Built-in test in RF circuits using non-intrusive sensors

Athanasios Dimakos

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Nano Electronique et Nano Technologies (NENT)**

Arrêté ministériel : 7 août 2006

Présentée par

Athanasios DIMAKOS

Thèse dirigée par **Haralampos-G STRATIGOPOULOS** et
codirigée par **Salvador MIR**

préparée au sein du **Laboratoire Techniques de l'Informatique et de la Microélectronique pour l'Architecture des systèmes intégrés (TIMA)**
dans **l'École Doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (E.E.A.T.S.)**

Test embarqué des circuits RF en utilisant des capteurs non-intrusifs

Thèse soutenue publiquement le **29 Mars 2016**,
devant le jury composé de :

M. Philippe DESCAMPS
Professeur, ENSICAEN/LaMIPS, Rapporteur
M. Hassan ABOUSHADY
Maître de conférences, UPMC/LIP6, Rapporteur
M. Philippe FERRARI
Professeur, UGA/IMEP-LACH, Examinateur
M. Matthias BUCHER
Associate Professor, Technical University of Crete, Examinateur
M. Emeric DE FOUCAULD
Ingénieur Senior, CEA-LETI, Invité
M. Salvador MIR
Directeur de recherche, CNRS/TIMA, Co-directeur de thèse
M. Haralampos-G. STRATIGOPOULOS
Chargé de recherche, CNRS/LIP6, Directeur de thèse

ACKNOWLEDGEMENTS

This work has been carried out under financial support of European Library-based flow of Embedded Silicon test Instruments (ENIAC ELESIS), Solutions for self Adaptation of Communicating Systems in Operation (ANR SACSO) and P-SOC CNRS/INS2I projects.

In the following paragraphs, I would like to acknowledge all the people, who helped me, each of them in his/her own unique way, to accomplish my thesis. First of all, I would like to acknowledge my supervisors Haralampos-G. STRATIGOPOULOS and Salvador MIR for the assignment of this thesis, for their continuous support and guidance, and for their ideas that, in essence, have been implemented in this project.

Also, I would like to thank all the members of the jury for their participation in the seven-party committee, including Mr. Philippe DESCAMPS from LaMIPS laboratory (Caen, France), Mr. Hassan ABOUSHADY from LIP6 laboratory (Paris, France), Mr. Philippe FERRARI from IMEP-LAHC laboratory (Grenoble, France), Mr. Matthias BUCHER from Technical University of Crete (Chania, Hellas), and Mr. Emeric DE FOUCAULD from CEA-LETI laboratory (Grenoble, France).

Special thanks should be directed to some people whose contribution to the overall work was significant. Many thanks should be also directed to the Reliable Mixed-signal Systems (RMS) group fellow students and friends, Martin ANDRAUD and Guillaume RENAUD as well as to our collaborator Alexandre SILIGARIS from CEA-LETI laboratory.

Furthermore, I would also like to acknowledge with much appreciation the crucial role of the kindest person I have ever met, Alejandro CHAGOYA, who was always available to solve any technical problem.

Moreover, I would like to thank my friends, Tasos PAGONAS, Giannis PAPADEDES, Lefteris IOANNIDIS, Savvina PARTHENI, and Chrysa VERGITSAKI for their psychological support, especially in times of utter desperation.

Above all, I want to thank my beloved parents, Panagiotis and Alexandra, and my siblings, Tasos and Betty to whom I owe my principles and values, for their love, sacrifices, companion, patience and support during my academic years.

Τούς μέν κενούς ασκούς η πνοή διίστησι, τούς δ' ανοήτους, τό οίημα.
Σωκράτης, 469-399 π.Χ.

Empty windbags are fulled by the breath, conceit does the fools.
Socrates, 469-399 B.C.

Στους γονείς μου

ABSTRACT

This thesis addresses the high-volume production test problem for Radio Frequency (RF) and millimeter-wave (mm-wave) circuits. Testing the RF/mm-wave functions of systems-on-chips (SoCs) incurs a very high cost. Built-in test (BIT) is a promising alternative to facilitate testing and reduce costs, but it is challenging since it should by no means degrade the performance of the Circuit Under Test (CUT). In this work, we study a BIT technique which is based on non-intrusive variation-aware sensors. The non-intrusive property is very appealing for designers since the sensors are totally transparent to the design and, thereby, the test is completely dissociated from the design. The non-intrusive sensors are dummy analog stages and single layout components that are copied from the topology of the CUT and are placed on the die in close physical proximity to the CUT. They simply offer an "image" of process variations and by virtue of this they are capable of tracking variations in the performances of the CUT. In essence, the technique capitalizes on the undesired phenomenon of process variations. The alternate test paradigm is employed to map the outputs of the non-intrusive sensors to the performances of the CUT, in order to replace the standard tests for measuring the performances directly. The proposed test idea is applied to two different CUTs, namely a 2.4GHz 65nm CMOS inductive degenerated Low-Noise Amplifier (LNA) and a wide-band mm-wave 60GHz 65nm CMOS 3-stage LNA. We demonstrate that by adding on-chip a few non-intrusive sensors of practically zero area overhead and by obtaining on these non-intrusive sensors DC or low-frequency measurements, we are able to track variations in all performances of the CUT with an average prediction error lower than one standard deviation of the performance and a maximum prediction error that is lower or at least comparable to the measurement and repeatability errors in a conventional Automatic Test Equipment (ATE) environment.

RÉSUMÉ

Cette thèse discute le problème de test de production en grand volume des circuits radio-fréquences (RF) et à ondes millimétriques (mm-wave). Le test des fonctionnalités RF et à ondes millimétriques est très onéreux. Le test intégré est une alternative prometteuse pour faciliter la procédure et réduire les couts, mais il est difficile à mettre en œuvre car il ne faut en aucun cas qu'il réduit la performance du circuit sous test (CUT). Dans cette thèse, nous étudions une technique du test intégré qui repose sur l'utilisation de capteurs non-intrusifs qui prend en compte la variabilité du procédé de fabrication. Cette technique est extrêmement intéressante pour les concepteurs des circuits RF et mm-wave car il leur permet de dissocier le test de la conception. Les capteurs non-intrusifs sont constitués d'étages analogiques triviaux et de composants simples qui sont copiés de la topologie du CUT et sont placés sur la puce à proximité du CUT. Ils offrent simplement une "image" des variations du procédé de fabrication, ce qui leur permet de suivre les variations de performance du CUT. En substance, cette technique tire parti des phénomènes non désirés de variabilité de procédé de fabrication. Le paradigme du test alternatif est utilisé pour estimer les performances du CUT à partir des mesures des capteurs non intrusifs, afin de remplacer les tests standards qui mesurent les performances directement. Ce principe de test est appliqué à deux différents CUTs, nommément un amplificateur à bas bruit à 2.4GHz réalisé en 65nm CMOS et un amplificateur à bas bruit large bande à 60GHz réalisé en 65nm CMOS. Nous démontrons qu'en ajoutant quelques capteurs non-intrusifs sur la puce, qui n'engendrent pratiquement pas de surcout de surface, et en obtenant de ces capteurs non-intrusifs certaines mesures dans le domaine continu et à basse fréquence, nous sommes capable de suivre les variations de toutes les performances du CUT avec une erreur de prédiction moyenne inférieure à l'écart-type de la performance, et une erreur de prédiction maximum qui est inférieure ou au moins comparable aux erreurs de mesure dans un équipement de test automatisé conventionnel.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

1.1 INTRODUCTION TO RF/MICRO- AND MM-WAVE ENGINEERING AND TESTING

Modern RF and micro-wave design is an challenging and dynamic field, due to the combination of recent advances in modern electronic device technology and the explosion in demand for voice, data, and video communication capacity that started in the 1990s and continues until today. RF and microwave technology has had a dramatic impact on wireless technologies. Nowadays, one can access voice and data in virtually any place in the world, from short range Bluetooth and WiFi networks, to cellular and satellite networks. These communication systems are employed across a wide range of environments, including industrial and manufacturing facilities, company offices, infrastructure for municipalities as well as private homes. The large variety of applications and operational environments has led, through the accompanying high production volumes, to colossal advances in cost-efficient manufacturing capabilities of RF and microwave products. This, consequently, has lowered the implementation cost of a host of new and cost-effective wireless as well as wired RF and microwave services. Inexpensive hand-held Global Positioning System (GPS) navigational aids, automotive collision-avoidance radar, and widely available broadband digital service access are among these. Microwave technology is well suited for these cutting edge applications in communications and sensing, since the high operational frequencies allow both large numbers of independent channels for the wide variety of uses envisioned as well as significant available bandwidth per channel for high-speed communication systems [1].

Moreover, over the last decade, (a) a continuous bandwidth of 7GHz range has been opened for unlicensed use at mm-wave frequencies around 60 GHz in the Unites States of America (57–64 GHz) and Japan (59–66 GHz) and (b) the need for designing high-data rate systems have both led industries to build a variety of applications, including gigabit/s point-to-point links, Wireless Local Area Networks (WLANs) with ultra-high capacity, short-range high data-rate Wireless Personal Area Networks (WPANs) and vehicular radar systems. Cost, size and power efficiency are amongst the most important design criteria that need to be accounted for when designing mm-wave circuits. This can be employed using compound semiconductor technologies. Comparing typical to mm-wave frequency spectrums, in the former case, the 3-5GHz spectrum is relatively crowded with many interferes in the WiFi band appearing also limited data rates as well as limited transmit power levels. In the latter case, the 60GHz band offers the same amount of spectrum (7GHz) without interferes and transmitted power up to 40dBm. Based on that, designing of wireless systems capable of supporting multi-Gb/s ultra high-data rates is envisioned [2].

Additionally, testing and characterization of RF and microwave/mm-wave circuits and systems is a challenging issue. Due to the high frequencies involved, the impedance and standing-wave effects associated with test cables and the parasitic capacitance of conventional test probes make the use of conventional low-frequency circuit characterization techniques impractical. Although advanced measurement techniques such as electro-optic sampling can sometimes be employed to circumvent the aforementioned difficulties, generally the loading effect of measurement equipment introduces significant measurement challenges for debugging and analyzing circuit performance, especially for nodes at the interior of the CUT [3].

Finally, cost of test is of major importance to semiconductor manufacturers. ATE vendors play a vital role in combating the increasing cost-of-test pressures of the market place. Modeling cost of test requires a fairly large number of parameters to be modeled. These parameters include: test time, yield, measurement accuracy, ATE equipment cost, lifetime, utilization, floor space, and maintenance [4], [5].

1.2 MOTIVATIONS

Today, high-volume production testing of RF circuits incurs a very high cost, which ends up being a large fraction of the overall manufacturing cost. Conventional test is carried out using sophisticated test instrumentation with advanced features. Different performances require sequential tests on different test configurations onto expensive ATE, which results in lengthy switching and settling times, other than the pure electrical test times. In addition, these tests need to be performed in an environment that is shielded from electromagnetic interference and noise. To overcome the key limitations of high test cost and lengthy test time, a very aggressive and promising test technique has been proposed, named *built-in test*. This technique consists of performing on-chip some of the test operations, e.g. stimulus generation, measurement extraction and analysis, etc. BIT can reduce dramatically the complexity of ATE since it provides digital, DC, or low frequency test signatures, it resolves issues related to electromagnetic interference, and it facilitates parallel testing to achieve a high throughput. However, there are several important criteria that need to be satisfied, in order to effectively replace the conventional test with the proposed BIT strategies. In particular, we need to (a) maintain a low area overhead, (b) avoid degrading the performances of the CUT, and (c) maintain the test accuracy of the standard test approach.

The goal of this work is to develop BIT solutions for RF and mm-wave circuits, in order to circumvent the need for elaborate test instrumentation and reduce the cost of test. The underlying idea is to build on-chip some auxiliary circuitry that extracts useful information about the CUT. The key characteristics of the proposed BIT solutions is that (a) they are non-intrusive, that is, they are transparent to the CUT and do not degrade its performances, (b) they incur low area overhead and (c) they employ DC or low-frequency signals. Recently, it was proposed to test RF circuits non-intrusively by employing variation-aware sensors. These sensors are basic analog stages that mimic part of the CUT architecture (we refer to them as dummy circuits) and Process Control Monitors (PCMs). They are placed in close proximity to the CUT on the same substrate without being electrically connected to it. Instead, their operation is based on the fact that they undergo the same process variations as the CUT. As a result, any degradation in the performances of the CUT is reflected to the sensors' outputs, which shift away from their nominal values. Hence, they can be used for enabling a low-cost alternate test approach where the performances of the CUT are inferred implicitly from their outputs. The alternate test paradigm is used, in order to predict the CUT's performances from the sensors' measurements.

1.3 THESIS OVERVIEW

The proposed test strategy is a transparent BIT approach based on non-intrusive sensors that are not electrically connected to the RF CUT. The non-intrusive sensors simply monitor process variations and by virtue of this they are capable of tracking variations in the performances of the RF circuit as well. The alternate test paradigm is employed to map the outputs of the sensors to the performances, in order to replace the standard tests for measuring the performances directly. The objective of this work is to demonstrate the non-intrusive BIT strategy for two case studies: (a) a typical inductive degenerated RF LNA operating at 2.4GHz, and (b) a 3-stage mm-wave LNA operating at 60GHz. Both circuits have been implemented in 65nm CMOS bulk technology provided by ST Microelectronics.

In the case of the 2.4GHz LNA case study, we observed that the variability in the inductors explains to a large degree the variability in the performances, thus it was deemed necessary to monitor the variability in the inductors. However, adding a dummy inductor PCM is not a smart choice since it incurs a large area overhead. For this reason, variability in the inductors was not considered at all in previous related work. In this work, we studied the sources of variability in the inductors and we found that most of the variability is due to the variability in the resistivity of the high-level metal and alucap layers that form the coils of the inductors and not due to the variability in the pure inductance of the coil. This led us to employ low area overhead metal and alucap resistor PCMs, instead of an area-hungry inductor PCM, in order to capture most of the variability in the inductors.

In the case of the 60GHz mm-wave LNA case study, we observed that the variability in MOS gate resistance explains to a large degree the variability in almost all the performances, including, input matching, noise figure (NF) and linearity. Thus it was deemed necessary to monitor the variability of this particular parasitic resistance. As will be described later, mm-wave circuits are implemented with transmission lines, which are used for impedance matching, interconnect wiring, and the bias networks. It was shown that variations in transmission lines are not the main source of variability in the performances of the LNA and, thereby, by not monitoring variations in the transmission lines does not affect the prediction results appreciably.

After this brief introductory chapter, the rest of the thesis is structured as follows. Chapter 2 presents the state of the art of testing RF/mm-wave circuits. In Chapter 3, we give a detailed introduction to RF/mm-wave circuits and we describe the two case studies. As discussed before, the proposed test idea capitalizes on the undesired phenomenon of process variations. Chapter 4 gives a brief description of process variations along with a few indicative examples of the technology used in this work. In Chapter 5, we discuss about the principle of operation of the non-intrusive variation-aware sensors for both case studies. Chapters 6 and 7 demonstrate, respectively the results of alternate test at simulation-level and with real measurements. Finally, Chapter 8 concludes the work and lists several perspectives for future work.

2

STATE-OF-THE-ART OF TESTING RF/MM-WAVE CIRCUITS

In this chapter, we briefly describe the state-of-the-art of RF/mm-wave circuits testing. Section 2.1 gives a short background on production testing of RF and mm-wave circuits. In Section 2.2, we discuss about the conventional test of RF/mm-wave circuits, including S-parameters, noise, and linearity measurements. The basic features of ATE are also presented since it is the standard way to characterize, test, and measure RF/mm-wave systems. Sections 2.3 and 2.4 present several alternative RF test techniques, including structural test, built-off test, RF/mm-wave transceiver built-in loop-back test, and the non-intrusive test technique. Next, in Section 2.5, we present the state-of-the-art of testing mm-wave circuits. Finally, Section 2.6 concludes this chapter.

2.1 PRODUCTION TEST OF RF DEVICES AND CIRCUITS

Up until the early 1980's, a few wireless consumer devices and systems existed. Testing those devices was a time-consuming procedure, which also required complex, sophisticated, and expensive test instrumentation. In the early 1990's, the RF technology came into the view of the consumer market in the form of cordless and wireless (cellular, mobile) phones. As a result, there was a subsequent market explosion and a massive production of mobile phones. It was obvious that the industry had expanded and, consequently, the prices of semiconductor devices had dropped drastically, following the famous Moore's law. [5].

Nowadays, devices and systems that are used in the frame of wireless communications are increasingly complex and more integrated than ever before. The low prices that consumers pay for wireless phones in a competitive market demand low-cost RF integrated circuits (RFICs). As it is critical to produce quality and properly working products, RF and SoC semiconductor devices and circuits are tested 100%, in order to ensure their intended functionality. The difficult task is to find and establish an efficient way to provide an adequate and comprehensive test methodology that can accurately distinguish good parts from defective ones at a low cost. The cost of test of modern RF devices and SoCs has become a significant part of the overall cost of producing these components. As a conclusion, production testing of RF devices and SoCs is the act of performing numerous tests in a short amount of time on high volumes of parts, where new test strategies need to be defined. The major objective is to have high throughput and low area overhead, or low cost of test, so as the production testing does not adversely impact the marketable value of the device.

The actual RFICs are electrically tested with either a rack-and-stack bench-top equipment connected to a Personal Computer (PC), or with commercially available ATE. Usually, the most costly and complex component in these systems are the RF transceiver, or spectrum analyzer/digitizer, and the RF sources [6]. Rack-and-stack is a suitable configuration for a production tester during the characterization and prototype stages of a device because the equipment contained on the rack can be quickly reconfigured to meet changing needs. Often, rack-and-stack configurations are customized to a specific part. On the other hand, ATE is a tester, which is designed as a complete stand-alone unit for optimal testing of devices produced in high-volume [5]. It requires computer control to ensure correct operation, record measurements, correlate vast amounts of measurement data, and present data in a form understandable by users who no longer directly control ATE features and most, if not all, functions are automated.

2.2 CONVENTIONAL RF TEST

Before describing the basic concepts of standard RF test methods, it would be helpful to give an overview of today's RFICs. There are three basic categories of ICs that are characterized and tested in industries. Table 1 summarizes these categories, depending on the type of the processed signal, as follows:

Type of processed signal	RFIC
RF	LNAs, PAs, Mixers, VCOs
RF/Mixed-Signal	WLAN, GSM, DAB transceivers
RF/Mixed-Signal and Digital Baseband	WLAN, GSM, GPS

Table 1: Three basic categories of ICs depending on the type of the processed signal.

where PA, VCO, GSM, and DAB correspond to Power Amplifier, Voltage-Controlled Oscillator, Global System for Mobile (communications), and Digital Audio Broadcasting, respectively.
The first category performs a single RF function, has a low pin count, and its testing includes several typical RF test setups, as shown in Table 2.

Performance	LNA	PA	Mixer	VCO
Gain	✓	✓	✓	
Impedance Matching	✓	✓	✓	
Standing Wave Ratio (SWR)	✓	✓	✓	
Noise Figure (NF)	✓		✓	
1-dB Compression Point (1-dB CP)	✓	✓	✓	
Third Order Input Intercept Point (IIP_3)	✓	✓	✓	
Adjacent Channel Power Ratio (ACPR)		✓		
Phase Noise				✓
Sensitivity				✓
Frequency Range				✓

Table 2: Typical RF test setups.

The second category performs several RF/Intermediate Frequency (IF) functions. In addition to the requirements of the first category, these ICs, usually have more digital pins and require complex programming, such as automatic gain control.

Finally, the third category is considered to be the most complex case, since it requires the combination of RF/IF functions with digital ones, such as quadrature I and Q base-band signals, Analog-to-Digital Conversion (ADC), Digital-to-Analog Conversion (DAC), and Digital Signal Processing (DSP). Additional test setups include Bit Error Rate (BER), Error Vector Magnitude (EVM), phase locking, jitter, and memory tests [6].

2.2.1 *Automatic Test Equipment (ATE)*

In an industrial production test environment, the population of RF circuits is tested using expensive and sophisticated ATE. ATE consists of the following parts:

- A central unit comprising the electronic resources and power tester instrumentation.
- Interface cards between the CUT and the tester.

- A test head, which includes all the electronic equipment that is needed to perform the most sensitive measurements. Examples of such devices include power resources, electronic contact pins, signal generators, etc. Furthermore, wafer-level test requires specific probes, which connect the inputs and outputs of the dies to the electrical test head resources. For testers dedicated to RF circuits, the cost increases significantly due to the more sophisticated features, which include RF generators, RF spectrum analyzers, Vector Network Analyzers (VNAs), modulation circuits and digital demodulation. There are different types of test heads. The choice of test head takes into account the complexity of the CUT in terms of number of pins required for its full characterization. Moreover, the complexity is directly related to its ability to test a plurality of circuits in parallel using several interface cards.
- A manipulator or articulated arm that physically position the CUTs. This arm is programmed to precisely deposit the CUT on the interface card, monitor test temperature, and finally select the placement of the circuit after the test (depending on whether the CUT is functional or defective)
- A workstation that provides the human interface with the ATE. Each tester is associated with specific tools to perform the test program. In a test program, the test engineer connects the input, output and power supply pins to the legs of the CUT, controls the operation of the manipulator, and defines the sequential tests that need to be performed, as well as the test stimuli, the specifications of the circuit, etc.
- A robot that controls the programming of different existing cards on the tester. Test engineers check every three months the programs installed on existing hardware cards in the tester using this robot. Additionally, other maintenance operations are necessary to ensure the reliability and proper functioning of the tester.
- The equipment needed to cool the test head and electronic cards. This is performed with water flowing in specific pipes. Usually two machines communicate with each other, in order to perform this task. The first injects water into the different pipes which pass through the tester, whereas the second one controls the water temperature.

The above description reveals the high complexity of today's standard ATE configurations. This explains their high cost and the cost of maintenance operations that need to be carried out frequently.

2.2.2 *S-parameters Measurements*

One of the most common techniques for characterizing the linear behavior of RF and microwave/mm-wave components is the use of S-parameters. While z-, y-, and h- parameter representations are commonly used at lower frequencies, these approaches can be problematic to implement at microwave/mm-wave frequencies. The use of S-parameters essentially captures the same information as the other parameter sets, but instead of directly measuring terminal voltages and currents, the forward and reverse traveling waves at the input and output ports are measured. This approach enables accurate characterization of components at very high frequencies to be performed with comparative ease. For a two-port network, the S-parameters are defined by [1]:

$$\begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix}$$

where the V^- terms are the wave components traveling away from the two-port, and the V^+ terms are the incident terms. These traveling waves can be thought of as existing on "virtual" transmission lines attached to the device ports. From this definition,we have:

$$\mathrm{S}_{11} = \left.\frac{V_1^-}{V_1^+}\right|_{V_2^+=0}, \mathrm{S}_{12} = \left.\frac{V_1^-}{V_2^+}\right|_{V_1^+=0}, \mathrm{S}_{21} = \left.\frac{V_2^-}{V_1^+}\right|_{V_2^+=0}, \mathrm{S}_{22} = \left.\frac{V_2^-}{V_2^+}\right|_{V_1^+=0}$$

To measure the S-parameters, the ratio of the forward and reverse traveling waves on the virtual input and output transmission lines is measured. To achieve the $V^+ = 0$, and $V^+ = 0$ conditions in these expressions,

the ports are terminated with the characteristic impedance, Z_0, of the virtual transmission lines. Although in principle these measurements can be made using directional couplers to separate the forward and reverse traveling waves and phase-sensitive detectors, in practice modern VNAs augment the measurement hardware with sophisticated calibration routines to remove the effects of hardware imperfections to achieve accurate S-parameter measurements.

2.2.3 *Noise Measurements*

Communications systems' noise figure is usually measured with three popular methods depending on the specifications. The first method requires a specialized *noise figure meter*. This is an expensive piece of hardware and has a limited upper frequency measurement capability (typically less than 3GHz). The second method is called the *gain method*, which has severe limitations in that we cannot measure low-noise figures observed in modern microwave/mm-wave devices and systems. For noise figure accuracy, this test also requires the Device Under Test (DUT) or the CUT, if a complete system or circuit is measured, to have a very high gain. The third technique is called the *Y-factor method* and is very accurate at measuring low-noise figures at high frequencies without requiring any high DUT gain limitations. The only specialized piece of necessary equipment for this test is an *Excess Noise Ratio (ENR) noise head*, which is a calibrated ultra wide-band noise generator.

The noise figure measurement setup is depicted in Figure 1, for the Y-factor method [7].

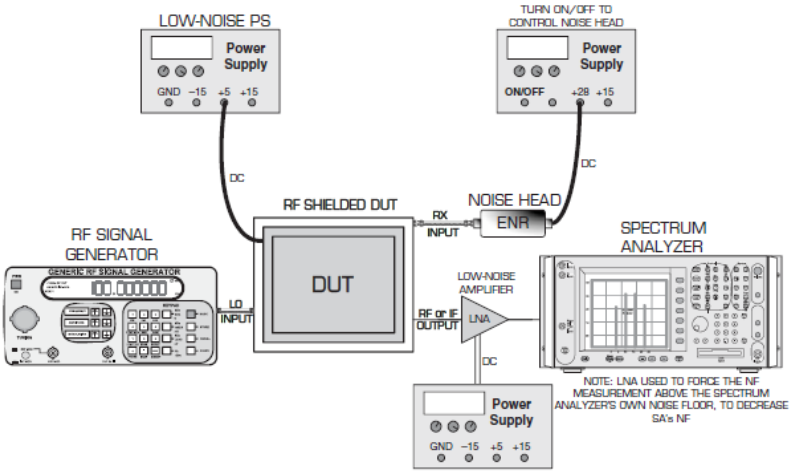


Figure 1: Noise figure measurement setup.

2.2.4 *Non-Linear Measurements*

Harmonic Distortion

A fundamental result of the distortion of non-linear circuits is that they generate frequency components in the output signal that are not present in the input signal. For sinusoidal inputs, the salient characteristic is *harmonic distortion*, whereby signal outputs consist of integer multiples of the input frequency.

While instruments are available at low frequencies to measure the total harmonic distortion (THD), the level of each harmonic is generally measured individually using a spectrum analyzer. Such a setup is shown in Figure 2.

Harmonic levels are usually measured in a relative manner by placing a marker on the fundamental signal and a delta marker at the n-th harmonic frequency. When measured in this mode, the harmonic level is expressed in dBc, which designates dB relative to carrier (i.e. the fundamental frequency) level. While it is convenient to set the spectrum analyzer sweep to include all harmonics of interest, it may be necessary to center a narrow

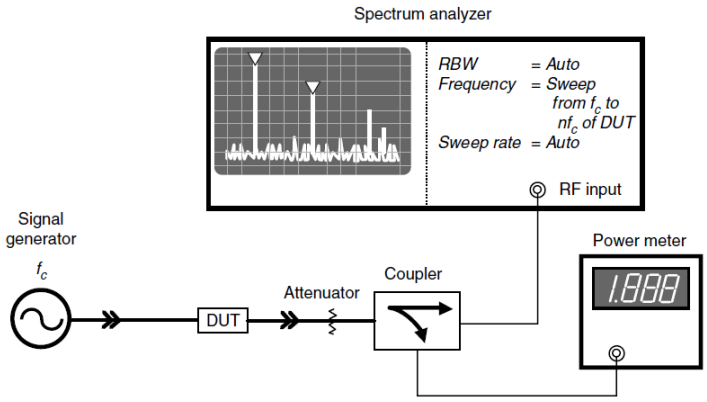


Figure 2: Harmonic distortion measurement setup.

span at the harmonic frequency, in order to reduce the *noise floor* on the spectrum analyzer. An attenuator may be needed to protect the spectrum analyzer from overload. Note that the power level present at the spectrum analyzer input includes all harmonics, not just the ones displayed on the screen. Finally, it is important to note that spectrum analyzers have their own non-linear characteristics that depend on the level input to the instrument. It is sometimes difficult to ascertain whether measured harmonic distortion is being generated within the device or with the test instrument. One method to do this is to use a step attenuator at the output of the device and step up and down. If distortion is being generated with the spectrum analyzer, the harmonic levels will change with different attenuator settings [1].

Gain Compression

A major result of changing impedances in RF and microwave/mm-wave circuits is signal gain and phase shift that depend on the input amplitude level. A change in signal gain between input and output may result from signal clipping due to device current saturation or cut-off. Insertion phase may change because of non-linear resistances in combination with a reactance. Though there are exceptions, signal gain generally decreases with increasing amplitude or power level. For this reason, the gain compression characteristics of microwave/mm-wave components are often characterized.

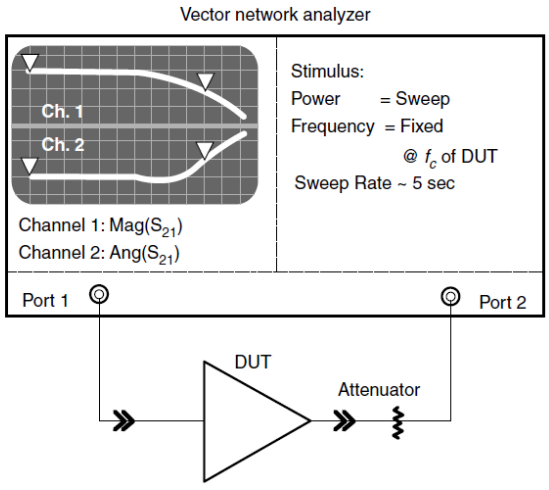


Figure 3: Gain compression measurement setup.

For band-pass components where the input frequency is equal to the output frequency, such as amplifiers, gain compression and phase deviation of a non-linear circuit are readily measured with a VNA in power sweep mode. Such a setup is shown in Figure 3. P_{1dB} is easily measured using delta markers by placing the reference

marker at the beginning of the sweep (i.e., where the DUT is not compressed) and moving the measurement marker where $\Delta Mag(S_{21})$ =-1 dB. Sweeping at too high a rate may affect the readings. The sweep must be slow enough so that steady-state conditions exist in both the thermal case and the DC bias network within the circuit. Sweeper retrace may also affect the first few points on the trace. These points must be neglected when setting the reference marker [1].

Intermodulation Distortion

When more than one frequency component is present in a signal, the distortion from a non-linear circuit is manifested as Intermodulation Distortion (IM). The IM performance of microwave/mm-wave circuits is important because it can create unwanted interference in adjacent channels. While band-pass filtering can eliminate much of the effects of harmonic distortion, intermodulation distortion is difficult to filter out because the IM components may be very close to the carrier frequency. A common figure-of-merit is two-tone intermodulation distortion.

IM is normally measured with two-signal generators and a spectrum analyzer. Such a setup is shown in Figure 4. Care must be taken to isolate the signal generators, as IM may result from one output mixing with signal from the opposing generator. Also, it is usually recommended that a power meter should be used to get an accurate reading of output power level from the DUT. Relative IM level is measured by placing a reference marker on one of the two carrier signals, and placing a delta marker at either sideband. Finally, the input level must be maintained well below the input IIP_3 of the spectrum analyzer to insure error free reading of the DUT [1].

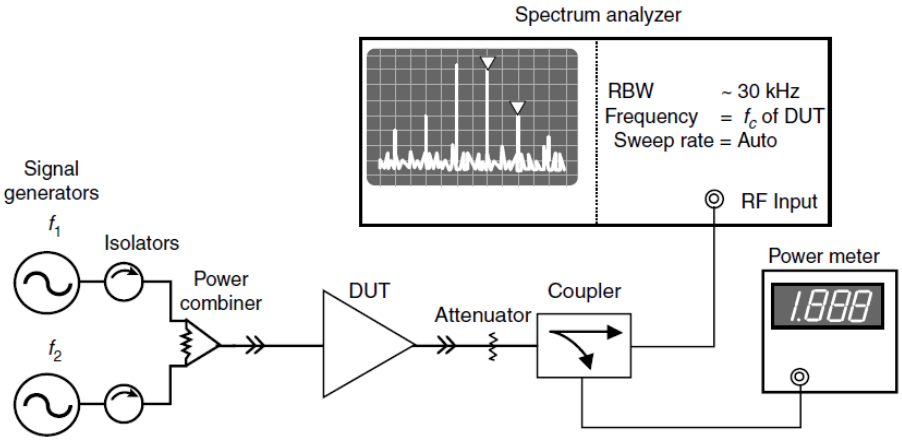


Figure 4: Intermodulation distortion measurement setup.

Noise figure, harmonic distortion, gain compression, and itermodulation distortion are figures of merit that describe the noise and linearity performances of a device or a complete system and are described later on in more details.

2.2.5 *Phase Noise*

Phase noise is a parameter that measures the spectral purity of a signal and it is particularly referenced to a sinusoidal waveform. In an ideal oscillator, the output of the circuit is a perfectly-periodic signal with the form $x(t) = Acos(\omega_c t)$, where ω_c is the oscillation frequency. The zero crossings of the output waveform occur at exact integer multiples of the oscillation period $T_c = {}^{2\pi}/_{\omega_c}$. Although, in real oscillators, the noise of the circuit itself randomly perturbs the zero crossings. In order to model this effect, we write $x(t) = Acos(\omega_c t + \phi_n(t))$, where $\phi_n(t)$ is a small random phase quantity that deviates the zero crossings from integer multiples of T_c. The term $\phi_n(t)$ is called as the "phase noise".

Figure 5 shows the measurement setup for testing the phase noise of a DUT. The measurement procedure requires a spectrum analyzer and a low-noise power supply [1].

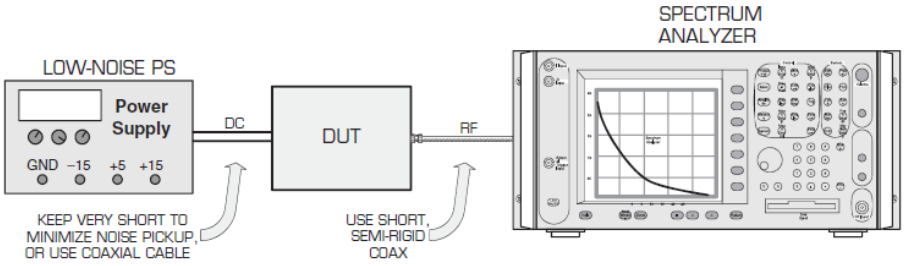


Figure 5: Phase noise measurement setup.

2.2.6 *Bit Error Rate (BER)*

BER is a specification that is needed, in order to determine the performance of a communication system in terms of transmitting and receiving of data, which corresponds to a certain type of digital modulation. In particular, for a random bit-sequence, BER represents the ratio between the number of incorrectly detected bits (e.g. due to noise) and the total number of bits that are sent. Based on this simple definition, we are able to derive a relationship between BER and Signal-to-Noise Ratio (SNR).

The conventional test of BER uses a pseudorandom bit sequence, which is modulated through a special signal generator onto an RF carrier, directly injected into the receiver's RF input port through a coaxial cable of known insertion loss. The RF input level is either controlled through an external adjustable attenuator or through the signal generator itself. The receiver's output base-band data is sent on to the BER test equipment, which continuously calculates the errors between the initially sent data and the currently received data. The amplitude of the RF input signal is then adjusted downward until the bit error rate display of the BER test equipment duplicates our required BER. Then, the level of the RF signal's power (in dBm) or voltage (in μV) at the receiver's RF input port is measured or calculated. The BER test equipment includes a Vector Signal Analyzer (VSA) and a spectrum analyzer. Like in cases of IM, measuring BER involves high test time and complicated measurement setups.

2.2.7 *Error Vector Magnitude (EVM)*

The EVM measurement quantifies the performance of a radio transmitter against an ideal reference. A signal sent by an ideal transmitter would have all points in the I–Q constellation and fall precisely at the ideal locations (i.e., magnitude and phase would be exact). Non-ideal behavior of the transmitter, however, causes the actual constellation points to fall in a slightly scattered pattern that only approximates the ideal I–Q location. EVM is a way to quantify how far the actual points are from the ideal locations. This is indicated in Figure 6.

Measurement of EVM is accomplished using a VSA. The equipment demodulates the received signal in a similar way to the actual radio demodulator. The actual I–Q constellation can then be measured and compared to the ideal constellation. EVM is calculated as the ratio of the Root Mean Square (RMS) power of the error vector over the RMS power of the reference [1].

2.3 ALTERNATIVE RF TEST TECHNIQUES

From the above discussions, it is clear that high-volume production testing of RF circuits based on the conventional methodology incurs a high cost due to the high complexity of test instrumentation and the necessity to rely on ATE with very advanced features. The test configurations are challenging because they require calibration, de-embedding, and the measurement environment has to be shielded from electromagnetic interference

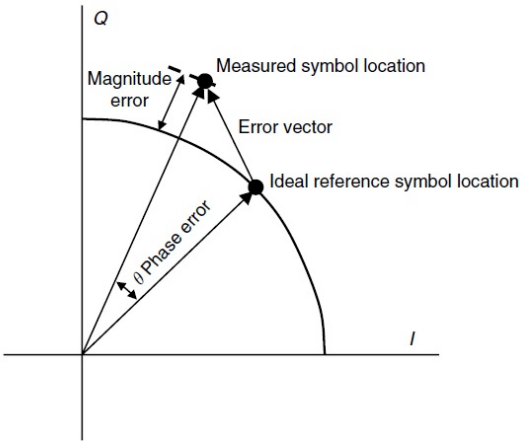


Figure 6: I–Q diagram indicating the error vector for EVM measurements.

and noise. In addition, system-level tests for RF transceivers, such BER and EVM, take up long times, which translates to high cost. As depicted in Figure 7, there are many different performances that need to be measured, requiring sequential tests on various test configurations, which result in lengthy switching and settling times, in addition to the pure electrical test times. For example, a t_1 time is needed in the first test configuration for the response to settle and for measuring performance 1 of. Thereafter, the data is loaded back to the ATE, where they will be analyzed. A switching time t_{12} is required to configure the second test configuration to measure performance 2 and so forth. In order to alleviate the burden of high test cost and long test times, researchers and practitioners are seeking to develop alternative test techniques while maintaining the same levels of test accuracy.

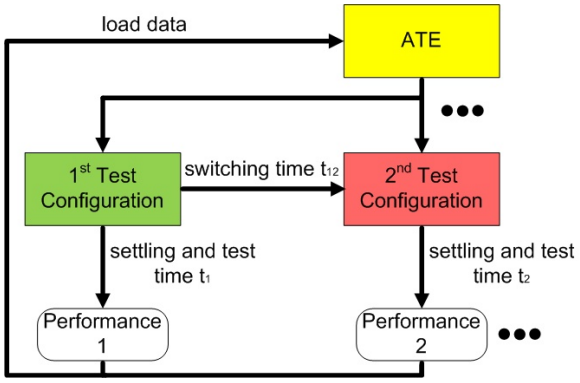


Figure 7: Automatic Test Equipment.

Akin to test practices in the digital domain, *structural test* techniques have been proposed aiming at developing tests that detect a list of the most probable faults. However, such techniques have not been materialized yet due to the difficulty in defining the fault model [9], [10]. Another technique is *built-off test*, which involves migrating some of the instrumentation in the ATE onto the test board [11], [12]. This approach can reduce drastically the test cost at the expense of devoting a significant effort to design a multi-purpose and versatile test board.

2.4 RF BUILT-IN TEST (BIT) TECHNIQUES

A technique with higher potential is *built-in test*, where the idea is to integrate on-chip some structures, so as to facilitate the test, such as a test stimulus generator and a response analyzer or on-chip sensors [13], etc., and reduce the external circuitry and the complexity of routing high-frequency signals out of the chip. The on-chip sensors track the electrical behavior of the CUT, generate low-frequency signals, and send them off-chip for further processing. Typical examples of BIT techniques include: (a) self-test, (b) loop-back test, (c)

sensor-based test, and (d) non-intrusive test. These techniques will be described in following paragraphs in more details.

2.4.1 *Built-In Test (BIT) for Self-Correction and Calibration*

BIT has been used in the context of self-correction/compensation of RF circuits. In [14], the authors present a reconfigurable LNA using a variable bias circuit implemented on the same chip, as shown in Figure 8. The proposed LNA topology can be reconfigured by tuning the bias voltage of the common-source transistor of the LNA, as shown in Figure 9, in order to compensate for gain, noise and power consumption depending on the runtime environment (temperature and input signal power). The authors show experimentally that the proposed LNA configuration achieves 73% of power reduction by controlling bias voltage. In addition, ΔIM_3 (the difference in the output power between the desired signal and the 3rd harmonic) and the output SNR are compensated simultaneously. Furthermore, the proposed LNA can improve the design cost because redundant design margin is not required.

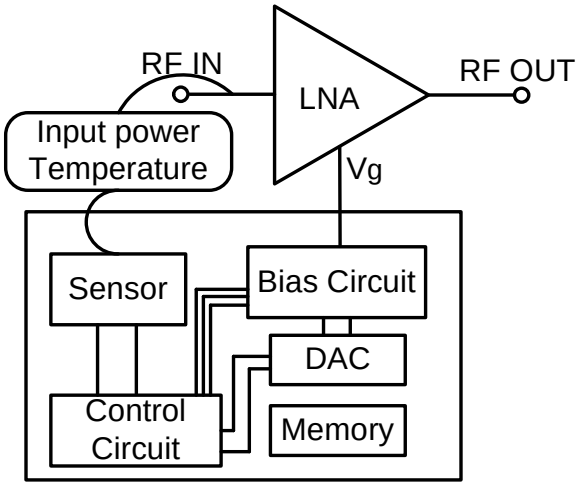


Figure 8: Schematic of the reconfigurable LNA.

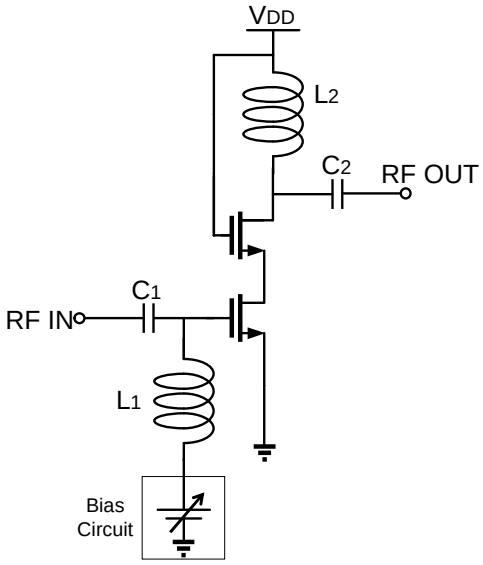


Figure 9: Schematic of the bias circuit used for compensating for the performances of the reconfigurable LNA.

BIT technique has also been applied in the context of on-chip calibration. In [15], the authors present an on-chip VCO design with digitally tunable resonator and a calibration system targeted for the DCS1800 and PCS1900 GSM transmitter bands. The design includes an accurate digital cell based divider, a counter running with 1GHz input for frequency measurement to enable on-chip calibration of the VCO, and the calibration system logic. As depicted in Figure 10, the only external signals that are needed include configuration signals, control signals, and a clock signal at a low frequency, so as to reduce the amount of ATE, which is required for characterizing and testing the whole system. The proposed circuit architecture operates in two modes. During calibration, the voltage control input of the VCO is connected to the calibration DAC output, the digital input of which comes from the calibration logic. In normal operation mode, when the VCO is part of the PLL, the control voltage is driven by the PLL loop filter. The digital calibration word of the VCO is controlled by the calibration logic in both modes.

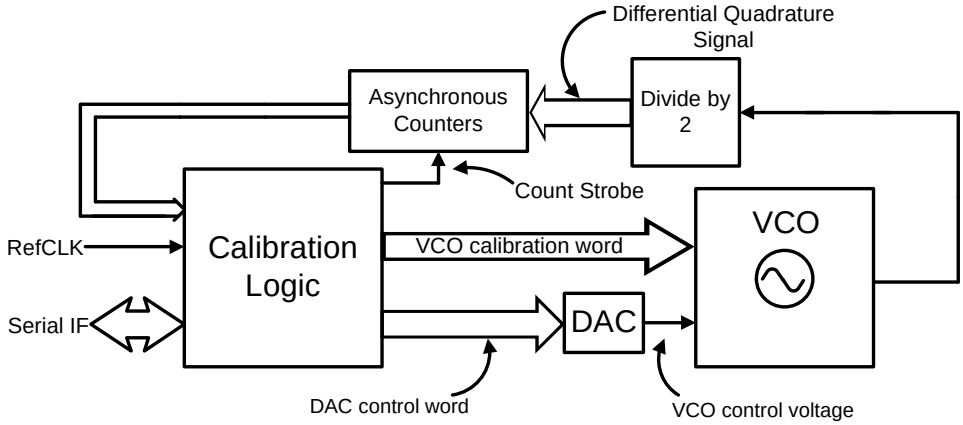


Figure 10: VCO calibration configuration.

2.4.2 *Loop-back Test Technique*

The most popular BIT approach for RF transceivers is the loop-back test where the test signals are generated in the base-band and the transmitter's output is switched to the receiver's input to analyze the test response also in the base-band [16], [17], [18], [19]. Loop-back connection requires the insertion of a switch and an attenuator and, for some types of receivers, even an extra mixer is inserted in the RF signal path. Figure 11 shows the loop-back test setup for IC transceivers. As described in [19], in order to enable the on-chip loop-back configuration, the test signal from the transmitter must be fed back to the receiver through a test attenuator (TA). This approach is consistent provided that both the transmitter and the receiver operate at the same frequency and the Local Oscillator (LO), only serves for up- and down-conversion, while the modulation process is performed at base-band. These conditions are sufficient unless the frequencies required for up- and down-conversion are different and the LO can provide only one of them at a time or the LO has one output, which in the normal transmit and receive mode toggles between the two frequencies. This would be typical of a one-step transmitter and a low-IF receiver [20], which operate at the same radio frequency, but the up-conversion proceeds from base-band to RF while the down-conversion from RF to non-zero IF. In this case, the loop-back test can be enabled in two ways. If the system is narrow-band (such as Bluetooth of 1 MHz bandwidth), then the transmitted base-band signal can be up-converted in the base-band (BB) processor to IF compensating thereby for the existing incompatibility between transmitter and receiver, and therefore, the setup shown in Figure 11 still holds. This is viable due to the fact that in low-IF receiver the IF is usually chosen half of the bandwidth. Using this technique for a wide-band system (such as WiFi of 20 MHz bandwidth) would impose more stringent requirements on the DAC due to much larger band required during test and power consumption as a consequence.

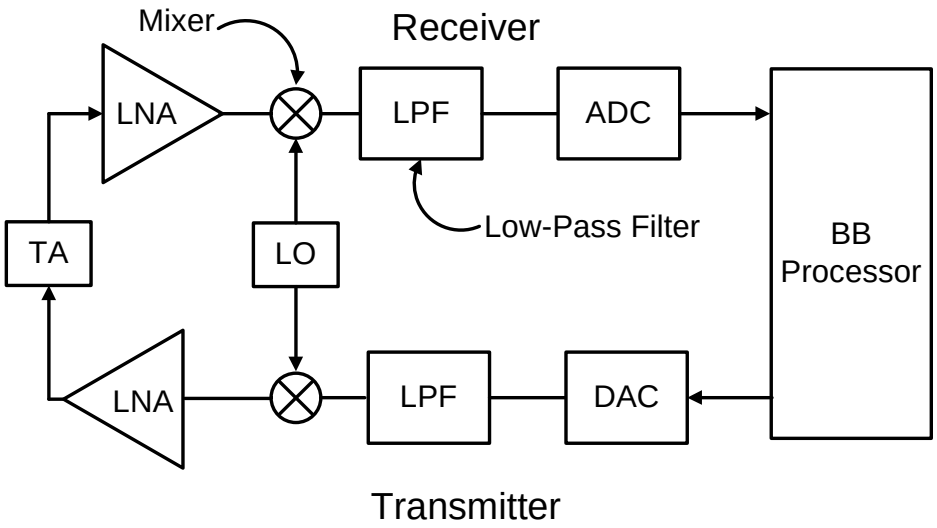


Figure 11: Loop-back test setup for IC transceivers.

2.4.3 *DC probes, Envelope Detectors, and Current Sensors*

Another popular BIT technique relies on the use of DC probes [21], [22], envelope detectors [21], [23], [24], [25], [26], [27] and current sensors [21], [28], [29], [30] to extract DC/low-frequency test signatures that nevertheless carry RF information.

A simple, yet efficient, way to detect random defects on a CUT is through DC probes [21], which can be set up by connecting a large resistor at a critical node, such that the current flow within the CUT is unaffected. DC probes should be placed, such that they can sense most of the degradation and defects in the topology. Their main drawback is that they are unresponsive to variations related to inductive and capacitive components. For this purpose, we inevitably must exercise the CUT at its operation frequency. The goal remains to obtain a DC or low-frequency signature that nevertheless carries RF information. We can achieve this by using an envelope detector or a current sensor.

The function of an envelope detector is to generate an output DC voltage proportional to the amplitude of the RF signal at the input of the sensor. Typically, a high input impedance is desirable for the detector, such that the loading effect and the influence on the performances of the CUT can be minimized. In order to serve as a general-purpose BIT module, the envelope detector needs to have appropriate transfer characteristics over a wide frequency band with sufficient dynamic range.

Moreover, current sensors can be used, in order to detect random defects and extract signatures of the power supply dynamic current. This test technique takes advantage of the small parasitic resistor (it can reach several Ohms) of the line connecting the CUT core to the power supply terminal.

2.4.4 *Alternate Test*

Alternate test is a generic technique that aims to map low-cost measurements to the RF performances [22] [31], [32], [33], [34] thus circumventing the need to measure directly the RF performances. Alternate test can also rely on information-rich, low-cost measurements obtained by built-in sensors [21]. Alternate test is described in Chapter 5 in more details.

