatique. La méthode vise à être aussi automatisée que possible de façon à pouvoir être appliquée à pratiquement n'importe quel circuit millimétrique. Elle exploite les modèles Monte Carlo du design kit et des informations de variations du BEOL pour proposer un jeu de capteurs non-intrusifs. Des mesures à basses fréquences permettent ensuite d'extraire des signatures qui contiennent des données pertinentes concernant la qualité des procédés de fabrication, et donc a fortiori de la performance du circuit. Cette méthode est supportée par des résultats expérimentaux sur des PAs fonctionnant à 65 GHz, conçus dans une technologie 55 nm de STMicroelectronics. Pour s'attaquer plus encore à la dégradation des performances induite par les variations des procédés de fabrication, nous nous sommes également penchés sur une procédure de calibration non-itérative. Nous avons ainsi présenté un PA à deux étages qui peut être calibré en post-fabrication. La méthode de calibration exploite une cellule de découplage variable comme moyen de modifier les performances de l'amplificateur. Des moniteurs de variation des procédés de fabrication, placés dans les espaces vides du circuit, sont utilisés afin de prédire la meilleure configuration possible pour les cellules de découplage variables. La faisabilité et les performances de cette approche ont été validés en simulation.

Mots-clés : BIST, apprentissage automatique, test indirect, ondes millimétriques, circuits intégrés, calibration

Title: Development of Built-In Self-Test solutions for RF/mm-wave integrated circuits

Abstract: Recent silicon technologies are especially prone to imperfections during the fabrication of the circuits. Process variations can induce a noticeable performance shift, especially for high frequency devices. In this thesis we present several contributions to tackle the cost and complexity associated with testing mm-wave ICs. In this sense, we have focused on two main topics: a) non-intrusive machine learning indirect test and b) one-shot calibration. We have in particular developed a generic method to implement a non-intrusive machine learning indirect test based on process variation sensors. The method is aimed at being as automated as possible and can be applied to virtually any mm-wave circuit. It leverages the Monte Carlo models of the design kit and the BEOL variability information to propose a set of non-intrusive sensors. Low frequency measurements can be performed on these sensors to extract signatures that provide relevant information about the process quality, and consequently about the device performance. The method is supported by experimental results in a set of 65 GHz PAs designed in a 55 nm technology from STMicroelectronics. To further tackle the performance degradation induced by process variations, we have also focused on the implementation of a one-shot calibration procedure. In this line, we have presented a two-stage 60 GHz PA with one-shot calibration capability. The proposed calibration takes advantage of a novel tuning knob, implemented as a variable decoupling cell. Non-intrusive process monitors, placed within the empty spaces of the circuit, are used for predicting the best tuning knob configuration based on a machine learning regression model. The feasibility and performance of the proposed calibration strategy have been validated in simulation.

Keywords: BIST, machine learning, indirect test, millimeter wave, integrated circuits, calibration

Thèse préparée au Laboratoire TIMA / Thesis prepared at TIMA Laboratory
Techniques de l'Informatique et de la Microélectronique pour l'Architecture des ordinateurs
Techniques of Informatics and Microelectronics for integrated systems Architecture
46 avenue Félix Viallet – 38031 GRENOBLE Cedex - France

ISBN: 978-2-11-129262-8