2.4.5 *Non-intrusive Test Approach*

For a successful implementation of a BIT technique, different criteria need to be satisfied simultaneously: (a) the test accuracy of the standard method is maintained; (b) the performances of the CUT are not affected; (c)

the pin and area overheads are acceptable; and (d) the achieved test cost reduction justifies the BIT development effort. It is noteworthy that none of the aforementioned state-of-the-art BIT techniques fully satisfies objective (b). For example, re-visiting loop-back BIT, this particular test technique requires the insertion of a switch and an attenuator and, for some types of receivers, even an extra mixer is inserted in the RF signal path. Envelope detectors and current sensors also tap into the RF signal path, as described in the previous paragraph. In general, adding components in the RF signal path degrades the impedance matching and adds parasitics, which inevitably shift the performances and unbalance the performance trade-offs achieved by design. To address this issue, the BIT circuitry needs to be co-designed with the CUT which increases design iterations to meet the target design specifications, if at all possible. For this reason, designers are rather reluctant to incorporate such BIT techniques since the design specifications are stringent and exploit the full capabilities of advanced technology nodes.

To this end, it has been proposed to rely on non-intrusive built-in sensors that have the comparative advantage that let the design intact [21], [35]. These non-intrusive sensors capitalize on the undesired phenomenon of process variations. The underlying idea is to monitor process variations instead of measuring directly the RF performances. For this purpose, we can employ PCMs, such as single layout components (e.g. transistors, capacitors, resistors, inductors), and dummy circuits that are extracted from the CUT topology (e.g. bias stages, current mirrors, gain stages, level-shifters, etc.). These sensors are placed in close physical proximity and are matched to identical structures in the CUT. For example, we can place a dummy bias stage next to the bias stage of the CUT, a dummy transistor next to a critical transistor in the CUT, etc. In this way, the sensors and the CUT "witness" the same die-to-die (D2D) and correlated within-die (WID) process variations and, as a result, the measurements obtained on the sensors will be correlated to the performances of the CUT to a very large extent. An indirect, low-cost test can be put in place by employing the alternate test paradigm to map the sensors' measurements to the performances. This test approach has been inspired by the PCMs typically placed in the scribe lines of a wafer to monitor variability and identify off-target process parameters [36], [37]. The idea of using die-level PCMs to extract information about performances has also been applied in the case of ADCs in [38].

It should be noted that with this test approach we can verify whether one or more performances violate their specifications due to excessive process variations, but we cannot detect defects within the CUT since the sensors are not electrically connected to it. Non-intrusive, defect-oriented BIT can be performed using either DC probes, envelope detectors, and current sensors or temperature sensors [39] as will be described in Section 2.4.6.

It should also be noticed that, in the case of a CUT that consists of a chain of blocks, for example an RF transceiver, this test approach will be able to monitor each individual block separately to predict its performances, for example gain, noise figure, etc., of the LNA, PA, etc., that comprise the RF transceiver. This may result in yield loss under the scenario where excessive variations within two blocks are being compensated and masked. One possibility is to use this test approach for predicting CUT-level performances, for example the EVM in the case of an RF transceiver [40]. Another possibility is to employ the predictions of performances of individual blocks to predict CUT-level performances from analytical mathematical expressions, for example the EVM in the case of an RF transceiver [41].

This type of non-intrusive BIT based on variation-aware sensors was demonstrated for a 0.25μm RF LNA in [35]. A potential problem of this technique, as we will explain in more details later on, is that the correlation between the sensors' measurements and the performances of the CUT and, ultimately, the success of alternate test, is affected by uncorrelated WID variations (e.g. mismatch). It is expected that this type of variations will become more prominent as we shrink transistor lengths.

2.4.6 *Detecting Random Defects*

In [21], authors present a variety of sensors that enable a low-cost, BIT of RF front-ends. DC probes, an envelope detector, and a current sensor are used, in order to detect random defects within the CUT. Figure 12

depicts these sensors attached to the RF front-end. However, as mentioned in Section 2.4.5, the main drawback of using DC probes, envelope detectors and current sensors is that they electrically connected to the CUT, thus they degrade its performances. In this case, the CUT must be re-designed to correct for the losses. To overcome this limitation, authors in [39] present a built-in, defect-oriented test approach for RF circuits that is based on thermal monitoring.

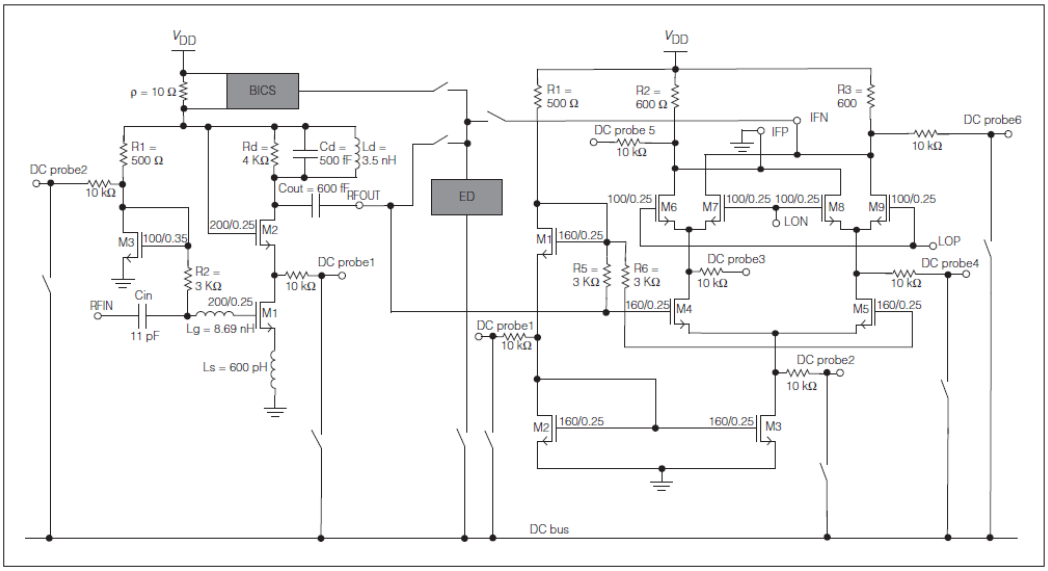


Figure 12: RF front-end schematic with embedded sensors for BIT. Embedded sensors include DC probes, an envelope detector, and a current sensor.

The underlying idea is that a defect will change the current that flows through a circuit branch and, thereby, the power dissipated by components in this branch. In turn, this will change the temperature profile on the die near a dissipating component. A temperature sensor, which consists of an operational amplifier built from bipolar transistors, captures the temperature change, thus indicating the presence of the defect. In this regard, a complete and comprehensive test methodology for detecting both excessive process variations and defects can be conceived based solely on non-intrusive sensors, which can replace altogether the standard RF test practice employed today in industry.

2.5 TESTING OF MM-WAVE CIRCUITS

The continuous scaling of channel lengths in CMOS processes down to 32nm and beyond has pushed the transition frequency (f_T) and maximum oscillation frequency (f_{max}) of silicon CMOS transistors in excess of 230GHz and higher [42], making it an attractive approach for designing mm-wave circuits and systems. Moreover, the need of designing high data-rate systems has led industry to develop several mm-wave applications, from block to system level, including amplifiers and filters [2], gigabit/s point-to-point links, WLANs with extraordinary capacity, short-range high data-rate WPANs [43], and vehicular radar. However, design of mm-wave circuits experience several challenges. Accurate modeling of active devices, and passive components, and interconnects is of major importance and often requires time-consuming iterations [44]. Accurate modeling of CMOS devices in mm-wave regime is pivotal due to low-resistivity substrates and parasitic source, drain, and gate resistances. In order to account for the aforementioned limitations, during the recent years, new methodologies have been proposed and developed, which include judicious modeling in several levels of abstraction, allowing trade-offs between flexibility, accuracy, and simulation time [45]. Furthermore, matching networks and resonators are often realized with transmission lines, which are inherently scalable in length, making them easier to model than lumped passive devices, such as inductors. Another issue that mm-wave design faces is its sensitivity to process, supply voltage, and temperature (PVT) variations. During the recent years, researchers

have been working towards the direction to increasing the yield of mm-wave circuits by adopting calibration methodologies, adaptive circuit strategies, and self-healing algorithms [42], [46], [47].

In [46], the authors present a novel calibration methodology suitable for mm-wave mixers. The idea relies on using a spectrum analyzer, in order to detect Intermediate Frequency (IF) components and measure the relative power at the IF port when the RF port is driven by a VNA, without using RF filters, since a suitable number of frequency points and a narrow frequency span can be selected on the VNA. [47] demonstrates a design and test strategy which adaptively configures mm-wave circuits for performance and/or yield enhancement by on-chip self-healing. The idea is based on indirect performance sensing using Bayesian Model Fusion (BMF) that enables sophisticated mm-wave performance prediction without demanding an integrated mm-wave circuit measurement infrastructure.

The conventional production test of mm-wave circuits is performed using expensive and complex VNAs as well as VNAs frequency extenders. The accuracy of the VNA measurements is achieved by using high quality signal sources and receivers, as well as sophisticated calibration methods that need to be developed, in order to eliminate systematic errors and remove the losses of the cables. On top of that, high-volume production test of mm-wave circuits and systems is a major challenge towards their widespread adoption. The only standard test approach would be to rely on ATE with advanced features and sophisticated calibration and de-embedding procedures and perform BER testing based on a built-in loop-back connection [48]. However, this test approach incurs a high cost. In addition, the built-in loop connection requires a large area overhead and is very challenging from a design perspective since it needs to be carefully co-designed with the mm-wave circuit to avoid performance degradation.

2.6 CONCLUSION

In this introductory chapter, we presented the state-of-the-art of testing RF and mm-wave circuits. In conventional high-volume production test floors, ATE with advanced features is required resulting in high test cost and lengthy test times, since different performances require sequential tests on different test configurations. In addition, these tests need to be performed in an environment that is shielded from electromagnetic interference and noise. To this end, many alternative test techniques have been proposed aiming at reducing the test cost and facilitating the test application. The built-in non-intrusive test is the most appealing to overcome the above key test limitations. Its key characteristic is that it is totally transparent to the design, so that it does not degrade the performances of the CUT. In mm-wave frequencies, tapping into the signal paths to extract useful information for testing is utterly intolerable and the non-intrusive test idea becomes the only possibility.

3

RF CIRCUITS: DESIGN AND CASE STUDIES

This chapter gives an introduction to RF circuits, the basic concepts in their design methodology, and the two cases studies in this work. Sections 3.1 and 3.2 describe the challenges and the basic concepts in RF design, respectively, including impedance matching, non-linearity, and noise. In Section 3.3, we present the first case study, which is a 2.4GHz inductive degenerated LNA. In Section 3.4, we present the second case study, which is a wide-band 60GHz CMOS 3-stage LNA along with some useful guidelines that are followed for designing circuits that operate in mm-wave frequency ranges. Finally, Section 3.5 concludes the discussions in this chapter.

3.1 CHALLENGES IN RF DESIGN

The design of analog and RF integrated circuits is a complex procedure in which many steps have to be carefully taken into account for a successful and efficient implementation. At first, the individual RFICs are components of entire transceivers. Hence, they have to comply with certain specifications. These result from the system level analysis of the transceiver and are determined by the application that the transceiver has to serve. After extracting the specifications, the appropriate circuit topology has to be chosen to achieve the desired functionality. The most important step is to properly bias and size the selected topology. The design choices are directly interdependent imposing trade-offs in analog/RF design, according to the RF design hexagon shown in Figure 13. This actually means that the individual figures of merit, such as gain, noise, linearity, power consumption are interdependent. Hence, methodologies for achieving the optimum performance have to be derived, rather than relying on the commonly used "trial & error" method. The latter has serious limitations regarding the time that is required for the circuit to converge to the expected functionality. Hand calculations can also be made prior to circuit simulations. For both hand calculations and simulations, a compact model describing the transistor operation is of major importance. The model should be characterized by simplicity, accuracy, and scalability over a large range of geometries, bias points and frequencies. In practice, a compact model acts as a bridge between the device and circuit [49].

With the aggressive technology scaling, the need for circuits operating under low voltage/low power becomes more and more imperative. In some recent designs, the operating point of MOS devices is moving from strong inversion (S.I.) to lower inversion levels, that is, to moderate inversion (M.I.) or even the sub-threshold region, also known as weak inversion (W.I.) - though, this is mostly the case for analog rather than RF. In weak inversion, specific problems occur in terms of leakage currents which dramatically increase as channel length decreases. This may have catastrophic consequences for circuit design since the gate current becomes comparable to the channel current [49].

This is why older technology nodes are often still preferable for analog design. Even in the case of integrated SoC, where the analog and digital parts should co-exist, the dictation of using modern CMOS technologies imposed by the digital domain, has resulted in employing design techniques with which the emerging disadvantages can be suppressed [50]. On the other hand, for RFIC design, operation under such small current level should generally be avoided since this dramatically decreases circuit's RF and noise performance through the decrease in the device's f_T. Recently, RFICs with exceptional performance operating within the M.I. region

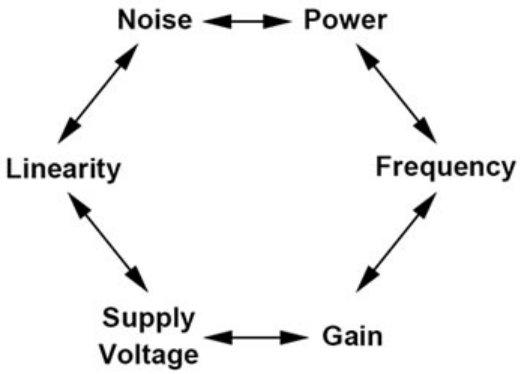


Figure 13: RF design hexagon.

have been reported [51]. Moving towards M.I. offers several advantages, some of which are presented below [52]:

- Increase of the transconductance efficiency which in turn results in minimizing power consumption.
- Lower electrical fields within the device due to the decrease in bias voltages. In this way, velocity saturation (VS) and the corresponding hot electron effects phenomena can be avoided. Thus, the scaling trend of f_T increases with the inverse of the square of channel length $f_T \sim 1/L^2$ rather than the inverse of channel length, which is the case when VS is present [53]. This has a direct implication in circuit design since MOS devices can be used as amplifiers in much higher frequencies.
- The combination of current efficiency with f_T and intrinsic gain of the device, results in figures of merit describing the overall performance of MOS device. It has been proven that the peak values of these figures of merit are moving towards inversion levels close to the center of M.I. with technology scaling.
- Reduced electron heating results in lower excess noise factor, meaning that the overall device noise levels, usually expressed via minimum noise figure (NF_{min}), will be decreased.
- With technology scaling, optimum performance regarding non-linearity metrics is shown to appear in the vicinity of the transition region from M.I. to S.I.

Therefore, M.I. represents the ideal trade-off between power consumption, noise and linearity, especially as technology scaling is driven towards the deca-nanometer regime.

3.2 BASIC CONCEPTS IN RF DESIGN

RF design relies on many concepts that come from several fields, including signals and systems theory, electromagnetics, microwave/mm-wave theory, and communications. This section presents these concepts, which play major role in analyzing and designing RF circuits. The most important RF design concepts are listed below and include impedance matching, non-linearity, and noise.

3.2.1 *Impedance Matching*

There are three different types of matching in an RF block, namely impedance matching, power matching and noise matching. The latter occurs when the source admittance is equal to the optimum admittance. Power matching maximizes power transfer to a load. Consider the case of a voltage source with a source impedance Z_S driving a load impedance Z_L. From basic circuit theory, the value of Z_L that maximizes the power dissipation in the load is achieved when $Z_L = Z_S^*$ [54]. In case of an LNA, for example, the input signal source is usually a band-select filter or an antenna. A band-select filter is typically designed and characterized with a standard

termination of 50Ω. If the load impedance seen by the filter deviates from 50Ω, then the filter may exhibit performance degradations such as loss and ripple [55]. In the absence of a filter, the antenna directly provides the incoming signal to the LNA. The antenna is also designed for a certain real load impedance, typically equal to 50Ω. When the LNA input impedance is also equal to 50Ω, the situation is called impedance matching. Note that in this case, power matching is identical to impedance matching. Poor matching at the receiver input causes reflections, loss and possibly voltage attenuation. The quality of the input matching is expressed by the input return loss, usually denoted by Γ. It is defined as the ratio of the reflected voltage to the incident voltage and is expressed as

$$\Gamma = \left| \frac{Z_{in} - R_S}{Z_{in} + R_S} \right|. \tag{1}$$

where Zin denotes the input impedance and R_S is the source impedance. Ideally, for $Z_{in} = R_S \Rightarrow \Gamma = 0$, and no reflection occurs.

3.2.2 *Noise*

Noise is defined as any random interference, which is unrelated to the signal under study. As all random processes or functions, noise is characterized and modeled by a Probability Density Function (PDF) and a Power Spectral Density (PSD). Noise is present in all electronic components and circuits and is generated by resistors, base and emitter of bipolar transistors as well as gate and channel resistance of MOS devices. In the following discussions, we give a brief overview of different sources of noise [56].

Thermal Noise in Resistors

The ambient thermal energy results in random disturbance of charge carriers in resistors and therefore noise is generated. This noise is modeled by a series voltage source with a PSD of $\bar{V}_n^2 = 4kTR_1$ (Thevenin equivalent in Figure 14(a)) or a parallel current source with a PSD of $\bar{I}_n^2 = \bar{V}_n^2/R_1$ (Norton equivalent in Figure 14(b)) [56], where k, R_1, and T are the Boltzmann constant, the resistor under study, and the temperature (in Kelvin degrees) respectively.

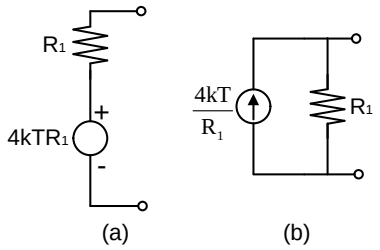


Figure 14: (a) Thevenin and (b) Norton models for resistor thermal noise.

Noise in MOSFETs

For a MOS transistor that operates in the saturation region, the thermal noise is approximated by a current source connected between the source and drain terminals [56], as depicted in Figure 15:

$$\bar{I}_n^2 = 4kT\gamma g_m, \tag{2}$$

where γ is the "excess noise coefficient" and g_m the transconductance of the MOS transistor. The value of γ is ⅔ for long-channel transistors and may rise to even 2 in short-channel devices [57].

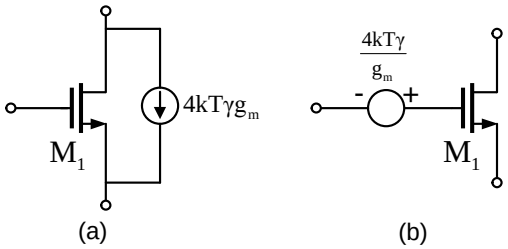


Figure 15: Thermal channel noise of MOS devise modeled as a (a) current source and (b) voltage source.

Another component of thermal noise arises from the gate resistance of MOS transistors. This effect tends to become increasingly more important as the gate length is scaled down. As illustrated in Figure 16(a), for a MOS transistor with width W and length L, this resistance is expressed as:

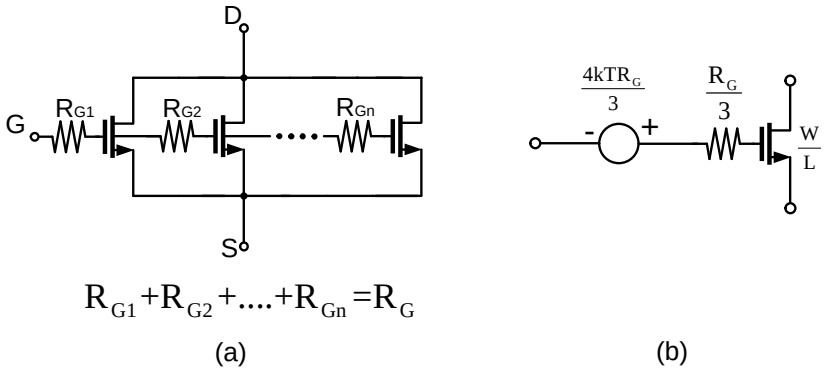


Figure 16: (a) equivalent circuit for noise calculations, (b) equivalent noise and resistance in lumped model.

$$R_G = \frac{W}{L} R_{\square}, \tag{3}$$

where $R_{\square}$ is the sheet resistance (resistance of one square) of the polysilicon gate. For example, if W = $1\mu m$, L = 45nm, and $R_{\square} = 15\Omega$, then $R_G = 333\Omega$. As shown in Figure 16(a), R_G is distributed over the width of the transistor, and for that reason, its noise must be calculated carefully. In Figure 16(b), one can see that the structure can be reduced to a lumped model having an equivalent gate resistance of $R_G/3$ with a thermal PSD of $4kTR_G/3$ [58]. In a good design, this noise must be less than that of the channel:

$$\frac{4kTR_G}{3} \ll \frac{4kT\gamma}{g_m}. \tag{4}$$

It should be noted that the gate and drain terminals also experience physical resistances, which are minimized through the use of multiple fingers. Several research studies have been done to investigate MOS transistor gate resistance since it tends to be more and more pronounced in advanced technologies and at very high frequencies. More details on that design parameter will be presented in Chapter 5.

MOSFET devices also experience "flicker" or "$1/f$" noise. This noise is modeled by a voltage source in series with the gate and exhibits the following PSD [56]:

$$\bar{V}_n^2 = \frac{K}{WLC_{OX}} \frac{1}{f}, \tag{5}$$

where K is a process-dependent constant which is given by

$$K = \frac{1}{2} \mu_n C_{OX} \frac{W}{L}, \tag{6}$$

where μ_n and C_{OX} are the carrier mobility are oxide capacitance, respectively.

For a given device size and bias current conditions, the $1/f$ noise PSD intercepts the thermal noise PSD at some frequency, which is called "$1/f$ corner frequency" (f_c). As depicted in Figure 17, f_c can be extracted by converting the flicker noise voltage to current and equating the result to the thermal noise current:

$$\frac{K}{WLC_{OX}}\frac{1}{f_c}g_m^2 = 4kT\gamma g_m, \tag{7}$$

which gives

$$f_c = \frac{K}{WLC_{OX}}\frac{g_m}{4kT\gamma}. \tag{8}$$

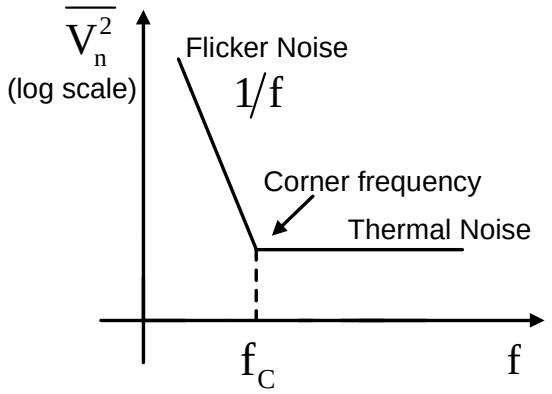


Figure 17: Flicker (-thermal) noise corner frequency.

Finally, another noise source in MOS transistors arise from the fluctuations of the channel charge in the inversion region of operation. These fluctuations usually induce a noisy current in the gate of the device due to capacitive coupling. This noise source is called induced gate noise. The concept of the induced gate noise has been used for decades and many researchers are still studying how to model it efficiently.

At very high frequencies, it is believed that the long channel voltage fluctuations due to thermal noise couple to the gate through the oxide capacitance and cause an induced gate noise current to flow. This noise current can be modeled by a noisy current source connected in parallel to the intrinsic gate-to-source capacitance C_{GS}. Usually, the induced gate noise is given by $\overline{i_{ng}^2}/\Delta f = 4kT\delta g_g$, where Δf is the frequency band, g_g is expressed as $g_g = \omega^2 C_{GS}^2/5g_{d0}$ and $\delta = 4/3$ for long-channel devises. g_{d0} is the zero-bias drain transconductance.

Noise Figure

In circuit and system design, SNR is of major importance for RF designers. The SNR is defined as the ratio of the signal power over the noise power. It is a useful figure-of- merit, which shows how the signal is degraded as it travels through a given circuit in presence of noise. If the circuit is noiseless, then the output SNR is equal to the input SNR even if the circuit acts like an attenuator. To quantify how noisy the circuit is, we define its NF as

$$NF = \frac{SNR_{in}}{SNR_{out}}. \tag{9}$$

Ideally, the noise figure of a noiseless circuit equals to unity. Since each quantity in this ratio has a dimension of power (or voltage squared), we express NF in decibels as

$$NF|_{dB} = 10log\frac{SNR_{in}}{SNR_{out}}. \tag{10}$$

NF plays an important role in designing RF circuits and in particular LNAs. The reason is that the noise contribution of an LNA has the greatest impact on the total receiver noise since the LNA is the first RF block in the

receiver chain.

3.2.3 *Non-Linearity*

One of the most important concepts in RF design is the non-linearity. Non-linear effects usually lead to important phenomena that can not be modeled by simple small-signal models. In this section, our analysis is limited to time-variant and memoryless systems. For simplicity, we assume that the input/output characteristic can be approximated by

$$y(t) \approx \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t). \tag{11}$$

The effects of non-linear and higher-order non-linear terms must be carefully studied, in order to justify that Equation 11 is a plausible representation. Non-linear effects are mostly due to the third-order term in Equation 11, while the second-order term manifests itself in certain types of transceivers [56].

Harmonic Distortion

If a sinusoid signal is applied to a non-linear system, the output, in general, experiences frequency components that are integer multiples of the input frequency. These multiples are called "harmonics". In Equation 11, if $x(t) = Acos(\omega t)$, then

$$\begin{aligned} y(t) &= \alpha_1 A cos\omega t + \alpha_2 A^2 cos^2\omega t + \alpha_3 A^3 cos^3\omega t \\ &= \alpha_1 A cos\omega t + \frac{\alpha_2 A^2}{2}(1 + cos2\omega t) + \frac{\alpha_3 A^3}{4}(3cos\omega t + cos3\omega t) \\ &= \frac{\alpha_2 A^2}{2} + \big(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\big) cos\omega t + \frac{\alpha_2 A^2}{2} cos2\omega t + \frac{3\alpha_3 A^3}{4} cos3\omega t. \end{aligned} \tag{12}$$

In Equation 12, the first term on the right-hand side is a DC quantity, which arises from second-order non-linear effects. The second term is called the "fundamental", the third term is the second harmonic, and the fourth term is the third harmonic [56].

From the above discussion, two observations can be made. First, even-order harmonics come from α_j with even j and disappear if the system has odd symmetry. In reality, random mismatches corrupt this symmetry, resulting in finite even-order harmonics. Second, in Equation 12, the amplitudes of the second and third harmonics are proportional to A^2 and A^3, respectively, and therefore, we say that the n-th harmonic is proportional to A^n.

Gain Compression

The small-signal gain of a circuit is often obtained assuming that harmonics are negligible. However, as the signal amplitude increases, the gain begins to vary. In fact, non-linearity can be viewed as variation of the small-signal gain with respect to the input level. This can be seen from the term $3\alpha_3 A^3/4$ added to $\alpha_1 A$ in Equation 12. In most RF circuits, the output is a "compressive" or "saturating" function of the input, which means that the gain approaches zero for considerably high input levels. In Equation 12 this occurs if $\alpha_1 < 0$. Written as $\alpha_1 + 3\alpha_3 A^2/4$, the gain is therefore a decreasing function of A. This effect is quantified by the 1-dB compression point (1-dB CP), defined as the input signal level that causes the small-signal gain to drop by 1 dB. If plotted on a log-log scale as a function of the input level, the output level falls below its ideal value by 1 dB at the 1-dB compression point [56], as depicted Figure 18.

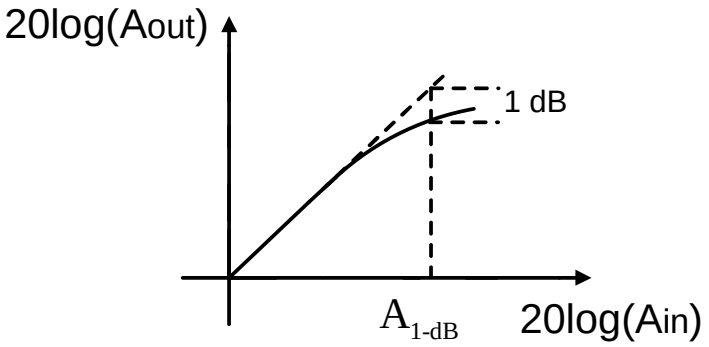


Figure 18: Definition of 1-dB compression point.

In order to calculate the 1-dB compression point, we can write from Equation 12

$$20\log\left|\alpha_1 + \frac{3}{4}\alpha_3 A_{1-dB}^2\right| = 20\log|\alpha_1| - 1dB, \tag{13}$$

which gives

$$A_{1-dB} = \sqrt{0.145\left|\frac{\alpha_1}{\alpha_3}\right|}. \tag{14}$$

The 1-dB compression point is a measure of the maximum input range of the circuit and occurs around -20 to -25 dBm (63.2 to 35.6 mV_{pp} in a 50Ω system) in typical RF amplifiers.

Intermodulation

Another interesting concept of non-linearity in RF design exists when two interferes (two different frequency tones) accompany the desired transmitted signal. If two frequency tones ω_1 and ω_2 are applied to a non-linear system, the output experiences components that are not harmonics of these frequencies. This phenomenon is called "intermodulation" (IM) and it arises from the "mixing" (multiplication) of the two components as their sum is raised to a power greater than unity. In order to understand how the non-linear input/output approximation of Equation 11 leads to the IM phenomenon, we fist assume that the input signal has a form of $x(t) = A_1 cos\omega_1 t + A_2 cos\omega_2 t$ and therefore we take

$$y(t) = \alpha_1(A_1 cos\omega_1 t + A_2 cos\omega_2 t) + \alpha_2 * (A_1 cos\omega_1 t + A_2 cos\omega_2 t)^2 + \alpha_3(A_1 cos\omega_1 t + A_2 cos\omega_2 t)^3 \tag{15}$$

After some long mathematical calculations, we obtain the following intermodulation products:

- at $2\omega_1 \pm \omega_2$: $\frac{3\alpha_3 A_1^2 A_2}{4} cos(2\omega_1 + \omega_2)t + \frac{3\alpha_3 A_1^2 A_2}{4} cos(2\omega_1 - \omega_2)t$
- at $2\omega_2 \pm \omega_1$: $\frac{3\alpha_3 A_1^2 A_2}{4} cos(2\omega_2 + \omega_1)t + \frac{3\alpha_3 A_1^2 A_2}{4} cos(2\omega_2 - \omega_1)t$

and the following fundamental components

- at ω_1, ω_2: $\left(\alpha_1 A_1 + \frac{3\alpha_3 A_1^3}{4} + \frac{3\alpha_3 A_1 A_2^2}{2}\right) cos\omega_1 t + \left(\alpha_1 A_2 + \frac{3\alpha_3 A_2^3}{4} + \frac{3\alpha_3 A_2 A_1^2}{2}\right) cos\omega_2 t$

Figure 19 offers a visualization of the aforementioned results, where the third-order IM products at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ are of major importance. This is because, if ω_1 and ω_2 are close to each other, then $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ appear in the vicinity of the fundamental frequencies ω_1 and ω_2. In an LNA, for example, if

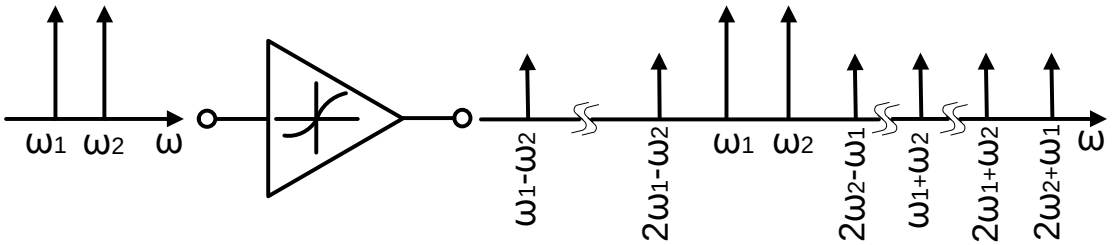


Figure 19: Intermodulation components in a two-tone test.

we assume that the interferer frequencies happen to satisfy $2\omega_1 - \omega_2 = \omega_0$, the intermodulation product at $2\omega_1 - \omega_2$ falls onto the desired channel, yielding a corrupted signal [56].

A usual way to characterize IM in an RF system is the "two-tone" test, where two sinusoid signals with equal amplitudes are applied to the input of interest. A measure of intermodulation is the third-order input intercept point (IIP_3). It is the point where the amplitude of the third-order intermodulation products (at $2\omega_1 - \omega_2$, $2\omega_2 - \omega_1$) becomes equal to that of the fundamental tones at the output. Figure 20 depicts on log-log scale the input level at which this occurs. Similarly, the corresponding output is represented by the third-order output intercept point (OIP_3).

To determine the IIP_3, we equate the fundamental and IM amplitudes as follows

$$|\alpha_1 A_{IIP_3}| = |\frac{3}{4}\alpha_3 A_{IIP_3}^3| \tag{16}$$

resulting in

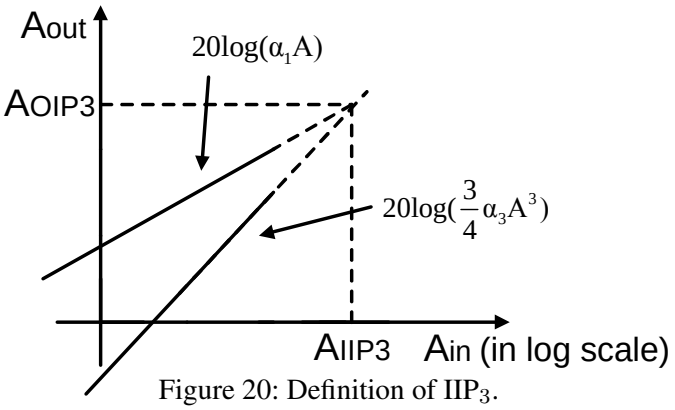


Figure 20: Definition of IIP_3.

$$A_{IIP_3} = \sqrt{\frac{4}{3}\left|\frac{\alpha_1}{\alpha_3}\right|}, \tag{17}$$

where A_{IIP_3} represents the input amplitude. Finally, it is interesting to observe that

$$\frac{A_{IIP_3}}{A_{1-dB}} = \sqrt{\frac{4}{0.435}} \approx 9.6dB. \tag{18}$$

It should be noted that the input and output third intercept points are used for receivers and transmitters, respectively.

3.3 CASE STUDY 1: 2.4GHZ CMOS INDUCTIVE DEGENERATED LNA

The first case study includes a typical 2.4GHz inductive degenerated LNA, as shown in Figure 21. Therefore, before we proceed, we present the design of the circuit since this will facilitate the following discussions.

The biasing stage of the circuit is formed by resistors R_{REF} and R_{BIAS} and transistor M_3. The inductors L_G and L_S provide the appropriate input matching. The gain stage is composed by M_1 and M_2. Transistor M_1 provides the high gain, whereas M_2 isolates the input from the output to compensate for the Miller effect and, thereby, to increase reverse isolation. At the circuit output, inductor L_D resonates with the gate-to-source capacitor of M_2 and the output capacitors C_{OUT} and C_D. This parallel LC resonator, as well as the transconductance of transistor M_1, are the main contributors for controlling the gain amplitude at the resonance frequency of 2.4GHz. Using this topology, we can match the circuit's input impedance to 50Ω and decrease the noise figure by carefully dimensioning inductors L_G, L_S, and transistor M_1. Transistors M_1 and M_3 are biased in moderate to strong inversion, in order to minimize drain current mismatch and achieve higher f_T. Finally, capacitor C_{EXT} and C_{IN} are used for fine tuning of the input matching and DC blocking, respectively.

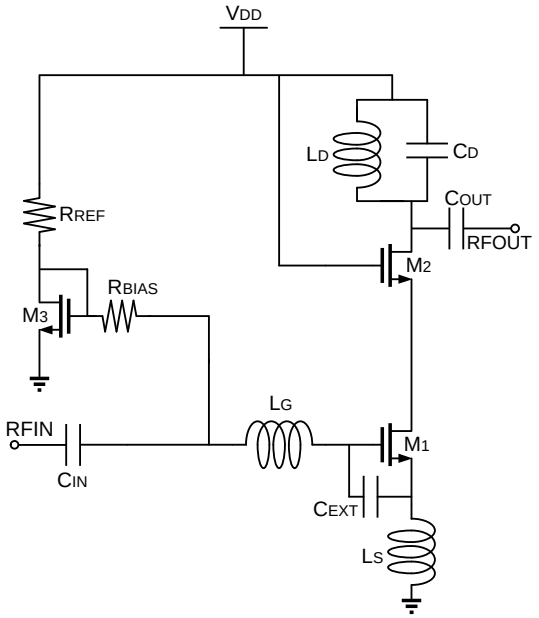


Figure 21: 2.4GHz CMOS inductive degenerated LNA schematic.

3.3.1 *Input Impedance*

We start the performance analysis of our LNA case study by first computing the input impedance of the circuit. For simplicity, we neglect the effects of C_{GD} and C_{SB} as well as the channel-length modulation and body effect. The high-frequency small-signal equivalent of the input loop of the LNA is depicted in Figure 22. In order to compute the input impedance of the circuit, we apply a test voltage V_x at the input, we compute the resultant test current I_x, and finally we estimate the ratio $Z_{in} = {}^{V_x}/_{I_x}$.

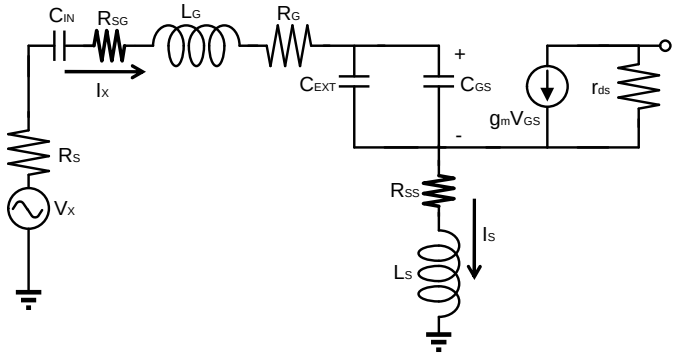


Figure 22: High-frequency small-signal equivalent of the input loop of the LNA.

Taking Kirchoff's Voltage Law (KVL) around the input loop and assuming that the test current I_x flows entirely through the parallel combination of C_{GS} and C_{EXT} capacitors, we have:

$$V_x = I_x \frac{1}{j\omega C_{IN}} + I_x Z_{L_G} + I_x R_G + I_x \frac{1}{j\omega(C_{GS} + C_{EXT})} + I_S Z_{L_S}, \tag{19}$$

where Z_{L_G} and Z_{L_S} correspond to the impedances of gate and source inductors L_G and L_S, respectively. We rewrite Equation 19 as

$$V_x = I_x \frac{1}{j\omega C_{IN}} + I_x(j\omega L_G + R_{SG}) + I_x R_G + I_x \frac{1}{j\omega(C_{GS} + C_{EXT})} + I_S(j\omega L_S + R_{SS}), \tag{20}$$

where R_{SG} and R_{SS} are the series resistances of inductors L_G and L_S, respectively.
Kirchoff's Current Law (KCL) at the source terminal of M_1 gives

$$I_S = I_x + g_m V_{GS}. \tag{21}$$

Substituting Equation 21 into Equation 20, we have

$$\begin{aligned} V_x = I_x \frac{1}{j\omega C_{IN}} + I_x j\omega L_G + I_x R_{SG} + I_x R_G + I_x \frac{1}{j\omega(C_{GS} + C_{EXT})} + (I_x + g_m V_{GS})(j\omega L_S + R_{SS}) \\ = I_x \frac{1}{j\omega C_{IN}} + I_x j\omega L_G + I_x R_{SG} + I_x R_G + \\ + I_x \frac{1}{j\omega(C_{GS} + C_{EXT})} + I_x j\omega L_S + I_x R_{SS} + g_m V_{GS} j\omega L_S + g_m V_{GS} R_{SS} \end{aligned} \tag{22}$$

Dividing Equation 22 at both sides by test current I_x, we have:

$$\begin{aligned} Z_{in} = \frac{V_x}{I_x} = \frac{1}{j\omega C_{IN}} + j\omega L_G + R_{SG} + R_G \\ + \frac{1}{j\omega(C_{GS} + C_{EXT})} + j\omega L_S + R_{SS} + \frac{g_m V_{GS} j\omega L_S}{I_x} + \frac{g_m V_{GS} R_{SS}}{I_x} \end{aligned} \tag{23}$$

Substituting $\frac{V_{GS}}{I_x} = \frac{1}{j\omega(C_{GS} + C_{EXT})}$ to Equation 23, we have:

$$\begin{aligned} Z_{in} = \frac{V_x}{I_x} = \frac{1}{j\omega C_{IN}} + j\omega L_G + R_{SG} + R_G + \frac{1}{j\omega(C_{GS} + C_{EXT})} + \\ + j\omega L_S + R_{SS} + \frac{g_m V_{GS} j\omega L_S}{I_x} + \frac{g_m V_{GS} R_{SS}}{I_x} \\ = \frac{1}{j\omega C_{IN}} + j\omega L_G + R_{SG} + R_G + \frac{1}{j\omega(C_{GS} + C_{EXT})} + \\ + j\omega L_S + R_{SS} + \frac{g_m L_S}{(C_{GS} + C_{EXT})} + \frac{g_m R_{SS}}{j\omega(C_{GS} + C_{EXT}} \\ = R_{SG} + R_G + R_{SS} + \\ + \frac{g_m L_S}{(C_{GS} + C_{EXT})} + j\left(\omega L_G + \omega L_S - \frac{1}{\omega C_{IN}} - \frac{1}{\omega(C_{GS} + C_{EXT})} - \frac{R_{SS}}{\omega(C_{GS} + C_{EXT})}\right) \end{aligned} \tag{24}$$

In order to achieve an impedance matching at 50Ω and at the operating frequency ω_0, we need to have

$$R_{SG} + R_G + R_{SS} + \frac{g_m L_S}{(C_{GS} + C_{EXT})} = R_S = 50 \tag{25}$$

$$j\left(\omega_0 L_G + \omega_0 L_S - \frac{1}{\omega_0 C_{IN}} - \frac{1}{\omega_0(C_{GS}+C_{EXT})} - \frac{R_{SS}}{\omega_0(C_{GS}+C_{EXT})}\right) = 0. \tag{26}$$

The MOS gate resistance R_G is usually negligible since it does not affect considerably the electrical behavior of amplifiers at the typical 2.4GHz frequency bandwidths.

Moreover, it must be noted that series resistances R_{SG} and R_{SS} are not known before we start designing the circuit and their values depend on the dimensions of the corresponding inductors. The dimensions of the inductors determine the inductance value and are chosen such that impedance matching of 50Ω is achieved. R_{SG} and R_{SS} play an important role in our proposed test technique, as it will be shown later on. For simplicity, they are neglected in the calculations of the rest of the performances in the following paragraphs.

3.3.2 *Noise Figure (NF)*

We continue with calculating the noise figure of the LNA, neglecting, for simplicity, the effects of channel-length modulation, body effect C_{GD}, DC blocker C_{IN}, and gate resistance R_G. As depicted in Figure 23, the noise introduced by transistor M_1 is represented by current source I_{n1}. Considering the output current of M_1, I_{out}, we have [56]:

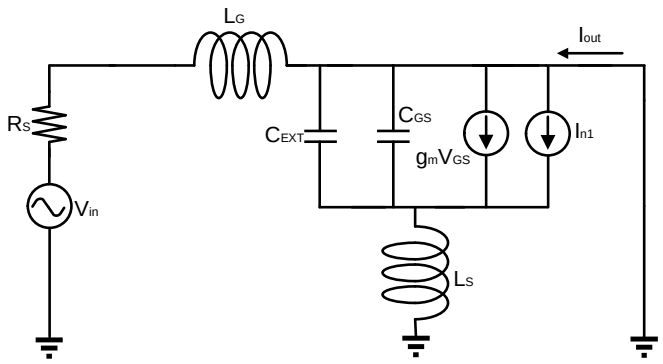


Figure 23: Equivalent circuit for calculating the NF of the LNA.

$$I_{out} = g_m V_{GS} + I_{n1}. \tag{27}$$

A KVL arround the input loop gives

$$V_{in} = (R_S + sL_G)V_{GS}s(C_{GS}+C_{EXT}) + V_{GS} + sL_S(I_{out} + V_{GS}s(C_{GS}+C_{EXT})) \tag{28}$$

Substituting for V_{GS} from Equation 27 gives

$$V_{in} = I_{out}sL_S + \frac{\left[(L_S+L_G)(C_{GS}+C_{EXT})s^2 + 1 + R_S s(G_{GS}+C_{EXT})\right]}{g_m}(I_{out} - I_{n1}) \tag{29}$$

The input network is designed to resonate at the frequency of interest ω_0 (2.4GHz in our case), that is, $(L_S + L_G)(C_{GS}+C_{EXT})s^2 + 1 = 0$ It follows that

$$V_{in} = I_{out}\left(jL_S\omega_0 + \frac{jR_S(C_{GS}+C_{EXT})\omega_0}{g_m}\right) - I_{n1}\frac{jR_S(C_{GS}+C_{EXT})\omega_0}{g_m}. \tag{30}$$

The coefficient of I_{out} represents the transconductance gain of the circuit including R_S

$$\left|\frac{I_{out}}{V_{in}}\right| = \frac{1}{\omega_0\left(L_S + \frac{R_S(C_{GS}+C_{EXT})}{g_m}\right)}. \tag{31}$$

As has already been stated, for input matching, we need to have $\frac{g_m L_S}{(C_{GS}+C_{EXT})} = R_S$ and since $\frac{g_m}{C_{GS}+C_{EXT}} \approx \omega_T$ (recall that $\omega_T = 2\pi f_T$), it follows that

$$\left|\frac{I_{out}}{V_{in}}\right| = \frac{\omega_T}{2\omega_0}\frac{1}{R_S}. \tag{32}$$

As long as the input is matched to 50Ω, the transconductance of the circuit remains independent of L_S, L_G, and g_m

Setting now $V_{in} = 0$ in Equation 30, we compute the output noise which is introduced by transistor M_1

$$|I_{n,out}|_{M_1} = |I_{n1}|\frac{R_S(C_{GS}+C_{EXT})}{g_m L_S + R_S(C_{GS}+C_{EXT})}, \tag{33}$$

which, for $g_m L_S/(C_{GS}+C_{EXT}) = R_S$, reduces to

$$|I_{n,out}|_{M_1} = \frac{|I_{n1}|}{2} \tag{34}$$

and therefore

$$\overline{I^2_{n,out}}|_{M_1} = kT\gamma g_m. \tag{35}$$

Dividing the output noise current by the transconductance of the circuit and by $4kTR_S$ and adding unity to the result, we have the noise figure of the circuit [60]

$$NF = 1 + g_m R_S \gamma \left(\frac{\omega_0}{\omega_T}\right)^2. \tag{36}$$

As mentioned in the beginning of our calculations, several effects and design parameters, such as channel-length modulation, body effect, DC blocker C_{IN}, gate resistance R_G, parasitic capacitance C_{GD}, and series resistances of the inductors' coils, respectively, were ignored, in order to simplify the analysis of NF.

It is of major importance to bear in mind that the above NF formula is valid only at the input resonance frequency and only when the input is matched.

3.3.3 *Gain*

The voltage gain of the LNA topology shown in Figure 21, is equal to the product of the transconductance of the circuit and the load resistance R, which includes: (a) the output resistance of the cascode stage seen at the drain of transistor M_2, (b) the series resistance of inductor L_D, and (c) capacitors C_D and C_{OUT}. These series resistances represent the losses of these passive components. Using Equation 32, the voltage gain of the circuit is estimated by [56]:

$$A_V = \frac{V_{out}}{V_{in}} = \frac{\omega_T}{2\omega_0}\frac{R}{R_S} = \frac{R}{2L_S\omega_0}. \tag{37}$$

Furthermore, the power gain is given by

$$A_P = \frac{P_{OUT}}{P_S} = \frac{V^2_{out}/R}{V^2_{in}/R_S} = \left(\frac{\omega_T}{\omega_0}\right)^2 \frac{R}{4R_S}. \tag{38}$$

The power gain increases with ω_T. Therefore, for a specific frequency of operation, moving towards deep-submicron technologies improves the gain of the LNA. The power gain can be also expressed as

$$A_P = Q^2_{in} g^2_m R R_S, \tag{39}$$

where Q_{in} is the effective quality factor of the amplifier's input network and is given by

$$Q_{in} = \frac{V_{GS}}{V_{in}} = \frac{1}{2\omega_0(C_{GS} + C_{EXT})R_S} \tag{40}$$

where the capacitance C_{IN} has not been taken into account for simplicity. However, it still affects the quality factor of the amplifier's input network, and thus the performances of the LNA.

We recall that for the derivation of voltage and power gains, it is assumed that the input and output of the LNA are matched to 50Ω.

3.3.4 *Third Order Input Intercept Point (IIP_3)*

As mentioned before, one of the main non-linearity problem of the LNA studied in this work is the third-order input intercept point. For that reason, we evaluate the intermodulation performance of a MOS transistor using the third-order intermodulation intercept voltage amplitude, which corresponds to the gate-source voltage for which the intermodulation drain current intercepts the fundamental drain current [54]. First, let us assume that the drain current of M_1 is the output of interest. For a simple square-law device, we have:

$$I_{DS} = \frac{1}{2}\mu C_{OX}\frac{W}{(L - \Delta L)}\frac{V_{OV}^2}{(1 + \theta V_{OV})} \tag{41}$$

where θ models the carriers mobility degradation due to VS effect and transverse electric field. V_{OV} corresponds to the overdrive gate voltage of transistor M_1, $V_{GS} - V_T$ and the factor $L - \Delta L$ models the channel length modulation.

Considering now small-signal operation, the output current is expressed as:

$$i_{ds} = \frac{1}{2}\mu C_{OX}\frac{W}{L}\left(\frac{(V_{OV} + v_{gs})^2}{(1 + \theta(V_{OV} + v_{gs}))} - \frac{V_{OD^2}}{(1 + \theta V_{OV})}\right) \tag{42}$$

where a constant drain-source voltage V_{DS} is assumed to be constant and the channel length modulations has been neglected.

Equation 42 can be expanded in a Taylor series [54] and after a few calculations, gives the following (based also on Equation 17):

$$IV3 = \sqrt{\frac{4}{3}\left|\frac{\alpha_1}{\alpha_3}\right|} = \sqrt{\frac{4}{3}\frac{V_{OV}(2 + \theta V_{OV})(1 + \theta V_{OV})^2}{\theta}}, \tag{43}$$

where IV3 is expressed in volts of amplitude.

In conclusion, for any NMOS transistor, the gate-source intercept voltage is estimated by the aforementioned equation. In essence, IV3 is the small-signal voltage amplitude at which the ratio of the amplitude of the third-order intermodulation signal over the amplitude of the fundamental one is equal to unity. Notice also that the IV3 is similar to the IIP_3 [61], [62].

Now, our focus is turned to estimating the IIP_3 of the LNA. First, the definition of the available source power, which is provided by V_{in} through the resistance of 50Ω, is given, as shown in Figure 24:

$$P_S = \frac{V_{in}^2}{4R_S} \tag{44}$$

Equations 44 and 40 are used to estimate the available source power of the circuit in correspondence to the small-signal v_{gs} voltage at which the third-order intermodulation ratio is equal to unity, as given by Equation 43.

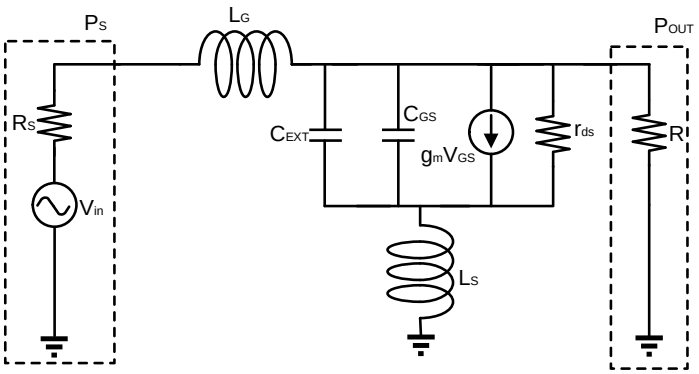


Figure 24: Small-signal equivalent circuit of the 2.4GHz LNA for calculating both voltage/power gain and IIP_3.

Thus, the IIP_3 of the circuit is expressed as:

$$IIP_3 = \frac{IV3^2}{2x4Q_{in}^2 R_S} = \frac{IV3^2 \omega_0^2 (C_{GS} + C_{EXT})^2 R_S}{2} \tag{45}$$

where the IIP_3 is expressed in Watts.
Substituting Equation 43 to Equation 45, we have:

$$IIP_3 = \frac{1}{8R_S} \frac{4}{3} \frac{V_{OV}(2 + \theta V_{OV})(1 + \theta V_{OV})^2}{\theta} \frac{1}{Q_{in}^2}, \tag{46}$$

which is also expressed in Watts.

Finally, IIP_3 can be expressed in units of dBm as follows:

$$IIP_3 = 5.25 + 10log\left(\frac{V_{OV}(2 + \theta V_{OV})(1 + \theta V_{OV})^2}{\theta}\right) - 20log(Q_{in}). \tag{47}$$

As it can be seen from Equation 47, the IIP_3 performance of the LNA is not only a function of the DC bias conditions of the circuit, but also depends on the quality factor of the amplifier's input network. Finally, according to the literature, it is noted that several analyses have been performed regarding the estimation of linearity performances of the LNA. For example, in [56], the author presents mathematically the calculation of the non-linear terms in the drain current of M_1 in the common-source configuration of Figure 24. It is shown that the drain current of M_1 exhibits third-order non-linear components, which are also caused by the degenerated inductor L_S. Moreover, another analysis in [56] shows that the series resistance of the inductors produce high-order non-linear effects that are directly related to the linearity performances (e.g. IIP_3) of the LNA. In particular, the series resistance of L_S is represented by a resistive degenerated common-source stage, in order to employ the analysis. All these observations will facilitate our discussions later on, where we show that variations in linearity performances of the LNA arise from both the quality factor of the amplifier's input network and the self-inductance and series resistances of the inductors' coils.

3.3.5 *Design Methodology*

Having understood how the degenerated CMOS LNA works, we proceed to the description of the design procedure. First, the procedure starts with five known design parameters, namely the operating frequency ω_0, the source and gate inductances L_S and L_G, the DC blocker C_{IN}, and the external capacitor C_{EXT}, which is connected between the gate and source terminal of transistor M_1 in parallel with MOS intrinsic capacitance C_{GS}. Apart from the operating frequency, each of the aforementioned design parameters are flexible during the procedure, and their values are selected such that the performances meet the initial design specifications after a number of necessary iterations. With ω_0 known, C_{GS} capacitance is estimated by Equation 26. ω_T

($= 2\pi f_T$) and g_m ($= \omega_T(C_{GS} + C_{EXT})$) are estimated by Equation 25. Next, we build the plots of g_m and f_T as functions of the overdrive voltage $V_{GS} - V_T$, in order to determine the width of transistor M_1, which can yield the estimated g_m and f_T simultaneously. Figure 25 plots the g_m and f_T of a transistor with W=20μm and L=0.06μm. Notice that the length of the transistors in the LNA is chosen to have the minimum value provided by the technology, in order to achieve as high f_T as possible.

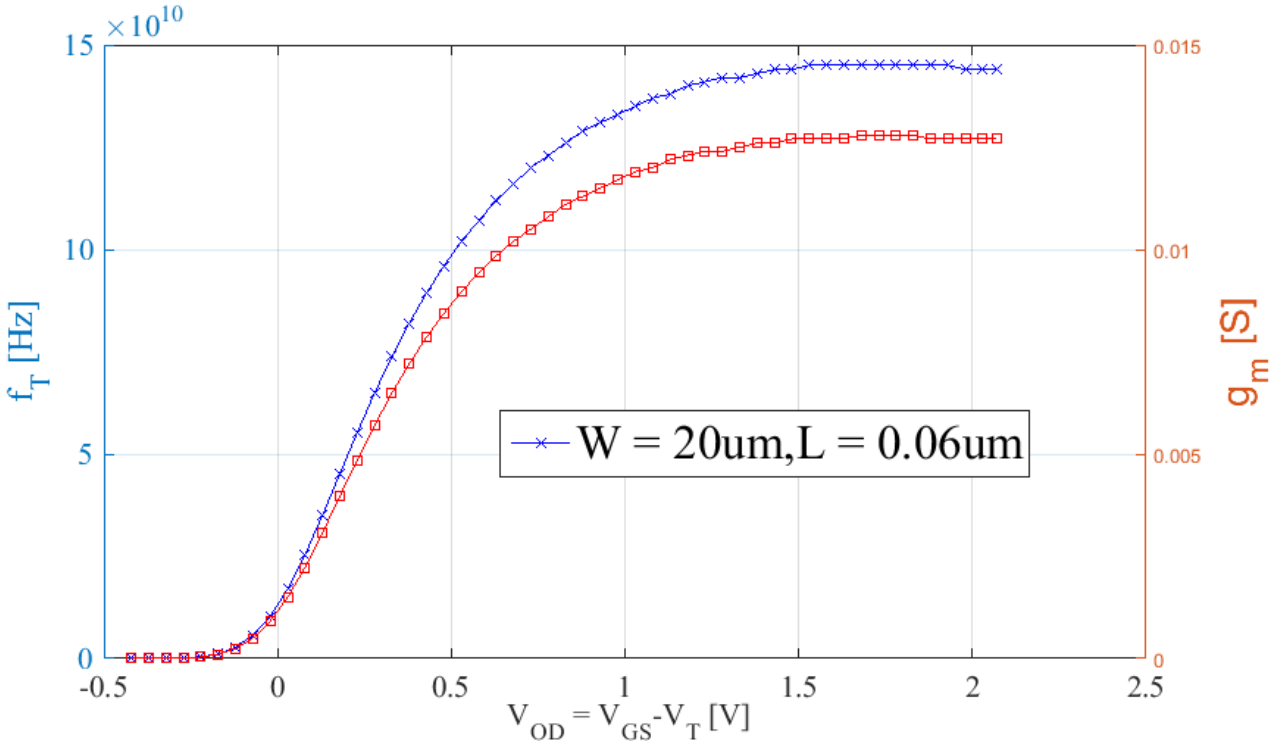


Figure 25: f_T and g_m versus overdrive voltage.

The next step is to size properly the cascode transistor. The width of the cascode transistor M_2 is chosen half to that of transistor M_1, in order to increase the isolation between the input and output of the LNA. The design procedure continues with selecting a proper value for inductor L_D, which resonates at ω_0 with capacitances C_D and C_{OUT}, drain-gate and drain-bulk capacitances of transistor M_2, and the parasitic capacitance and series resistance of inductor L_D. The next step of the design procedure includes the investigation of the input matching. Due to the Miller gate-drain capacitance of transistor M_1, the real and imaginary parts of the input impedance could deviate from their ideal values, requiring a proper tuning of L_G and L_S. Finally, knowing the DC bias conditions (see Figure 25) and the dimensions of transistor M_1, the drain current is estimated and a proper current mirror configurations is selected for providing the appropriate bias to transistor M_1. Alternatively, an external DC bias can be used instead, increasing the number of Input/Output (I/O) pads of the design.

3.4 CASE STUDY 2: 60GHZ MM-WAVE CMOS 3-STAGE LNA

The second case study is a wide-band mm-wave CMOS 3-stage LNA operating at 60GHz. In this section, some helpful guidelines for designing mm-wave amplifiers will be presented first. Apart from MOS devices, capacitors, and resistors, these circuits are implemented with transmission lines, which are used for impedance matching, interconnect wires, and bias networking. The schematic of the 60GHz LNA as well as its design methodology are discussed briefly at the end of the section.

3.4.1 *Guidelines for Designing mm-wave LNAs*

The list of the most important performances of LNAs that operate at mm-wave frequencies include the power gain, impedance matching, noise figure, stability, linearity, power dissipation and frequency bandwidth. The design methodologies that are considered in designing mm-wave amplifiers are very similar to them that are used for designing and implementing amplifiers that operate at typical much lower frequency ranges. However, the circuit topologies are properly chosen, in order to account for the following design challenges [63]: (a) transistors are designed so that they operate closer to their cut-off frequency, (b) signals with small wavelengths are used resulting in distributed effects within the actual elements of the circuit, and (c) parasitic components represent a much larger portion of the total impedance on a given node. These design challenges are discussed in the following paragraphs.

First, transistors that operate close to their cut-off frequency have lower gain and higher noise figure. For that reason, multi-stage topologies are usually selected, in order to provide enough gain and minimize the noise contribution from the mixer and the subsequent stages of the receiver. However, a disadvantage of using multi-stage amplifiers is that they (a) dissipate more power, (b) occupy more area within the chip, (c) introduce more harmonics, and (d) include more nodes that require stability checks during the design steps.

Moreover, in LNAs that operate at mm-wave frequencies, when signals pass through devices that have a size comparable to the wavelength, they actually behave as waves rather than simple stationary voltages or currents. This leads to a appreciable phase delay across these component, and as a result, distributed effects need to be taken into account and accurately modeled before designing a mm-wave circuit.

Finally, parasitic components in circuits that operate in mm-wave frequency ranges include shunt capacitances, series inductances, and series resistances which depend directly on frequency. For instance, a capacitance value at 60GHz is approximately 30fF. A parasitic capacitance of 3fF represents about 10% of the total capacitance on the node of interest resulting in shifting the resonance frequency by 5%.

Common-source and cascode stages are usually used, in order to design mm-wave amplifiers. Several mm-wave amplifiers have been designed using cascode topologies [2], [64], [65]. The advantage of using cascode configurations is that they can be made unconditionally stable at the operating frequency. Consequently, the design becomes more robust and the matching networks simpler. On the other hand, the noise contribution of a cascode is higher compared to common-source stages when operating close to f_T due to the reduced degeneration on the cascode stages at high frequencies. For this reason, common-source stages are preferred for low noise designs.

Microstrip and Coplanar Waveguide (CPW) transmission lines are used in the design for impedance matching, interconnect wiring, and bias networks. The length of the lines is usually much shorter than $\lambda/4$, in order to reduce losses and minimize noise contributions at the input of the amplifiers. Finally, Metal-Oxide-Metal (MOM) capacitors are used for (a) AC-coupling of the input and output as well as intermediate stages, and (b) as bypass capacitors for the DC feed lines.

Before presenting the second case study of this work, it would helpful to give a brief description of transmission lines, which are one of the most critical components in designing mm-wave circuits.

3.4.2 *Transmission Lines*

Transmission lines play an important role in the range of mm-wave frequencies. Long structures such as quarter wave can be realized on-chip due to the relatively small wavelength. They are components suitable for designing circuits that operate at high frequencies since there is no ambiguity in how the reference planes are defined. Furthermore, it is easy to connect a transmission line structure at any point in the circuit and predict the resulting reactances. Typically, in mm-wave frequencies, the reactive components that are needed for matching networks and resonators become significantly small, requiring inductance values in the order of 50-250 pH. Given the quasi transverse electromagnetic (quasi-TEM) mode of propagation, transmission lines are scalable in length and are capable of employing small values of reactances very precisely [2].

A transmission line is characterized by its equivalent frequency-dependent RLGC distributed model shown in Figure 26, where R, L, C, and G correspond to the series resistance per unit length in Ω/m, series inductance per unit length in H/m, shunt capacitance per unit length in F/m, and shunt conductance per unit length in S/m, respectively. The model is related to the characteristic impedance (Z_0) and complex propagation constant (γ) of the line [66] as follows:

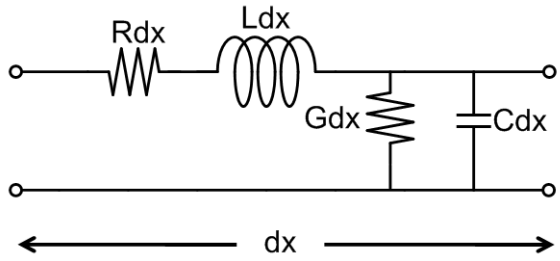


Figure 26: Distributed RLGC lossy transmission line model.

$$Z_0 = \sqrt{\frac{R + j\omega_0 L}{G + j\omega_0 C}} \tag{48}$$

$$\gamma = \sqrt{(R + j\omega_0 L)(G + j\omega_0 C)} \tag{49}$$

In the state-of-the-art, there are two options for realizing transmission lines on-chip, namely (a) microstrip and (b) CPW transmission lines.

Microstrip lines on silicon are implemented using the top layer metal as the signal path, and the bottom layer metal for the ground plane. The metal shield protects the substrate from any electric field penetration and as a result, the shunt loss G is only due to the loss tangent of the oxide, yielding a capacitive quality factor of 30 at mm-wave frequencies [2]. The most important disadvantage of employing microstrip lines on CMOS silicon substrate is the close physical proximity of the ground planes to the signal path, generating a very distributed inductance and thus yielding a low inductive quality factor.

CPW transmission lines are implemented with one signal line surrounded by two adjacent grounds. The width of the signal path is used to minimize the conductor's loss, whereas the signal-to-ground spacing controls Z_0 and the trade-off between capacitive and inductive quality factors. In general, CPW lines are preferred in design of amplifiers because of their significantly higher inductive quality factor compared to that of microstrip transmission lines.

3.4.3 *60GHz LNA Design*

The second case studying of this work is a wide-band mm-wave 3-stage LNA operating at 60GHz. The schematic of the LNA is as depicted in Figure 27. The first stage uses a common-source topology, in order to improve the NF, whereas the last two identical stages use a cascode topology, so as to achieve high gain and isolation. In addition, the two cascode amplification stages ensure an unconditional stability for the LNA in the 0-70GHz frequency range. Transistors in both the common-source and the cascode amplification stages are biased at the strong-inversion side of moderate inversion, so as to achieve high transconductance efficiency, high f_T, and high bandwidth. Micro-strip transmission lines and short ended stubs are used for impedance matching, interconnect wiring, and the bias networks. The transmission lines at the gate and drain terminals of the transistors are used to supply bias and are also incorporated into the matching networks. Finally, the input and output of the LNA are AC-coupled and have been designed along with Ground-Signal-Ground (GSG) RF pads, in order to be matched at an impedance of 50Ω.

All the performances of the LNA design are a strong function of several design parameters, including MOS transconductances, bias voltages, MOS capacitances, MOS gate and substrate resistances, MOM capacitors, resistors and, of course, characteristic impedances of transmission lines. As it will be shown later, MOS gate

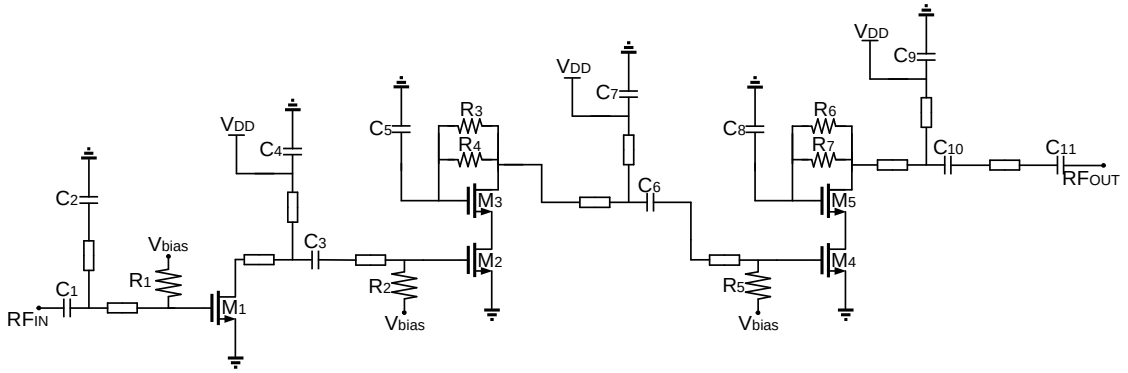


Figure 27: 60GHz mm-wave 3-stage LNA schematic implemented in 65nm CMOS bulk technology.

resistance is one of the most critical design parameters that affect significantly the electrical behavior of the 60GHz LNA and in general, of amplifiers that operate in the range of mm-wave frequencies.

In the rest of this section, we provide a brief theoretical analysis of the most important performances of the 60GHz LNA case study, including the noise figure and the small-signal gain.

Starting with the noise figure, Figure 28 depicts the high-frequency small-signal equivalent of a common-source stage where 1) the C_{GD} effect has been neglected and 2) four noise sources are considered, including:

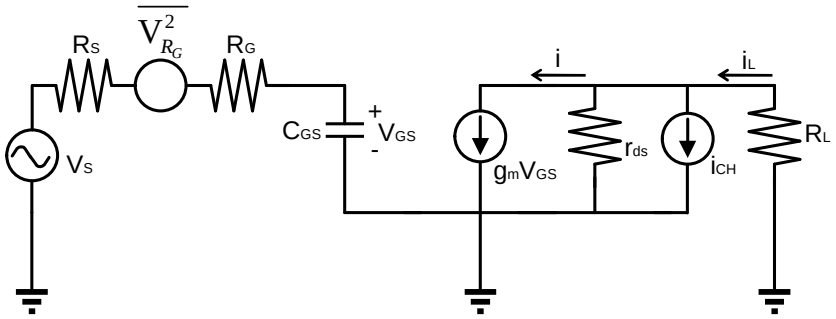


Figure 28: High-frequency small-signal equivalent of a common-source stage for noise calculations.

- R_S: $\overline{V_S^2} = 4kTR_SB$
- R_G: $\overline{V_{R_G}^2} = 4kTR_GB$
- R_{CH}: $\overline{I_{CH}^2} = 4kT\gamma g_{d0}B$
- R_L: $\overline{I_L^2} = 4kTG_LB$

where B is the bandwidth, G_L is the load transconductance, and g_{d0} corresponds to the channel transconductance when $V_{DS} = 0$ in case of a long transistor. Assuming first low-frequency operation, summing all the noise at the drain terminal of the transistor, we have [67]:

$$\overline{i^2} = \overline{I_{CH}^2} + \overline{I_L^2} + (\overline{V_S^2} + \overline{V_G^2})g_m^2, \tag{50}$$

which results in the estimation of the noise figure as follows:

$$NF = 1 + \frac{\overline{V_G^2}}{\overline{V_S^2}} + \frac{\overline{I_{CH}^2} + \overline{I_L^2}}{g_m^2\overline{V_S^2}} = 1 + \frac{R_G}{R_S} + \frac{g_{d0}\gamma + G_L}{R_S g_m^2}. \tag{51}$$

Next, assuming that $g_m = g_{d0}$ for a long device, noise figure is estimated by:

$$NF = 1 + \frac{R_G}{R_S} + \frac{\gamma}{R_S g_m} + \frac{G_L G_S}{g_M^2}. \tag{52}$$

Finally, we estimate the noise figure of the common-source stage in higher frequencies by taking into account the frequency dependence of the circuit transconductance G_m, which is calculated as follows:

$$i = g_m V_{GS} = g_m \frac{V_S}{R_S + R_G + \frac{1}{j\omega C_{GS}}} \left(\frac{1}{j\omega C_{GS}}\right) =$$
$$= V_S \frac{g_m}{1 + j\omega C_{GS}(R_S + R_G)} \approx V_S \frac{g_m}{j\omega C_{GS}(R_S + R_G)} \tag{53}$$

This equation can be written as $i = G_m V_S$, where:

$$G_m = -j\frac{\omega_T}{\omega}\frac{1}{R_S + R_G}, \tag{54}$$

which consolidates the noise calculations since the total noise is given by

$$\overline{i^2} = G_m^2(\overline{V_G^2} + \overline{V_S^2}) + \overline{I_{CH}^2}. \tag{55}$$

Hence, the noise figure is estimated as follows:

$$NF = 1 + \frac{\overline{V_G^2}}{\overline{V_S^2}} + \frac{\overline{I_{CH}^2}}{G_m^2 \overline{V_S^2}}. \tag{56}$$

Next, we substitute all the noise sources to the above equation and we have:

$$NF = 1 + \frac{R_G}{R_S} + \frac{(\frac{\gamma}{\alpha})g_m}{R_S}\left(\frac{\omega}{\omega_T}\right)^2 (R_S + R_G)^2. \tag{57}$$

Finally, assuming that $R_S \gg R_G$, we have:

$$NF = 1 + \frac{R_G}{R_S} + \left(\frac{\gamma}{\alpha}\right)\left(\frac{\omega}{\omega_T}\right)^2 (g_m R_S). \tag{58}$$

In following chapters, we will see how R_G affects the RF behavior of the LNA and how it explains to a large degree the variations in the performances of the circuit as well as the accuracy of our proposed test approach. Finally, notice that Equation 58 is very similar to Equation 36.

Continuing the theoretical analysis of the performances of the 60GHz LNA case study, we now give an approximated mathematical model of the small-signal gain. Figure 29 depicts the high-frequency small-signal equivalent circuit of a common-source stage. The C_{GD} effect has also been neglected, in order to simplify the calculations.

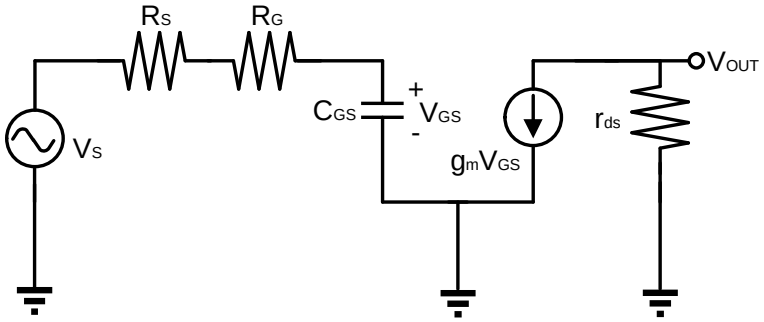


Figure 29: High-frequency small-signal equivalent of a common-source stage for gain calculations.

First, for the output voltage V_{OUT}, we can write:

$$V_{OUT} = -g_m V_{GS} r_{ds}. \tag{59}$$

Also, the voltage divider, which is formed in the loop from the input voltage V_S until the MOS intrinsic capacitance C_{GS}, gives:

$$V_{GS} = \frac{\frac{1}{j\omega C_{GS}}}{R_G + R_S + \frac{1}{j\omega C_{GS}}} V_S. \tag{60}$$

Combining the above equations, we have:

$$\frac{V_{OUT}}{V_S} = Gain = -g_m r_{ds} \frac{1}{1 + j\omega C_{GS}}. \tag{61}$$

A similar analysis can be done in the case of the cascode stage for its noise contribution and small-signal gain.

3.5 CONCLUSIONS

In this chapter, a brief introduction to RF circuits as well as the basic concepts in their design methodology were given. We presented our two case studies, where the proposed test approach will be demonstrated, namely a 2.4GHz CMOS inductive degenerated LNA and a wide-band 60GHz mm-wave CMOS 3-stage LNA. We carried out a theoretical performance analysis to derive equations that express the performances as a function of the most important design parameters.

4

PROCESS VARIATIONS

As mentioned in Chapter 2, BIT using non-intrusive variation-aware sensors is a test approach that capitalizes on the undesired phenomenon of process variations ans is employed to minimize the cost of RF test. Therefore, before discussing the principle of operation of non-intrusive variation-aware sensors in Chapter 5, it would be helpful to re-visit some of the classification schemes of process variations and variability sources in the fabrication line. Section 4.1 gives the definition of process variations as well as a brief description of their sources which are covered in Section 4.2 in more details. In Section 4.3, we classify the different types of process variations based on different criteria and based on this classification, we present a brief overview of process variations in Section 4.4. Process variations related to geometry and material variations are described. Some of the most important process parameters which are responsible for device and interconnect variations in 65nm CMOS bulk technology are discussed in Section 4.5. Finally, Section 4.6 concludes the chapter.

4.1 DEFINITION OF PROCESS VARIATIONS

Variation is the deviation from intended or designed values for a structure or circuit parameter of concern. During the IC fabrication steps, various process parameters, such as dopant concentration, effective channel length, oxide thickness, carrier mobility, transistor width, inter-layer dielectric thickness, sheet resistance, etc., can shift from their nominal values due to different variation sources, including lens imperfection, proximity effects, temperature gradients, misalignment masks, etc. Variability in active and passive devices as well as interconnects tend to be more critical as CMOS technologies scale down. For that reason, accurate characterization, analysis, and modeling of process variations is deemed necessary, in order to maintain the production yield as high as possible.

4.2 SOURCES OF VARIATIONS

As technology scales down, variability in transistors and interconnects performance will continue to increase, making these devices less reliable. This creates several challenges in building reliable systems, from the unpredictability of delay to increasing leakage currents. During the last decades, process engineers have been working towards the direction to variability, in order to tackle, control, and model all the sources of variations that are responsible for device failure and yield loss. Variability [68], single-event upsets (soft errors), and device degradation are some of the most important challenges that process engineers and designers must deal with, in order to prevent the sudden death of Moore's law [69].

There is an increasing variety of effects that leads to an increase of variability with respect to the continuation of scaling. [70] gives a summary of challenges in terms of variability for the very advanced 45nm technology. The most important sources of variations in the circuit's performances include: (a) random dopant fluctuations [71], (b) line-edge roughness [72], (c) variations of oxide thickness [73], (d) non-uniform threshold voltage by fixed charge [74], (e) defects and traps [75], (f) patterning proximity effects [76], (g) polishing [77], (h) strain-induced variation [78], (i) variations in implant and annealing process [79], (j) variation of temperature in operation, (k) aging and wear-out, (l) signal coupling and cross-talk, and (m) supply voltage and package noise.

4.3 PROCESS VARIATIONS CLASSIFICATION

Process variations can be classified into different categories [80], as shown in Figure 30. Systematic variations are variations due to factors causing the parameter of interest to be shifted away from the expected value in a given direction and come from errors in lithography or pattern proximity. These variations can be determined before manufacturing. Once the physical synthesis is finished, systematic variations can be measured and modeled with fixed values. A typical example of systematic variations is the randomness of interconnect metal thickness. After layout and routing, the patterns of interconnects can be accurately analyzed. Therefore, the layout-related metal thickness variations in different areas can be predicted. With this information, the resistance and capacitance of interconnects can be modeled more accurately in sign-off analysis [81]. In case of MOSFET devices, gate length is affected by variations in lithography for mask optimization. These variations can be determined by computing the post-Optical Proximity Correction (OPC) gate lengths on the critical path to achieve more accurate timing analysis [82]. In both cases, systematic variations are represented using fixed values instead of statistical variables. This is more accurate than simply analyzing circuit performance assuming random variations in metal thickness and from lithography. Both effects, although, can be incorporated only after physical synthesis. During the first iteration of logic synthesis, the circuit can only be optimized corresponding to the performance by modeling systematic variations as random variables.

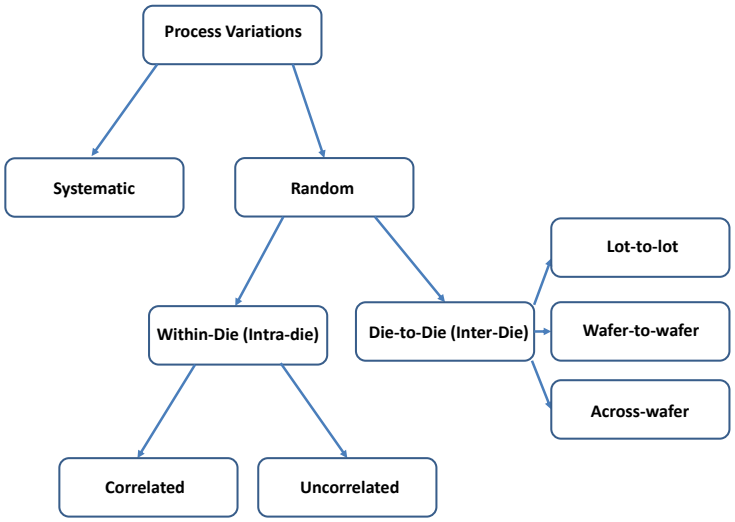


Figure 30: Process variations classification.

Random variations are due to factors that affect a parameter in a random manner and unlike systematic variations, they cannot be determined before manufacturing. These variations result from the inaccuracy of process control and are independent of the circuit design. They are always associated with a probability distribution that may either be explicit, in terms of a large number of samples provided from fabrication line measurements, or implicit, in terms of a known probability density function (such as a Gaussian or a log-normal distribution) that has been fitted to the measurements. Examples of random variations are variations in doping density and in layout-independent metal thickness of interconnects.

According to their spatial characteristics (their physical range on a die or wafer), random variations are further classified into D2D (or inter-die) and WID (or intra-die) variations. D2D variations refer to smooth and slow-varying variations that affect all devices on a die in the same way. For example, they cause the gate lengths of identical MOS devices to be larger or smaller than the nominal by the same amount. They include lot-to-lot variations, wafer-to-wafer variations, and across-wafer variations [83], as depicted in Figure 31(a).

WID variations refer to variations that rapidly vary over distances smaller than the die dimensions, as depicted in Figure 31(b) . Thus, WID variations may affect differently identical devices that are placed on the same die causing, for example, some devices to have larger gate lengths and others smaller gate lengths than the nominal.

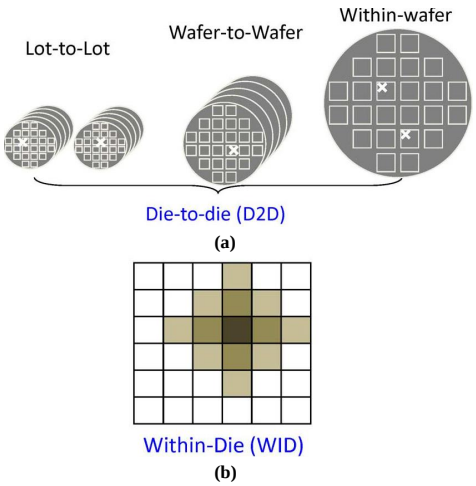


Figure 31: (a)Die-to-Die (D2D) variations and (b) Within-Die (WID) variations

Another criterion which is used for classifying process variations is based on environmental (e.g. environment temperature during circuit's operation) and intrinsic factors (e.g. variations in metal resistivity, metal thickness, etc.) that cause these process deviations, whereas, as shown in Table 3, process variations are often classified into spatial and temporal. Spatial variations concern deviations in space (e.g. variation of line resistivity across a wafer), whereas, temporal variations refer to variations in time (e.g. degradation of conductivity due to electromigration conditions) [81].

Table 3: Comparison between spatial and temporal variations.

Spatial variations	Temporal variations
Variations in space	Variations in time
e.g., variation of line resistivity across a wafer	e.g., degradation of conductivity due to electromigration

Re-visiting D2D and WID variations, and based on the classification of Table 3, it should be noted that, on one hand, D2D variations show a large degree of spatial correlation, which means that neighboring devices on the same die are affected in the same way. On the other hand, for some process parameters, such as effective channel length, WID show a large degree of local spatial correlation, while for some other process parameters, such as oxide thickness and dopant concentration, WID variations are uncorrelated.

Apart from the correlated variation component, WID variations still exhibit a purely random effect. The purely random variations come from the random fluctuation during manufacturing process, which thus imposes its effect on each device without correlation. Because of the inaccuracy of manufacturing equipments and process control, purely random variations exist in nearly every processing step. Examples are the random distortion of the lens used during the photolithography step and the purely random variation of the doping.

In design terms, the uncorrelated WID variations are typically referred as mismatch. Minimizing mismatch is a challenging issue in analog/RF design where it is often needed to have well-matched active or passive components, for example, the transistors in current mirrors, the transistors in differential stages, the sampling and integrating capacitors in a pipeline ADC, etc. To minimize the mismatch effects, there are specific rules during the design steps at both schematic and layout levels that must be followed. For example, at schematic level, transistors in current mirrors are biased in strong inversion to provide better current matching while transistors in differential stages operate in weak inversion to achieve better threshold voltage matching. At layout level, several recommendations for achieving good matching between two components exist, including using the same orientation, placing dummy components to create identical environments, minimize proximity, use non-minimum geometry, layout in common centroid, etc [84].

4.4 OVERVIEW OF PROCESS VARIATIONS

Based on the classification of process variations that was shown before, we can give now a brief overview of some of particular variations that should be taken into account during characterizing and measuring devices and interconnects in integrated circuit design [85], [86].

4.4.1 *Device Geometry Variations*

First, process variations related to the physical geometric structure of MOS transistors and other devices, including resistors, capacitors, and inductors are presented. They include:

- *Variations in thickness*: The gate oxide thickness is a critical process parameter. Variations in this parameter occur from one wafer to another with good across-wafer and across-die control. Also, other intermediate process thickness variations such as poly or spacer thickness can impact channel length, but are rarely directly modeled.
- *Variations in lateral dimensions*: Variations in lateral dimensions (length, width) come from photolithography proximity effects, lens imperfection, mask misalignments, and plasma etch dependencies.

Is is well known that MOS transistors are highly sensitive to variations in effective channel length (and therefore to poly gate length), threshold voltage, and gate oxide thickness and to some degree to variations in channel width. Among these parameters, channel length and threshold voltage variations require major attention, due to their direct impact on the output current characteristics.

4.4.2 *Device Material Parameter Variations*

Another set of process variations in MOS devices is related to internal material parameters.

- *Variations in doping*: junction depth and dopant profiles can be affected by implant dose, energy, or angle variations, which may also impact effective channel length as well as other electrical parameters such as threshold voltage. In deep sub-micron devices, drain engineering (e.g., halo implants to reduce short-channel effects) further increases the impact of implant and diffusion variations. Variations in thermal anneal and gate doping also change the degree of gate depletion in an active device and cause variations in the gate oxide thickness. Depending on the gate technology, these variations can lead to some loss in the matching of NMOS versus PMOS devices even in the case where within-wafer and within-die variations are very small.

4.4.3 *Device Electrical Parameter Variations*

Apart from variations in device geometry and material parameters, it is helpful to focus on variations that occur in several key electrical parameter parameters, including:

- *Threshold voltage variations*: One of the most important process parameters for characterizing a MOS transistor is the threshold voltage (V_T). In addition to various geometric sources, mobile charge in the gate oxide can introduce a bias dependent variation, which sometimes is approximated to be about 10% of the V_T of the smallest device in a given technology [86].
- *Variations in number of dopant atoms*: Another source of V_T variations that tends to become more and more important as technology scales down is related to random dopant fluctuations, which result from discreteness of dopant atoms in the channel of a transistor [87]. Transistor's channels are doped with dopant atoms to control their threshold voltage. The decrease in transistor size of each technology generation

reduces its area by half, and thus the number of dopant atoms in the channel decreases exponentially over generations [88]. Notice that in the 1-micron technology generation, there were thousands of dopant atoms, whereas in a 32nm to 16nm generation there are only tens of dopant atoms left in the channel and the law of large numbers no longer applies. Therefore, two transistors sitting side-by-side will have different electrical characteristics because of randomness in a few dopant atoms, resulting in variability.

- *Leakage currents*: Sub-threshold leakage currents may vary considerably, and can be impacted by Shallow Trench Isolation (STI) structure and stress imperfections due to oxidation and Chemical-Mechanical Planarization (CMP) step.

4.4.4 *Interconnect Geometry Variations*

Like in cases of devices, vertical and lateral dimensions as well as material property deviations can also be important sources of variations in interconnect structures. These variations include:

- *Line width and line space variations*: Variations in the width of patterned lines also stem from photolithography and etch dependencies. At the shortest dimensions (lower metal levels), proximity and lithographic effects may be important, whereas at other levels etch effects, which depend on line width and layout configurations, can be significant. Deviations in line width may impact line resistance as well as the capacitance from one layer to another. Deviations in line width can also result in line space differences affecting the magnitude of line-to-line coupling within the layer.

- *Metal thickness variations*: In metal interconnects, the thickness of sputtered or deposited metal films and barriers is often a well-controlled parameter, but can vary from wafer-to-wafer and across-wafer. On the other hand, in processes, where copper is used as an additive pattern (e.g. damscene), dishing and erosion steps can impact the final thickness of patterned lines considerably [89], with line thickness losses of 10% to 20% depending on the particular patterns.

- *Dielectric thickness variations*: The thickness of deposited and polished oxide films can also suffer substantial deviations. While wafer-level deposition thickness can vary, more troublesome are pattern-dependent aspects of such deposition. For instance, the deposition profile using High Density Plasmas (HDP) can depend strongly on the width or size of the feature being deposited over. Furthermore, the CMP process can introduce strong across-die variations, which arise from the effective density of raised topography in different regions across the chip [90].

- *Contact and via size variations*: Contact and via size are usually affected by etch process variations as well as systematic layer thickness dependencies. For example, the etch depth can be significantly different, depending on the location of the via or contact. Consequently, different degrees of lateral opening are observed. Such size differences can directly impact the resistance of contacts and vias.

4.4.5 *Interconnect Material Parameter Variations*

Finally, some other material parameter variations might also be important in interconnects, including:

- *Contact and via resistance variations*: Contact and via resistance related to good ohmic contact can be sensitive to etch and clean processes, with substantial wafer-to-wafer and random components.

- *Metal resistivity and dielectric constant variations*: Metal resistivity usually varies significantly from one wafer to another and is usually a well-controlled parameter. The same stands for dielectric constant, which may vary depending on the deposition process and is often a well-controlled parameter as well. It is possible that pattern-dependent and directional effects can be important for some low-K dielectrics considering also future interconnect technologies.

4.5 VARIABILITY IN 65NM CMOS BULK TECHNOLOGY - AN OVERVIEW

In this section, we present the most important process parameters that are responsible for device (active and passive) and interconnect variations in 65nm CMOS bulk technology provided by STMicroelectronics. Our investigation focuses on variability in MOS transistors, spiral inductors, poly resistors, and MOM capacitors.

4.5.1 *Variations in MOS Transistors*

Figure 32 depicts the correlation plots between the drain current of a short-channel MOS device and some process and design parameters, extracted based on a Monte Carlo simulation campaign with 1000 samples where both D2D and WID variations were considered. The list of the most important process and design parameters includes the threshold voltage, effective channel length, channel width, carriers mobility and oxide thickness. The device under study is a low-voltage/low-power NMOS transistor (this type of transistors has been used in both case studies of this work) with an aspect ratio of W/L = $10\mu m/0.06\mu m$ and 10 fingers. The device has been biased in strong inversion and saturation region of operation. A short-channel device has been simulated, since in RF design, MOS transistors with a short channel are usually chosen, so that a higher f_T is achieved.

As illustrated in Figure 32, the variability of threshold voltage and channel length explains to a large degree the variability of the drain current showing a correlation of 98.9% and 82.4%, respectively. On the other hand, it seems that channel width, oxide thickness and carriers mobility do not play any significant role in the variability of the drain current. Furthermore, we can see the strong dependence of threshold voltage on the channel length, which is due to the major impact of short channel effects on this particular parameter. For example, for long-channel MOSFETs, gate has control over the channel and supports most of the charge. As channel length becomes shorter, the device threshold voltage begins to decrease as the charge in the depletion region is now supported by the drain and the source terminals. Therefore, the gate needs to support less charge in this region and consequently the threshold voltage decreases. This phenomenon is known as the Charge Sharing effect.

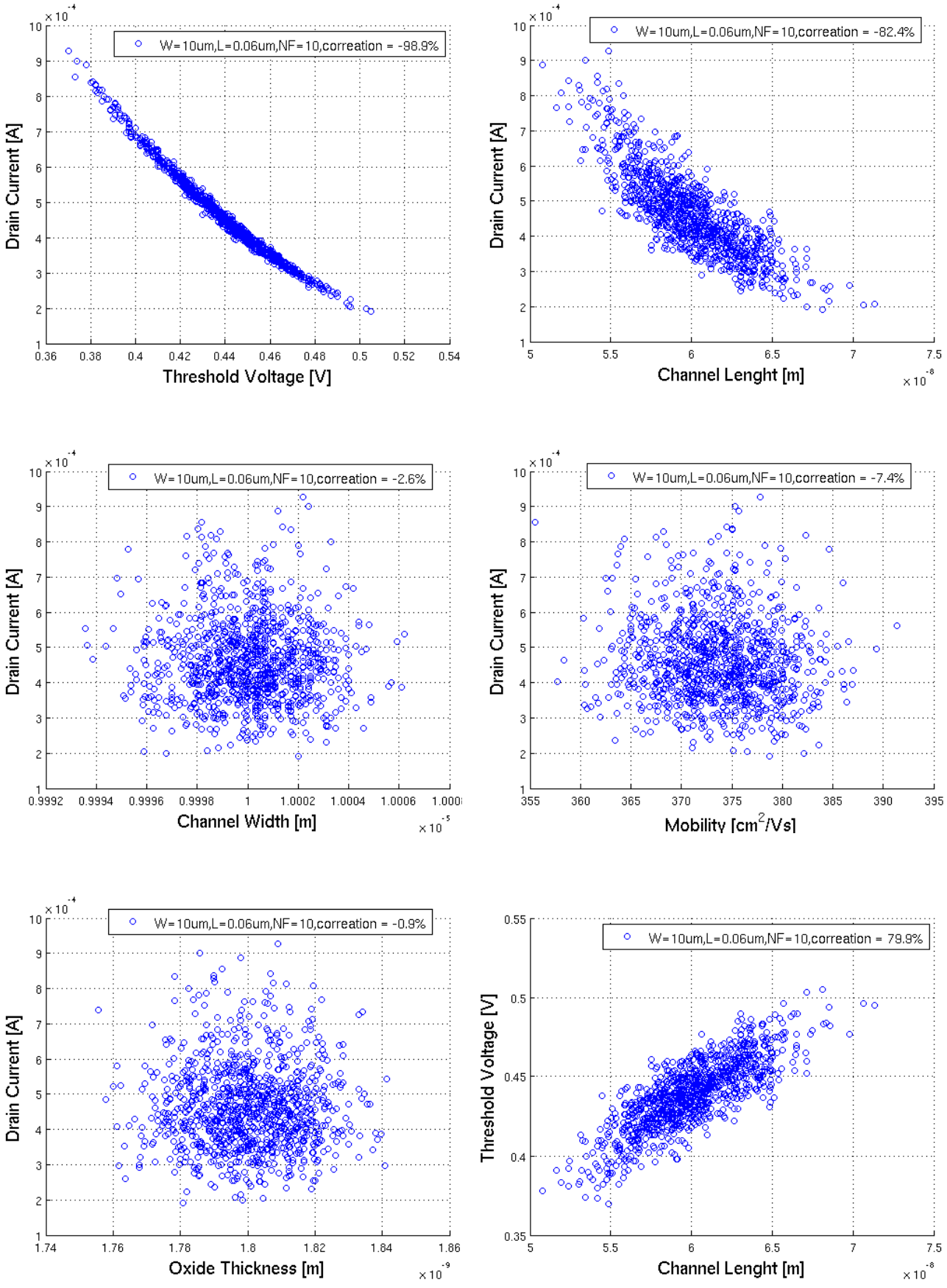


Figure 32: Correlation plots for a short-channel NMOS transistor.

4.5.2 *Variations in Spiral Inductors*

Similar observations are made in case of a spiral inductor with a patterned ground shield. The equivalent model of a spiral inductor constructed with a patterned ground shield is shown in Figure 33. The spiral inductors, which are used in this work, are composed of pure coil inductances L_{s1} and L_{s2}, series resistances of the coil R_{s1} and R_{s2}, parasitic and oxide capacitances C_{OXm}, C_{OX1} and C_{OX2}, and resistances R_{pattm}, R_{patt1} and R_{patt2} that model the patterned ground shield. In general, inductors are realized by stacking several metal levels available in the 65nm technology and an alucap layer to further reduce the coil resistance. Alucap consists of aluminium structures that are placed at the top metal layer and are electrically floating. They are used, so as to a) avoid fractures-scrapes on the oxide during passivation deposition process and b) keep constant the metal vs. dielectric density all across the wafer. The series resistances in the model correspond to these metal and alucap layers. In order to investigate the variability effects in these particular passive components, a set of 1000 Monte Carlo statistical samples was generated, including D2D and WID variations. The list of the most important process parameters includes metal and alucap resistivity, $\Delta\rho$, low-K dielectric thickness, and metal barrier thickness. Figure 34 depicts the correlation plots between the series resistance of the coils that form the inductor and metal 5 low-K dielectric thickness, alucap resistivity, metal 1, 5 and 7 resistivity, and metal 7 barrier thickness. As it can be seen, the variation in the series resistance of the inductor is mostly correlated to the resistivity of metal layers 5, 7, and alucap showing a correlation of 68%, 61% and 36.1%, respectively, whereas is uncorrelated to the low-K and barrier dielectric thickness.

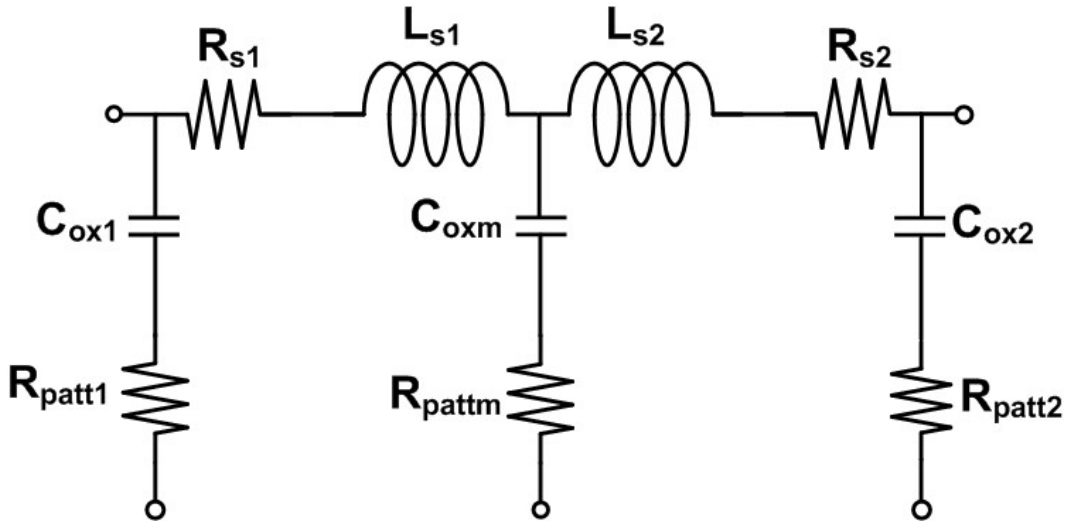


Figure 33: Lumped 2π- model of spiral inductor with patterned ground shield.

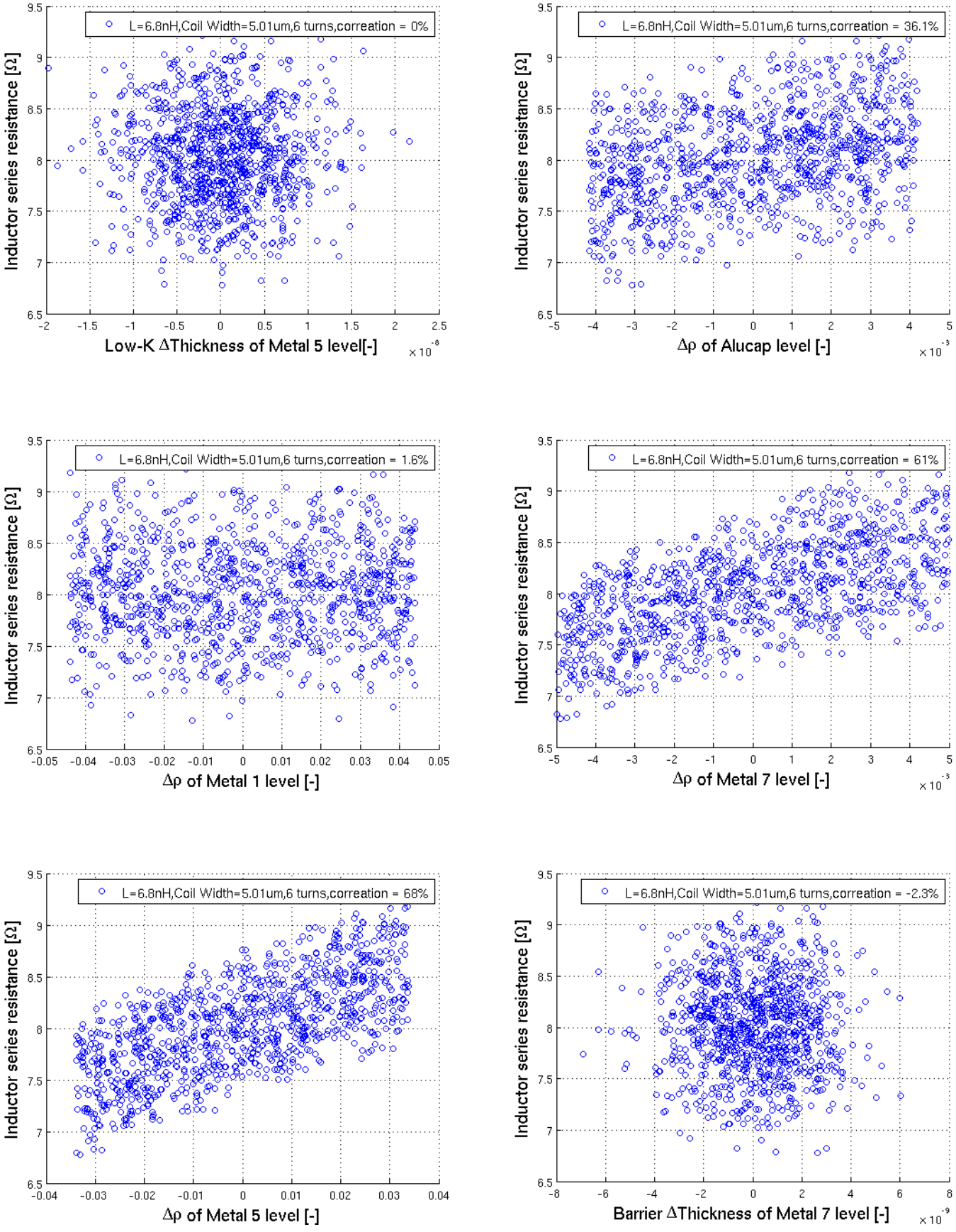


Figure 34: Correlation plots for a spiral inductor with patterned ground shield.

4.5.3 *Variations in Resistors*

In both case studies of this work, $N+$ poly unsilicided resistors have been used. In this section, we present the sources of variability in this particular component. According to the results from the characterization flow, variations in this kind of resistors are due to $N+$ poly silicided sheet resistance, $N+$ poly contact resistance, $N+$ poly unsilicided sheet resistance, poly/bulk specific capacitances, and poly/bulk fringe capacitances. Figure 35 shows the correlation plots between the resistor value and the aforementioned process parameters that define the electrical behavior of the selected resistor. A set of 1000 Monte Carlo statistical samples was generated, including D2D and WID variations. The analysis was carried out on a 722Ω $N+$ poly resistor with length of $9.3\mu m$ and width of $0.8\mu m$.

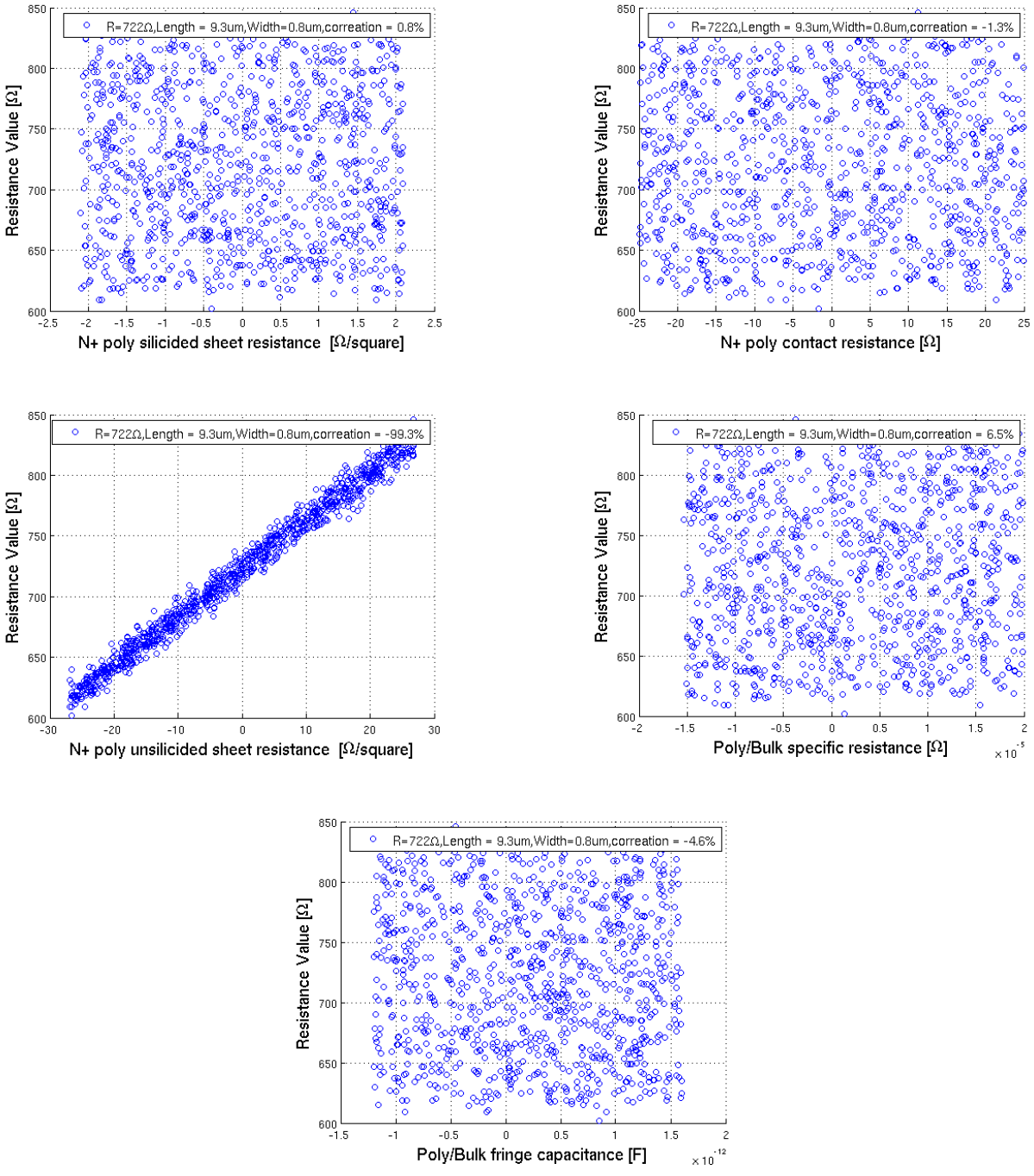


Figure 35: Correlation plots for a $N+$ poly resistor.

As it can be seen, variations in the $N+$ poly resistors mostly correlate to variations in unsilicided sheet resistance, showing a correlation of 99.3%, and are uncorrelated to variations in contact and poly/bulk resistances.

4.5.4 *Variations in Capacitors*

RF MOM capacitors have been used in our designs. Typically, these devices are built by stacking the metal layers provided by the process technology. Variations in these devices are due to several process parameters, including metal 1 and metal X thickness, Inter-Metal Dielectric (IMD) metal 1 and metal X thickness, and metal X capacitance density, where metal X corresponds to metal layers 2, 3, 4, 5, and 6. IMD refers to the dielectric of the material that is used for insulating the layers of electrical conductors in an integrated circuit. In order to investigate the variations sources of these devices, a set of 1000 Monte Carlo statistical samples was generated, including D2D and WID variations. The analysis has been carried out using a 4.3pF MOM capacitor that has 229 fingers in the X direction and 200 fingers in the Y direction. Figure 36 depicts the correlation plots between the capacitance value and the aforementioned process parameters.

Looking at Figure 36, it can be seen that variations in MOM capacitors are correlated to variations in metal X capacitance density and thickness, showing correlations of 84.7% and 57.9%, respectively and are uncorrelated to the other process parameters.

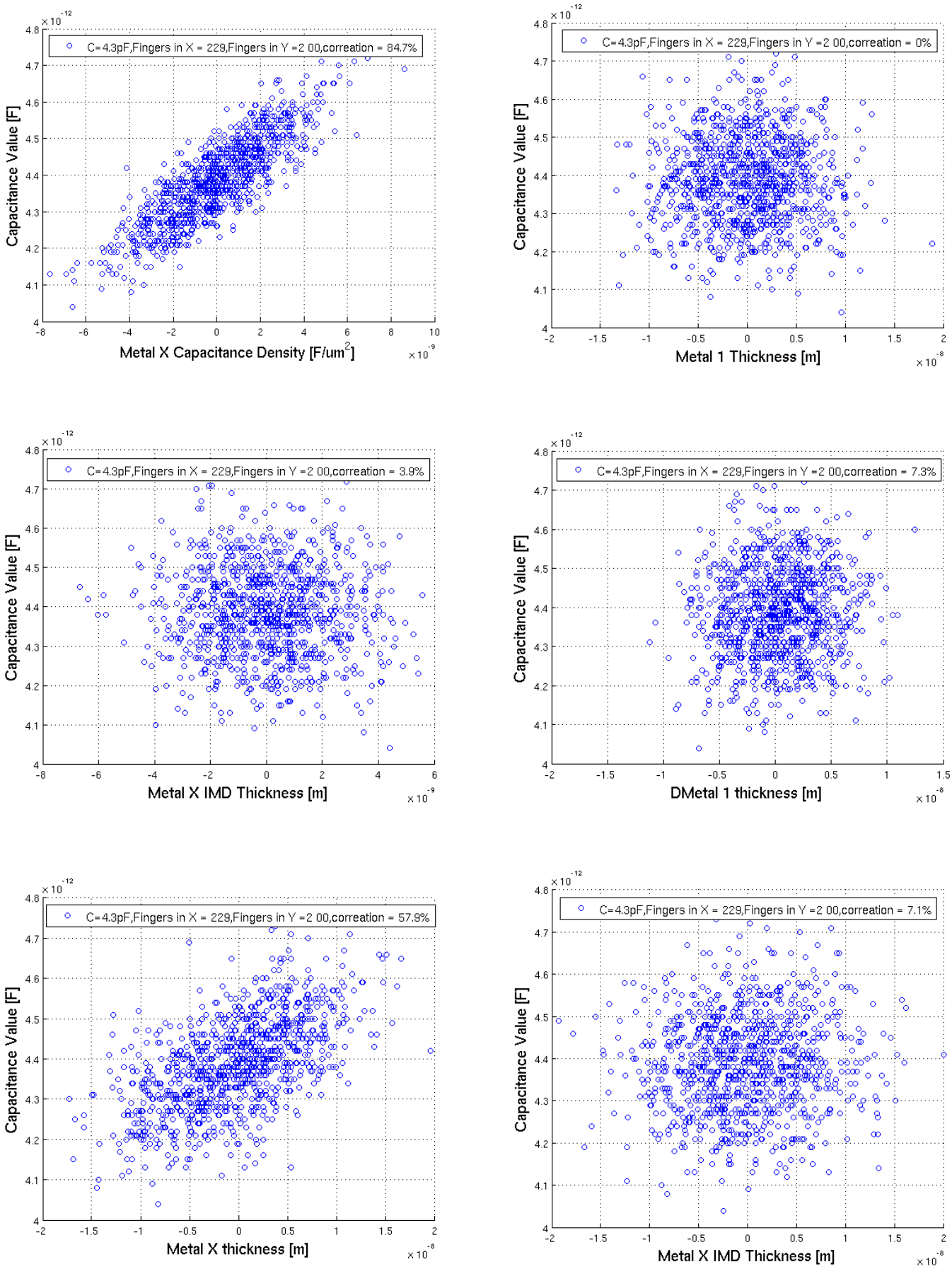


Figure 36: Correlation plots for a MOM capacitor.

4.6 CONCLUSIONS

BIT using non-intrusive variation-aware sensors is a test technique which capitalizes on the undesired phenomenon of process variations and it is used to minimize the RF test cost. In this chapter, we provided a brief description of process variations and their classification into different categories according to several criteria. We, also investigated the sources of variability in four different components,namely low-voltage/low-power MOS transistors, spiral inductors, poly resistors, and MOM capacitors. These components have been used, in order to implement the case studies of this work. After understanding the term and definition of process variations, next chapter describes the principle of operation of the variation-aware sensors, which will help us to facilitate our discussions regarding the results of the proposed test technique.

5

NON-INTRUSIVE VARIATION-AWARE SENSORS

In this chapter, we discuss the principle of operation of non-intrusive variation-aware sensors, the overall benefits of the proposed test approach, and the procedure to select these sensors for our case studies. In particular, Section 5.1 gives an overview of the proposed BIT strategy. Section 5.2 describes the alternate test paradigm. Section 5.3 demonstrates the benefits and disadvantages of our test technique. Section 5.4 points to the challenges of the proposed test technique providing also some guidelines. Finally, in Section 5.5, the selection of non-intrusive variation-aware built-in sensors for our two case studies will be presented in details. Section 5.6 concludes the chapter.

The basic characteristic of non-intrusive variation-aware sensors is that they capitalize on the undesired phenomenon of process variations. The underlying idea is to monitor process variations instead of measuring directly the RF performances of the CUT. These sensors are placed in close physical proximity and are matched to identical structures in the CUT. In this way, the sensors and the CUT experience the same D2D and correlated WID process variations and, as a result, the measurements obtained on the sensors will be correlated to the performances of the CUT to a very large extent. An indirect, low-cost test can be put in place by employing the alternate test paradigm to map the sensors' measurements to the performances. This test approach has been inspired by the PCMs typically placed in the scribe lines of a wafer to monitor variability and identify off-target process parameters [36], [37]. Essentially, non-intrusive variation-aware sensors include two types of circuits, namely dummy stages and PCMs. PCMs refer to basic layout components (e.g. transistors, capacitors, resistors, inductors), while dummy circuits are simple analog stages that mimic part of the CUT architecture (e.g. bias stages, current mirrors, gain stages, level-shifters, etc.).

We refer to the dummy structures (dummy circuits and PCMs) as *non-intrusive variation-aware sensors* or *non-intrusive sensors* or simply *non-intrusive embedded sensors*.

5.1 NON-INTRUSIVE BUILT-IN TEST (BIT): AN OVERVIEW

Considering any analog circuit, the underlying idea is to decompose it into "primitive" analog stages and critical single components and/or assemble basic analog stages from identical components that can be found in the circuit's topology. Analog stages may include bias stages, gain stages, current mirrors, level-shifters, etc., whereas single components may include transistors, resistors, capacitors, inductors, transmission lines, etc. Such analog stages and single components are then placed as dummy structures on the same die together with the CUT. In particular, they are placed in close physical proximity to the corresponding analog stages and components of the CUT that they are mimicing, such that corresponding analog stages and components are well matched following specific layout rules.

The underlying principle of operation is that thanks to this physical proximity, the non-intrusive sensors will "witness" the same D2D and correlated WID process variations as the corresponding structures in the CUT. As a result, we expect that measurements on the non-intrusive sensors (e.g. bias voltages, DC gain, transistor transconductance, resistance, capacitance, etc.) will offer an "image" of process variations within the CUT and, thereby, we expect that they will correlate to a large degree with the performances of the CUT. If such a correlation exists, then we expect that we will be able to predict the performances indirectly solely from the measurements on the non-intrusive sensors based on the alternate test paradigm [91].

5.2 ALTERNATE TEST

Alternate test refers to a generic test approach that consists of inferring the performances of the CUT through unconventional, low-cost alternative measurements (e.g. the measurements provided by the non-intrusive sensors in our case). Essentially, it is used to predict the outcome of a standard test approach based solely on a measurement pattern [35], [91], [92] as shown in Figure 37. As it can be seen, the alternate test flow consists of two phases, namely the training and the test phase. In the training phase, we collect N circuit instances from different lots, wafers, and sites on the same wafer, such that they are representative of the process variations that may occur. Formally, let P_j denote the j-th performance of the CUT and let X denote the sensors' measurements. The pairs $\{X, P_j\}$ are split into training and test sets. The variation of the performance P_j, denoted by ΔP_j, and the variation of sensors' measurements, denoted by ΔX, are linked to the D2D and correlated WID variations, denoted by Δp, through some non-linear functions f_{1j} and f_2:

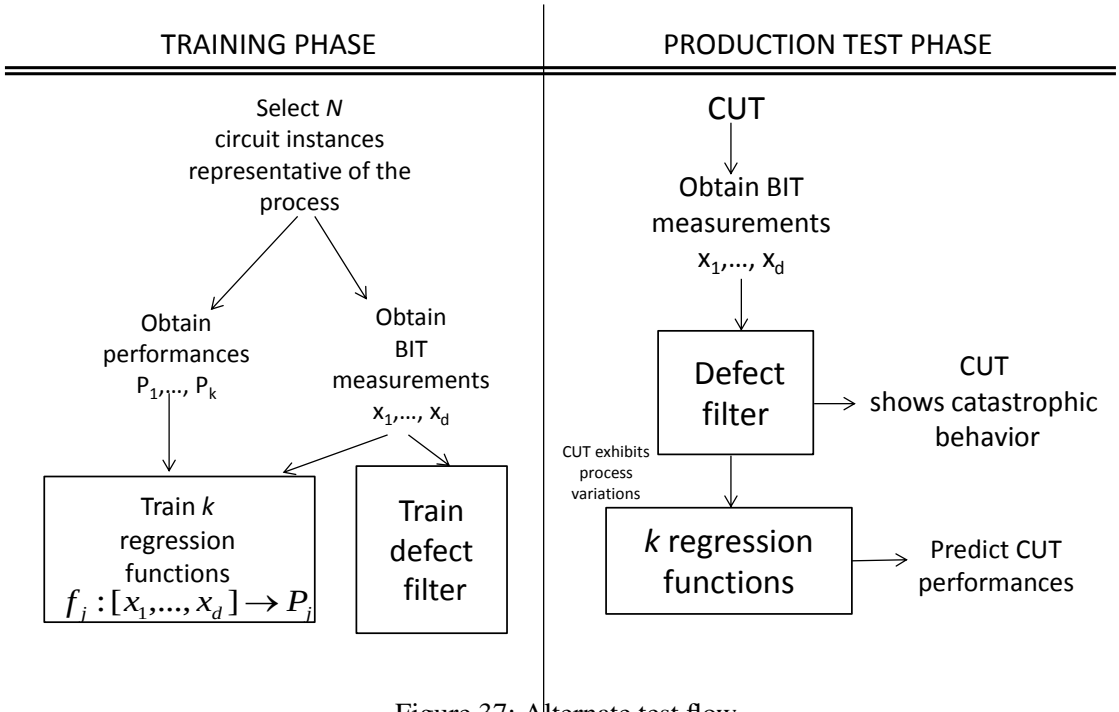


Figure 37: Alternate test flow.

$$\begin{aligned} \Delta P_j &= f_{1j}(\Delta p) + r_1 \\ \Delta X &= f_2(\Delta p) + r_2, \end{aligned} \tag{62}$$

where the parameters r_1 and r_2 represent the uncorrelated WID variations. From Equation (62), we can write

$$\Delta P_j = f_{1j}\left(f_2^{-1}(\Delta X - r_2)\right) + r_1. \tag{63}$$

Therefore, by tracking variations on the sensors' measurements, we obtain information about the variations on the performances.

To this end, we can use the alternate test paradigm to infer implicitly the performances from the sensors' measurements [22], [31], [32], [33]. Specifically, in an off-line preparatory training phase where we employ a large set of circuit instances with process variations, we learn one regression function per performance that maps the sensors' measurements to the performances:

$$g_j : X \rightarrow P_j. \tag{64}$$

The training set is used to learn the regression function g_j while the test set is used as an independent set to assess the error in predicting the performance P_j from X using g_j . If the set of measurements X yields a low prediction error, the learned regression functions can be used to provide a fast, low-cost test for new instances coming out of the production line.

However, it should be noted that the regression functions should be applied only to instances that exhibit process variations and, in particular, to instances whose measurement pattern is within the range defined by the measurement patterns of the instances in the training set. Otherwise, the prediction will be somewhat random. Thus, a defect filter is used to distinguish instances that exhibit process variations from instances that contain defects. In [93], a defect filter based on non-parametric density estimation is proposed. It has the attractive property that it operates as an one-class classifier, that is, it is trained using the alternate measurement patterns of the instances in the training set and no information about defects is required. The instances with process variations constitute the majority of instances that will be seen in production, hence the large benefit of alternate test.

Once the training phase is completed, the regression functions can be readily used in the test phase to predict the performances of any CUT simply from its sensors' measurements pattern. In short, as also stated in the introduction of the chapter, the non-intrusive BIT strategy capitalizes on the undesired phenomenon of process variations to provide low-cost alternative measurements from which performances can be predicted implicitly with accuracy. For various practical ideas towards enhancing the efficiency of alternate test, the interested reader is referred to [33], [34], [91], [94], [95].

The time-consuming part of the training phase is the collection of datasets from different wafers and lots such that they are representative of the fabrication process. Once the datasets collection is completed, learning the regression functions is a step that can be performed in seconds using modern non-linear regression tools, such as feed-forward neural networks, multivariate adaptive regression splines, support vector machines, etc. [96], [97]. In the interest of expediting the training phase and avoiding having to collect large datasets, we can rely on the BMF technique that achieves efficient learning of the regression functions across the feasible design space by combining intelligently readily available large simulation datasets with small silicon datasets [34]. Another possibility is to follow an adaptive approach, where the regression functions are learned based on limited silicon datasets, but for a given CUT, a decision is made whether there is confidence in using these regression functions [33]. If not, then the standard test procedure is applied to this CUT. As we keep collecting datasets in production, the decision block is updated periodically, and after a certain point in time, all CUT with process variations are forwarded irrespectively to alternate test. The drawback of this approach is that for some time a two-tier test flow needs to be put in place, which may complicate considerably the high-volume production test flow.

The accuracy of the drawn correlation and, thereby, the predictions of alternate test, is negatively affected by three factors. First, the sensors' measurements might not capture all the sources of performance variations. Second, the uncorrelated WID variations can introduce noise in the correlation and the existing trend may be obscured or even eclipse. Third, inaccuracies of sensors' measurements introduce noise in the drawn correlations during the training phase and result in offsets in the predictions during the test phase. The effect of sensors' measurements inaccuracies is rather insignificant since the sensors' measurements are either DC or low-frequency and an excellent repeatability can be achieved through averaging. Furthermore, during the training phase, the regression, by default, averages the noise in the data.

5.2.1 *Artificial Neural Networks and Regression Analysis*

As mentioned before, the training set of the alternate test is used to learn the regression function g_j and build the mappings between the sensors' measurements and the performances of the CUT. Regression analysis is a statistical technique, which is used for investigating and analyzing the relationship between dependent variables (responses) and one or more independent variables (regressors or predictors). Essentially, regression analysis describes how the typical values of the dependent variables change when one or more independent variables are

varied, while the other independent variables are kept fixed. The performance of regression analysis methods depends on the way that the data are generated, and how it relates to the regression approach being used. Since the true form of the process that produces the data under study, is not known a priori, regression analysis usually depends to some extent on making assumptions about this process. These assumptions are sometimes testable if a sufficient size of data is available. Therefore, regression models are built, where several assumptions are made. One way to implement a regression model is to use artificial neural networks by employing feed-forward and back-propagation techniques. Before describing these techniques, we give a brief overview of artificial neural networks.

Artificial neural networks are structures that are designed to solve specific types of problems by trying to emulate the way the human brain would solve the problem. The general form of a neural network is a "black box" structure that is often used to model and analyze high-dimensional, non-linear data. Typically, most of the neural networks are used to solve prediction problems for complex systems [98].

Moreover, in most of the case studies, multilayer feed-forward artificial neural networks are employed. Multilayer feed-forward artificial neural networks are multivariate statistical models, which are used to relate p predictor variables $x_0, x_2, ..., x_{p-1}$ to n response variables $y_1, ..., y_n$. In our case, the predictors are the alternative measurements captured from the sensors and the responses are the performances of the CUT. The model has several layers, each of one including either the original or some constructed variables. The most common structure consists of three layers: (a) the inputs, which are the original predictors, (b) the hidden layer, which is a set of constructed variables, and (c) the output layer, which provides the responses. Each variable in a layer is called a node. Figure 38 depicts a typical neural network with three layers.

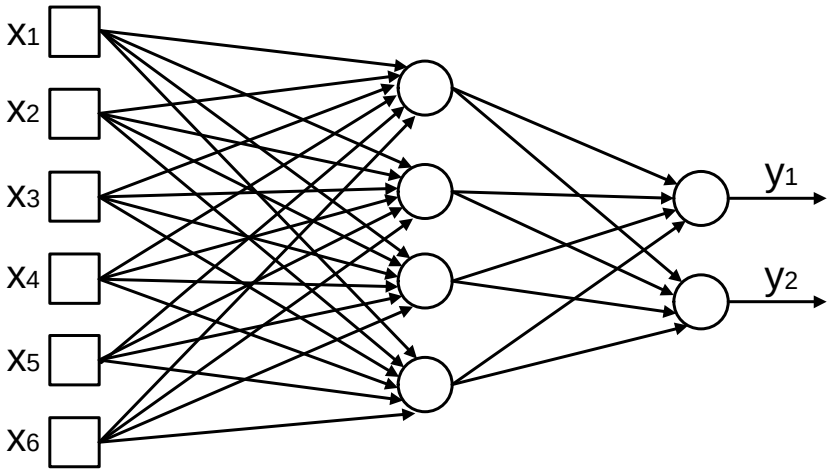


Figure 38: Artificial neural network with one hidden layer.

A node takes as input a transformed linear combination of the outputs from the nodes in the layer below it. Then, it sends as an output a transformation of itself that becomes one of the inputs to one or more nodes in the next layer. The transformation functions are usually sigmoid or linear and are called activation functions or transfer functions. Let each of the k hidden layer nodes α_u be a linear combination of the input variables,

$$\alpha_u = \sum_{j=1}^{p} w_{1ju} x_j + \theta_u$$

where w_{1ju} are called weights that are unknown parameters that need to be estimated (essentially, they denote the layer of the neural network). θ_u is called bias node and is a parameter that plays the role of an intercept in linear regression.

Each node is transformed by an activation function $h(\cdot)$. Let the output of node α_u be denoted as $Z_u = h(\alpha_u)$. Now, we form a linear combination of these outputs, $b_n = \sum_{u=0}^{k} w_{2un} z_u$, where $z_0 = 1$. Finally, the n-th response

y is a transformation of the b, $y_n = \overline{h}(b_n)$, where $\overline{h}(\cdot)$ is the activation function for the response. All the above can be combined to give:

$$y_n = \overline{h}\left[\sum_{u=1}^{k} w_{2un} h\left(\sum_{j=1}^{p} w_{1ju} x_j + \theta_u\right) + \theta_{2u}\right] \tag{65}$$

The response y_n is a transformed linear combination of the original predictors. Finally, it should be noted that the activation function is often chosen to be either the logistic function or the hyperbolic tangent function. The choice of the activation function for the output layer depends on the nature of the response.

In our case, the regression models are build implementing a neural network, which employs both feed-forward and back-propagation techniques. Both of them are described in the next sessions.

5.2.2 *Training Phase: Building Regression Models*

The training of the neural network is achieved by using the early stopping technique, which allows to control the complexity of the regression function. In particular, the back-propagation algorithm first uses training data to adjust the weights and at the end of each iteration the prediction error is calculated on the validation data. If the error does not reduce to a number of consecutive iterations, this implies that the network is at the limit of "over-fitting" and therefore, the training stops. Finally, the model that gives the smallest prediction error on the validation data is used in the testing phase. The disadvantage of neural networks is that the optimal number of neurons in the intermediate layers is not known a priori and it takes an approach to "trial & error" to be fixed. Therefore, during the training phase, we repeat learning and validation, several times, for each model and we choose the model that has the smallest error. Another disadvantage is that this algorithm does not include any enabling method generally used when the statistical sample is limited and non-representative. At simulation-level, we used a sample of 800 instances in the training phase, and hence, we assume that this sample is statistically representative. In the case of a small number of samples, k-fold cross-validation can be used to obtain trustworthy prediction results.

In order to build our regression models, a neural network of four layers is used. The network has one layer at the input, one at the output, one hidden layer that employs a sigmoid activation function with a specific number of neurons, and one layer that performs a linear activation function. The network is trained using a Monte Carlo sample set. In fact, neural networks with one hidden layer can approximate any non-linear mapping. The sensors' measurements are presented to the input of the neural network and through the hidden layer are mapped to the performances of the CUT at the output (feed-forward technique). The network is trained with the **net** and **train** functions in Matlab.

5.2.3 *Test Phase: Performance Prediction*

Once the regression functions are built (trained neural network), another independent Monte Carlo sample set is used to evaluate the mappings of Equation 64 and estimate the error in predicting the performances of the CUT in the validation set.

The prediction error is expressed in terms of several error metrics, including: (a) the average RMS error across the validation in %, (b) the absolute average RMS error across the validation set in units of each performance, (c) the maximum error observed in the validation set in units of each performance, and (d) the correlation coefficient between the simulated (e.g. "true") and predicted performances (Pearson correlation coefficient is used in this work). The average RMS error and the absolute average RMS error demonstrate the overall accuracy of alternate test. The former is defined as

$$\epsilon_i = \frac{100}{P_{i,nom}} \sqrt{\frac{\sum_{j=1}^{N} (P_{i,j} - \hat{P}_{i,j})^2}{N}} \tag{66}$$

where $P_{i,nom}$ is the nominal value of the i-th performance, $P_{i,j}$ is the simulated (actual) i-th performance of the j-th instance, $\hat{P}_{i,j}$ is the predicted i-th performance of the j-th instance, and N is the total number of instances in the validation set.

The maximum error, which is larger than the absolute average RMS error, demonstrates the worst case scenario observed in the validation set and should be smaller or at least comparable to measurement and repeatability errors in a conventional ATE.

Finally, the correlation coefficient is a qualitative measure that shows the linear fit between the true and predicted performances. The closer the correlation coefficient to 100% is, the better the prediction results are. However, it should be stressed that a correlation coefficient considerably lower than 100% does not necessarily lead to unsatisfactory prediction results.

The **Postreg** Matlab functions post-process the trained network response and perform a linear regression between the network response and the target values, in order to compute their correlation. The mathematical form of the **postreg** function is given by

$$[m, b, r] = postreg(Y, T), \tag{67}$$

where Y refers to the network responses and T represents the target values. m, b, and r correspond to the slope, intercept, and correlation coefficient of the linear regression function, respectively.

It should be noted that, in our case study, random defects are not captured by the proposed BIT structures. The defect filter, shown in Figure 37, is not used and the learned regression functions are employed only for circuits instances that exhibit process variations.

5.3 BENEFITS AND DISADVANTAGES

The non-intrusive BIT strategy presents several advantages:

- The non-intrusive variation-aware sensors occupy a very small area on the die. They are in fact laid out and placed in areas that are left void so as to respect electromagnetic design rules.
- It is low-cost since the test stimuli applied to the non-intrusive sensors and the measurements provided by the non-intrusive sensors are DC or low-frequency. In some cases, the non-intrusive sensors are self-biased. Thus, high-cost ATE with advanced features is no longer needed and, moreover, the inputs and outputs of the non-intrusive sensors can be multiplexed in a test bus so as to use a maximum of two pins for test.
- It is particularly attractive for RF and mm-wave circuits since the BIT circuitry is totally transparent to the CUT and does not degrade the performances of the CUT. In fact, design and test are completely dissociated.
- The CUT does not need to be powered-on during testing, which largely facilitates parallel multi-site testing.
- It is generic since virtually it is applicable to any analog circuit.
- Finally, since the non-intrusive variation-aware sensors offer an “image" of process variations within the CUT, it has been shown that they can be used for one-shot calibration where test and calibration are performed together in one single step [99].

The disadvantages of the non-intrusive BIT strategy are the following:

- It can only detect failures due to excessive process variations, thus it needs to be complemented with another test strategy that is capable of detecting defects. Recently, it was shown that defects can also be detected non-intrusively based on temperature sensors [39].

- As stated before, the CUT does not need to be powered-on during testing which largely facilitates parallel multi-site testing. On the other hand, the negative aspect of the CUT being powered-off during test is that failures due to coupling effects and electromagnetic interference between internal blocks of the CUT or between the CUT and other neighboring blocks cannot be detected. However, such types of failures are typically circumvented by taking special precautions at the design stage.

- It cannot deal with uncorrelated WID variations (e.g. mismatch) that become important for advanced technology nodes from 65nm and below. Roughly speaking, the approach will fail if the uncorrelated WID variations start being comparable to D2D and correlated WID variations. It should be noticed, however, that the performances of RF and mm-wave circuits do not depend on matching unless we consider fully differential topologies. Furthermore, mismatch appears as “noise" in the training set used to learn the regression functions in alternate test and, by default, the regression operation averages out this noise.

5.4 CHALLENGES AND GUIDELINES

The main challenges of the non-intrusive BIT strategy are the following:

- Identify the main process parameters (i.e. sheet resistance, oxide thickness, dopant concentration, junction capacitances, etc.) whose variation explains the variation in the performances of the CUT.

- Design non-intrusive sensors that capture and track variations in these process parameters.

- In addition, the non-intrusive sensors should be tested at DC or low-frequency and should incur low area overhead, in order to achieve a low cost.

There are several guidelines that can be followed towards addressing these challenges:

- A straightforward approach that has shown to work very well is to decompose the CUT into analog stages and single components and use as non-intrusive sensors identical dummy analog stages and single components.

- Process Design Kits (PDKs) for advanced technology nodes have hundreds of process parameters and it is not an easy task to dig into and fully understand the intricate relationships between the process parameters and the performances of a circuit. For this reason, we can rely on the main design parameters (transistors’ geometry, transistors’ transconductance, resistors’ and capacitors’ values, bias currents and voltages, etc.), which, in turn, are intricate functions of the process parameters, and choose to design non-intrusive sensors that capture effectively these design parameters. The designer’s insights and input can be very valuable for performing this task.

- Rely on previous experience gained by applying the non-intrusive BIT strategy to other circuit classes and/or other circuit architectures within the same circuit class. For example, the non-intrusive test strategy has been applied to a 2.4GHz RF LNA [35], [100] and conclusions from this case study can be explored towards applying the concept to a 60GHz mm-wave LNA, which is the second case study of this work.

- Include an initial set of non-intrusive sensors that is the most straightforward from a conceptual point of view and perform an extensive simulation campaign to examine whether this set is largely sufficient. If not, the problem of selecting non-intrusive sensors can be revisited and studied more throughly.

- Include a large set of non-intrusive sensors and then perform an extensive simulation campaign to study redundancy and keep the optimal set.

5.5 NON-INTRUSIVE VARIATION-AWARE SENSORS DESIGN AND ALTERNATIVE MEASUREMENTS

5.5.1 *2.4GHz Inductive Degenerated LNA Case Study*

The non-intrusive variation-aware sensors include PCMs (e.g. transistors, capacitors, resistors, inductors) and dummy analog stages (e.g. bias stages, current mirrors, gain stages, level-shifters, etc.). In the case of the LNA of Figure 21, an example of a dummy analog stage that we can derive from its topology, as shown in Figure 39(a), is a current mirror composed of transistors M_1, M_2, and M_3, and resistor R_{REF} (this current mirror actually results from the LNA topology by short-circuiting inductors L_G, L_S, and L_D and removing the capacitors C_{IN} and C_{EXT}). As PCMs we use identical components that exist in the topology of the CUT, including MOM capacitors, spiral inductors with patterned ground shield, and diode-connected transistors, as shown in Figure 39(b). The dummy analog stages are also built from such identical components. The non-intrusive variation-aware sensors are placed in close physical proximity to the corresponding structures in the LNA that they are mimicing and are well-matched together. The underlying idea is that the non-intrusive variation-aware sensors and the LNA will be subject to the same D2D and correlated WID variations. Therefore, measurements on the non-intrusive variation-aware sensors (i.e. bias voltages, gains, capacitances, inductances, MOS transconductances, etc.) will be correlated to the performances of the LNA (i.e. S_{21}, S_{11}, S_{22}, NF, 1-dB compression point, IIP_3, etc.).

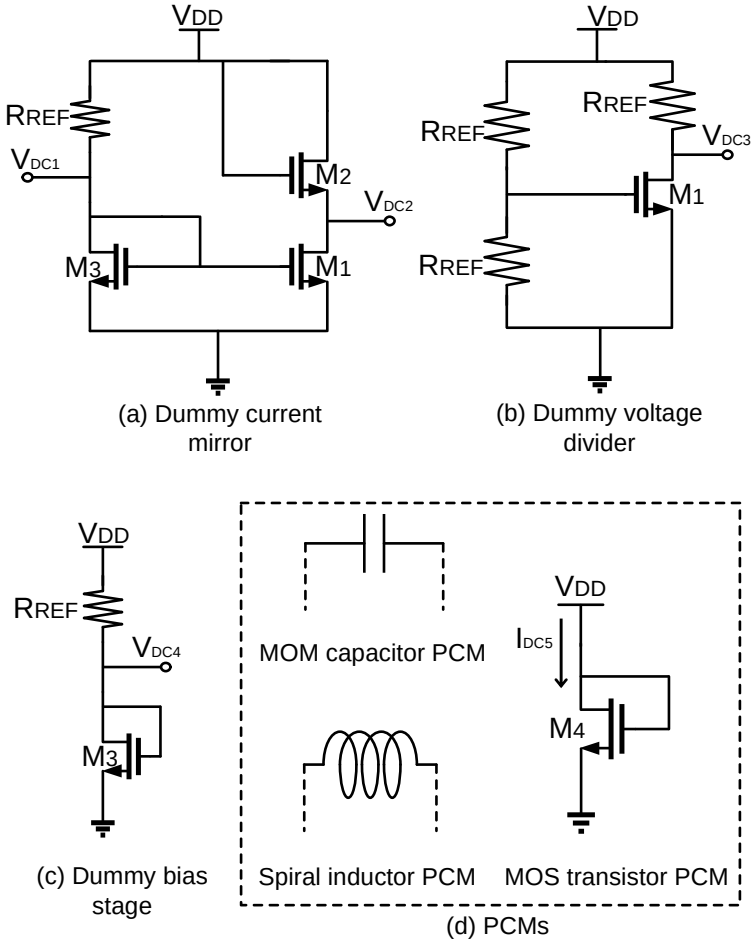


Figure 39: Variation-aware sensors built from identical components that exist in the topology of the 2.4GHz LNA.

To decide what type of non-intrusive variation-aware sensors to use, first it is needed to identify the main process parameters on which the performances strongly depend. For this purpose, we use the closed-form mathematical expressions that relate the performances with the design parameters. These expressions are developed

during the design stage so as to guide appropriately the sizing of the components. In turn, each design parameter is a function of process parameters. Thus, once we write down the relationships between performances and design parameters, we can identify the main process parameters we should seek to monitor.

LNA Performance	Design Parameters
S_{21} (dB)	g_{m1},g_{m2},g_{ds1},g_{ds2} Z_{L_S},Z_{L_G},Z_{L_D},C_{IN},C_{EXT},C_D
S_{11} (dB)	g_{m1},Z_{L_S},Z_{L_G},C_{IN},C_{EXT},C_{GS1},C_{GD1}
S_{22} (dB)	g_{m1},g_{ds1},g_{m2},g_{ds2},Z_{L_S},Z_{L_D},C_D,C_{OUT},C_{GS2},C_{GD2}
NF (dB)	g_{m1},Z_{L_S},Z_{L_G},C_{IN},C_{EXT},C_{GS1},C_{GD1}
1-dB CP (dBm)	V_{OV},g_{m1},Z_{L_S},Z_{L_G},C_{IN},C_{EXT},C_{GS1}
IIP_3 (dBm)	V_{OV},g_{m1},Z_{L_S},Z_{L_G},C_{IN},C_{EXT},C_{GS1}

Table 4: Dependence of 2.4GHz LNA performances on design parameters.

Table 4 shows the main performances of the LNA in the first column and the design parameters on which they depend in the second column. g_{m1}, g_{m2} and g_{ds1}, g_{ds2} are the gate and drain transconductances, respectively, of transistors M_1 and M_2. C_{GS1}, C_{GD1} are the gate-to-source and gate-to-drain capacitances of transistor M_1, respectively, and V_{OV} is the overdrive voltage of M_1, which is fixed by the bias stage that consists of M_3, R_{REF}, and R_{BIAS}. The other design parameters are capacitors and inductors that can be identified in the LNA schematic in Figure 21. Table 4 gives a simplified and brief representation of the theoretical equations that have been derived in Section 3.3.

Based on the dependencies in Table 4, we expect to be able to obtain information on the process parameters that define the gate and drain transconductances and the overdrive voltage of transistor M_1 by monitoring the DC voltages in a dummy current mirror and the drain current of a MOS PCM, as shown in Figure 39. The dummy current mirror and the MOS PCM are built from transistors identical to those in the LNA topology and are biased in moderate to strong inversion and saturation region identically to the biasing conditions of the LNA.

Although there are numerous process parameters that are inter-dependent and intricately related, as we will see in Chapter 6, the selected MOS PCM and dummy current mirror suffice to implicitly track the variations in MOS transistors since, overall, the measurements from the selected non-intrusive variation-aware sensors are shown to correlate very well to the LNA performances. The only design parameters in Table 4 related to MOS transistors that are not covered from a first glance are the intrinsic capacitances C_{GS1} and C_{GD1}, since they require to be excited at a high frequency, which we want to avoid, in order to maintain a low test cost. However, the satisfactory correlation between sensors' measurements and performances observed in Chapter 6 implies that any process parameters that are not monitored have a secondary effect on the variability of performances. If the correlation results are deemed insufficient in the end of the analysis, then one should revisit the selection of sensors to try to capture the variations in all influential design and process parameters.

Variations in the capacitors of the LNA are covered by placing a dummy capacitor PCM. The capacitance value of the capacitor PCM is extracted by S-parameters at 100MHz using the expression:

$$C = \frac{\mathrm{Im}(Y_{11})}{\omega}. \tag{68}$$

Regarding variations in the inductors of the LNA, from our analysis in Chapter 6, it turns out that they explain a large fraction of the variability in the LNA performances. Thus, on one hand, we need to monitor variations in the inductors of the LNA, however, on the other hand, the straightforward inclusion of a dummy inductor PCM in the set of sensors is not a smart choice since an inductor is an area-hungry component and it would increase the area overhead significantly making the BIT approach less appealing.

Therefore, we set out to study the sources of variability in the inductors. For this purpose and as shown in Section 4.5.2, we employed the equivalent lumped 2π-model of the inductors, shown in Figure 33, which is

composed of pure coil inductances L_{s1} and L_{s2}, series resistances of the coil R_{s1} and R_{s2}, parasitic and oxide capacitances C_{OXm}, C_{OX1} and C_{OX2}, and resistances R_{pattm}, R_{patt1} and R_{patt2} that model the patterned ground shield. We recall that inductors in the 65nm technology are realized by stacking several metal levels, while an alucap layer is used to further reduce the coil resistance. The series resistances in the model correspond to these metal and alucap layers.

First, we included an inductor PCM and we extracted the impedance of the inductance by S-parameters at 100MHz:

$$Z_L = R_s + j\omega L_s, \tag{69}$$

where the real R_s and imaginary L_s parts correspond to the series resistance of the metal and alucap layers that form the coil of the inductor and the pure inductance of the coil, respectively. Alternatively, R_s can be extracted by biasing the inductor with a DC voltage source and measuring the current that flows through the coil. Notice also that R_s is expected to have a low value which makes the measurement rather challenging. We added R_s and L_s in the set of measurements from the other types of sensors mentioned above (e.g. dummy current mirror, MOS PCM, capacitor PCM) and, as demonstrated in Chapter 6, the correlation between the sensors' measurements and the LNA performances, not only is significantly improved, but it becomes near perfect, resulting in very accurate alternate test prediction results. Next, we replaced R_s and L_s in the set of measurements with the parameters of the lumped 2π-model, shown in Figure 33, and we obtained very similar results. We experimented by removing the parameters of the lumped 2π-model from the set of measurements one by one and we concluded that we obtain similar results by just keeping the series resistances of the coil R_{s1} and R_{s2}. We observed that the correlation between the series resistance of the coil and the resistivity of the high-level metal (in particular, metal 5 and metal 7) and alucap layers that form the coil reaches 68%, 61%, and 36%, respectively as shown in Figure 34. Also, using the resistivities of metal and alucap layers as alternative measurements, the correlation between the series resistance of the coils and the resistivities reaches 100%. This observation implies that the variability of the series resistance of the coil is principally due to variability of the resistivity of the high-level metal and alucap layers and that the variations in the geometry, thickness, and dielectric of the layers play a less important role. We proceeded by keeping only the resistivity of the high-level metal and alucap layers in the set of measurements and we observed only a small and affordable degradation of the correlation which, in terms of alternate test prediction accuracy, has a very little and rather negligible effect.

This observation points us in monitoring only the variations in the resistivity of the high-level metal and alucap layers of the inductors since this explains to a very large degree the variability in the inductors and, in turn, it suffices to obtain an excellent correlation between the sensors' measurements and LNA performances. To monitor the variations in the resistivity, we no longer need to include an area-hungry inductor PCM, as this can be done with DC measurements on simple metal and alucap resistor PCMs that can laid out in the form of snakes such that they occupy much less area. Furthermore, resistor PCMs of several tens of Ohms can be laid out, thus allowing an accurate measurement, while, as already mentioned above, a direct measurement of R_s may have limited accuracy due to the expected small value of R_s.

5.5.2 *60GHz mm-wave 3-stage LNA Case Study*

Figure 27 shows the topology of the 60GHz mm-wave 3-stage LNA design of our second case study. Looking at this schematic, we can directly identify two analog stages which we can replicate on the die as dummy and use them as non-intrusive sensors. In particular, we can use (a) a dummy common-source stage composed of transistor M_1 and resistor R_1, as shown in Figure 40(a) and (b) a dummy cascode stage composed of transistors M_2 and M_3 and resistors R_2, R_3 and R_4, as shown in Figure 40(b). Notice, as mentioned in Section 3.4, that the LNA consists of one common-source stage and two identical cascode stages. These dummy analog stages aim at monitoring the process variations that influence the gate and drain transconductances of the transistors of the LNA. We use DC currents I_{DC1} and I_{DC2} as alternative measurements for the dummy common-source and the

dummy cascode stages, respectively, based on the test configurations shown in Figure 40(a)-(b). Furthermore, we can use a dummy $N+$ poly resistor R_1 and a dummy MOM capacitor C_1, as shown in Figure 40(c)-(d), in order to directly monitor variations in the $N+$ poly resistors and MOM capacitors of the 60GHz mm-wave LNA. For the dummy resistor we use the DC current I_{DC3} as alternative measurement based on the Ohm's law test configuration shown in Figure 40(c). For the dummy capacitor, we extract the dummy capacitance C_1 from S-parameters at 100MHz using Equation 68. This straightforward approach of extracting non-intrusive sensors from the topology of the CUT has shown to work very well for the 2.4GHz RF LNAs as reported in [35], [100].

As mentioned in the case of the 2.4GHz LNA, to decide what type of non-intrusive variation-aware sensors to use, first it is needed to identify the main process parameters on which the performances strongly depend. We follow the same idea for the case of the 60GHz LNA.

Table 5 shows the main performances of the LNA in the first column and the design parameters on which they depend in the second column, where Z_0 corresponds to the characteristic impedance of the transmission lines, R_G corresponds to the gate resistance of the MOS transistors. Table 5 gives a simplified and brief representation of the theoretical equations that have been derived in Section 3.4. Based on the dependencies in the same table, we expect to be able to obtain information on the process parameters that define the gate and drain transconductances, gate resistances, resistors, capacitors, transmission lines, and bias voltages.

LNA Performance	Design Parameters
S_{21} (dB)	R_G,g_{mX},g_{dsX},C_{GSX},C_{GDX},tranmission lines' Z_0 C_1,C_2,C_3,C_4,C_5,C_6,C_7,C_8,C_9,C_{10},C_{11}
S_{11} (dB)	R_G,g_{m1},g_{ds1},C_1,C_2,C_{GS1},C_{GD1},input loop Z_0
S_{22} (dB)	R_G,g_{mX},g_{dsX},C_9,C_{10},C_{11},C_{GS4},C_{GD4},C_{GS5},C_{GD5},output loop Z_0
NF (dB)	R_G,g_{m1},g_{ds1},C_{GS1},C_{GD1},C_1,C_2,C_3,C_4,C_5,C_6,C_7,C_8,C_9,C_{10},C_{11}
IIP_3 (dBm)	R_G,V_{bias},R_1,R_2,R_5,g_{m1},g_{m2},g_{m3},g_{m4},g_{m5}, transmission lines' Z_0

Table 5: Dependence of 60GHz LNA performances on design parameters.

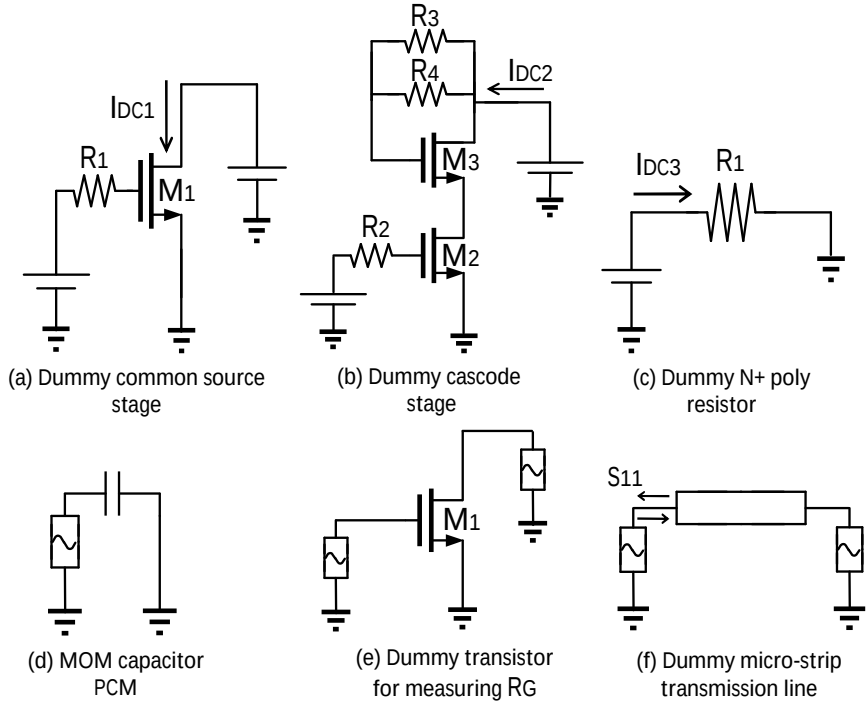


Figure 40: Variation-aware sensors built from identical components that exist in the topology of the 60GHz LNA.

Concerning variations in gate resistance, it turns out that it explains to a large degree the variability in almost all the performances, as it will be shown in Chapter 6. It is known that gate resistance has a major impact on impedance matching and noise performance of MOS transistors that operate at high frequencies. In principle, it is a bias-independent component at DC and low frequencies, but may include an additional component with bias dependence in high frequencies due to two additional effects, namely the distributed transmission line effect on the gate and the Non-Quasi Static (NQS) effect in the channel [101], [102]. A simple expression of gate resistance R_G, based on that at DC and low frequencies, has been used to estimate the value of gate resistance with the influence of the distributed transmission line effect on the gate at high frequencies. Although, a factor α is introduced, which is $\frac{1}{3}$ or $\frac{1}{12}$ depending on the layout structures of the gate connection to account for the distributed RC effects, as given by the following estimation [103]:

$$R_{G,poly} = \frac{R_{Gsh}}{N_f L_f}\left(W_{ext} + \frac{W_f}{\alpha}\right) \tag{70}$$

where R_{Gsh} is the gate sheet resistance, W_F is the channel width per finger, N_f the number of fingers, and W_{ext} is the extension of the polysilicon gate over the active region. The NQS effect in the channel is another effect that needs to be accurately modeled for the high frequency behavior of a MOSFET, where additional bias and geometry dependences need to be taken into account [101], [102]. In other words, when a MOSFET operates at high frequencies, its gate resistance depends not only on the physical gate electrode resistance, but also on the distributed channel resistance, which is "seen" by the signal applied to the gate. Hence, gate resistance consists of two parts:

$$R_G = R_{G,poly} + R_{G,NQS} \tag{71}$$

where $R_{G,poly}$ is the distributed gate electrode resistance [104], [105], from the polysilicon gate material and $R_{G,NQS}$ is the NQS distributed channel resistance "seen" from the gate and is a function of both bias and geometry [101], [102].

Electrically speaking, as frequency of operation rises, the impedance of the parasitic capacitances in MOS devices decrease allowing extra leakages to flow to the ground. Consequently, any small changes at the input of a MOS transistor will be reflected to the output of interest affecting significantly its performances, such as the impedance matching and noise figure.

The theoretical analysis of the input matching network of the 2.4GHz LNA case study led us to (a) investigate why gate resistance plays a dramatically more significant role in mm-wave frequencies compared to the typical frequency ranges of 2.4GHz and, in turn, (b) find an efficient way to monitor variations in the gate resistance. In the aforementioned analysis, one can see the dependence of the input impedance of the circuit on gate resistance. Figure 41 shows the high-frequency small-signal equivalent of the input loop of the 2.4GHz LNA (inductors have been neglected), where the gate resistance has been included. In such frequencies, the impedance of the intrinsic capacitances is much higher than the corresponding impedances in an LNA that operates in mm-wave frequencies. Figure 41 shows the high-frequency small-signal equivalent model of a MOS transistor along with

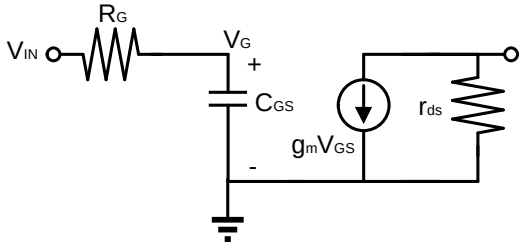


Figure 41: High-frequency small signal equivalent circuit of a MOS device including the effect of gate resistance R_G.

its resistance "seen" from the gate as well as the intrinsic C_{GS} capacitance. R_G and C_{GS} form a voltage divider, where the effect of capacitance C_{GD} can be neglected. Voltage V_G is given by:

$$V_G = \frac{Z_{C_{GS}}}{Z_{C_{GS}} + R_G} V_{IN}, \tag{72}$$

where $Z_{C_{GS}}$ is the impedance of C_{GS} capacitance. As frequency of operation decreases, impedance $Z_{C_{GS}}$ increases and, as a result, R_G becomes negligible compared to $Z_{C_{GS}}$. Hence, voltage V_G is reduced to $V_G = Z_{C_{GS}} V_{IN}$. On the other hand, as frequency increases up to the mm-wave range, impedance $Z_{C_{GS}}$ becomes comparable to the gate resistance R_G and the gate voltage is estimated by Equation 72, implying that any small changes at the input of the MOS device will be reflected to the output of interest affecting significantly its performances, such as impedance matching and noise figure. This is just a simplified explanation why MOS gate resistance becomes to contribute to the performances of circuits that operate in the mm-wave frequency ranges.

Re-visiting our case study, the expression of gate resistance can be extracted by S-parameters using the high-frequency small-signal equivalent of MOS devices, as shown in Figure 42. The procedure is based on the calculation of Y_{11} parameter [105], neglecting the effects of substrate resistances and parasitic gate-to-bulk capacitance (C_{GB}) and setting $V_S = V_D = 0$,

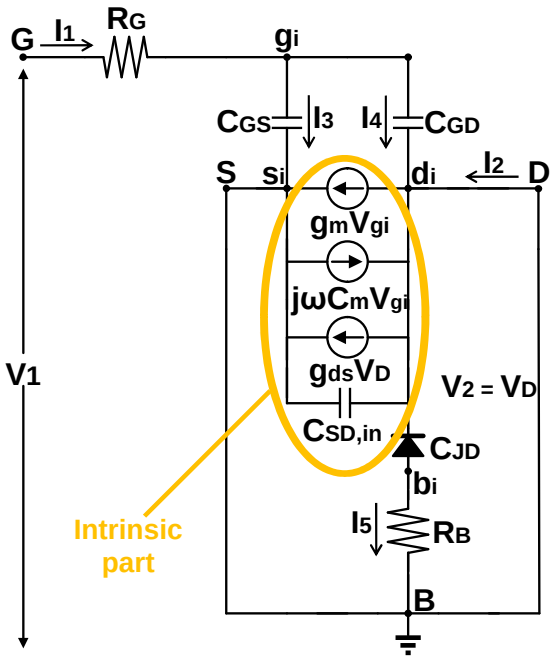


Figure 42: Small-signal equivalent of MOS transistor used for extracting Y_{11} parameter with $V_D = 0$. Both extrinsic and intrinsic part are included.

$$Y_{11} = \left.\frac{I_1}{V_1}\right|_{V_S = V_D = 0} \tag{73}$$

Applying KCL and Ohm's law at node g_i, and branches g_i-s_i and g_i-d_i, G-g_i, respectively, the Y_{11} parameter is derived as follows:

$$I_1 = I_3 + I_4 \tag{74}$$

Next, applying Ohm's law at branches g_i - s_i and g_i-d_i, we have:

$$I_3 = (V_{gi} - V_S) j\omega C_{GS} = V_{gi} j\omega C_{GS} \tag{75}$$

and

$$I_4 = (V_{gi} - V_D)j\omega C_{GD} = V_{gi}j\omega C_{GD} \tag{76}$$

Hence Equation 74, using Equation 75 and Equation 76, gives:

$$I_1 = V_{gi}j\omega C_{GS} + V_{gi}j\omega C_{GD} \Rightarrow I_1 = V_{gi}j\omega(C_{GS} + C_{GD}) \tag{77}$$

Applying Ohm's law at branch G-g_i,

$$I_1 = \frac{V_1 - V_{gi}}{R_G} \tag{78}$$

Therefore, Equation 77, using Equation 78, gives:

$$V_1 = V_{gi}[1 + j\omega(C_{GS} + C_{GD})R_G] \tag{79}$$

Substituting Equation 77 and Equation 79 in Equation 73, we have:

$$Y_{11} = \frac{\omega^2(C_{GS} + C_{gd})^2 R_G + j\omega(C_{GS} + C_{GD})}{1 + \omega^2(C_{GS} + C_{GD})^2 R_G^2} \tag{80}$$

Assuming that $\omega^2(C_{GS} + C_{GD})^2 R_G^2 << 1$, Equation 80 gives,

$$Y_{11} \approx \omega^2(C_{GS} + C_{GD})^2 R_G^2 + j\omega(C_{GS} + C_{GD}) \tag{81}$$

Finally, the expression that estimates the gate resistance is given by:

$$R_G = \frac{Re[Y_{11}]}{(Im[Y_{11}])^2} \tag{82}$$

Variations in gate resistance are covered by placing a dummy transistor PCM, which is identical to one of the corresponding MOS transistors in the topology of the LNA. The value of the gate resistance is extracted by Equation 82 at 100MHz and the DC biasing conditions of the MOS PCM are exactly the same as in the corresponding M_1 device in the LNA schematic.

Finally, we need also to account for variations in the transmission lines. We adopt a straightforward solution, which is to include in the set of non-intrusive sensors a dummy transmission line, as shown in Fig. 40(f). As alternative measurement we use the S_{11} extracted at a low frequency 100MHz, for which both the inductance and capacitance of the transmission line are sensitized.

5.6 CONCLUSIONS

In this chapter, the principle of operation of non-intrusive variation-aware sensors, the overall benefits of the proposed test approach and the procedure to select such sensors for our case studies were discussed. In short, the basic characteristic of the non-intrusive variation-aware sensors is that they capitalize on the undesired phenomenon of process variations. The underlying idea is to monitor process variations instead of measuring directly the RF performances of the LNA. These sensors are placed in close physical proximity to the corresponding structures in the LNA, so that they both "witness" the same D2D and correlated WID variations and, as a result, the measurements obtained on the sensors will be correlated to the performances of the LNA to a very large extent. The alternate test is used, so as to map the sensors' measurements to the performances of the LNA, providing an indirect low-cost test. Low area overhead, low test cost, non-intrusiveness and genericness are some of the most important benefits of the proposed BIT approach. Finally, the procedure of selecting

the appropriate non-intrusive variation-aware built-in sensors for our two case studies was presented. In next chapter, the results of alternate test based on the previously presented embedded sensors are demonstrated and discussed in details.

6

SIMULATION RESULTS

In this chapter, the non-intrusive idea will be demonstrated for our two case studies. In Section 6.1, we present the alternate test experiment results for the 2.4GHz LNA based on both schematic-level and post-layout simulations. In Section 6.2, we present the alternate test experiment results for the 60GHz LNA based on post-layout simulations. Finally, Section 6.3 concludes the chapter.

6.1 2.4GHZ INDUCTIVE DEGENERATED LNA CASE STUDY

6.1.1 *Simulation Results at Schematic Level*

The topology of the 2.4GHz LNA is shown in Figure 21. The amplifier has been designed using the 65nm CMOS065 bulk technology provided by ST Microelectronics. Figures 43 and 44 illustrate the nominal simulation results for the main performances, including S-parameters, noise figure, and linearity, while Figure 45 shows the stability performances, K_f, and B_{1f}. Table 6 summarizes the values of the performances of the circuit.

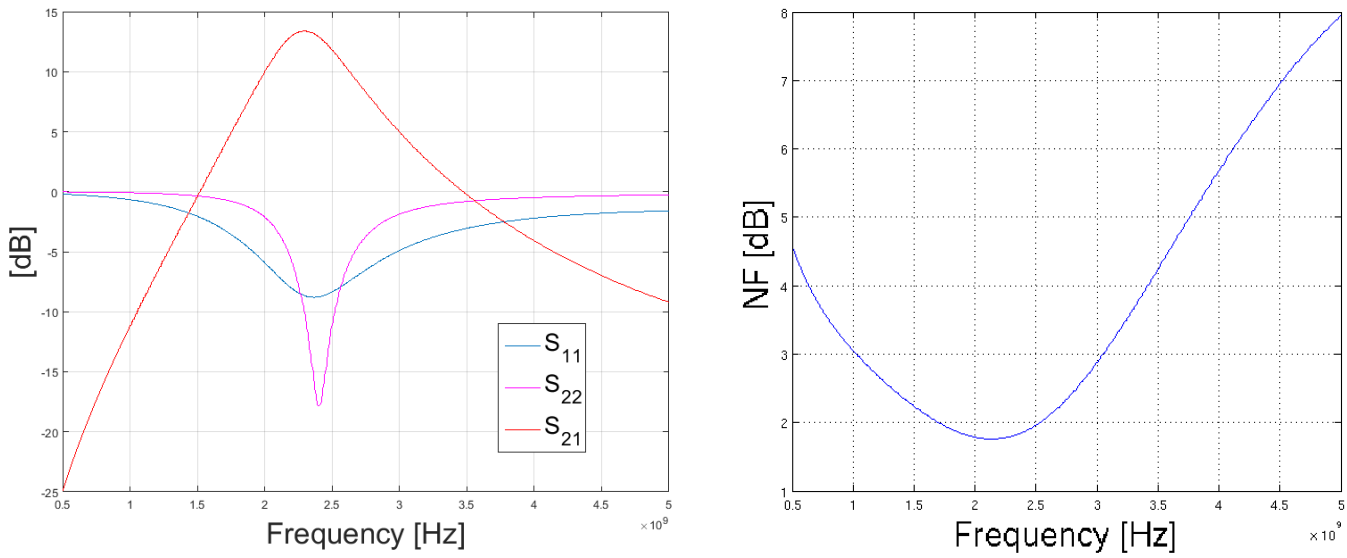


Figure 43: S-parameters and NF of the 2.4GHz LNA.

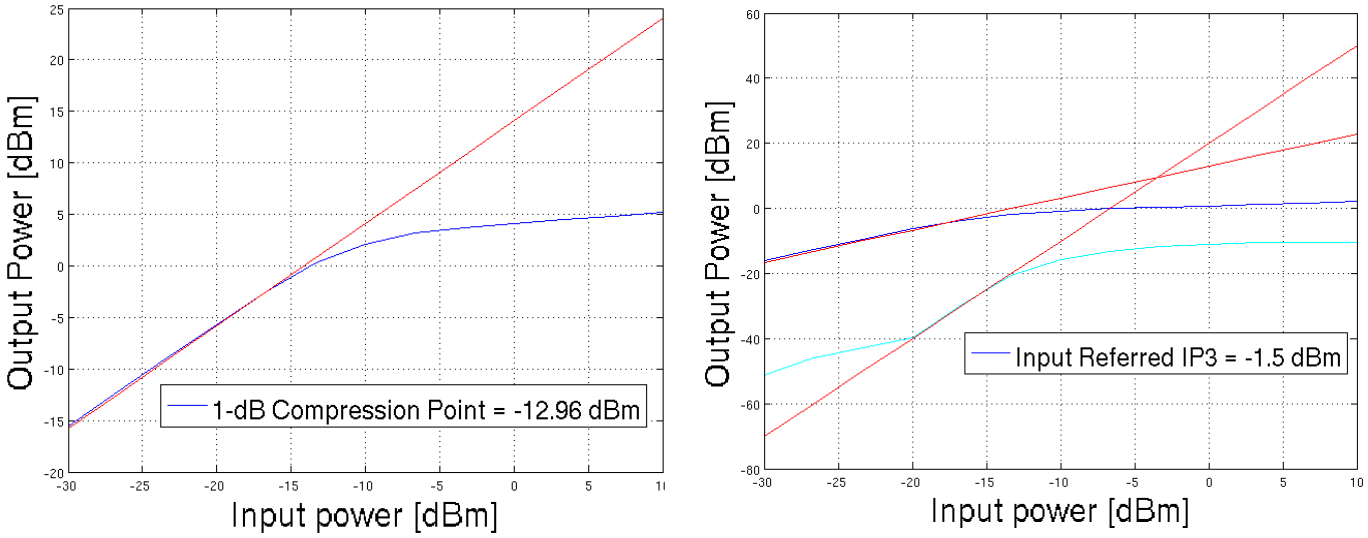


Figure 44: Linearity performances of the 2.4GHz LNA.

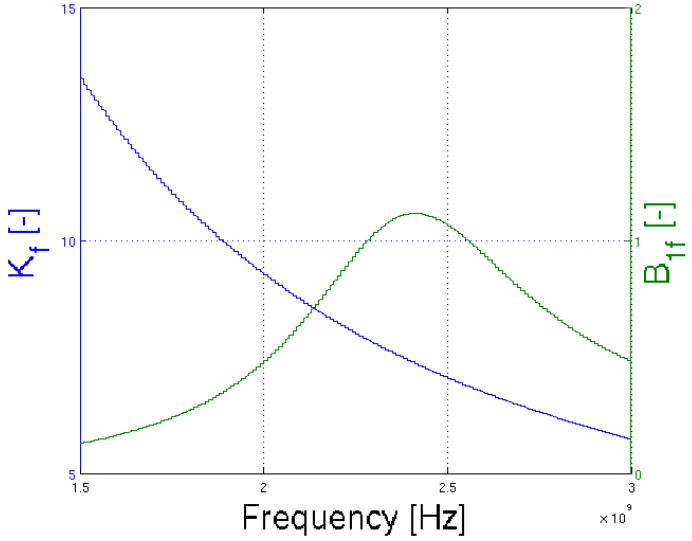


Figure 45: K_f and B_{1f} stability factors of the 2.4GHz LNA.

Summarized Results	
Frequency [GHz]	2.4
Supply Voltage [V]	1.2
DC Power Dissipation [mW]	9.6
S_{21} [dB]	14.45
S_{11} [dB]	-8.76
S_{22} [dB]	-17.81
NF [dB]	1.86
K_f [-]	7.4
B_{1f} [-]	1.1
1-dB CP [dBm]	-12.9
IIP_3 [dBm]	-1.55

Table 6: Summary of 2.4GHz LNA performances using schematic-level simulations.

The schematic view of the non-intrusive variation-aware sensors is shown in Figure 39. The embedded sensors that we have employed in the analysis, include: (a) a dummy current mirror, (b) a diode-connected MOS PCM, (c) a MOM capacitor PCM, and (d) an inductor PCM, as well as (e) high-level metal and alucap resistor PCMs. The set of alternative measurements include: (a) DC voltages V_{DC1}, V_{DC2} from the dummy current mirror, (b) DC current I_{DC3} from the diode-connected MOS PCM, (c) capacitance from the MOM capacitor PCM, denoted by C_{PCM}, (d) inductance and resistance from the inductor PCM (e.g. real R_s and imaginary L_s parts of inductor impedance), and (e) resistivity from high-level metal and alucap resistor PCMs, denoted by ρ_{metal} and ρ_{alucap}, respectively, obtained by dividing a DC voltage applied at their terminals with the DC current that flows through them. The capacitance and the inductance are extracted by S-parameters at a low frequency in the order of a few tens of MHz. In the following alternate test experiments, we employ different groups of non-intrusive variation-aware sensors and their corresponding measurements, as shown in Table 7.

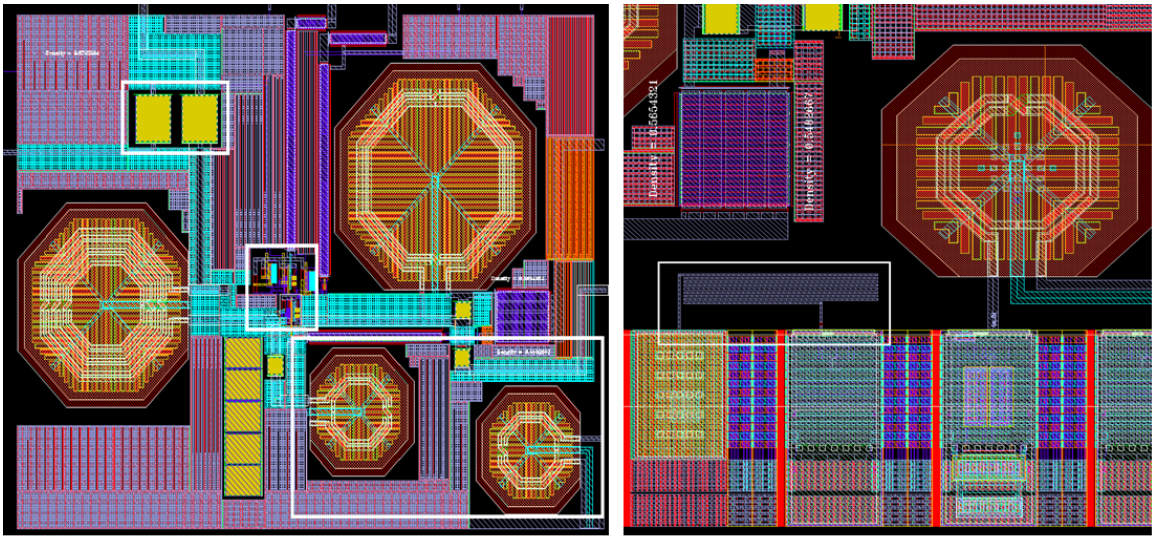
It should be noted that only measurements V_{DC1}, V_{DC2}, and I_{DC3} from the dummy current mirror and MOS PCM have been used in the analysis, since thy prove to be sufficient in providing, statistically, reliable prediction results. Measurements from dummy voltage divider and dummy bias stages are not considered, since, statistically speaking, they do not offer better prediction results for two reasons: (a) the large Monte Carlo sample set is proved to be sufficient and instances that are used in the training set are representative of the fabrication process and (b) simulation experiments have shown that monitoring variations in voltages and currents of the dummy voltage divider and dummy bias stages are already covered by tracking variations in measurements V_{DC1}, V_{DC2}, and I_{DC3}, where MOS transistors' and poly resistors' variability is considered. Furthermore, notice that the aspect ratio of the MOS PCM is W/L = $20\mu m/0.06\mu m$, where the width of the MOS PCM is not the same as in transistor M_1. The width of the MOS PCM has been selected to be lower than the width of M_1, so as to: (a) show that variations in this parameter are not as critical as in case of length (as already shown in Session 4.5) and (b) ensure that the MOS PCM is able to drive a sufficient current without causing any breakdown effects during its DC operation. On the other hand, as discussed in Chapter 7 later on, all measurements from the dummy stages are considered in the analysis, since a small sample set of fabricated instances is used for proving the non-intrusive test idea.

	Dummy current mirror	MOS PCM	MOM PCM	Inductor PCM	Inductor PCM	Resistor PCMs
	V_{DC1}, V_{DC2}	I_{DC3}	C_{PCM}	R_s	L_s	$\rho_{metal}, \rho_{alucap}$
Group 1	✓	✓				
Group 2	✓	✓	✓			
Group 3	✓	✓	✓	✓	✓	
Group 4	✓	✓	✓	✓		
Group 5	✓	✓	✓		✓	
Group 6	✓	✓	✓		✓	✓
Group 7	✓	✓	✓			✓

Table 7: Groups of non-intrusive variation-aware sensors and the corresponding groups of measurements used in the analysis.

Figure 46(a) illustrates the layout view of the LNA with the embedded non-intrusive variation-aware sensors, including the ground planes which are used to guarantee proper return ground paths. Post-layout simulation results will be presented in Section 6.1.2. We illustrate the layout view herein, in order to give an idea of the area overhead of the non-intrusive variations-aware sensors. More details on the design of the layout will be presented in Section 6.1.2.

The highlighted area in the upper left of Figure 46(a) encloses the capacitor C_{IN} of the LNA and an identical dummy MOM capacitor PCM that has been placed in close proximity to it. The highlighted area in the middle of Figure 46(a) encloses the transistors of the LNA, as well as all the dummy stages (dummy current mirror,



(a) LNA with non-intrusive variation-aware sensors

(b) Metal 7 resistor PCM

Figure 46: Layout view of (a) 2.4GHz LNA with embedded non-intrusive variation-aware sensors and (b) metal 7 resistor PCM.

dummy voltage divider, dummy bias stage, and diode-connected MOS PCM). The highlighted area in bottom-right of Figure 46(a) encloses the inductor L_S of the LNA, which is the smaller inductor of the LNA as compared to L_G and L_D, and an identical dummy inductor PCM placed in close physical proximity to it. Figure 46(b) illustrates the placement of a dummy metal 7 resistor PCM close to the dummy inductor L_S. Finally, Figure 47 depicts the placement of a dummy metal 5 resistor PCM outside the I/O pad ring. The full layout view of the design is illustrated in Chapter 7.

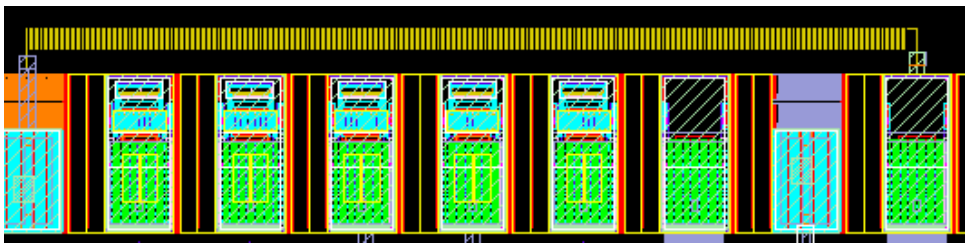


Figure 47: Layout view of metal 5 resistor PCM.

The LNA occupies an area of $442141\mu\text{m}^2$, which is mainly determined by the three inductors, whereas the high-level metal and alucap resistor PCMs, the MOM PCM, and the dummy circuits occupy, respectively, an area of $2400\mu\text{m}^2$, $2332\mu\text{m}^2$, and $7410\mu\text{m}^2$. However, these non-intrusive variation-aware sensors practically have a zero area overhead since they are placed in areas that otherwise are left void to respect electromagnetic rules. It should be noticed that the areas of the high-level metal and alucap resistor PCMs increase with the resistance value that is implemented. We should choose the smaller possible resistance value that can be measured reliably in a real environment. Specifically, in our experiment, we used 30Ω resistances. In any case, the areas of the high-level metal and alucap resistor PCMs can be made considerably smaller than the area of the smallest inductor L_S. The only sensor for which we cannot claim zero area overhead is the inductor PCM, but, as it will be demonstrated, it can be replaced effectively by the high-level metal and alucap resistor PCMs, without degrading significantly the prediction accuracy of alternate test.

The LNA and the non-intrusive variation-aware sensors have been simulated in the Cadence Virtuoso environment using the SpectreRF simulator. The actual distributions of process parameters defined in the PDK were used during Monte Carlo analysis to model D2D and WID variations. Table 8 lists the LNA performances and the standard deviation observed in a Monte Carlo sample set of 1000 circuit instances considering (a) only D2D and (b) both D2D and WID variations.

Performance	Standard Deviation D2D	Standard Deviation D2D and WID
S_{21} (dB)	0.34	0.34
S_{11} (dB)	0.3	0.3
S_{22} (dB)	2.53	2.54
NF (dB)	0.07	0.07
1-dB CP (dBm)	0.44	0.42
IIP_3 (dBm)	0.28	0.27

Table 8: List of 2.4GHz LNA performances showing their standard deviation from a Monte Carlo analysis considering (a) only D2D variations and (b) both D2D and WID variations.

The same sample set was used to learn the regression functions in alternate test by employing different groups of non-intrusive variation-aware sensors, as shown in Table 7. Out of the 1000 instances, we used 800 circuit instances in the training set and the rest of 200 instances were used as a validation set to provide an unbiased estimate of the prediction error of alternate test.

As mentioned in Chapter 5, the prediction error of alternate test is expressed in terms of several error metrics, including the (a) average RMS error expressed in %, (b) absolute average RMS error in units of each performance, (c) maximum error in units of each performance, and (d) Pearson correlation coefficient between the simulated (e.g. “true") and predicted performances expressed in %.

These metrics should be used with caution and judiciousness. The average RMS error metric offers an unspecific indication of prediction quality. It is commonly reported in alternate test experiments and we also include it in our analysis for the purpose of completeness. The average absolute RMS error and maximum error are the most informative error metrics since they allow us to make direct judgments about the quality of pass/fail test decisions after comparing the predicted performance with its specification. To speak of acceptable prediction results, the average absolute RMS error is recommended to be much smaller than one standard deviation of the performance, whereas the maximum error is recommended to be smaller than the measurement and repeatability errors in an ATE environment. Furthermore, we need to take into consideration the specification for each performance and the margin between the performance value with respect to the specification. No matter how large these error metrics are, if they are smaller than this margin, then the alternate test will not lead to any erroneous pass/fail test decisions. Finally, the Pearson correlation coefficient should be viewed with a pinch of salt. It offers a metric of linear correlation and a poor value does not necessarily imply poor prediction results that lead to erroneous pass/fail decisions. It is a qualitative measure that shows the linear fit between the true and predicted performances. The closer the correlation coefficient to 100% is, the better the prediction results are. However, it should be stressed that a correlation coefficient considerably lower than 100% does not necessarily lead to unsatisfactory prediction results. As a first step, in the following discussions that concern the 2.4GHz LNA case study, we give a simple, yet sufficient, way of interpreting the prediction results. The interpretation is mostly based on the overall statistical behavior of the prediction errors as well as the correlation coefficients. The analysis of the prediction results becomes more challenging in case of the 60GHz LNA case study, where the error metrics are compared to the standard deviation of the performances and the measurement repeatability error in a conventional ATE environment. This comparison is essentially one efficient way to evaluate the success of the alternate test based on the proposed non-intrusive test idea.

Table 9 shows the prediction results using Group 1 of sensors considering only D2D variations. As it can be seen, the prediction results are not satisfactory, implying that monitoring capacitance and inductance is required to increase the prediction accuracy.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	2.28 %	0.33 dB	30.65 %	1.13 dB
S_{11}	3.28 %	0.28 dB	28.59 %	0.80 dB
S_{22}	13.63 %	2.15 dB	35.6 %	6.58 dB
NF	3.82 %	0.07 dB	27.57 %	0.23 dB
1-dB CP	3.35 %	0.43 dBm	24.6 %	1.31 dBm
IIP_3	17.7 %	0.27 dBm	26.64 %	0.96 dBm

Table 9: Alternate test prediction results using Group 1 of non-intrusive variation-aware sensors considering only D2D variations.

Table 10 shows the prediction results using Group 2 of sensors where the MOM capacitor PCM is added to Group 1 of sensors. As it can be seen, the prediction results are slightly improved in case of S-parameters and NF, whereas they are significantly improved in case of 1-dB CP and IIP_3, which justifies the theoretical approach of Section 3.3.4. Moreover, correlation coefficient is appreciably higher compared to that in Table 9 for NF, 1-dB CP, and IIP_3, while for S_{21}, S_{11}, and S_{22}, the correlation is slightly improved. In any case, the prediction results are still not satisfactory, implying that monitoring inductance is required to increase the prediction accuracy.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.95 %	0.28 dB	56.65 %	0.69 dB
S_{11}	3.19 %	0.28 dB	34.95 %	0.66 dB
S_{22}	13.48 %	2.13 dB	30.03 %	6.05 dB
NF	3.09 %	0.06 dB	63.14 %	0.14 dB
1-dB CP	2.01 %	0.26 dBm	81.12 %	0.93 dBm
IIP_3	12.15 %	0.18 dBm	74.92 %	0.64 dBm

Table 10: Alternate test prediction results using Group 2 of non-intrusive variation-aware sensors considering only D2D variations.

Table 11 shows the prediction results using Group 3 of sensors where the inductor PCM is added to Group 2 of sensors using both R_s and L_s in the list of measurements. As it can be seen, the prediction results are almost perfect considering any error metric. This result demonstrates that, when considering only D2D variations, a low-cost alternate test based on Group 3 of sensors is capable of replacing the standard specification tests for measuring the performances while maintaining the same test accuracy.

The next step is to investigate whether we can avoid using an inductor PCM since it incurs a large area overhead. Table 12 shows the prediction results using Group 4 of sensors where the inductor PCM is added to Group 2 of sensors using only R_s in the list of measurements, while Table 13 shows the prediction results using Group 5 of sensors where the inductor PCM is added to Group 2 of sensors using only L_s in the list of measurements. Comparing Tables 12 and 13 to Table 11, we observe that the results deteriorate when we do not include R_s in the list of measurements, while, if we do not include L_s in the list of measurements, the results show the same test accuracy in practice, with the exception of performance S_{22} for which the correlation coefficient drops to 54.89%. However, according to anecdotal evidence from industry, the test for performance S_{22} is known to be redundant, since if a circuit fails the specification of S_{22}, then it is guaranteed to fail the

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.61 %	0.09 dB	96.66 %	0.34 dB
S_{11}	1.11 %	0.1 dB	94.37 %	0.28 dB
S_{22}	3.14 %	0.5 dB	97.67 %	1.3 dB
NF	1.29 %	0.02 dB	94.59 %	0.06 dB
1-dB CP	0.72 %	0.09 dBm	97.76 %	0.23 dBm
IIP_3	3.67 %	0.06 dBm	97.97 %	0.13 dBm

Table 11: Alternate test prediction results using Group 3 of non-intrusive variation-aware sensors considering only D2D variations.

specification of another performance too. In any case, despite the rather low correlation coefficient, the average absolute RMS error is not prohibitive and is comparable to the repeatability measurement error on an ATE. We conclude that we can safely avoid employing measurement L_s without affecting significantly the test accuracy.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.06 %	0.15 dB	89.3 %	0.41 dB
S_{11}	1.39 %	0.12 dB	91.32 %	0.3 dB
S_{22}	12.23 %	1.93 dB	54.89 %	4.81 dB
NF	1.29 %	0.02 dB	94.54 %	0.05 dB
1-dB CP	1.23 %	0.16 dBm	93.3 %	0.47 dBm
IIP_3	8.22 %	0.12 dBm	89.3 %	0.38 dBm

Table 12: Alternate test prediction results using Group 4 of non-intrusive variation-aware sensors considering only D2D variations.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.79 %	0.26 dB	65.37 %	0.66 dB
S_{11}	3.07 %	0.27 dB	43.84 %	0.65 dB
S_{22}	7.16 %	1.13 dB	87.23 %	3.8 dB
NF	3.07 %	0.06 dB	63.41 %	0.14 dB
1-dB CP	1.8 %	0.23 dBm	84.89 %	0.68 dBm
IIP_3	9.93 %	0.15 dBm	83.92 %	0.45 dBm

Table 13: Alternate test prediction results using Group 5 of non-intrusive variation-aware sensors considering only D2D variations.

A further step is to try to find alternative sensors to replace the measurement R_s. As explained in Chapter 5, the idea is to place high-level metal and alucap resistor PCMs and circumvent the need to include an inductor PCM. Table 14 shows the prediction results using Group 6 of sensors where in addition to Group 2 of sensors we have included the measurement L_s and the high-level metal and alucap resistor PCMs. Comparing Table 14 to Table 11, we observe that the prediction accuracy is practically the same, thus, we can replace the measurement R_s with the resistivity of high-level metal and alucap lines. Table 15 shows the prediction results using Group 7 of sensors where, compared to Group 3 the inductor PCM is eliminated and the high-level metal and alucap

resistor PCMs are added. Comparing Table 15 to Table 12, we can observe that the prediction accuracy is practically the same.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.53 %	0.08 dB	97.42 %	0.1 dB
S_{11}	1.11 %	0.1 dB	94.39 %	0.26 dB
S_{22}	2.62 %	0.41 dB	98.38 %	1.43 dB
NF	1.31 %	0.02 dB	94.41 %	0.06 dB
1-dB CP	0.68 %	0.09 dBm	98.06 %	0.22 dBm
IIP_3	3.31 %	0.05 dBm	98.33 %	0.15 dBm

Table 14: Alternate test prediction results using Group 6 of non-intrusive variation-aware sensors considering only D2D variations.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.09 %	0.16 dB	88.62 %	0.44 dB
S_{11}	1.37 %	0.12 dB	91.4 %	0.31 dB
S_{22}	12.41 %	1.96 dB	53.59 %	4.86 dB
NF	1.3 %	0.02 dB	94.49 %	0.05 dB
1-dB CP	1.22 %	0.16 dBm	93.39 %	0.49 dBm
IIP_3	8.38 %	0.13 dBm	88.83 %	0.38 dBm

Table 15: Alternate test prediction results using Group 7 of non-intrusive variation-aware sensors considering only D2D variations.

Finally, we repeated the analysis for Groups 6 and 7 of sensors considering this time both D2D and WID variations. The prediction results are shown in Tables 16 and 17. Comparing Tables 16 and 17 to Tables 14 and 15, respectively, where only D2D variations are taken into account, we observe that WID variations degrade slightly the correlations, yet the correlations remain strong and the alternate test still offers excellent results.

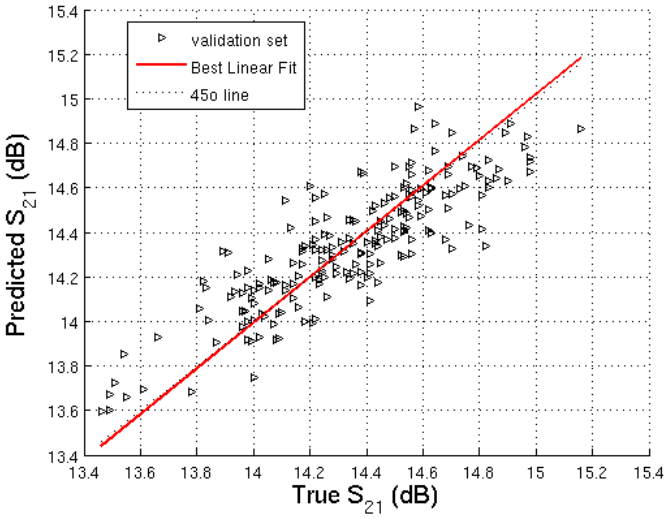
Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.72 %	0.1 dB	94.98 %	0.31 dB
S_{11}	1.35 %	0.12 dB	92.78 %	0.34 dB
S_{22}	3.03 %	0.48 dB	98.13 %	2.01 dB
NF	1.18 %	0.02 dB	95.14 %	0.05 dB
1-dB CP	1.09 %	0.14 dBm	94.14 %	0.39 dB
IIP_3	5.62 %	0.08 dBm	94.79 %	0.25 dB

Table 16: Alternate test prediction results using Group 6 of non-intrusive variation-aware sensors considering both D2D and WID variations.

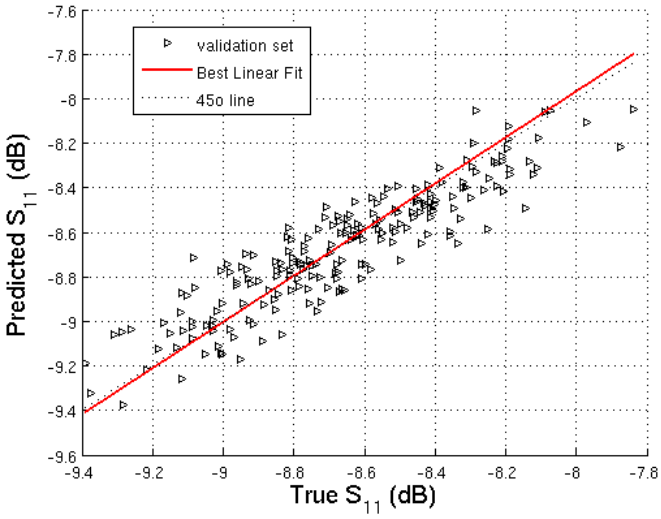
Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.24 %	0.18 dB	84.25 %	0.5 dB
S_{11}	1.58 %	0.14 dB	89.95 %	0.37 dB
S_{22}	13.12 %	2.1 dB	55.96 %	5.26 dB
NF	1.16 %	0.02 dB	95.30 %	0.05 dB
1-dB CP	1.57 %	0.2 dBm	87.7 %	0.57 dBm
IIP_3	9.87 %	0.15 dBm	83.09 %	0.45 dBm

Table 17: Alternate test prediction results using Group 7 of non-intrusive variation-aware sensors considering both D2D and WID variations.

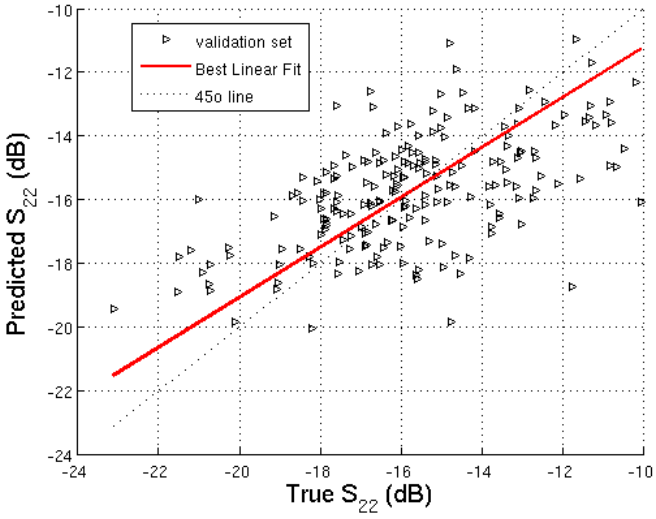
Figure 48 depicts the scatter plots of true vs. predicted performances for Group 7 of sensors considering both D2D and WID variations. Figure 48 offers a visualization of the prediction results in Table 17. Each point corresponds to one circuit instance in the validation set. As it can be seen, the points are distributed across the diagonal for all performances except S_{22} for which the correlation coefficient from Table 17 is 55.96%.



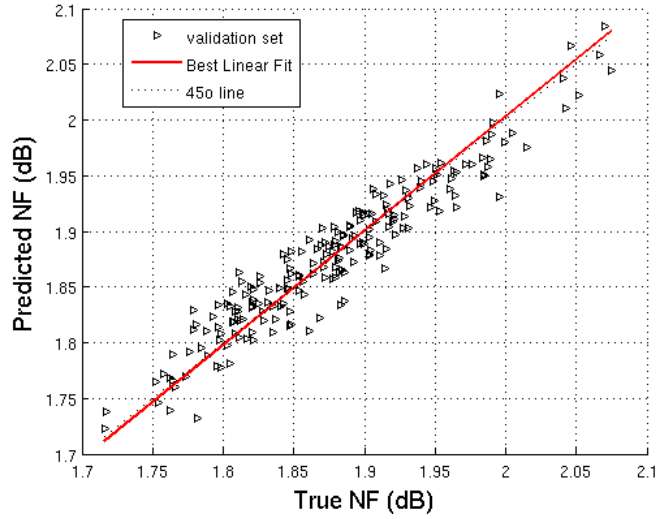
(a) S_{21}



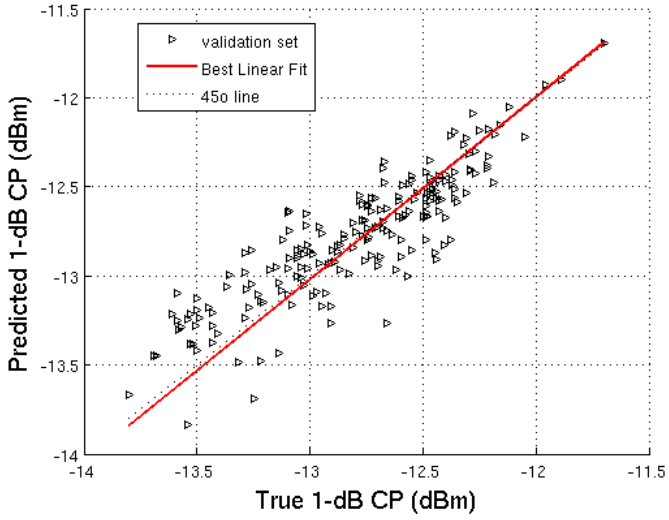
(d) S_{11}



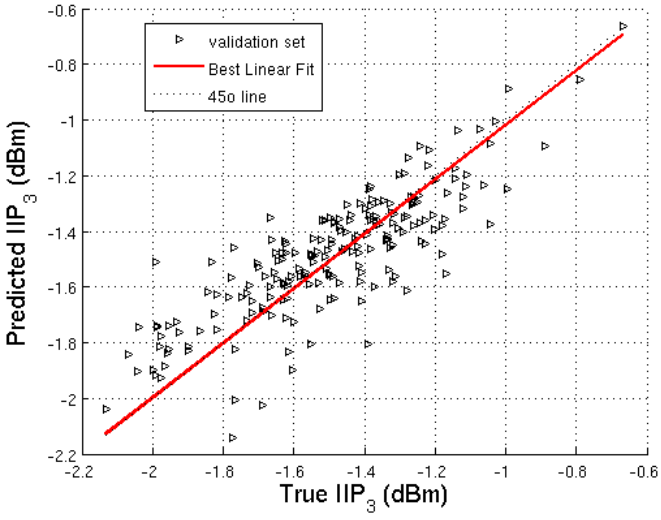
(b) S_{22}



(e) NF



(c) 1-dB CP



(f) IIP_3

Figure 48: Scatter plots of true vs. predicted values for all the performances of the LNA considering Group 7 of sensors. Both D2D and WID variations are considered.

6.1.2 *Post-Layout Simulation Results*

Figure 49 shows the layout of the 2.4GHz LNA along with the embedded non-intrusive variation-aware sensors, where compared to Figure 46, we have removed the ground planes. The ground planes have been removed so as to be able to perform post-layout simulations in a time-efficient manner without convergence issues.

Design Rule Check (DRC) and Layout Versus Schematic (LVS) checks have been performed using Calibre tools provided by Mentor Graphics. DRC is performed, so as to check and verify that the design rules of the 65nm technology are followed. These rules include minimum width of metal layers, distance between the two metals of the same level, distance between two wells of the same doping, minimum metal density, minimum distance between contacts and vias, etc. The design check rules are included in the Design Rules Manual (DRM). After employing the DRC check, LVS simulation is performed, in order to verify that the components sizing and wiring in the layout view meet the corresponding elements in the schematic view. In addition, Electrical Rule Check (ERC) rules are part of the LVS checks, such that electrical rules (e.g. maximum voltage tolerated on the gate of a MOS transistor, etc.) are also validated. The final step before running post-layout simulations consists of components' parasitic extraction. Calibre tools provide the PEX (Parasitic EXtraction) emulator, which estimates the parasitics that are introduced in the layout. In our design, we have performed four different extraction modes at the same time. These modes include: (a) parasitic net resistances extraction (mode R), (b) parasitic net-to-ground coupling capacitances extraction (mode C), (c) parasitic net-to-net coupling capacitances extraction (mode Cc), and (d) parasitic net self-inductances extraction. Mode C is more suitable for digital applications, whereas mode Cc is very often used in RF applications.

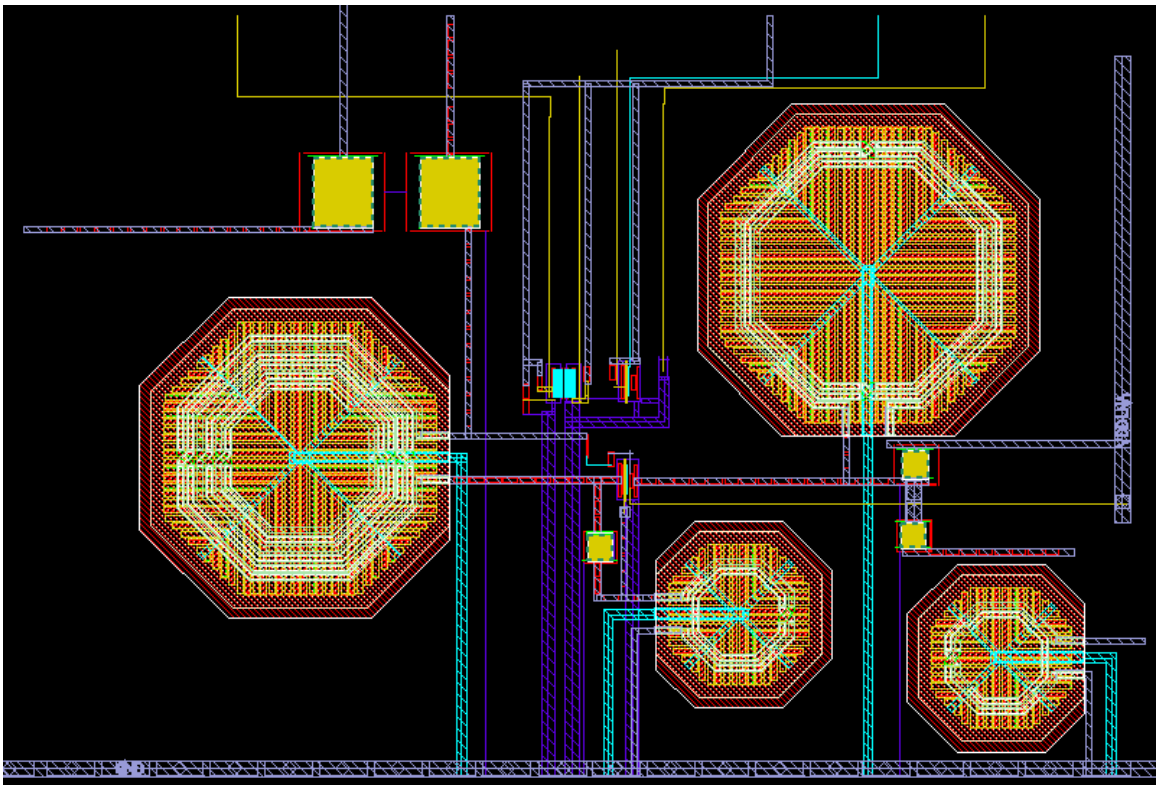
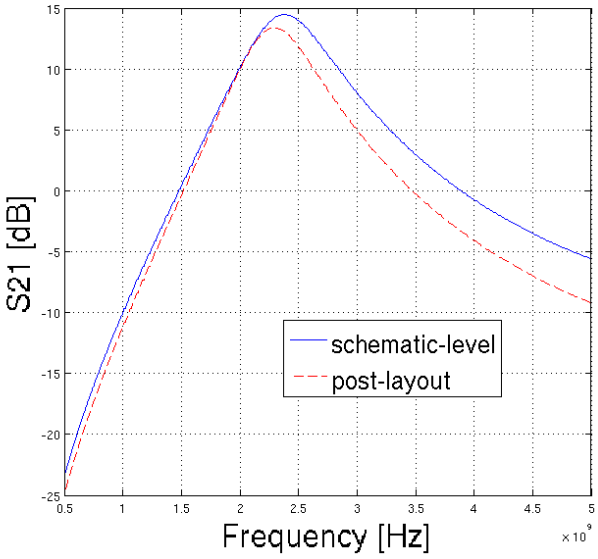
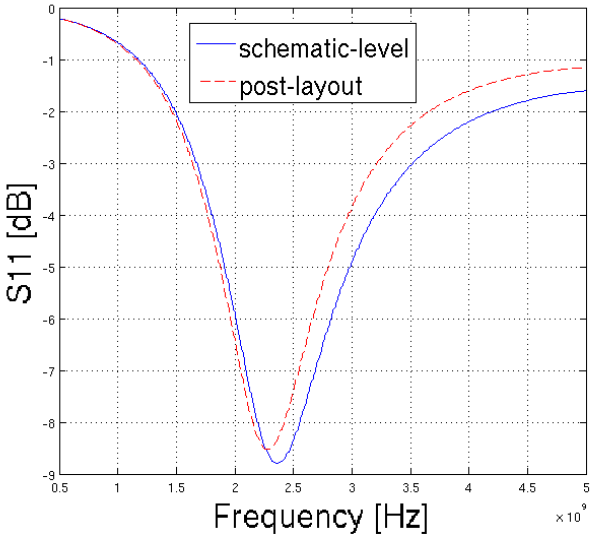


Figure 49: Pre-silicon layout view of the 2.4GHz LNA along with the embedded non-intrusive variation-aware sensors.

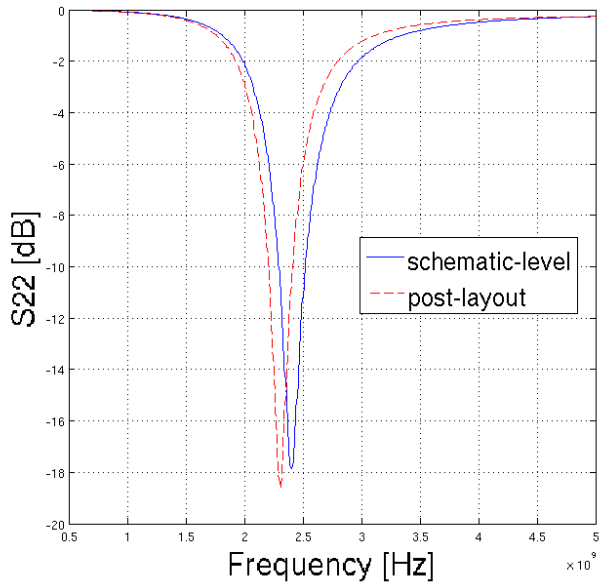
Figure 50 plots S_{21}, S_{11}, S_{22}, and NF using both schematic-level (in blue solid line) and post-layout (in red dashed line) simulations. Figure 51 illustrates 1-dB CP and IIP_3 using both schematic-level (in blue solid line) and post-layout (in red dashed line) simulations. As it can be seen from Figure 50, there is a frequency shift of about 100MHz in all the performances due to parasitic net resistances, net-to-ground, net-to-net capacitances and net self-inductances. Table 18 summarizes the performances.



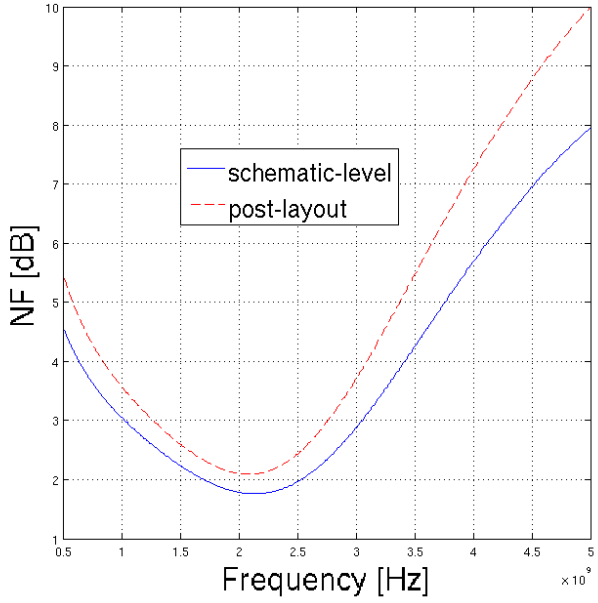
(a) S_{21}



(b) S_{11}



(b) S_{22}

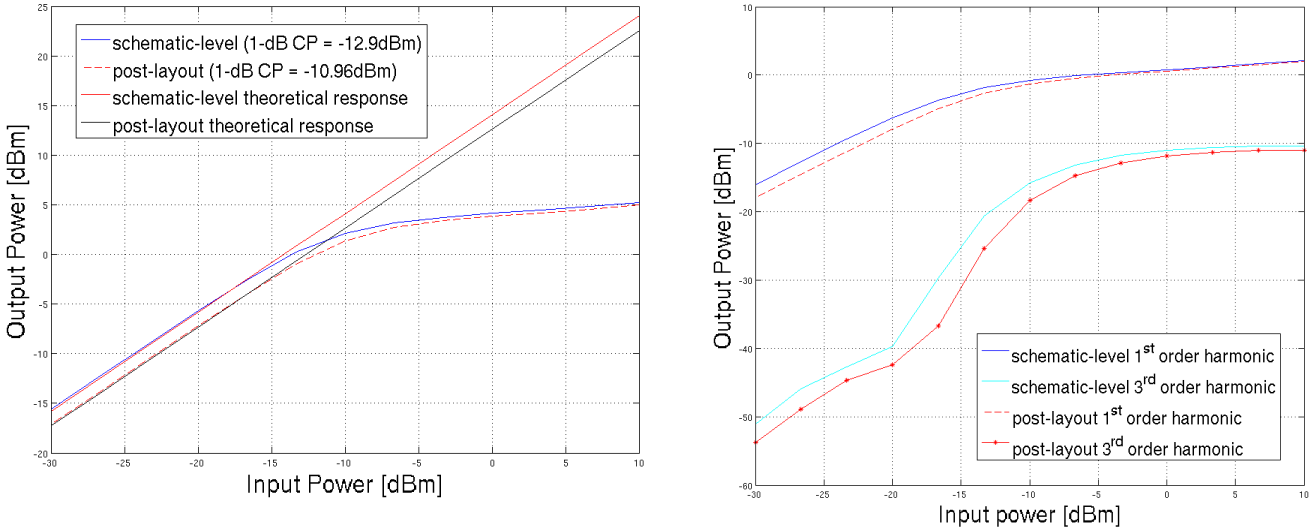


(d) NF

Figure 50: Comparison between schematic-level and post-layout simulation results of the 2.4GHz LNA for S-parameters and NF.

The alternate test technique is demonstrated next based on the post-layout simulation results of the 2.4GHz LNA circuit. Table 19 lists the standard deviation of the post-layout simulated performances observed in a Monte Carlo sample set of 1000 circuit instances considering only D2D variations. The same sample set was used to learn the regression functions in the alternate test by employing the same groups of non-intrusive variation-aware sensors, as shown in Table 7. Out of the 1000 instances, we used 800 circuit instances in the training set and the rest of 200 instances were used as a validation set to provide an unbiased estimate of the prediction error of alternate test.

The list of the performances include the S-parameters, NF, and IIP_3. The 1-dB CP has not been simulated, since it requires large memory storage and is time-consuming due to the simulation methodology. In principle,



(a) 1-dB CP.

(b) IIP_3.

Figure 51: Comparison between schematic-level and post-layout simulation results of the 2.4GHz LNA for 1-dB CP and IIP_3.

Summarized Results	
Frequency [GHz]	2.4
Supply Voltage [V]	1.2
DC Power Dissipation [mW]	8.76
S_{21} [dB]	12.89
S_{11} [dB]	-8.13
S_{22} [dB]	-10.67
NF [dB]	2.28
1-dB CP [dBm]	-10.96
IIP_3 [dBm]	-0.4

Table 18: Summary of 2.4GHz LNA performances using post-layout simulations.

1-dB CP is simulated by sweeping the input power and observing the output power. On the other hand, IIP_3 has been simulated rapidly and accurately following the method proposed in [106].

Performance	Standard Deviation D2D
S_{21} (dB)	0.44
S_{11} (dB)	0.34
S_{22} (dB)	3.2
NF (dB)	0.08
IIP_3 (dBm)	0.35

Table 19: List of 2.4GHz LNA performances showing their standard deviation from a Monte Carlo analysis considering only D2D variations. Post-layout simulations have been performed.

Tables 20 - 25 show the prediction results of the alternate test using the group of non-intrusive sensors shown in Table 7.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	2.53 %	0.32 dB	70.85 %	0.95 dB
S_{11}	3.69 %	0.3 dB	42.25 %	0.81 dB
S_{22}	24.04 %	2.6 dB	48.68 %	6.66 dB
NF	2.87 %	0.06 dB	67.18 %	0.16 dB
IIP_3	55.97 %	0.19 dBm	83.51 %	0.46 dBm

Table 20: Alternate test prediction results using Group 2 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.72 %	0.09 dB	97.95 %	0.31 dB
S_{11}	0.86 %	0.07 dB	97.7 %	0.18 dB
S_{22}	2.94 %	0.32 dB	99.43 %	1.16 dB
NF	1.09 %	0.03 dB	95.85 %	0.06 dB
IIP_3	16.99 %	0.06 dBm	98.6 %	0.16 dBm

Table 21: Alternate test prediction results using Group 3 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.62 %	0.21 dB	89.34 %	0.74 dB
S_{11}	2.45 %	0.19 dB	79.2 %	0.42 dB
S_{22}	24.33 %	2.67 dB	47.54 %	6.78 dB
NF	1.33 %	0.03 dB	93.85 %	0.07 dB
IIP_3	37.79 %	0.13 dBm	92.86 %	0.33 dBm

Table 22: Alternate test prediction results using Group 4 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	2.09 %	0.27 dB	81.37 %	0.80 dB
S_{11}	3.14 %	0.25 dB	64.71 %	0.67 dB
S_{22}	6.28 %	0.69 dB	97.41 %	2.15 dB
NF	2.77 %	0.06 dB	68 %	0.16 dB
IIP_3	44.52 %	0.15 dBm	89.92 %	0.43 dBm

Table 23: Alternate test prediction results using Group 5 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.63 %	0.08 dB	98.44 %	0.29 dB
S_{11}	0.84 %	0.07 dB	97.8 %	0.17 dB
S_{22}	2.49 %	0.27 dB	99.59 %	0.99 dB
NF	1.12 %	0.02 dB	95.71 %	0.06 dB
IIP_3	15.64 %	0.05 dBm	98.81 %	0.14 dBm

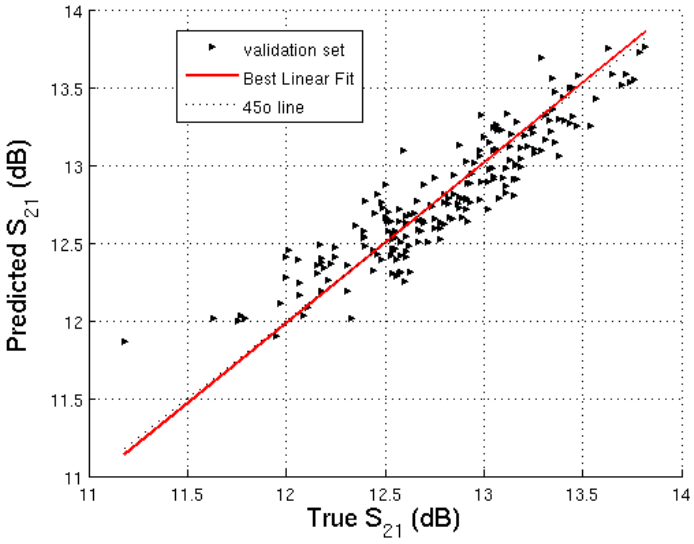
Table 24: Alternate test prediction results using Group 6 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.53 %	0.19 dB	90.5 %	0.72 dB
S_{11}	2.48 %	0.2 dB	78.55 %	0.45 dB
S_{22}	24.93 %	2.74 dB	46.98 %	6.82 dB
NF	1.31 %	0.03 dB	94.03 %	0.07 dB
IIP_3	36.37 %	0.13 dBm	93.42 %	0.29 dBm

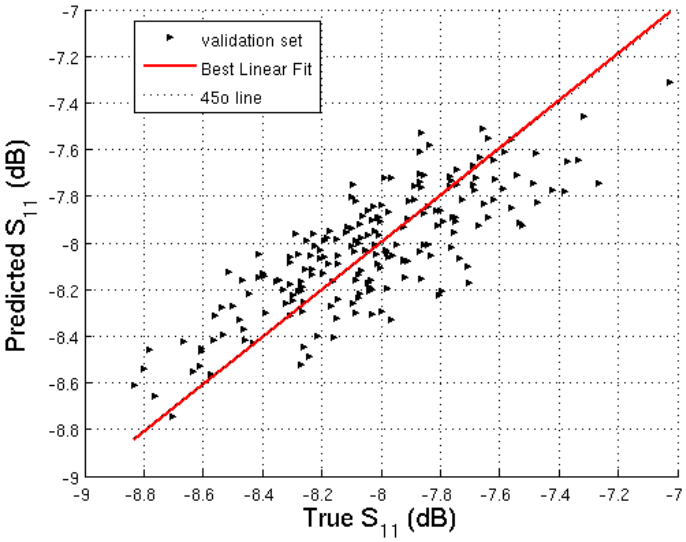
Table 25: Alternate test prediction results using Group 7 of non-intrusive variation-aware sensors considering only D2D variations. Post-layout simulations have been performed.

Comparing Table 25 with Table 15, we observe that from a statistical point of view, the prediction results are very similar. The only difference concerns the average RMS error for the IIP_3. The large value observed in Table 25 is due to the nominal value of IIP_3 being close to zero. Otherwise, we observe that the average absolute RMS error is well below the one standard deviation of 0.35dBm (see Table 19) and that the correlation is excellent above 90%.

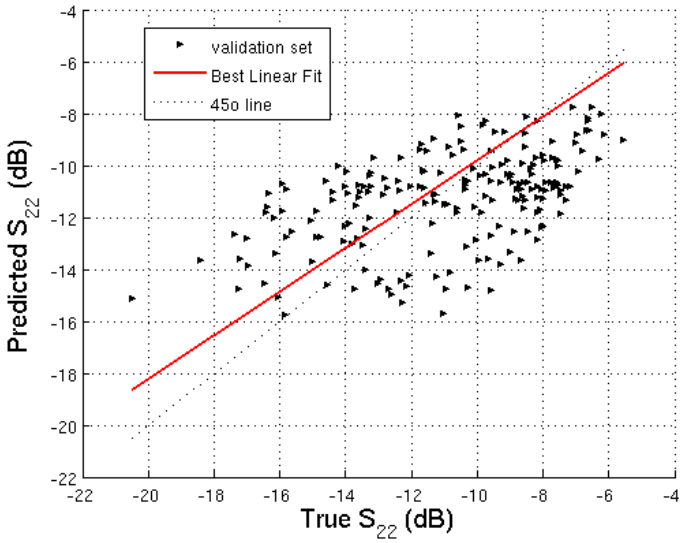
Figure 52 depicts the scatter plots of true vs. predicted performances for Group 7 of sensors considering only D2D. We recall that the prediction results in Section 6.1.1 show that WID variations affect slightly the accuracy of alternate test, which is expected also in case of the experiment that considers post-layout simulations. Therefore, since post-layout Monte Carlo simulations are time-consuming, we did not perform the experiment a second time considering WID variations in addition to D2D variations. Figure 52 offers a visualization of the prediction results in Table 25 considering only D2D variations. Each point corresponds to one circuit instance in the validation set. As it can be seen, the points are distributed across the diagonal for all performances except S_{22} for which the correlation coefficient from Table 25 is 46.98%.



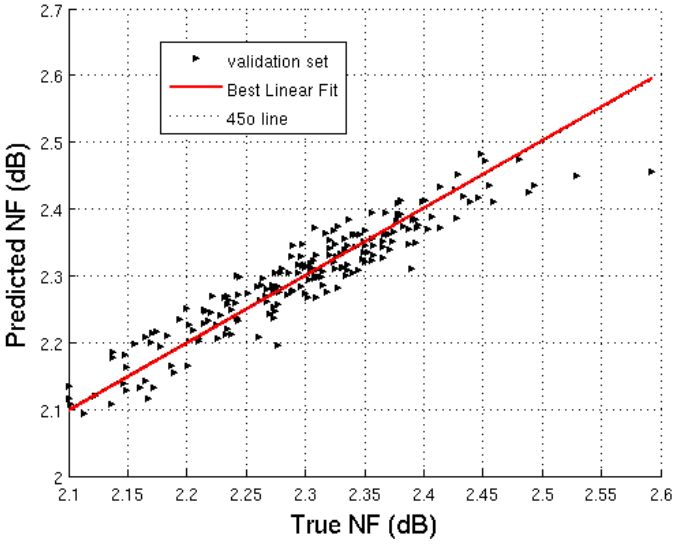
(a) S_{21}



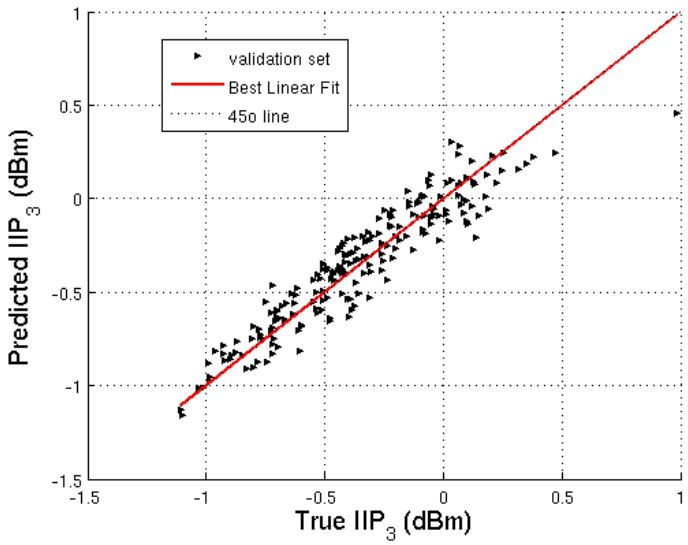
(d) S_{11}



(b) S_{22}



(e) NF



(c) IIP_3

Figure 52: Scatter plots of true vs. predicted values for all the performances of the LNA considering Group 7 of sensors. Only D2D are considered. Post-layout simulations have been performed.

6.2 60GHZ MM-WAVE 3-STAGE LNA CASE STUDY

6.2.1 *Post-layout Simulation Results*

The topology of the 60GHz mm-wave LNA is shown in Figure 27. Figure 53 shows the layout pointing to the placement of the LNA and the various non-intrusive sensors on the die. Both the LNA and the non-intrusive variation-aware sensors have been implemented and simulated in the Cadence Virtuoso environment using SpectreRF simulator using the 65nm CMOS065 bulk technology provided by ST Microelectronics. Table 26 summarizes the performances of the LNA.

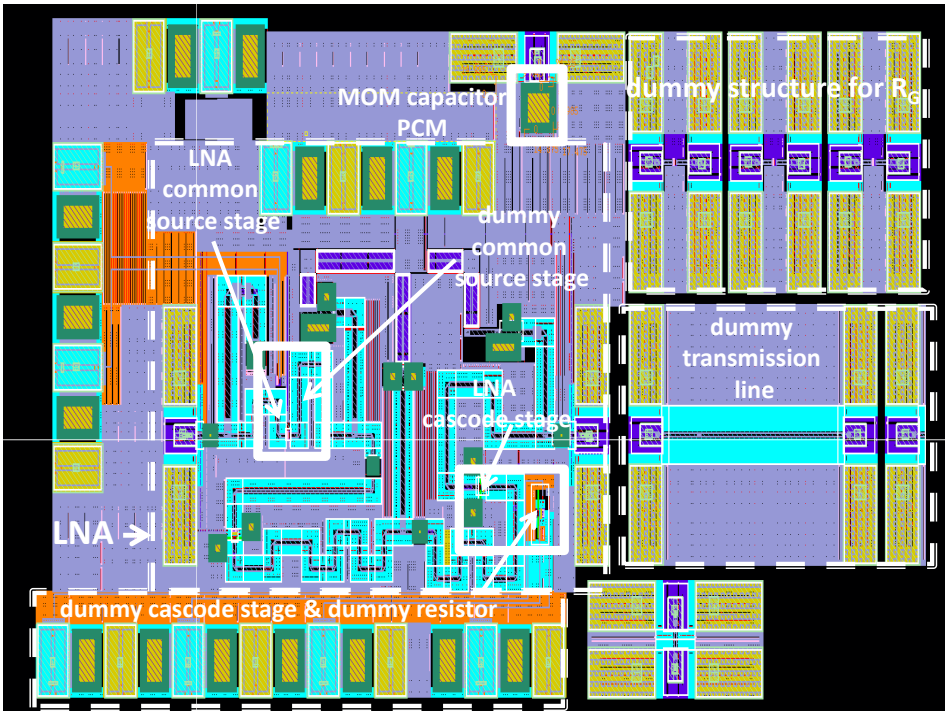


Figure 53: Layout view of the 60GHz mm-wave LNA and the embedded non-intrusive variation-aware sensors.

Summarized Results	
Frequency [GHz]	60
Supply Voltage [V]	1.2
DC Power Dissipation [mW]	16
S_{21} [dB]	16.41
S_{11} [dB]	-15.40
S_{22} [dB]	-17.98
NF [dB]	5.68
IIP_3 [dBm]	-12.24

Table 26: Summary of 60GHz LNA performances.

The schematic view of the non-intrusive variation-aware sensors is shown in Figure 40. Concerning again the layout in Figure 53, the dummy common-source stage, the dummy cascode stage, and the dummy resistor are placed in close physical proximity to corresponding structures in the LNA. Excepting their pads, they can be considered as having zero area overhead since their layout can be mingled with the layout of the LNA and placed in areas that are left void so as to respect electromagnetic design rules. The MOM capacitor PCM has a relatively small value of 4.3pF and is placed close to its GSG RF pads so as to perform a measurement that is as accurate as possible. Notice that in this design each of these dummy structures has its own dedicated pads. However, given that they are tested at DC or low-frequency, nothing would have prevented us from multiplexing

their inputs and outputs so as to use only two pads and reduce the pad-dependent area overhead. On the other hand, the dummy structure for measuring R_G and the dummy transmission line occupy a large area on the die. A large fraction of this area is devoted to the de-embedding structures that should be used so as to estimate and remove the parasitics of the GSG RF pads. As we will see later on, while R_G is very useful in predicting variations in the performances, including the dummy transmission line in the set of non-intrusive sensors does not offer a significant advantage and, thus, we may as well choose to remove it.

Table 27 lists the LNA performances, the nominal design values, and the standard deviation observed in a Monte Carlo sample set of 1000 circuit instances considering both D2D and WID variations when statistical variations in transmission lines are (a) disabled and (b) enabled. Statistical models for D2D and WID variations are included into the PDK of the technology. The library of components for this technology does not include transmission lines and, therefore, transmission lines are custom-built. The statistical model for the transmission lines is also custom-built and is rather pessimistic.

Performance	Standard Deviation (a)	Standard Deviation (b)
S_{21} (dB)	0.87	0.97
S_{11} (dB)	1.04	1.09
S_{22} (dB)	2.49	2.4
NF (dB)	0.23	0.27
IIP_3 (dBm)	0.64	0.69

Table 27: Standard deviation of 60GHz LNA performances when statistical variations in transmission lines are (a) disabled and (b) enabled.

In the following analysis, we employ different sets of non-intrusive sensors and for each set we present the results of the alternate test. The sets of sensors and their corresponding groups of alternative measurements are shown in Table 28. The regression functions in alternate test are learned using a training set of 800 Monte Carlo circuit instances and their generalization ability is evaluated on an independent validation set of 200 Monte Carlo circuit instances.

Set of Sensors	Alternative Measurements
A	$I_{DC1}, I_{DC2}, I_{DC3}$
B	$I_{DC1}, I_{DC2}, I_{DC3}, C_1$
C	$I_{DC1}, I_{DC2}, I_{DC3}, C_1, R_G$
D	R_G
E	$I_{DC1}, I_{DC2}, I_{DC3}, C_1, R_G, S_{11}$

Table 28: Sets of non-intrusive sensors and corresponding groups of alternative measurements.

Table 29 shows the prediction results when using set A of non-intrusive sensors, which includes the dummy common-source stage, the dummy cascode stage, and the dummy resistor. As it can be seen, the prediction results are not very satisfactory, but already we observe that the average absolute RMS error is smaller or comparable to one standard deviation for almost all the performances. The maximum error however, can be as high as three standard deviations and is clearly higher than the measurement and repeatability errors for some performances, such as S_{21} and S_{22}.

Table 30 shows the prediction results when using set B of non-intrusive sensors, which includes the dummy common-source stage, the dummy cascode stage, the dummy resistor, and the MOM capacitor PCM. In fact, this set contains the sensors with the minimum area overhead possible. Comparing Table 30 and Table 29, the prediction results are still not very satisfactory, but we observe that the average absolute RMS error is smaller or comparable to one standard deviation for each performance. The maximum error, however, can be as high as three standard deviations and is clearly higher than the measurement and repeatability errors for some

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	4.78 %	0.78 dB	64.1 %	2.91 dB
S_{11}	7.25 %	1.12 dB	13.53 %	3.62 dB
S_{22}	9.55 %	1.72 dB	65.6 %	5.55 dB
NF	4.75 %	0.27 dB	10.1 %	0.75 dB
IIP_3	5.09 %	0.62 dBm	36.33 %	1.79 dBm

Table 29: Alternate test prediction results using Set A of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	4.83 %	0.79 dB	64.07 %	3.08 dB
S_{11}	6.41 %	0.98 dB	47.82 %	2.41 dB
S_{22}	9.68 %	1.74 dB	64.31 %	4.16 dB
NF	4.74 %	0.27 dB	12.55 %	0.62 dB
IIP_3	5.11 %	0.63 dBm	34.55 %	1.45 dBm

Table 30: Alternate test prediction results using Set B of non-intrusive sensors.

performances, such as S_{21} and S_{22}. In case of S_{11} performance, the prediction results as well as the correlations coefficient both improve, implying that the MOM capacitor explains the variability of this performance to a, statistically, non-negligible degree.

Table 31 shows the prediction results when using set C of non-intrusive sensors, where compared to set B of non-intrusive sensors we have added the dummy structure to measure R_G. As it can be seen, the prediction results have improved drastically. This is justified by the fact that variations in R_G explain to a large degree the variations of most performances, as can be attested by Table 32, where the set D of sensors contains only the dummy structure to measure R_G. As it can be seen from Table 31, for each performance the average absolute RMS error is smaller than one standard deviation and for some performances is close to half standard deviation. The maximum error, is smaller or at least comparable to the measurement and repeatability errors for all performances except S_{22}. Overall, the predictions are deemed excellent for all performances except S_{22}.

Table 33 shows the prediction results when using set E of non-intrusive sensors, where compared to set C of non-intrusive sensors we have added the dummy transmission line. We observe that the prediction results compared to set C have improved appreciably but not dramatically, implying that variations in the transmission lines are not the main source of performance variability. This can be also justified by Table 27 where we observe that the standard deviation does not vary appreciably when statistical variations in transmission lines are disabled or enabled. We argue that it is perhaps a wiser choice not to deploy a dummy transmission line for the following reasons: (a) the prediction results for set C of non-intrusive sensors are already deemed excellent for all performances except S_{22}; (b) the dummy transmission line does not improve the error metrics for S_{22}; (c) the statistical model for the transmission line is rather pessimistic as mentioned above, implying that the prediction results for set C of non-intrusive sensors are expected to be better; and (d) the dummy transmission line occupies a large area overhead.

Regarding the unsatisfactory maximum error for S_{22}, it is clear that the main process parameters that affect S_{22} are not well covered by the set of non-intrusive sensors. We studied the process parameters on which S_{22} strongly depends and we believe that the error metrics for S_{22} can be improved if we monitor variations in the silicon substrate of transistors. It is well known that, unlike at DC and low frequencies, substrate resistance begins to contribute at high frequencies, where the signal at the drain couples to the source and bulk terminals through the source/drain junction capacitance and the substrate itself. As reported in [107], [108], the substrate

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	3.77 %	0.62 dB	80.03 %	2.36 dB
S_{11}	3.42 %	0.53 dB	88.39 %	1.37 dB
S_{22}	9.68 %	1.74 dB	63.78 %	5.19 dB
NF	2.71 %	0.15 dB	81.40 %	0.41 dB
IIP_3	3.13 %	0.38 dBm	81.09 %	1.07 dBm

Table 31: Alternate test prediction results using Set C of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	5.60 %	0.92 dB	46.65 %	4.35 dB
S_{11}	5.09 %	0.78 dB	71.82 %	2.42 dB
S_{22}	12.68 %	2.28 dB	8.25 %	7.24 dB
NF	2.69 %	0.15 dB	81.73 %	0.41 dB
IIP_3	3.54 %	0.43 dBm	75.01 %	0.93 dBm

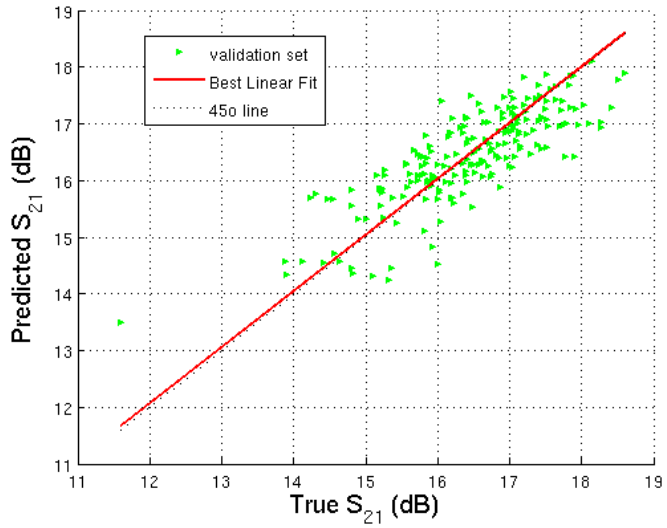
Table 32: Alternate test prediction results using Set D of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	3.16 %	0.52 dB	86.40 %	1.69 dB
S_{11}	3.14 %	0.48 dB	90.25 %	0.95 dB
S_{22}	9.40 %	1.69 dB	65.76 %	4.11 dB
NF	2.29 %	0.13 dB	87.14 %	0.38 dB
IIP_3	2.39 %	0.29 dBm	89.63 %	0.64 dBm

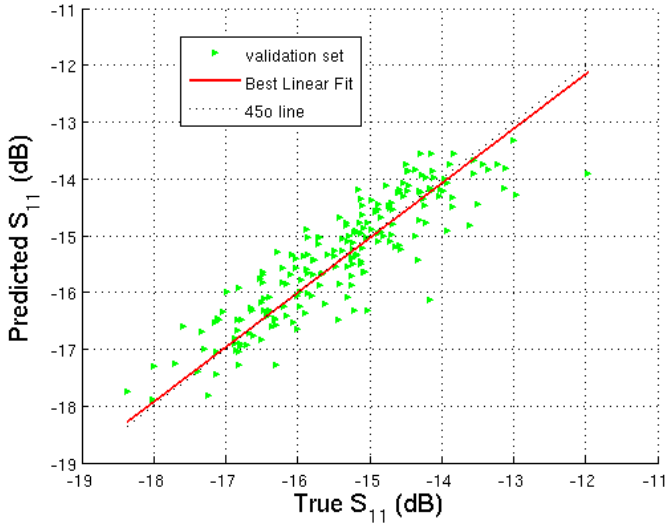
Table 33: Alternate test prediction results using Set E of non-intrusive sensors.

resistance affects mainly the output characteristics and can contribute as much as 20% or more to the total output admittance.

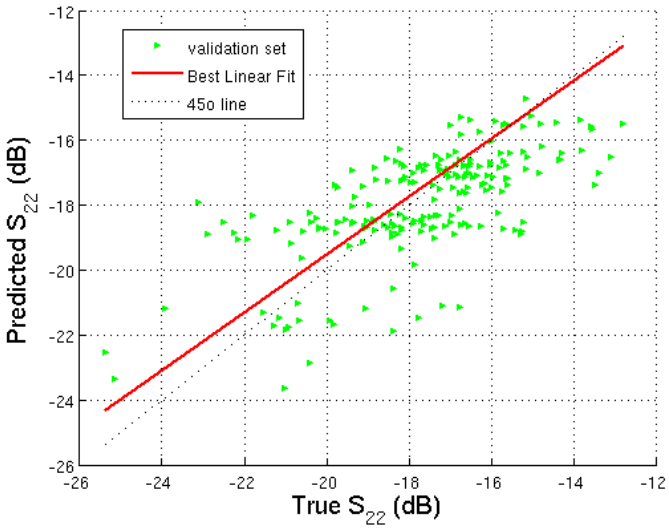
Finally, Figure 54 offers a visualization of the prediction results in Table 31. Each point corresponds to one circuit instance in the validation set. As it can be seen, the points are distributed across the diagonal for all performances. In addition, since MOS gate resistance is responsible for a large fraction of variability in the performances of the 60GHz LNA, in Figure 55 we demonstrate a visualization of the prediction results in Table 32. Again, as shown, the points are distributed across the diagonal for all performances except S_{22} for which the correlation coefficient from Table 32 is 8.25%. On the other hand, NF shows the highest correlation coefficient of 81.73%, which manifests the great impact of gate resistance on the noise performance of MOS devices and, in turn, the overall amplifier.



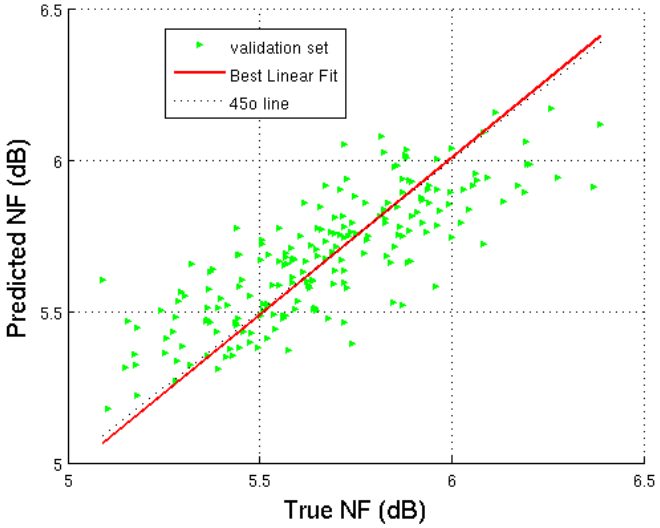
(a) S_{21}



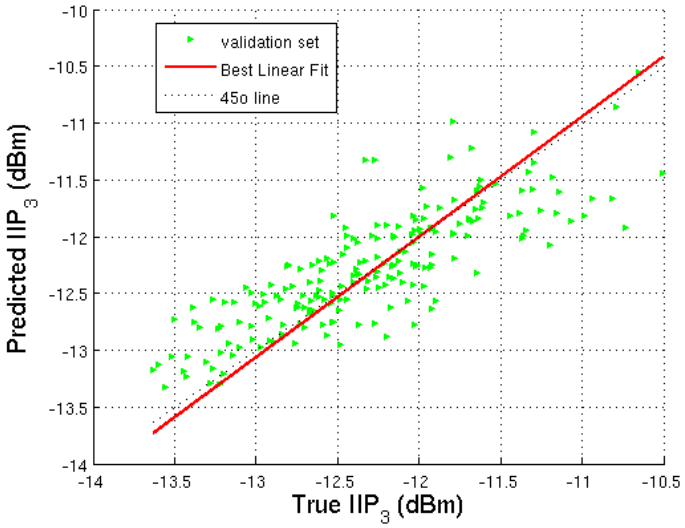
(d) S_{11}



(b) S_{22}

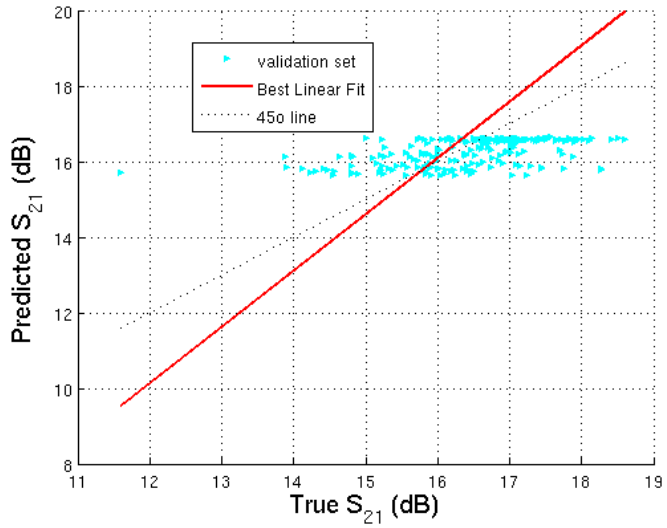


(e) NF

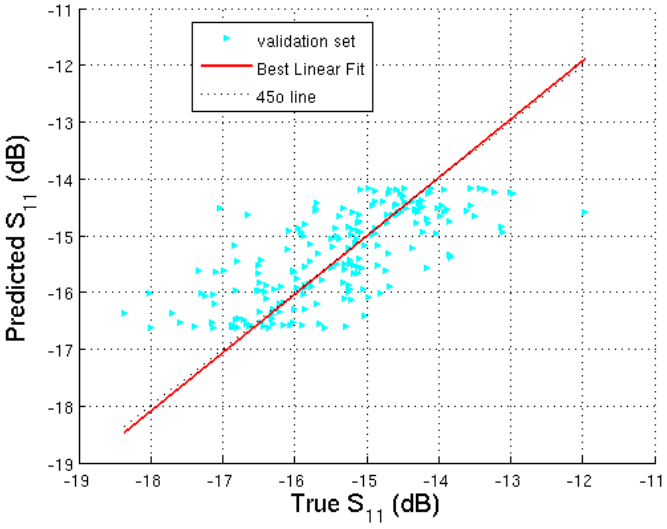


(c) IIP_3

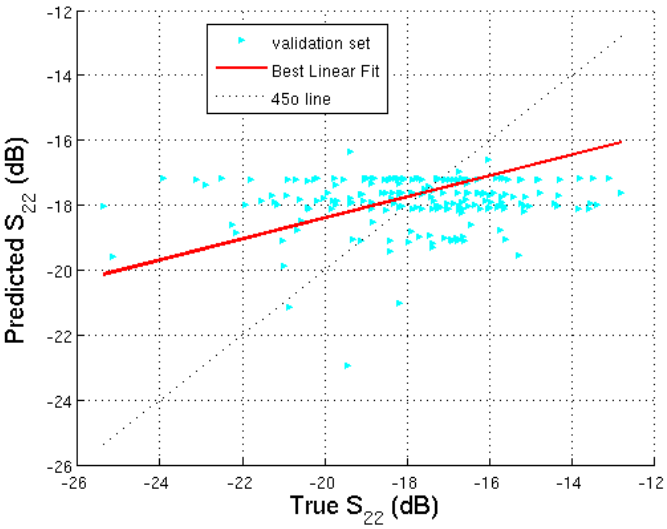
Figure 54: Scatter plots of true vs. predicted values for all the performances of the LNA considering Set C of sensors.



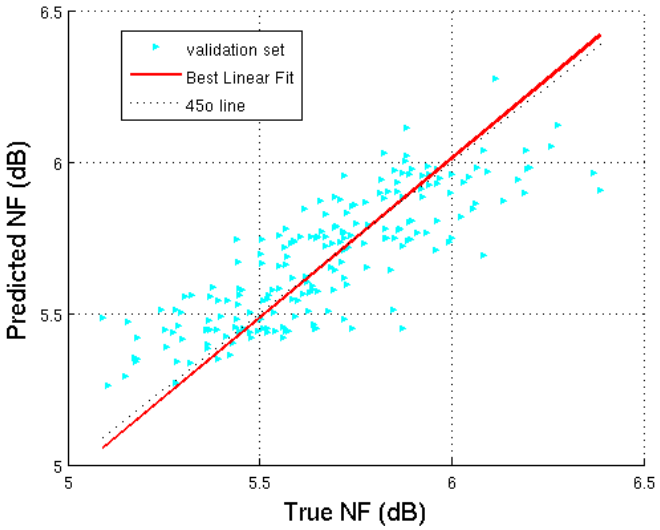
(a) S_{21}



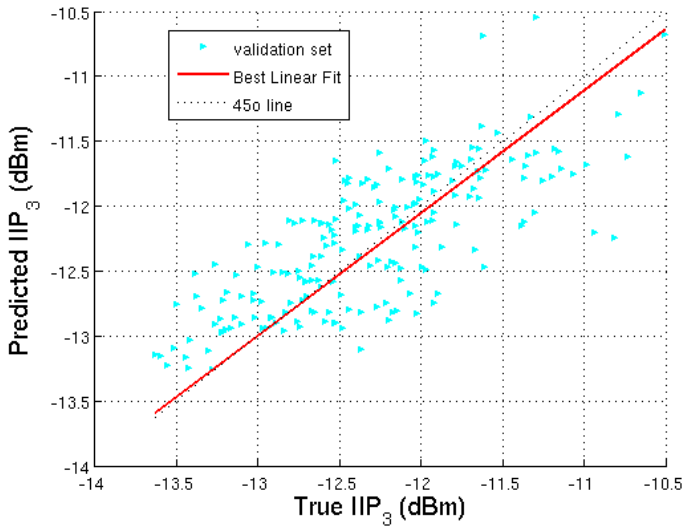
(d) S_{11}



(b) S_{22}



(e) NF



(c) IIP_3

Figure 55: Scatter plots of true vs. predicted values for all the performances of the LNA considering Set D of sensors where only gate resistance, R_G is used as alternative measurement.

6.3 CONCLUSIONS

In this chapter, we demonstrated through simulations the non-intrusive built-in test strategy based on variation-aware sensors for two case studies, namely a 65nm 2.4GHz inductive degenerated LNA and a 65nm 60GHz mm-wave 3-stage LNA.

In the former case, by laying out a low area overhead dummy analog stage and a few low area overhead PCMs in close physical proximity to the LNA, and by using DC/low-frequency measurements obtained on these sensors as alternative measurements, we were able to predict the performances of the LNA with an acceptable error based on the alternate test paradigm. We found that the variability in the inductors of the LNA explain to a large degree the variability in the performances of the LNA. We also found that the variability in the inductors stems mostly from the variability in the resistivity of the high-level metal and alucap layers that form the coil of the inductors. To this end, to circumvent using a dummy inductor PCM that is area-hungry, we demonstrated that by using instead low area overhead high-level metal and alucap resistor PCMs we are able to obtain practically the same prediction results as if a dummy inductor PCM was used. We also demonstrated that WID variations, which cannot be handled by the non-intrusive variation-aware sensors, do not affect appreciably the prediction results.

In the latter case, tapping into the signal paths of mm-wave circuits, so that to extract useful information for testing, is utterly intolerable from a design point of view. Therefore, a standard loop-back test strategy for the complete transceiver or a non-intrusive test strategy, such as the one presented for the first time in this work, for the complete transceiver or for its individual blocks, are the only candidate solutions. Compared to the loop-back test strategy, the non-intrusive test strategy has the comparative advantage that it incurs a much lower test cost at the expense of predicting the performances rather than measuring them directly. We found that gate resistance is responsible for a large fraction of the variability in the performances of the LNA and that variations in transmission lines are less important. We demonstrated that by placing simple dummy analog stages and dummy single components (PCMs) on the die and by obtaining on these dummy structures simple DC or low-frequency measurements that offer an “image" of process variations, we are able to track variations in all performances except S_{22} with an average error smaller than one standard deviation and a maximum error smaller or at least comparable to the measurement and repeatability errors in an ATE environment.

7

MEASUREMENT RESULTS

In this chapter, the non-intrusive test idea will be demonstrated for the 2.4GHz LNA case study based on real measurements. In Section 7.1, we discuss the chip fabrication. Section 7.2 describes the design of the Printed Circuit Board (PCB). Section 7.3 presents the measurement setup and Section 7.4 presents the results of the alternate test. Finally, Section 7.5 concludes the chapter.

7.1 CHIP FABRICATION

Figure 56 shows the layout view of the 65nm 2.4GHz LNA, including the embedded non-intrusive sensors and the I/O pad ring. The chip also includes a PA, which is part of another project and will not be discussed herein. The LNA and PA are electrically separated and share only the same ground line on the PCB. The I/O pad ring of the design has been provided by the I/O pad library of STMicroelectronics. The I/Os include: (a) I/O analog pads which are used to capture DC values from the embedded non-intrusive sensors that provide DC measurements, (b) GND pads that are used for the ground paths of the design, (c) RF pads which are used for measuring the RF performances of the LNA as well as the MOM capacitor and dummy inductor PCMs, and (d) power supply V_{DD} pads, which are used for supplying both the LNA and the DC non-intrusive sensors with DC power. In addition, Electro-Static Discharge (ESD) clamps are placed across power supply and ground rails, in order to protect the internal circuits against ESD damage.

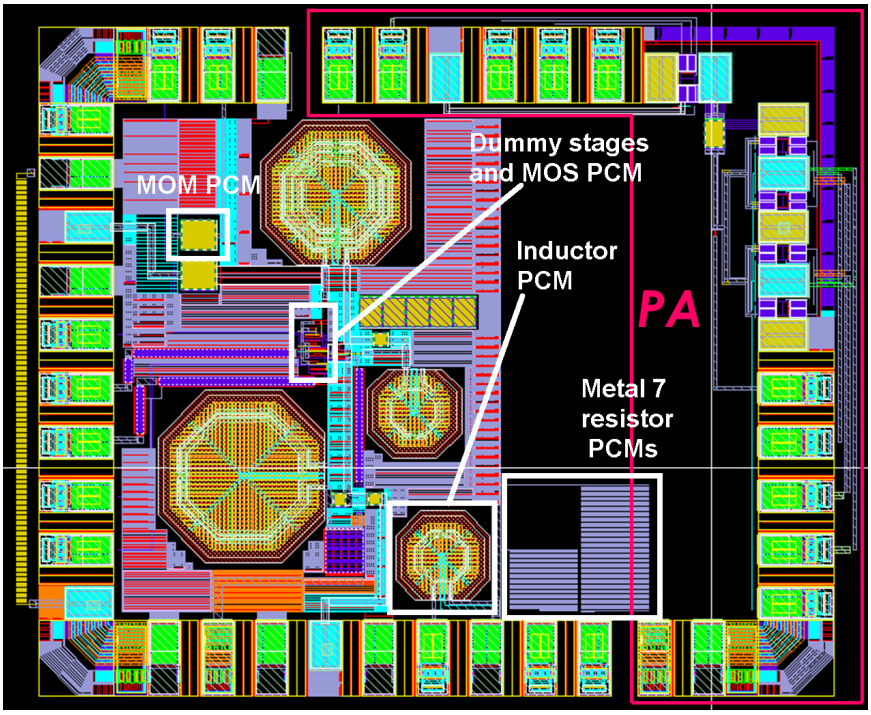


Figure 56: Layout view of the 2.4GHz LNA, including the embedded non-intrusive sensors and the I/O pad ring.

Figure 57 shows the photo of the fabricated chip. In total, we fabricated seventy-five chips coming from three wafers that were produced in three different lots. The first wafer is typical. The second wafer is a corner wafer that has slow transistors and maximum poly resistance. The third wafer is also a corner wafer that has fast transistors and maximum poly resistance. STMicroelectronics could produce corner wafers by assuming fast/slow transistors and minimum/maximum poly resistance. We chose the above combination, since based on our simulations, it resulted in the maximum variation of performances. Obtaining a sample with significant variation is important, so as to be able to perform statistical analysis and prove the alternate test concept.

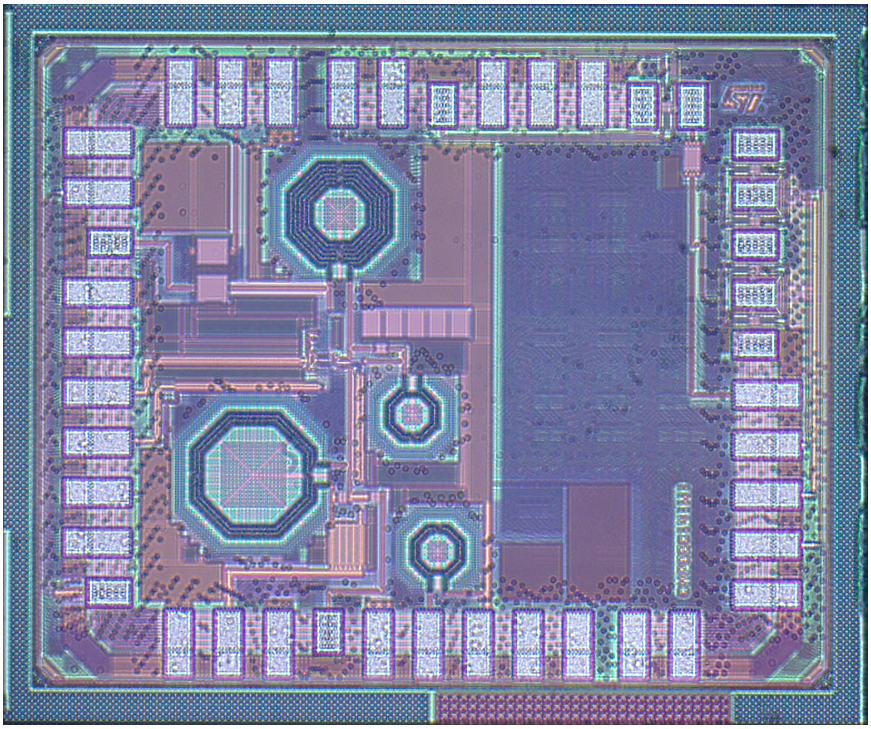
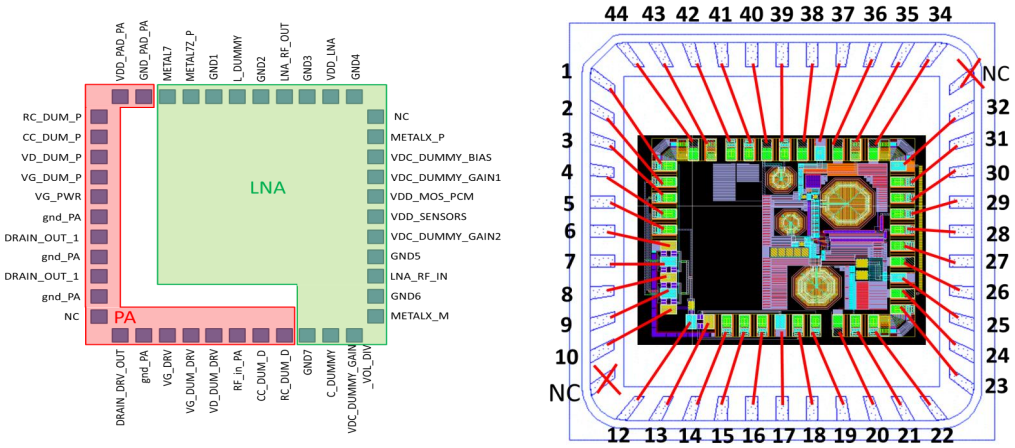


Figure 57: Photograph of the fabricated chip.

The chips were packaged using a QFN44 package. Out of the forty-four pins of the QFN44 package, only forty-one are used, and three are not connected. Figure 58 depicts the location of the various pins of the QFN44 package as well as the plan of the wire bonding between the layout of the RF chip and the QFN44 package.

Table 34 lists the number, name and type of pads/pins which have been used in this work. LNA_RF_IN and LNA_RF_OUT correspond to the I/O RF pads that are used for measuring the RF performances of the LNA. C_DUMMY and L_DUMMY correspond to the pads that are used for measuring the MOM capacitor PCM and dummy inductor PCMs, respectively. VDC_DUMMY_GAIN1 (V_{DC2} in Figure 39), VDC_DUMMY_GAIN2 (V_{DC1} in Figure 39), VDC_DUMMY_GAIN_VOL_DIV (V_{DC3} in Figure 39), VDC_DUMMY_BIAS (V_{DC4} in Figure 39), METALX_P, META7L_P, and METAL7 correspond to the I/O analog pads that are used for capturing the DC measurements from the dummy current mirror, dummy voltage divider, dummy bias stage, and metal 5 and 7 resistor PCMs, respectively. Two different sizes of metal 7 resistor PCMs have been laid out. META7L and META7L_P pads correspond to these two metal 7 resistor PCMs, which have, respectively, an area equal and half of that of the inductor PCM. Finally, VDD_SENSORS, VDD_MOS_PCM, and VDD_LNA correspond to the V_{DD} pads that are used for supplying the non-intrusive sensors and the LNA with DC power.



(a) (b)

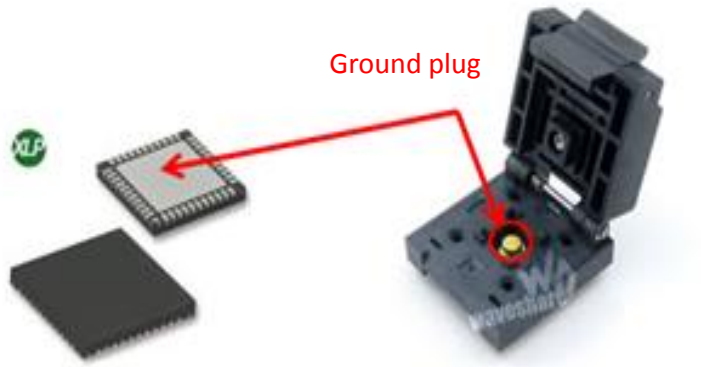
Figure 58: (a) Plan of the different pins on the complete layout of the RF chip (the PA is also included), (b) Plan of the wire bonding between the layout and the QFN44 package.

Pin number	Pin name	Pin type
20	GND7	GND
21	C_DUMMY	I/O Analog
22	VDC_DUMMY_GAIN_VOL_DIV	I/O Analog
23	METALX_M (No Connection)	*****
24	GND6	GND
25	LNA_RF_IN	RF
26	GND5	GND
27	VDC_DUMMY_GAIN2	I/O Analog
28	VDD_SENSORS	VDD
29	VDD_MOS_PCM	VDD
30	VDC_DUMMY_GAIN1	I/O Analog
31	VDC_DUMMY_BIAS	I/O Analog
32	METALX_P	I/O Analog
33	No Connection	*****
34	GND4	GND
35	VDD_LNA	VDD
36	GND3	GND
37	LNA_RF_OUT	RF
38	GND2	GND
39	L_DUMMY	RF
40	GND1	GND
41	META7L_P	I/O Analog
42	METAL7	I/O Analog

Table 34: List of pin/pad types which are used in the 2.4GHz LNA case study.

7.2 PRINTED CIRCUIT BOARD (PCB) DESIGN

A PCB has been designed for characterizing and testing the RF chips. The PCB includes a socket for mounting and unmounting the chips. Figure 59 depicts the socket which has been used, in order to test the population of the fabricated chips. The choice of the socket was left free, but two criteria had to be fulfilled: (a) it is compatible to RF characterization and testing, (b) it has a ground plug which is connected to the bottom of the QFN44 package, and (c) it introduces as low interference as possible.



General view of the QFN44 package

General view of the socket

Figure 59: General view of the QFN44 package (left) and the socket (right).

In addition, the fabricated PCB includes three different types of connectors which are necessary for the interface between the fabricated chips and the test instrumentation. These types of connectors have been chosen depending on the type of the measured signal and they include:

- Sub-Miniature version A (SMA) connectors for measuring RF signals. Transmission lines, matched at an impedance of 50Ω, have been appropriately designed on the PCB paths that connect the QFN44 package and the SMA connectors.

- 4mm banana connectors for power, GND and I/O analog supply rails.

- Test probes Single-In-Line (SIL) connectors for measuring analog signals.

The input and output transmission lines on the PCB, including the I/O of the LNA, the MOM capacitor and the inductor PCM, have been appropriately sized to have an impedance of 50Ω. The general block diagram of the fabricated PCB is shown in Figure 60.

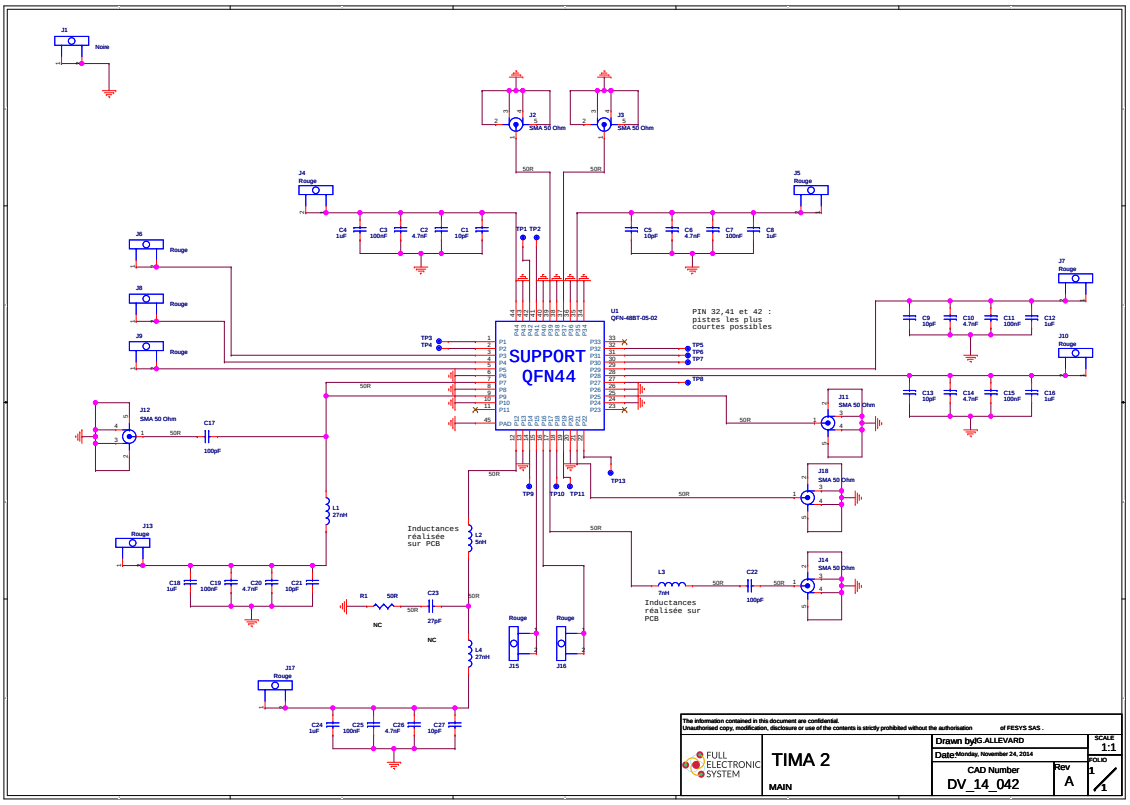


Figure 60: General block diagram of the fabricated PCB.

7.3 MEASUREMENT SETUP

Out of the seventy-five chips, sixty-seven were characterized using benchtop equipment. The rest of them were not measured, since some of them were destroyed due to handling and some others were not operating properly due to possible fabrication or packaging faults. The S-parameters of the LNA were measured using a VNA from Agilent Technologies (ENA series, E5061B), where the input power had been set at -30dBm, in order to avoid any compression effect at the output of the LNA circuit. The 1-dB CP performance was measured using the E5061B VNA. The value of the 1-dB CP is the input power, where the gain S_{21} is 1 dB lower than the gain at low power. The input power sweeps from -40dBm to 10dBm and S_{21} had been plotted on the screen of the VNA. Furthermore, NF has been measured using a spectrum analyzer from Rohde&Schwarz (Rohde&Schwarz FSL). A noise source was connected between the input of the LNA and the supply of the spectrum analyzer. Finally, a low-noise pre-amplifier was connected between the output of the LNA and the port of the spectrum analyzer. The pre-amplifier is used for improving spectrum analyzer's sensitivity and achieving higher measurement accuracy. The outputs of the sensors were measured using DC multimeters from Agilent Technologies (34405A) and DC supplies from ISO-TECH (IPS 3303D). The measurement procedure was automated through LabView software using Universal Serial Bus (USB) interface between a general-purpose PC and the measurement equipment. Measuring the performances of the LNA requires three different test configurations, including the S-parameters, NF, and 1-dB CP, while the sensors' outputs are extracted using three single test configurations, including the dummy structures that provide DC measurements, the MOM capacitor, and the dummy inductor, which both provide low frequency measurements. Figures 61, 62, and 63 show the DC and RF measurement setups.

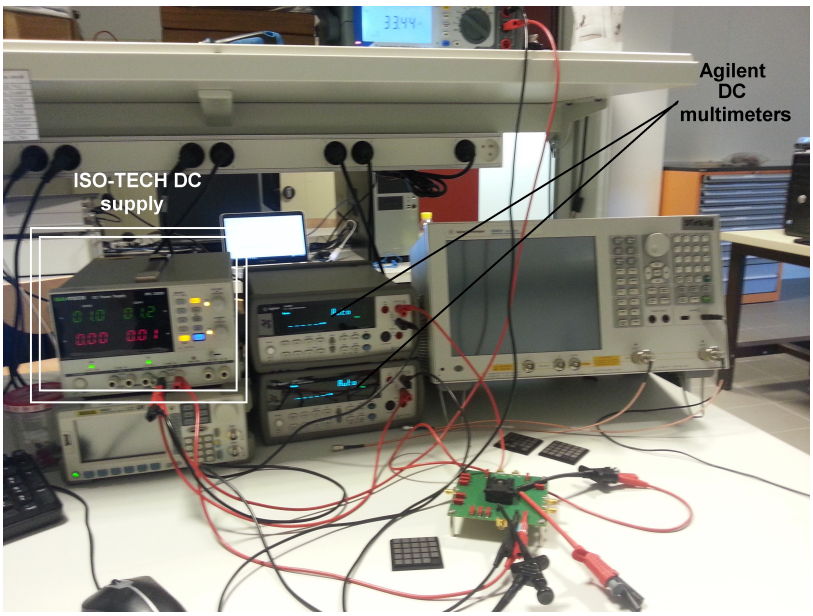


Figure 61: Photograph of the DC setup measuring the DC non-intrusive sensors.

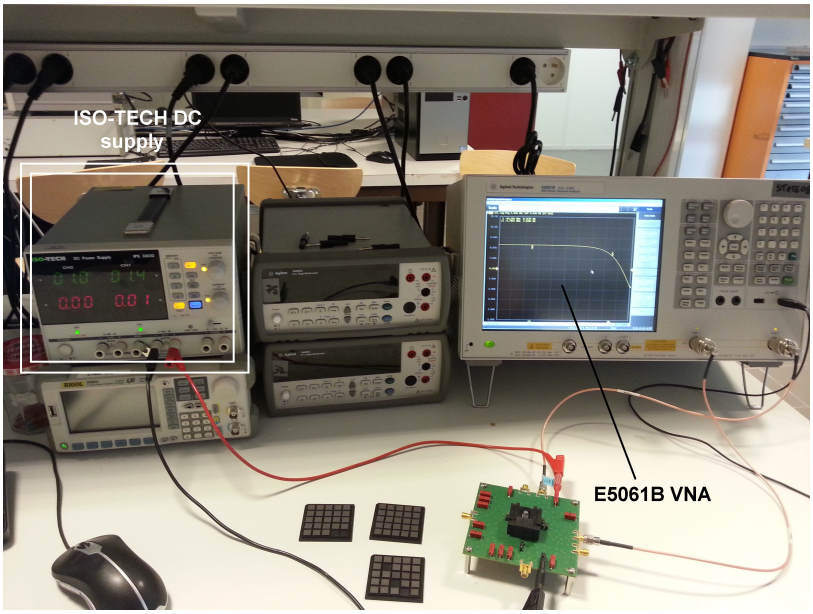


Figure 62: Photograph of the RF setup measuring the S-parameters of the 2.4GHz LNA.

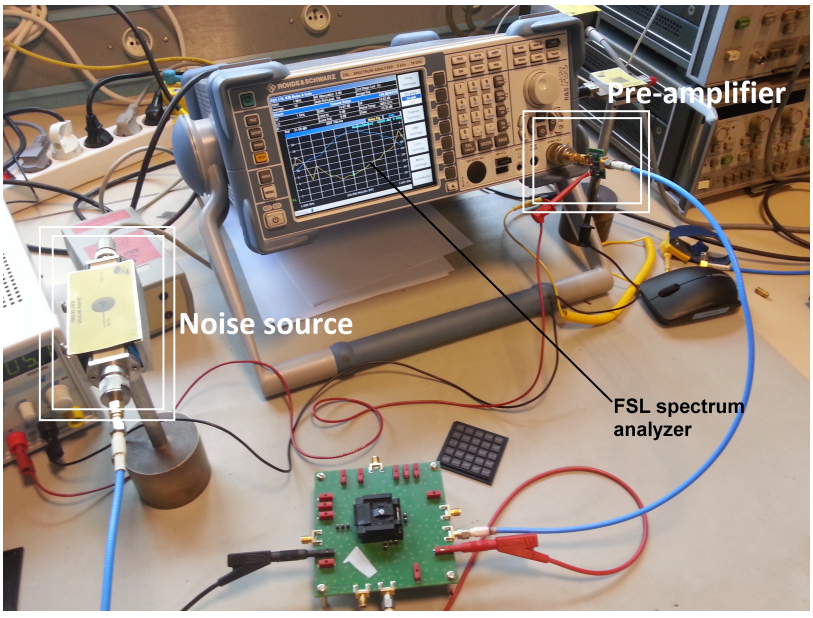


Figure 63: Photograph of the RF setup measuring the NF of the 2.4GHz LNA.

7.4 RESULTS

Figure 64 shows the measured RF performances of a chip from the typical wafer, including the input matching, given by S_{11}, voltage gain, given by S_{21}, output matching, given by S_{22}, and NF. As it can be seen, there is a frequency shift of about 300MHz away from the 2.4GHz operating frequency due to parasitics introduced by the LNA design, the wire bonding and the transmission lines of the PCB.

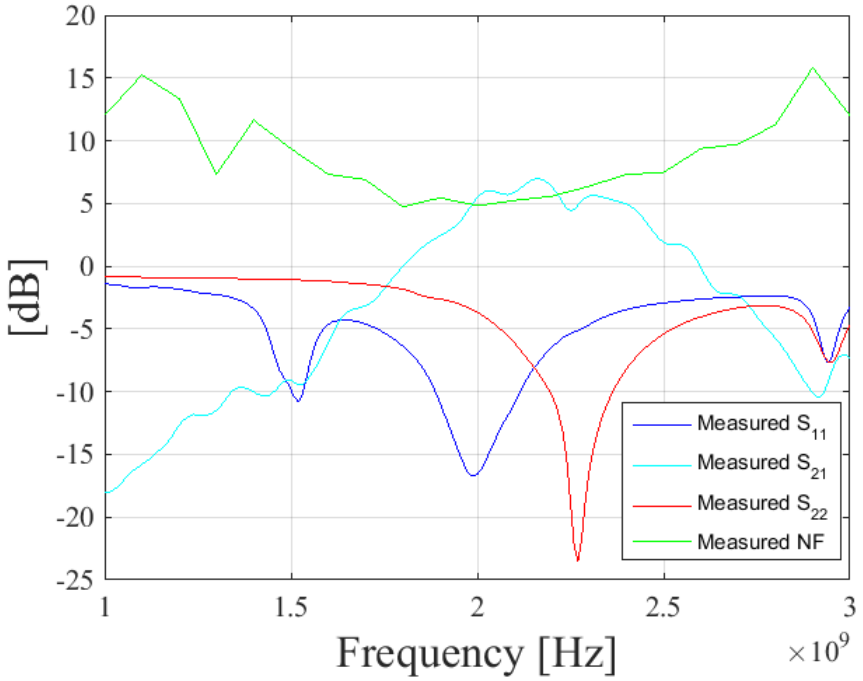


Figure 64: Measured S-parameters of a 2.4GHz LNA chip from the typical wafer.

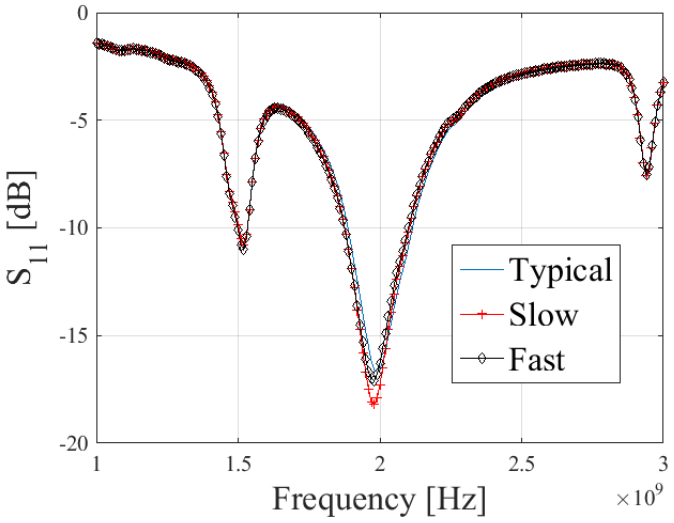
Table 35 summarizes the values of the performances of the chip from the typical wafer. The S-parameters and NF have been extracted in different frequencies, in particular at frequencies where resonance occurs. Specifically, S_{21} has been extracted at 2.16GHz, S_{11} has been extracted at 1.98GHz, S_{22} has been extracted at 2.26GHz, and NF has been measured at 2GHz. Finally, 1-dB CP was measured at 2.4GHz by sweeping the input power from -40dBm to 10dBm and plotting the S_{21}. In the case of S-parameters, the input power had been set to -30dBm, in order to avoid any compression.

Summarized Results	
Frequency range [GHz]	1.98 - 2.4
Supply Voltage [V]	1.2
DC Power Dissipation [mW]	8
S_{21} at 2.16GHz [dB]	7.22
S_{11} [dB] at 1.98GHz	-16.7
S_{22} [dB] at 2.26GHz	-22.6
NF [dB] at 2GHz	4.47
1-dB CP [dBm] at 2.4GHz	-4.58

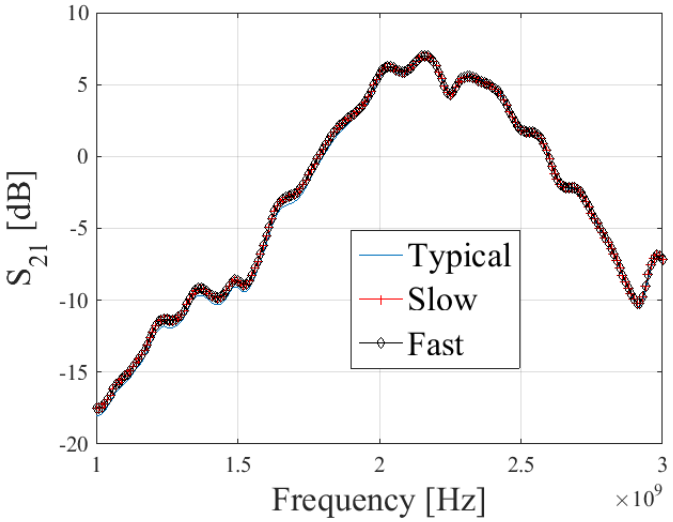
Table 35: Summary of 2.4GHz LNA chip from the typical wafer.

Figure 65 depicts the comparison of S-parameters, 1-dB CP, and NF plots between slow, typical, and fast wafers.

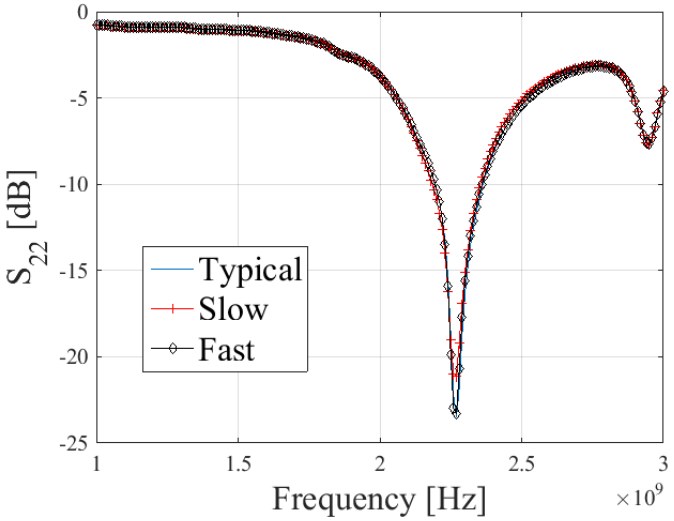
Since the sample of sixty-seven chips is small, and therefore, probably, unrepresentative of the fabrication process, specific statistical methods have been used, so as to present reliable and trustworthy prediction results. As demonstrated in Chapter 6, a simple approach that was used during the simulation step was to estimate the



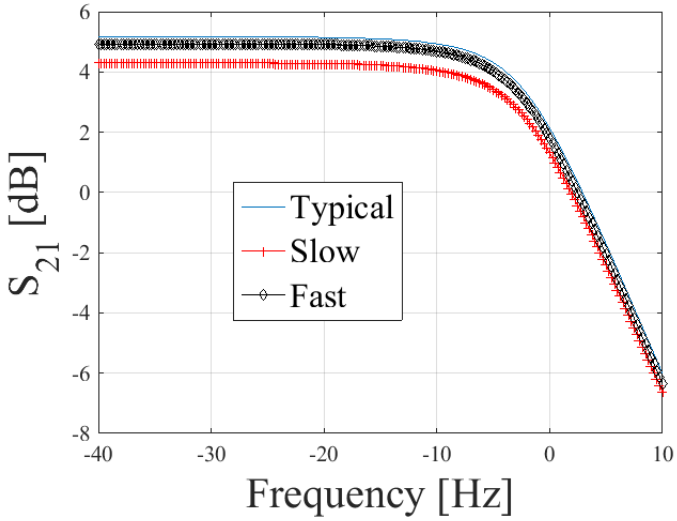
(a) S_{11} from slow, typical, and fast wafers



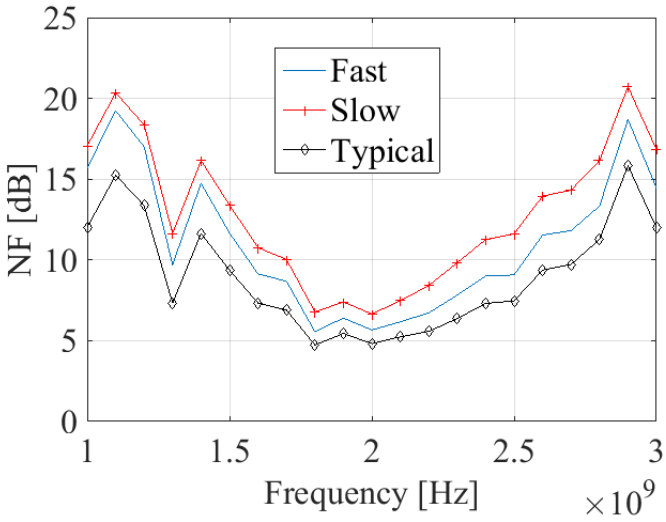
(b) S_{21} from slow, typical, and fast wafers



(c) S_{22} from slow, typical, and fast wafers



(d)1-dB CP from slow, typical, and fast wafers



(d) NF from slow, typical, and fast wafers

Figure 65: Comparison of S-parameters, 1-dB CP, and NF of the 2.4GHz LNA case study between slow, typical, and fast wafers.

prediction error by splitting the data into two groups, e.g. 80% for the training phase and 30% for the test phase. This approach is based on the assumption that the statistical sample is large and representative of the fabrication process. However, in case where the size of the sample is limited, the specific data decomposition for both the training and test phases begins to impact the estimation of the prediction error, which means that the prediction error will show an appreciable variance. In order to report the actual degree of correlation between performances and alternative measurements, we used *cross-validation*. Cross-validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set and is mainly used in settings where the goal is prediction, especially in case of statistically non-representative populations. Cross-validation techniques include two cases, namely k-fold cross-validation and Leave-One-Out Cross Validation (LOOCV):

- k-fold cross-validation is employed by dividing the statistical sample set in k groups (or folds), where k-1 groups are used in the training set and the remaining group is used in the validation set. We then calculate the error on the validation set and we repeat k times, such that each fold is used exactly once as the validation set. The average error of the k errors is finally reported. We also report the maximum error observed.
- LOOCV method is a special case of k-fold cross-validation, where $k = n$ and n is the size of the statistical sample (e.g. the number of chips). The idea is that all chips are used both for training and validation and thus we do not bias the reported errors by a random split of chips into training and validation sets.

The procedure is depicted in Figure 66. We randomly shuffle the chips and we perform k-fold cross validation. For each step of the cross validation, regression learning is repeated Y times. We repeat for Z random shuffles, and we finally report the average error and the maximum error.

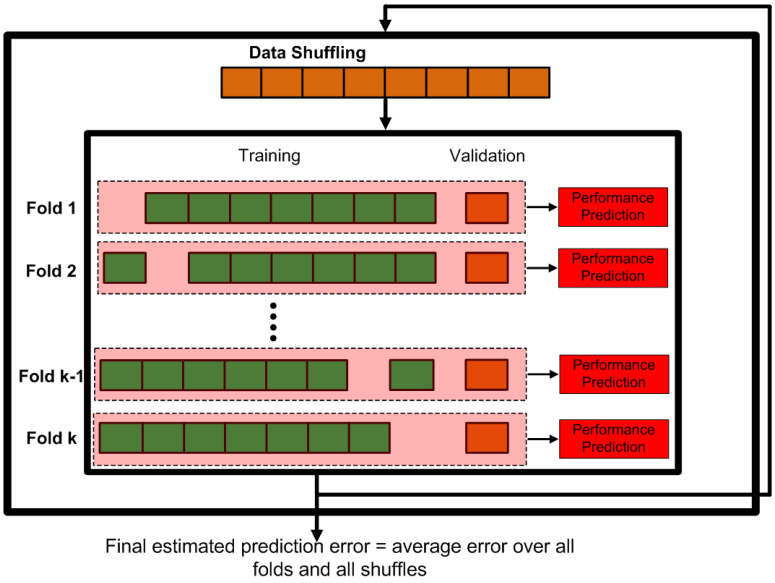


Figure 66: Averaging for reporting trustworthy prediction results.

The prediction error of alternate test is expressed in terms of several error metrics, including: (a) the absolute average RMS error in units of each performance, (b) the maximum error in units of each performance, and (c) the Pearson correlation coefficient between the simulated (e.g. “true") and predicted performances expressed in %. Furthermore, a figure of merit is introduced, namely prediction figure-of-merit, which is defined as the ratio of the absolute average RMS error over the standard deviation of each performance. To speak of acceptable prediction results, the prediction figure-of-merit should be smaller than 1 and the maximum error is recommended to be smaller than the measurement and repeatability errors in an ATE environment.

The schematic and layout view of the non-intrusive variation-aware sensors is depicted in Figure 39. The set of sensors include: (a) a dummy current mirror, (b) a dummy voltage divider, (c) a dummy bias stage, (d) a diode-connected MOS PCM, (e) a MOM capacitor PCM, (f) an inductor PCM, and (g) high-level metal resistor

PCMs. The set of alternative measurements include: (a) DC voltages V_{DC1}, V_{DC2} from the dummy current mirror, (b) DC voltage V_{DC3} from the dummy voltage divider, (c) DC voltage V_{DC4} from the dummy bias stage, (d) DC current I_{DC3} from the diode-connected MOS PCM, (e) capacitance from the MOM capacitor PCM, denoted by C_{PCM}, (f) inductance from the inductor PCM (e.g. imaginary part L_s of inductor impedance), and (g) DC currents, $I_{resistorPCM_1}$, $I_{resistorPCM_2}$, $I_{resistorPCM_3}$, that flow through the high-level metal resistor PCMs (these DC currents are directly dependent on the resistivity of the metal layers). Indexes 1,2,3 correspond to metal 7 resistor PCM with area equal to that of the inductor's PCM area, metal 7 resistor PCM with area half of that of inductor's PCM area and metal 5 resistor PCM, respectively. The capacitance and the inductance are extracted by S-parameters at a low frequency in the order of a few tens of MHz. In the following alternate test experiments, we employ different groups of non-intrusive variation-aware sensors and their corresponding measurements, as shown in Table 36.

	Dummy current mirror, dummy voltage divider, and dummy bias stage	MOS PCM	MOM PCM	Inductor PCM	Resistor PCMs
	$V_{DC1}, V_{DC2}, V_{DC3}, V_{DC4}$	I_{DC3}	C_{PCM}	L_s	$I_{resistorPCM_X}$
Group A	✓	✓			
Group B	✓	✓	✓		
Group C	✓	✓	✓	✓	
Group D	✓	✓	✓	✓	✓

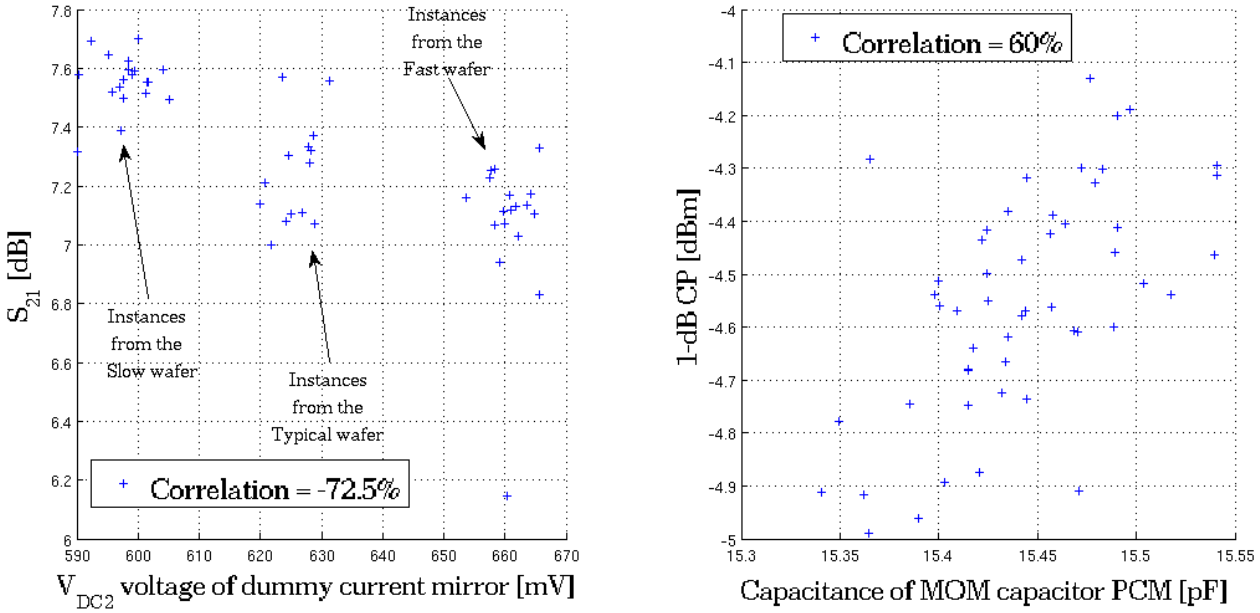
Table 36: Groups of non-intrusive variation-aware sensors and the corresponding groups of measurements used in the analysis. X corresponds to indexes 1,2, and 3, as described in the text.

Table 37 lists both the LNA performances and sensors' measurements along with their nominal value and standard deviation as well as the measurement repeatability. As it can be seen, the repeatability variance is much lower than the standard deviation of both the LNA performances and sensors' measurements. The success of alternate test depends on the fact that the standard deviation in both the LNA performances and sensors' measurements is considerably higher than the variance of the repeatability error in a benchtop or ATE environment.

LNA Performances/ Sensors' Measurements	Nominal Value	Standard Deviation	Measurement repeatability
S_{21} (dB)	7.22	0.32	0.008
S_{11} (dB)	-16.7	0.51	0.01
S_{22} (dB)	-22.6	1.06	0.007
NF (dB)	4.47	0.13	0.01
1-dB CP (dBm)	-4.58	0.21	0.005
V_{DC1} (mV)	520.72	21.59	0.13
V_{DC2} (mV)	627.77	27.25	0.03
V_{DC3} (mV)	41.33	7.42	0.04
V_{DC4} (mV)	434.92	27.28	0.02
I_{DC3} (mA)	13.75	1.99	0.003
C_{PCM} (pF)	15.44	0.05	0.002
L_s (nH)	25.79	0.11	0.01
$I_{resistorPCM_1}$ (mA)	10.67	0.07	0.001
$I_{resistorPCM_2}$ (mA)	5.69	0.06	0.001
$I_{resistorPCM_3}$ (mA)	2.35	0.02	0.0002

Table 37: List of nominal value and standard deviation of the measured 2.4GHz LNA performances and sensors' measurements. Measurement repeatability is also shown.

We first show one-to-one correlations between the performances of the LNA and sensors' measurements. Figure 67(a) depicts the correlation between S_{21} and V_{DC2} voltage, which is captured by the dummy current mirror. This correlation is expected since the gain of the LNA depends on the DC operating point of transistor M_1. Figure 67(b) shows the correlation between 1-dB CP and the capacitance of MOM capacitor PCM. This correlation is expected since 1-dB CP depends on the quality factor of the input network of the LNA. Furthermore, as depicted in Figure 67(a), we can see the clustering of the fabricated instances into three different groups, which correspond to the typical wafer and the two corner wafers.



(a) (b)

Figure 67: One-to-one correlation between (a) S_{21} and V_{DC2} and (b) 1-dB CP and capacitance of MOM capacitor PCM.

We built our regression model using a feed-forward neural network with four layers. The network has one layer at the input, one at the output, one hidden layer that employs a sigmoid activation function with a specific number of neurons, and one layer that performs a linear activation function, as depicted in Figure 68. The input of the neural network consists of a specific number of alternative measurements, which are selected based on Table 36. The optimal number of hidden layers was selected after a "trial & error" approach.

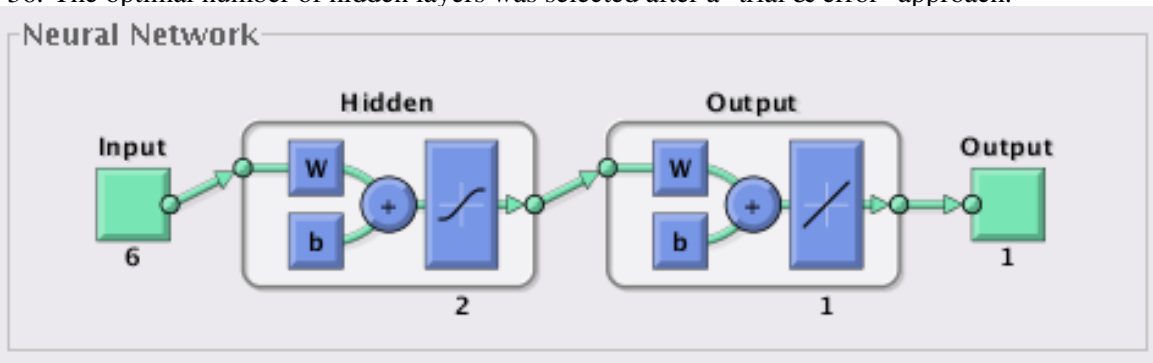


Figure 68: A simple graphical representation of the neural network used for building our regression model. The network includes two layers of adaptive weights.

Table 38 shows the prediction results using Group A of sensors. As it can be seen, for all the performances apart from NF, the prediction figure-of-merit is lower than unity, indicating that the absolute average RMS error is lower than one standard deviation. Moreover the maximum error is lower or at least comparable to the measurement and repeatability errors in a conventional ATE environment. Also, the correlation coefficient

is higher than 45% for all the performances, apart from NF. Results for NF are not as good, but they are still deemed satisfactory. Possible reasons are: (a) there is no significant variation in NF across the samples and, thereby, we do not expect to see a correlation, (b) there are sources of variations affecting the NF that are not captured by the sensors of Group A, (c) the NF measurement is not as accurate, thus observing the correlations, and (d) the limited size of the sample set does not provide enough information for making the mappings between the alternative measurements and the NF performance.

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.21	0.77	66.14 %	1.03 dB
S_{11}	0.39	0.83	63.72 %	1.34 dB
S_{22}	0.67	0.69	72.89 %	2.10 dB
NF	0.09	1.13	25.37 %	0.57 dB
1-dB CP	0.19	0.93	45.03 %	0.53 dBm

Table 38: Alternate test prediction results using Group A of non-intrusive variation-aware sensors.

Table 39 shows the prediction results using Group B of sensors, where compared to Group A we have added the MOM PCM. Comparing Table 39 to Table 38, the predictions results as well as the correlation coefficient are not improved except in case of 1-dB CP performance. By monitoring variations in the MOM capacitors of the LNA, 1-dB CP shows a correlation as high as 62.2%, which means that variations in the capacitors C_{IN} and C_{EXT} play a appreciable role in the variability of this performance. This is also justified by the theoretical analysis that is demonstrated in Section 3.3.4.

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.24	0.87	59.96 %	1.15 dB
S_{11}	0.45	0.94	57.34 %	1.91 dB
S_{22}	0.64	0.67	74.71 %	1.99 dB
NF	0.12	1.43	14.46 %	0.72 dB
1-dB CP	0.17	0.80	62.2 %	0.63 dBm

Table 39: Alternate test prediction results using Group B of non-intrusive variation-aware sensors.

Table 40 shows the prediction results using Group C of sensors, where compared to Group B we have added the inductor PCM. As it can be seen, measurement L_s does not offer any improved prediction results, implying that variations in inductance are not responsible for the variability of the performances of the LNA.

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.28	0.99	53.39 %	1.30 dB
S_{11}	0.46	0.97	57.82 %	2.06 dB
S_{22}	0.61	0.63	77.93 %	1.8247 dB
NF	0.12	1.33	12.14 %	0.65 dB
1-dB CP	0.18	0.85	59.46 %	0.66 dBm

Table 40: Alternate test prediction results using Group C of non-intrusive variation-aware sensors.

Table 41 shows the prediction results using Group D of sensors, where compared to Group C we have added the resistor PCMs, which include the resistances of two metal 7 layers of different sizes and one metal 5 layer.

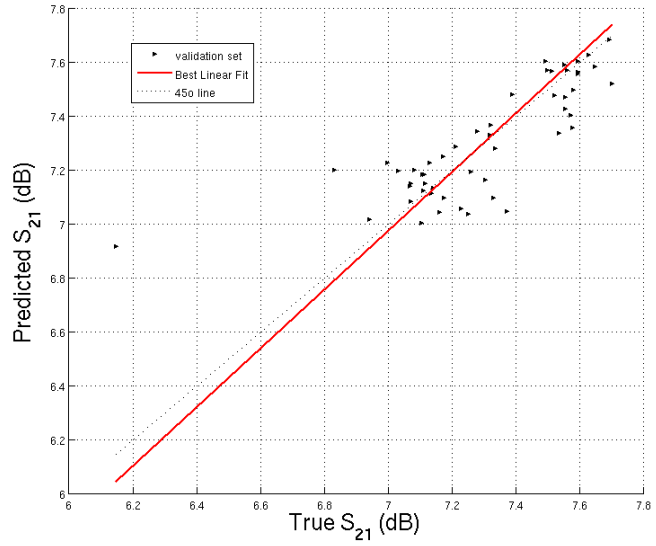
As it can be seen, the resistor PCMs do not improve the prediction results, which means that the variability of the resistivity of the metal layers, and thereby the resistivity of the inductor's coils, do not explain the variability of the performances of the LNA.

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.27	0.97	55.23 %	1.26 dB
S_{11}	0.45	0.95	53.97 %	1.79 dB
S_{22}	0.68	0.71	73.16 %	2.15 dB
NF	0.10	1.15	17.4 %	0.54 dB
1-dB CP	0.19	0.88	55.59 %	0.64 dBm

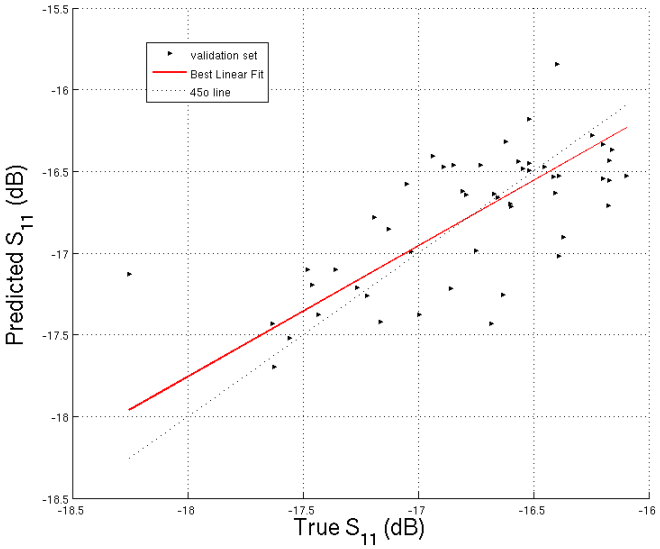
Table 41: Alternate test prediction results using Group D of non-intrusive variation-aware sensors.

In contrast to our observations based on the simulation campaign, we can explain the fact that variability in the inductors does not explain the variability in the performances by the fact that the corner wafers have been produced by varying the speed of the transistor (e.g. fast or slow) and the poly resistance.

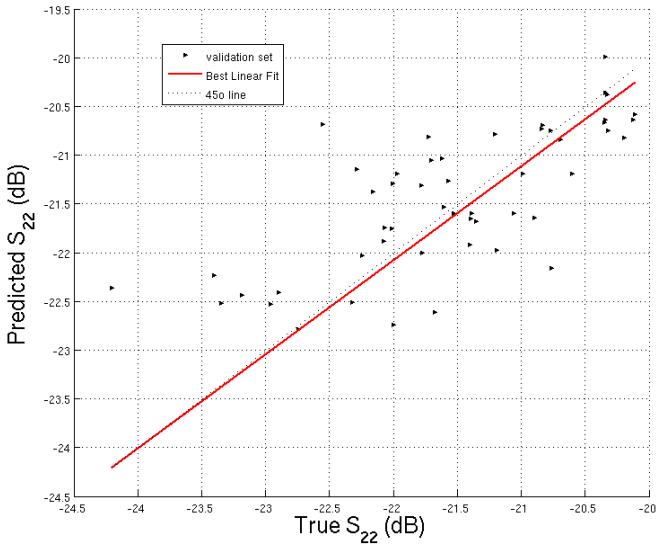
Figure 69 offers a visualization of the prediction results in Table 39. Each point corresponds to one circuit instance in the validation set. Based on the observations on the tables above, one can simply say that Group A could be chosen for predicting performances S_{21}, S_{11}, S_{22}, and NF and Group B for 1-dB CP.



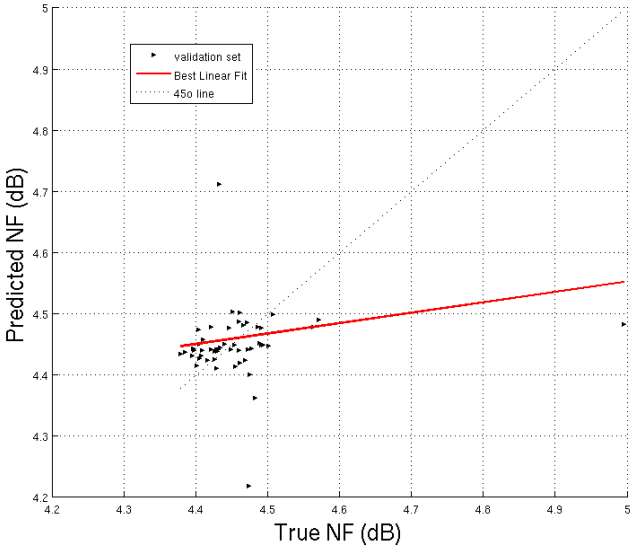
(a) S_{21}



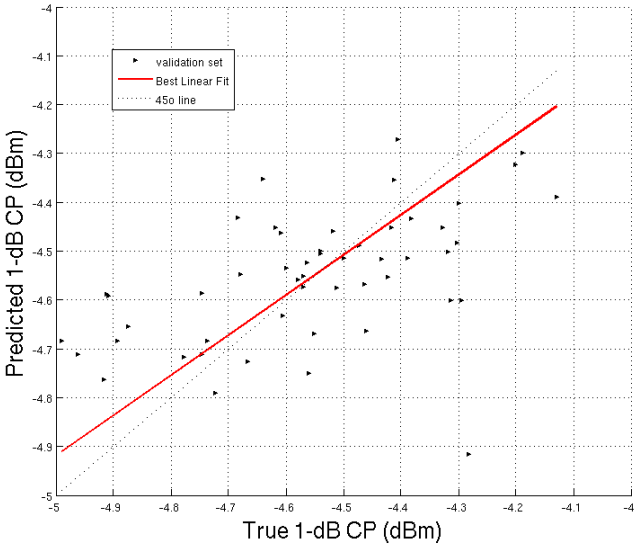
(d) S_{11}



(b) S_{22}



(e) NF



(c) 1-dB CP

Figure 69: Scatter plots of true vs. predicted values for all the measured performances of the LNA considering Group B of sensors.

Finally, one way to improve the prediction results of our technique is to remove from the dataset a few outlier chips, which give an unacceptable maximum prediction error. Outlier chips could be due to outlier sensors' measurements that are different from the bulk of the sensors' measurements and, thereby, are poorly mapped to performances, or poorly predicted performances due to insignificant regression training locally.

We consider removing the first two outliers. The same two outliers were observed by training the regression multiple times and by randomly shuffling the chips. Table 42 shows the prediction results using Group B of sensors. Comparing Table 42 to Table 39, the technique of removing the outliers works well only for S_{11} and NF performances. In case of S_{11}, the maximum error is reduced from 1.91dB to 0.87dB, while in case of NF, the maximum error is reduced from 0.72dB to 0.12dB

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.26	0.95	55.91 %	1.19 dB
S_{11}	0.32	0.77	66.41 %	0.87 dB
S_{22}	0.63	0.74	70.86 %	1.97 dB
NF	0.04	1.09	37.12 %	0.12 dB
1-dB CP	0.18	0.89	53.32 %	0.63 dBm

Table 42: Alternate test prediction results using Group B of non-intrusive variation-aware sensors. The first two outliers have been removed.

On the other hand, if we remove the first ten outliers from the data set, we obtain the predictions results of Table 43. Comparing Table 43 to Table 39, the technique works for: (a) the S_{22} performance, where the maximum error is reduced from 1.99dB to 1.40dB, (b) NF, where the maximum error is reduced from 0.72dB to 0.08dB, and (c) 1-dB CP, where the maximum error is reduced from 0.63dBm to 0.47dBm. Although, in the case of NF, the correlation coefficient never exceeds the value of 50% and, hence, statistically, there is no correlation between this performance and sensors' measurements. This is, probably, due to the limited size of the sample set, where there is no enough information for making the mappings between the alternative measurements and the performances of the LNA.

Performance	Average Absolute RMS Error	Prediction figure-of-merit	Correlation Coefficient	Maximum Error
S_{21}	0.25	0.75	44.58 %	1.40 dB
S_{11}	0.31	1.03	41.32 %	0.87 dB
S_{22}	0.49	0.75	72.27 %	1.40 dB
NF	0.03	1.08	22.09 %	0.08 dB
1-dB CP	0.15	1.02	38.22 %	0.47 dBm

Table 43: Alternate test prediction results using Group B of non-intrusive variation-aware sensors. The first ten outliers have been removed.

It should be noted that in both Tables 42 and 43, the prediction figure-of-merit is higher compared to the rest of the tables, since the standard deviation is lower after removing the instances that present an unacceptable high maximum error.

7.5 CONCLUSIONS

In this chapter, we demonstrated, through real measurements, the non-intrusive BIT strategy based on variation-aware sensors for the 2.4GHz CMOS inductive degenerated LNA case study. Seventy-five chips were fabricated

from three process corners, namely slow, typical and fast. The fabricated samples were placed in the PCB and the measurement procedure was carried out using LabView environment. Due to the small sample set, random shuffling, multiple regression learning and cross-validation were used, in order to report reliable and trustworthy average prediction results. By using a few zero area overhead dummy structures and PCMs, and by obtaining on these dummy structures simple DC or low-frequency measurements that offer an "image" of process variations, we were able to track variations in almost all performances of the LNA with a prediction error lower than one standard deviation and a maximum error lower or at least comparable to the measurement and repeatability errors in a conventional ATE environment. NF had shown the lowest standard deviation compared to the rest of the performances, and hence, it was not correlated to the measurements of the non-intrusive sensors.

8

CONCLUSIONS AND PERSPECTIVES

8.1 CONCLUSIONS

High-volume production testing of RF and microwave/mm-wave circuits incurs a very high cost and is responsible for the largest fraction of the overall manufacturing cost. Standard test functions are employed using complex test instrumentation with advanced features, where different performances are characterized using sequential tests on different test configurations onto sophisticated and expensive ATE. Consequently, this results in lengthy switching and settling times, other than the pure electrical test times. Moreover, these tests need to be performed in an environment that is shielded from electromagnetic interference and noise. To overcome the key limitations of high test cost and lengthy test times, BIT has been proposed and used during the past few years. This test technique consists of performing on-chip some of the test operations such as stimulus generation, measurement extraction and analysis, etc. BIT can reduce dramatically the complexity of ATE since (a) it provides digital, DC, or low frequency test signatures, (b) it resolves issues related to electromagnetic interference, and (c) it facilitates parallel testing to achieve a high throughput. Although, in order to replace effectively the conventional test with the proposed BIT strategies, several important criteria need to be satisfied. Specifically, we need to (a) avoid degrading the performances of the CUT, (b) maintain a low area overhead, and (c) maintain the test accuracy of the standard test approach.

The goal of this work was to develop BIT solutions for RF and mm-wave circuits based on non-intrusive variation-aware sensors, in order to circumvent the need for elaborate test instrumentation and reduce the test cost. The key characteristics of the proposed BIT solutions are that (a) they are non-intrusive, that is, they do not tap into the RF signal path of the CUT, and therefore they do not degrade its performances, (b) they incur low area overhead and (c) they employ DC or low-frequency signals. The non-intrusive variation-aware sensors include dummy analog stages and PCMs, which both mimic part of the CUT architecture. They are placed n the die in close proximity to the CUT without being electrically connected to it. Their operation is based on the fact that they experience the same process variations as the CUT. In essence, the technique capitalizes on the undesired phenomenon of process variations. As a result, any degradation in the performances of the CUT is reflected to the sensors' outputs, which shift away from their nominal values. Hence, they can be used for enabling a low-cost alternate test approach where the performances of the CUT are inferred implicitly from their outputs. The alternate test paradigm is employed to map the outputs of the non-intrusive sensors to the performances of the CUT, in order to replace the standard tests for measuring the performances directly.

In this thesis, we demonstrated the BIT paradigm based on non-intrusive variation-aware sensors for two case studies, namely a 2.4GHz 65nm CMOS inductive degenerated LNA and a wide-band mm-wave 60GHz 65nm CMOS 3-stage LNA.

In the case of the 2.4GHz LNA, we found, through simulations, that the variability in the inductors of the LNA explains to a large degree the variability in its performances. We also showed that the variability in the inductors stems mostly from the variability in the resistivity of the high-level metal and alucap layers that form the coil of the inductors. To this end, to circumvent using a dummy inductor PCM that is area-hungry, we demonstrated that by using instead low area overhead high-level metal and alucap resistor PCMs, we are able to obtain practically the same prediction results as if a dummy inductor PCM was used. We also showed that WID variations, which cannot be handled by the non-intrusive variation-aware sensors, do not affect appreciably

the prediction results. The proposed test idea has given very promising prediction results using real silicon measurements. Seventy-five chips were fabricated from three corner wafers, namely slow, typical and fast and measured using benchtop equipment. The fabricated samples were placed in PCB and the measurement procedure was automated using LabView environment.

In the case of the 60GHz LNA, we demonstrated that variations in MOS transistor gate resistance explains to a large degree the variations in most of the performances of the LNA, while variations in the transmission lines are not the main source of performance variability.

In both cases, we demonstrated that by adding on-chip a few non-intrusive sensors of practically zero area overhead and by obtaining on these non-intrusive sensors DC or low-frequency measurements, we were able to track variations in all performances of the CUT with an average prediction error lower than one standard deviation of the performance and a maximum prediction error that is lower or at least comparable to the measurement and repeatability errors in a conventional ATE environment.

Finally, Table 44 summarizes the advantages and disadvantages of BIT intrusive and non-intrusive test approaches that have been reported in the literature including our proposed technique.

Test Idea	Non-Intrusiveness	Cost	Area Overhead	Detect Random Defects	Parallel Test
DC probes	x	low	zero	✓	x
Envelope Detectors	x	medium	medium	✓	x
Loop-back	x	high	medium	✓	x
Proposed test idea	✓	low	zero	✓	✓

Table 44: Advantages and disadvantages of BIT intrusive and non-intrusive test approaches.

8.2 PERSPECTIVES

In terms of perspectives, there is a list of several future tasks that could be carried out, including:

- fabricating more chips of the 2.4GHz LNA case study, in order to obtain a statistically significant sample and, thereby, more trustworthy prediction results.
- proving that the non-intrusive test paradigm can be applied to CUTs that are designed using very advanced technology nodes (i.e. 45nm, 28nm, etc.), where process variations become more prominent.
- showing that the proposed test idea could be applied in designs that are implemented in more sophisticated technologies like Silicon-On-Insulator (SOI) or Fully-Depleted SOI (FD-SOI), where the design procedures, the type of process variations, the sources of process variability, the parametric yield, etc., differ compared to the in typical CMOS bulk technology nodes. It would be interesting to prove that the proposed idea of monitoring variations non-intrusively still holds the same independently of the technology.
- proposing new variation-aware sensors that are able to handle intra-die variations.
- enhancing the library of non-intrusive sensors with smart yet simple-to-implement sensors. The type of sensors that are needed strongly depends on the circuit type, in particular on its architecture, its subcircuits, its devices, etc. For example, the library of sensors should expand in the case of mm-wave blocks (mixers, resonators, VCOs, etc), which are designed using mostly transmission lines. Monitoring variations in these passive devices could lead us to come up with some smart ways to predict the performances of these CUTs, where tapping into the signal paths is prohibitive.

- applying the non-intrusive test paradigm based on variation-aware sensors to full RF and mm-wave transceivers, where several blocks are involved.
- characterizing and testing the 60GHz LNA. There is on-going work regarding this perspective. A chip has been taped-out and measurements will be performed within the following months .

9

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. Golio, J. Golio, *RF and Microwave Circuits, Measurements and Modeling*, CRC Press, 2008.

[2] C.H. Doan, S. Emami, A.M. Niknejad, R. Brodersen, "Millimeter-Wave CMOS Design," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 1, pp. 144–155, 2005.

[3] M. Golio, *The RF and Microwave Handbook*, CRC Press, 2001.

[4] R. Garcia, "Redifining Cost of Test in an SoC World," *Evaluation Engineering, at www.evaluationengineering.com/archive/0603ic.htm*, accessed 2003.

[5] K.B. Schaub, J. Kelly, *Production Testing of RF and System-on-a-chip Devices for Wireless Communications*, Artech House, 2004.

[6] D. Gizopoulos, *Editor, Advances in Electronic Testing: Challenges and Methodologies*, Springer, pp. 337–369, 2006.

[7] C. Sayre , *Complete Wireless Design, Second Edition*, McGraw-Hill, 2008.

[8] E. Newman, *Optimizing receiver performance through error vector analysis*, Microwave ProductDigest,www.mpdigest.com October 2004.

[9] E. Acar and S. Ozev, "Defect-oriented testing of RF circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 5, pp. 920–931, 2008.

[10] H.-G. Stratigopoulos and S. Sunter, "Efficient Monte Carlo-based analog parametric fault modelling," in *Proc. IEEE VLSI Test Symposium*, 2014.

[11] J. Ferrario, R. Wolf, S. Moss, and M. Slamani, "A low-cost test solution for wireless phone RFICs," *IEEE Communications Magazine*, vol. 41, no. 9, pp. 82–88, 2003.

[12] D. Brown, J. Ferrario, R. Wolf, J. Li, J. Bhagat, and M. Slamani, "RF testing on a mixed signal tester," *Journal of Electronic Testing: Theory and Applications*, vol. 23, no. 1, pp. 85–94, 2007.

[13] M. G. Méndez-Rivera, A. Valdes-Garcia, J. Silva-Martinez, and E. Sánchez-Sinencio, "An on-chip spectrum analyzer for analog built-in testing," *Journal of Electronic Testing: Theory and Applications*, vol. 21, no. 3, pp. 205–219, 2005.

[14] D. Kawazoe, H. Sugawara, T. Ito, K. Okada, K. Masu, "Reconfigurable CMOS low-noise amplifier for self-compensation," in *Proc. IEEE International Symposium on Circuits and Systems*, pp. 4, 2006

[15] P. Väännen, N. Mikkola, P. Heliö, "VCO Design with On-Chip Calibration System," *IEEE Transactions on Circuits and Systems*, vol. 53, no. 10, pp. 2157–2166, 2006.

[16] M. Jarwala, L. Duy, and M. S. Heutmaker, "End-to-end test strategy for wireless systems," in *Proc. IEEE International Test Conference*, 1995, pp. 940–946.

[17] J.-S. Yoon and W. R. Eisenstadt, "Embedded loopback test for RF ICs," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 5, pp. 1715–1720, 2005.

[18] A. Valdes-Garcia, J. Silva-Martinez, and E. Sanchez-Sinencio, "On-chip testing techniques for RF wireless transceivers," *IEEE Design & Test of Computers*, vol. 23, no. 4, pp. 268–277, 2006.

[19] 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.

[20] B. Razavi, *RF Microelectronics*, Englewood Cliffs, NJ: Prentice-Hall, 1998.

[21] L. Abdallah, H.-G. 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.

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

[23] S. Bhattacharya and A. Chatterjee, "A DFT approach for testing embedded systems using DC sensors," *IEEE Design & Test of Computers*, vol. 23, no. 6, pp. 464–475, 2006.

[24] Y.-C. Huang, H.-H. Hsieh, and L.-H. Lu, "A built-in self-test technique for RF low-noise amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 2, pp. 1035–1042, 2008.

[25] 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.

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

[27] C. Zhang, R. Gharpurey, and J. A. Abraham, "Built-in self test of RF subsystems with integrated sensors," *Journal of Electronic Testing: Theory and Applications*, vol. 28, no. 5, pp. 557–569, 2012.

[28] A. Gopalan, M. Margala, and P. R. Mukund, "A current based self-test methodology for RF front-end circuits," *Microelectronics Journal*, vol. 36, no. 12, pp. 1091–1102, 2005.

[29] M. Cimino, H. Lapuyade, Y. Deval, T. Taris, and J.-B. Bégueret, "Design of a 0.9V 2.45 GHz self-testable and reliability-enhanced CMOS LNA," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 5, pp.1187–1194, 2008.

[30] Y. Maidon, "3.3 V CMOS Built-in Current Sensor," *Electronics Letters*, vol. 33, no. 5, pp. 345-346,1997.

[31] R. Voorakaranam, S. Cherubal, and A. Chatterjee, "A signature test framework for rapid production testing of RF circuits," in *Proc. Design, Automation and Test in Europe Conference*, 2002, pp. 186–191.

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

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

[34] J. Liaperdos, H.-G. Stratigopoulos, L. Abdallah, Y. Tsiatouhas, A. Arapoyanni, and X. Li, "Fast deployment of alternate analog test using bayesian model fusion," in *Proc. Design, Automation and Test in Europe Conference*, 2015.

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

[36] M. Bhushan, A. Gattiker, M. B. Ketchen, and K. K. Das, "Ring oscillators for CMOS process tuning and variability control," *IEEE Transactions on Semiconductor Manufacturing*, vol. 19, no. 1, pp. 10–18, 2006.

[37] L.-T. Pang and B. Nikolic, "Measurements and analysis of process variability in 90 nm CMOS," *IEEE Journal on Solid-State Circuits*, vol. 44, no. 5, pp. 1655–1663, 2009.

[38] A. Zjajo, M. Barragan Asian, and J. Pyneda de Gyvez, "BIST method for die-level process parameter variation monitoring in analog/mixed-signal integrated circuits," in *Proc. Design, Automation & Test in Europe Conference*, 2007, pp. 1301–1306.

[39] L. Abdallah, H.-G. Stratigopoulos, S. Mir, and J. Altet, "Defect-oriented non-intrusive RF test using on-chip temperature sensors," in *Proc. IEEE VLSI Test Symposium*, 2013.

[40] A. Serhan, L. Abdallah, H.-G. Stratigopoulos, and S. Mir, "Low-cost EVM built-in test of RF transceivers," in *Proc. IEEE International Design & Test Symposium*, 2014, pp. 51–54.

[41] A. Nassery, S. Ozev, M. Verhelst, and M. Slamani, "Extraction of EVM from transmitter system parameters," in *Proc. IEEE European Test Symposium*, 2011, pp. 75–80.

[42] C. Maxey, S. Raman, K. Groves, T. Quach, L. Orlando, A. Mattamana, G. Creech, J. Rockway, "Mixed-Signal SoCs With In-Situ Self-Healing Circuitry," *IEEE Design and Test of Computers*, vol. 29, no. 6, pp. 27–39, 2013.

[43] A. Siligaris, C. Mounet, B. Reig, P. Vincent, A. Michel, "CMOS SOI technology for WPAN. Application to 60 GHz LNA ," *IEEE Integrated Circuit Design and Technology and Tutorial*,2008, pp. 17–20.

[44] D. Heo, J. Kim, "Guest Editors' Introduction: Design and Testing of Millimeter-Wave/Subterahertz Circuits and Systems," *IEEE Design and Test*, vol. 31, no. 6, pp. 6–7, 2014.

[45] C.H. Doan, S. Emami, A.M. Niknejad, R. Brodersen, "Millimeter-wave CMOS device modeling and simulation," in *Proc. IEEE International Symposium on Circuits and Systems* , 2004, pp. 524–527.

[46] A. Kazemipour, M. Salhi, T.K. -Ostmann, T. Schrader, "A simple new method to calibrate millimeter-wave mixers," *IEEE Design and Test*, vol. 21, no. 3, pp. 205–219, 2014.

[47] J.-O. Plouchart, B. Parker, B. Sadhu, A. Valdes-Garcia, D. Friedman, F. Wang, X. Li, M. Sanduleanu, A. Balteanu, "Adaptive circuit design methodology and test applied to millimeter-wave circuits," *IEEE Design and Test*, 2014.

[48] R.-L. Schmid, P. Song, C.-T Coen, A. Ulusoy, and J.-D. Cressler, "A W-band integrated silicon-germanium loop-back and front-end transmit-receive switch for built-in-self-test," *IEEE International Microwave Symposium*, 2015, pp. 1–4, 2015.

[49] A. Antonopoulos, *Nanoscale RF CMOS Transceiver Design*, PhD thesis, School of Electronic and Computer Engineering, Technical University of Crete, Chania, Hellas, February 2013.

[50] P. Kinget, "Device mismatch and tradeoffs in the design of analog circuits," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 6, pp. 1212–1224, 2005.

[51] A. Heiberg, T. Brown, K. Mayaram, T. Fiez, "A 250 mV, 352 uW low-IF quadrature GPS receiver in 130 nm CMOS," *VLSI Circuits (VLSIC), 2010 IEEE Symposium on*, pp. 135–136, 2010.

[52] C. Enz and E. Vittoz, *Charge-based MOS transistor modeling: the EKV model for lowpower and RF IC design*, Chichester, England: John Wiley, 2006.

[53] A.S. Roy, *Noise and small-signal modeling of nanoscale MOSFETs*, PhD thesis, STI, Lausanne, 2007.

[54] P. Leroux, M. Steyaert, *LNA-ESD Co-Design for Fully Integrated CMOS Wireless Receivers*, Springer, 2005.

[55] D. Allstot, X. Li, S. Shekhar, "Design considerations for CMOS low-noise amplifiers," *IEEE Radio Frequency Integrated Circuits (RFIC) Digest of Papers Symposium*, pp. 97–100, 2004.

[56] B. Razavi, *RF Microelectronics*, Pearson, 2011.

[57] A.A. Abidi, "High-frequency noise measurements on FETs with small dimensions," *IEEE Trans. Electron Devices*, vol. 33, pp. 1801-1805, 1986.

[58] B. Razavi, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Trans. Circuits and Systems-Part 1*, vol. 41, pp. 750–754, Nov. 1994.

[59] B. Razavi, *Design of analog CMOS integrated Circuits*, Boston: McGraw-Hill, 2001.

[60] D.K. Shaeffer, T.H. Lee, "A 1.5-V,1.5-GHz CMOS Low Noise Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 745–759, May 1997.

[61] P. Leroux, M. Steyaert, *LNA-ESD Co-Design for Fully Integrated CMOS Wireless Receivers*, Springer, 2005.

[62] J. Janssens, *Deep Submicron Cellular Receiver Front-Ends*, PhD thesis, K.U.Leuven, Belgium, July 2001.

[63] A. Niknejad, H. Hashemi (Editors), *mm-Wave Silicon Technology - 60GHz and Beyond*, Springer, New York, 2008.

[64] T. Yao, M. Gordon, K. Yau, M.T. Yang, S.P. Voinigescu, "60-GHz PA and LNA in 90-nm RF-CMOS," *IEEE RFIC Symposium Diges*, pp. 147–150, 2006.

[65] C.M. Lo, C.S. Lin, H. Wang, "A Miniature V-band 3-stage Cascode LNA in 0.13um CMOS," *IEEE ISSCC Dig. Tech. Papers*, pp. 322–323, 2006.

[66] D.M. Pozar, *Microwave engineering*, Wiley, New York, 1998.

[67] http://rfic.eecs.berkeley.edu/ niknejad/ee142$_f a05 lects/pdf/lect14.pdf$

[68] S. Borkar, "Parameter Variations and Impact on Circuits and Microarchitecture," *Proc. 40th Design Automation Conf. (DAC 03)*, IEEE CS Press, pp. 338-342, 2003.

[69] G.E. Moore, *Cramming more components onto integrated circuits*, Electronics 38, 114 ff, 1965.

[70] C. Kenyon, A. Kornfeld, K. Kuhn, M. Liu, A. Maheshwari, W. Shih, S. Sivakumar, G. Taylor, P. VanDerVoorn, K. Zawadzki, "Managing process variation in Intel's 45nm CMOS technology," *Intel Technology Journal 12(2), 2008*, URL http://www.intel.com/technology/itj/ 2008/v12i2/3-managing/1-abstract.htm

[71] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1um MOSFET's: A 3D "atomistic" simulation study," *IEEE Transactions on Electron Devices*, 45(12), pp. 2505–2513, 1998. DOI 10.1109/16.735728

[72] C.H. Diaz, H.J. Tao, Y.C. Ku, A. Yen, K. Young, "An experimentally validated analytical model for gate line-edge roughness (LER) effects on technology scaling," *IEEE Electron Device Letters*, 22(6), pp. 287—289, 2001. DOI 10.1109/55.924844

[73] A. Asenov, S. Kaya, J.H. Davies, "Intrinsic threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations," *IEEE Transactions on Electron Devices*, 49(1), pp. 112—119, 2002. DOI 10.1109/16.974757

[74] V.S. Kaushik, B.J. O'Sullivan, G. Pourtois, N. Van Hoornick, A. Delabie, S. Van Elshocht, W. Deweerd, T. Schram, L. Pantisano, E. Rohr, L.A. Ragnarsson, S. De Gendt, M. Heyns, "Estimation of fixed charge densities in hafnium-silicate gate dielectrics," *IEEE Transactions on Electron Devices*, 53(10), pp. 2627—2633 (2006). DOI 10.1109/TED.2006.882412

[75] G. Lucovsky, "Intrinsic limitations on the performance and reliability of high-k gate dielectrics for advanced silicon devices," *Proc. IEEE Int. Integrated Reliability Workshop Final Report*, 2005, DOI 10.1109/IRWS.2005.1609592

[76] L. Capodieci, "From optical proximity correction to lithography-driven physical design (1996- 2006): 10 years of resolution enhancement technology and the roadmap enablers for the next decade," *Proceedings of SPIE*, vol. 6154 (3), 2006

[77] S. Nag, A. Chatterjee, K. Taylor, I. Ali, S. O'Brien, S. Aur, J.D. Luttmer, I.C. Chen, "Comparative evaluation of gap-fill dielectrics in shallow trench isolation for sub-0.25 /spl mu/m technologies," *Proc. Int. Electron Devices Meeting IEDM '96*, pp. 841—845, 1996. DOI 10.1109/IEDM.1996.554111

[78] Tsang, Y.L., Chattopadhyay, S., Uppal, S., Escobedo-Cousin, E., Ramakrishnan, H.K., Olsen, S.H., O'Neill, A.G., "Modeling of the threshold voltage in strained si/si1-x gex/si1-ygey(x-y) cmos architectures," *IEEE Transactions on Electron Devices*, 54(11), pp. 3040—3048, 2007. DOI 10.1109/TED.2007.907190

[79] A. Al-Bayati, H. Graoui, J. Spear, H. Ito, Y. Matsunaga, K. Ohuchi, K. Adachi, K. Miyashita, T. Nakayama, M. Oowada, Y. Toyoshima, "Advanced CMOS device sensitivity to USJ processes and the required accuracy of doping and activation," *Proc. 14th Int. Conf. Ion Implantation Technology*, 2002, pp. 185–188, 2002. DOI 10.1109/IIT.2002.1257969

[80] D. Blaauw, K. Chopra, A. Srivastava, L. Scheffer, "Statistical timing analysis: From basic principles to state of the art," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 27(4), pp. 589—607, 2008. DOI 10.1109/TCAD.2007.907047

[81] M. Dietrich,J. Haase *Process Variations and Probabilistic Integrated Circuit Design*, Springer, 2012

[82] J. Yang, L. Capodieci, D. Sylvester, "Advanced timing analysis based on post-OPC extraction of critical dimensions," *ACM/IEEE Design Automation Conference (DAC)*, pp. 359–364, 2005

[83] S.S. Sapatnekar, "Overcoming Variations in Nanometer-Scale Technologies," *IEEE Transactions on Emerging Topivs in Circuits and Systems*, vol. 1, No. 1, March 2005

[84] Y. Tsividis, *Mixed Analog-Digital VLSI Devices and Technology*, World Scientific, 2002

[85] D. Boning, S. Nassif , *Models of Process Variations in Device and Interconnect*, Massachusetts Institute of Technology, IBM Austin Research Laboratory, 2000

[86] K. Bernstein, K. Carrig, C. Durham, P. Hansen, D. Hogenmiller, E. Nowak, N. Rohrer, *High Speed CMOS Design Styles*, Kluwer, Boston, 1998

[87] X. Tang, V. De, and J. Meindl, "Intrinsic MOSFET Parameter Fluctuations Due to Random Dopant Placement," *IEEE Trans. Very Large Scale Integration (VLSI) Systems*,vol. 5, no. 4, pp. 369–376, Dec. 1997

[88] S. Borkar, "Designing Reliable Systems from Unreliable Components: The Challenges of Transistor Variability and Degradation," *IEEE Computer Society*, 2005

[89] T. Park, T. Tugbawa, D. Boning, J. Chung, S. Hymes, R. Muralidhar, B. Wilks, K. Smekalin, G. Bersuker, "Electrical Characterization of Copper Chemical Mechanical Polishing," *CMP-MIC*, pp. 184-191, Santa Clara, CA, Feb. 1999

[90] B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung, D. L. Hetherington, I. Ali, G. Shinn, J. Clark, O.S. Nakagawa S.-Y. Oh, "A Closed-Form Analytic Model for ILD Thickness Variation in CMP Processes," *CMP-MIC*, pp. 266-273, Santa Clara, CA, Feb. 1997.

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

[92] S.S. Akbay, J.L. Torres, J.M. Rumer, A. Chatterjee, J. Amtsfield, "Alternate Test of RF Front Ends with IP Constraints: Frequency Domain Test Generation and Validation," *IEEE International Test Conference*,2006, pp. 1–6.

[93] H.-G. Stratigopoulos, S. Mir, E. Acar, S. Ozev, "Defect filter for alternate RF test," in IEEE Eur. Test Symp.,2009,pp.101–106.

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

[95] M. J. Barragan and G. Léger, "A procedure for alternate test feature design and selection," *IEEE Design & Test of Computers*, vol. 32, no. 1, pp. 18–25, 2015.

[96] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

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

[98] D.C. Montgomery,E.A. Peck, G.G. Vining, *Introduction to Linear Regression Analysis*, Wiley, 2012.

[99] M. Andraud, H.-G. Stratigopoulos, and E. Simeu, "One-shot calibration of RF circuits based on non-intrusive sensors," in *Proc. Design Automation Conference*, 2014.

[100] A. Dimakos, H.-G. Stratigopoulos, A. Siligaris, S. Mir, and E. De Foucauld, "Parametric built-in test for 65nm rf lna using non-intrusive variation-aware sensors," *Journal of Electronic Testing: Theory and Applications*, 2015.

[101] X. Jin, "An effective gate resistance model for CMOS RF and noise modeling," in *IEDM Tech. Dig,* 1998, pp. 961–964

[102] Y. Cheng, "High frequency characterization of gate resistance in RF MOSFETs," *IEEE Electron Device Letters*, vol. 22, no. 2, pp. 98–100, Feb. 2001.

[103] Y. Cheng, M.J. Deen, C.-H. Chen, "MOSFET Modeling for RF IC design," *IEEE Transactions on Electron Devices*, vol. 52, no. 7, pp.1286–1303, Jul. 2005.

[104] B. Razavi, R.-H. Yan, K.-F. Lee, "Impact of distributed gate resistance on the performance of MOS devices," *IEEE Transactions Circuits and Systems I*, vol. 41, pp. 750–754, Nov. 1994.

[105] 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, Feb. 2000.

[106] http://www.designers-guide.org/analysis/intercept-point.pdf

[107] S. H. Jen, C. Enz, D. R. Pehlke, M. Schroter, and B. J. Sheu, "Accurate MOS transistor modeling and parameter extraction valid up to 10GHz," in *European Solid-State Device Research Conference*, 1998, pp. 484–487.

[108] C. Enz and Y. Cheng, "MOS transistor modeling for RF IC design," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 2, pp.186–201, 2000.

LIST OF PUBLICATIONS

1. **Athanasios Dimakos**, Matthias Bucher, R.K. Sharma, Ilias Chlis, "Ultra-low voltage drain-bulk connected MOS transistors in weak and moderate inversion", 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Seville, Spain, 2012.

2. **Athanasios Dimakos**, Haralampos-G. Stratigopoulos, Alexandre Siligaris, Salvador Mir, Emeric De Foucauld, "Non-intrusive built-in test for 65nm RF LNA", 9th International Mixed-Signals, Sensors and Systems Test Workshop (IMS3TW), Porto Alegre, Brazil, 2014.

3. **Athanasios Dimakos**, Haralampos-G. Stratigopoulos, Alexandre Siligaris, Salvador Mir, Emeric De Foucauld, "Parametric built-in test for 65nm RF LNA using non-intrusive variation-aware sensors", Journal of Electronic Testing: Theory and Application (JETTA), August 2015.

4. **Athanasios Dimakos**, Martin Andraud, Louay Abdallah, Haralampos-G. Stratigopoulos, Emanuel Simeu, Salvador Mir, "Test and Calibration of RF Circuits Using Built-in Non-intrusive Sensors", IEEE Computer Society Annual Symposium on VLSI, (ISVLSI), Montpellier, France, July 2015.

5. **Athanasios Dimakos**, Haralampos-G. Stratigopoulos, Alexandre Siligaris, Salvador Mir, Emeric De Foucauld, "Built-in Test of Millimeter-Wave Circuits Based on Non-intrusive Sensors" Design, Automation and Test in Europe, (DATE), Dresden, Germany, March 2016 (to appear).

10

RÉSUMÉ EN FRANÇAIS

Cette thèse discute le problème de test de production en grand volume des circuits radio-fréquences (RF) et à ondes millimétriques (mm-wave). Le test des fonctionnalités RF et à ondes millimétriques est très onéreux. Le test intégré est une alternative prometteuse pour faciliter la procédure et réduire les couts, mais il est difficile à mettre en œuvre car il ne faut en aucun cas qu'il réduit la performance du circuit sous test (CUT). Dans cette thèse, nous étudions une technique du test intégré qui repose sur l'utilisation de capteurs non-intrusifs qui prend en compte la variabilité du procédé de fabrication. Cette technique est extrêmement intéressante pour les concepteurs des circuits RF et mm-wave car il leur permet de dissocier le test de la conception. Les capteurs non-intrusifs sont constitués d'étages analogiques triviaux et de composants simples qui sont copiés de la topologie du CUT et sont placés sur la puce à proximité du CUT. Ils offrent simplement une "image" des variations du procédé de fabrication, ce qui leur permet de suivre les variations de performance du CUT. En substance, cette technique tire parti des phénomènes non désirés de variabilité de procédé de fabrication. Le paradigme du test alternatif est utilisé pour estimer les performances du CUT à partir des mesures des capteurs non intrusifs, afin de remplacer les tests standards qui mesurent les performances directement. Ce principe de test est appliqué à deux différents CUTs, nommément un amplificateur à bas bruit à 2.4GHz réalisé en 65nm CMOS et un amplificateur à bas bruit large bande à 60GHz réalisé en 65nm CMOS. Nous démontrons qu'en ajoutant quelques capteurs non-intrusifs sur la puce, qui n'engendrent pratiquement pas de surcout de surface, et en obtenant de ces capteurs non-intrusifs certaines mesures dans le domaine continu et à basse fréquence, nous sommes capable de suivre les variations de toutes les performances du CUT avec une erreur de prédiction moyenne inférieure à l'écart-type de la performance, et une erreur de prédiction maximum qui est inférieure ou au moins comparable aux erreurs de mesure dans un équipement de test automatisé conventionnel.

10.1 INTRODUCTION

Le test des circuits Radiofréquences (RF) produits en grande quantité requiert des équipements automatiques de test (Automatic Test Equipment ATE) avec des caractéristiques avancées. Les configurations de test représentent un défi car elles requièrent une calibration et un de-embedding, ainsi qu'un blindage de l'environnement de mesures par rapport aux interférences électromagnétiques et au bruit. De plus, les métriques de test d'émetteurs /récepteurs RF au niveau système, comme le taux d'erreur de bits (Bit-Error Rate BER) ou l'ErrorVector Magnitude (EVM) prennent beaucoup de temps pour être calculées. Pour ces raisons, les procédures de test standards sont très couteuses. Ces dernières années, les chercheurs et professionnels du test cherchent à développer des techniques de test alternatives qui visent à réduire le coût du test tout en pouvant maintenir les mêmes niveaux de précision. De manière analogue aux techniques de test développées pour les circuits numériques, le test structurel a été proposé dans le but de développer des tests qui permettent de détecter une liste des fautes les plus probables. Cependant, ces techniques n'ont pas été matérialisées à cause des difficultés à construire un modèle de fautes précis. Une autre technique de test est le built-off test, qui implique de faire migrer certaines fonctionnalités de l'ATE sur une carte de test. Cette approche permet de réduire significativement le coût du test en contrepartie d'un effort sur la conception d'une carte de test qui doit pouvoir s'adapter à différentes applications. Une technique avec un potentiel plus important est le test in-situ, ou built-in test. Le principe consiste à intégrer dans la puce certaines structures pour faciliter le test comme un générateur de signal, un analyseur de réponse. L'approche la plus répandue pour le test in-situ d'émetteurs/récepteurs RF est le test par boucle de

retour, ou loop-back test. Les signaux de test sont générés en bande de base, et la sortie du transmetteur est connectée à l'entrée du récepteur pour analyser les résultats de test en bande de base, le tout en restant au sein la puce. Une autre technique se base sur des détecteurs d'enveloppe et des capteurs de courant pour extraire des signatures de test continues ou basses-fréquences qui sont néanmoins porteuses d'informations en RF. Enfin, le test alternatif, ou alternate test, est une méthode générique qui élimine le besoin de mesurer directement les performances RF. Le test alternatif se base également sur des mesures bas coût, mais riches en informations, réalisées avec des capteurs in-situ.

Pour l'implémentation réussie d'une technique de test in-situ, plusieurs critères doivent être satisfaits simultanément: (a) le niveau de précision des méthodes de test standard est maintenu ; (b) les performances du circuit sous test (Circuit Under test CUT) ne sont pas affectées; (c) le nombre de broches supplémentaires du circuit et l'augmentation de sa surface sont acceptables ; et (d) la réduction du coût de test réalisée justifie l'effort de développement de la technique de test in-situ. Il est à noter que l'ensemble des techniques de l'état de l'art précédemment évoquées ne satisfont pas pleinement l'objectif (b). Par exemple, la connexion de la boucle de retour dans le test loop-back requiert un interrupteur et un atténuateur, ainsi qu'un autre mélangeur pour certains types de récepteurs RF qui est inséré dans le chemin du signal. Les détecteurs d'enveloppe et les capteurs de courant sont aussi insérés dans le chemin du signal. En général, cette insertion de composant sur le chemin du signal RF dégrade l'adaptation d'impédance et ajoute des parasites, ce qui va inévitablement modifier les performances et déséquilibrer les compromis de performances établis pendant la conception. Pour résoudre ce problème, la circuiterie de test in-situ doit être conçue avec le CUT ce qui augmente le temps de conception nécessaire à obtenir les performances désirées, si elles peuvent être obtenues. Pour ces raisons, les concepteurs de circuits sont souvent peu enclins à incorporer ces techniques de test in-situ, puisque les spécifications sur les circuits sont déjà très serrées et exploitent au mieux les possibilités des nœuds technologiques avancés.

Dans ce but, il a été proposé de se baser sur des capteurs in-situ non intrusifs qui comparativement laissent le circuit intact. Ces capteurs tirent profit du phénomène indésirable des variations du procédé de fabrication. L'idée sous-jacente est de monitorer les variations du procédé au lieu de mesurer directement les performances RF. Dans ce but, on peut utiliser des moniteurs de procédé (Process Control Monitors PCMs), comme des composantsseuls (transistors, capacités, résistances…) ainsi que des circuits factices, ou "dummy" quisontextraits de la topologie du circuit (miroirs de courants, étages de gainsou de polarisation…). Ces capteurs sont placés à proximité physique des strcutures identiquesqu'ilsmonitorent. Par exemple, i lest possible de placer un étage de polarisation dummy à côté de l'étage de polarisation réel du CUT, un transistor dummy à côt du transistor réel du CUT, etc. De cettemanière, les capteurs et le CUT sont soumis aux mêmes variations de procedé puce à puce (die-to-die D2D) et auxvariations intra-pucecorrélées (Within-Die WID)et, au final, les mesures prises par les capteurs seront largement corrélées aux performances réelles du CUT. Une technique indirecte et à bas coût peut ensuite être mise en place, basée sur le modèle du test alternatif, pour faire correspondre les mesures des capteurs aux performances du CUT. Cette technique de test a été inspirée par les PCMs typiquement placées dans les lignes entre les puces des wafers lors de la fabrication utilisés pour monitorer la variabilité et identifier les paramètres du procédés trop déviants. L'idée d'utiliser des PCMs pour extraire des informations à propos des performances a été aussi appliquée dans le cas de convertisseurs analogiques-numériques (CAN).

Il peut être signalé qu'avec cette approche de test il est possible de vérifier si une ou plusieurs performances violent leurs spécifications à cause de variations de procédé excessives, mais que nous ne pouvons pas détecter les défauts dans le CUT car les capteurs ne sont pas électriquement connectés à celui-ci. Pour cela, un test non-intrusif peut-être réalisé avec un capteur de température. L'idée sous-jacente est qu'un défaut changerale courant qui circule à travers une branche du circuit et , de ce fait, changera également la puissance dissipée par les composants dans cette branche ainsi que leprofil de température dans la puce à proximité. Uncapteur de température peut capturerce changement de température, ce qui indique ainsi la présence du défaut. À cet égard, une méthodologie complète des essais pour détecter les variations de procedé et les défauts peut être conçu uniquement basée sur des capteurs non intrusifs ,quipeuvent remplacer totalement les pratiques de test RF standard utilisées aujourd'hui dans l'industrie.

Il convient également de noter que, dans lecas d' un CUT qui se composed'unechaîne de blocs, par exemple, unémetteur-récepteur RF, aveccette approche de test, nous allons suivre chaque bloc individuellement pour

prédireses performances, comme par exemplelegain, de bruit, etc., de l' amplificateur à faiblebruit (LNA), de l'amplificateur de puissance (PA), etc., quicomposentl'émetteur-récepteur RF. Cela peut entraînerune perte de rendement dans le scénario où les variations excessives dans les deux blocs sont compensées et masquées. Une possibilité est d'utiliser cette approche de test pour prédire les performances au niveau du CUT, par exemple l' EVM dans lecas d' unémetteur-récepteur RF. Uneautrepossibilitéestd'employer les prédictions de performances des blocs individuels pour prédire les performances de niveau CUT à partir d'expressions mathématiques analytiques, par exemple l' EVM dans lecas d' un émetteur-récepteur RF.

En plus de cela, l'évolution continuelle des technologies CMOS jusqu'à 32nm et au-dessous a poussé la fréquence de transition et de la fréquence d'oscillation maximale de transistors au delà de 230GHz, ce qui les rend attractifs pour la conception de circuits et systèmes millimétriques, ou mm-Wave. La nécessité de concevoir des systèmes à très haut débit de donnéesa conduit à plusieurs applications millimétriques, comme les amplificateurs et les filtres, des liaisons point à pointgigabit / s, les réseaux locaux sans fil (WLAN) d'une capacité extraordinaire, réseaux personnels haut débit de données sans fil à courte portée (WPAN), et radar dans les véhicules. La conception de circuits et systèmes mm-Wave présente plusieurs défis car elle nécessite une modélisation précise des dispositifs actifs, composants passifs, des interconnexions et des lignes de transmission, et sont soumis à de grandes variations du procédé, de la tension d'alimentation, et de température. Les efforts récents visent à accroître le rendement des circuits mm-Wave grâce à l'étalonnage, les stratégies de circuits adaptatifs, et des algorithmes d'auto-guérison ou self-healing. Le test de circuits et systèmes mm-Wave haut volume de production est un enjeu majeur en vue de leur adoption généralisée. La seule approche de test standard serait d'utiliser un ATE avec des fonctions très avancées et une calibration sophistiquée,et des procédures de de-embedding, ansi que d'effectuer des tests de taux d'erreur avec un structure de test en loop-back. Cependant, cette approche de test entraîne un coût élevé. En outre, la connexion loop-back nécessite une augmentation de la surface du circuit, et est très difficile du point de vue de la conception, car elle doit être soigneusement conçu conjointement avec le circuit mm-Wave pour éviter la dégradation des performances dans le circuit mm-Wave.

Dans ce travail, nous démontrons la stratégie de test non-intrusif in-situ intégré sur deux cas d'études, à savoir un LNA RF en technologie 65nm et un LNA mm-Wave à 3 étages à 60GHz. Dans le premier cas, nous montrons que la corrélation reste forte dans cette technologie avancée spécifique. Nous avons observé que la variation des inductances explique dans une large mesure la variabilité des performances, il a donc été jugé nécessaire de surveiller la variation des inductances. Cependant, l'ajout d'une inductance PCM « dummy »n'est pas un choix judicieux car il encourt une augmentation trop importante de la surface. Pour cette raison, la variabilité dans les inductances n'avait pas été considérée du tout. Dans ce travail, nous avons étudié les sources de variabilité dans les inductances et nous avons constaté que la majeure partie de la variabilité est due à la variabilité de la résistivité du métal des couches supérieures et de la couche alucap qui forment les bobines des inductances, et ne sont pas dues à la variabilité de l'inductance pure de la bobine. Cela nous a conduit à utiliser des lignes de métal en faible surface et une résistance alucap en tant que PCMs, au lieu d'une inductance PCM qui est couteuse en terme de surface, afin de capturer la plus grande partie de la variabilité dans les inductances. Dans le dernier cas d'étude, nous montrons que la résistance de grille des transistors MOS explique dans une large mesure la variabilité des performances du LNA.

10.2 TEST IN-SITU NON-INTRUSIF

Considérant tout circuit analogique, l'idée sous-jacente est de le décomposer en étages analogiques "primitifs" et des composants critiques et/ou assembler des étages analogiques de base à partir de composants identiques qui peuvent être trouvés dans la topologie du circuit. Les étages analogiques peuvent être des étages de polarisation, des étages de gain, des miroirs de courant, etc., tandis que les composants individuels peuvent comprendre des transistors, des résistances, des condensateurs, des inductances, des lignes de transmission, etc. Ces étages analogiques et ces composants individuels sont ensuite placés comme des structures "dummy" dans la puce, à côté du circuit sous test (CUT). En particulier, ils sont placés à proximité physique des étages analogiques

correspondants et des composants du CUT qu'ils copient, de telle sorte que les deux soient bien matchés suivant des règles de layout spécifiques. Nous appelons ces structures « dummy » comme capteurs non intrusifs. Le principe sous-jacent de l'opération est que, grâce à cette proximité physique, les capteurs et le CUT sont soumis aux mêmes variations de procedé puce à puce (die-to-die D2D) et auxvariations intra-pucecorrélées (Within-Die WID) que le CUT. En conséquence, nous nous attendons à ce que les mesures sur les capteurs non intrusifs (tensions, par exemple de polarisation, le gain DC, le transistor transconductance, résistance, capacité, etc.) offrent une «image» des variations du procédé au sein du CUT et, par conséquent, nous nous attendons à que ces mesures soient largement corrélées avec les performances du CUT. Si une telle corrélation existe, alors nous serons en mesure de prédire les performances indirectement uniquement à partir des mesures sur les capteurs non intrusifs basé sur le modèle du test alternatif. Le test alternatif se réfère à une approche générique de test qui consiste à déduire les performances de la CUT par le biais de mesures alternative à faible coût (par exemple les mesures fournies par les capteurs non intrusifs dans notre cas). Formellement, M et Pj désignent respectivement les mesures alternatives et les performances du j-ème, j = 1 ... n. Les n correspondances entre M et P_j, j ..n, ne peuvent pas facilement être calculées sous forme analytique, elles sont modélisées par n fonctions de régression n f_j: M -> P_j. Les n fonctions de régression sont apprises pendant une phase d'entraînement hors-ligne qui utilise un échantillon de CUT fabriqués, qui doit idéalement couvrir tous les cas de processus et doit être globalement très représentatif du processus de fabrication. Une fois les n fonctions de régression sont apprises, elles peuvent être facilement utilisées pour déduire les n performances de toute instance du CUT simplement en effectuant les mesures alternatives. Pour diverses idées concrètes en vue de renforcer l'efficacité du test alternatif, le lecteur intéressé peut se reporter à la bibliographie relative à ce travail. En bref, la stratégie de test intégré non-intrusive capitalise sur le phénomène indésirable des variations de procédé de fabrication pour fournir des mesures alternatives à faible coût à partir de laquelle les performances peuvent être prédites implicitement avec précision.

10.2.1 *Avantages et inconvénients*

La stratégie de test in-situ non-intrusif présente plusieurs avantages:

- Il est générique car il est virtuellement applicable à tout circuit analogique.
- Il est peu coûteux étant donné que les stimuli de test appliqué aux capteurs non intrusifs et les mesures fournies par les capteurs non intrusifs sont continus ou à basse fréquence. Dans certains cas, les capteurs non intrusifs sont auto-polarisés. Ainsi, un ATE avec des fonctionnalités avancées n'est plus nécessaire, et les entrées et les sorties des capteurs non intrusifs peuvent être multiplexées dans un bus de test de manière à utiliser uniquement deux broches pour le test.
- Il est particulièrement intéressant pour les circuits RF et mm-Wave car le test in-situ est totalement transparent pour le CUT. En fait, la conception et le test sont complètement dissociés.
- Le CUT n'a pas besoin d'être alimenté sur au cours du test qui facilite grandement les tests multi-site parallèles.

Les inconvénients de la stratégie de test in situ non-intrusif sont les suivants:

- On ne peut détecter que les défaillances dues à des variations excessives de processus, ainsi il doit être complété par une autre stratégie de test qui est capable de détecter des défauts. Il a été montré récemment que des défauts peuvent également être détectés de manière non intrusive en utilisant des capteurs de température.
- - Il ne peut pas faire face à des variations WID décorrélés (par exemple le mismatch) qui deviennent importants pour les nœuds technologiques avancés à partir de 65nm et au-dessous. En bref, l'approche ne marchera pas si les variations WID décorrélés deviennent comparable aux variations D2D et WID

corrélées. Il faut noter cependant que les performances des circuits RF et mm-Wave ne dépendent pas dumismatch, sauf si nous considérons topologies entièrement différentielles. En outre, l'offset apparaît comme un«bruit» dans les algorithmes utilisés pour apprendre les fonctions de régression lors de l'entraînement et, par défaut, la régression apparaît comme une moyenne sur ce bruit.

10.2.2 *Défis et directives*

Les principaux défis de la stratégie de test in-situ non-intrusif sont les suivants:

- Identifier les principaux paramètres du procédé (à savoir la sheet résistance, l'épaisseur d'oxyde, la concentration de dopant, les capacités de jonction, etc.) dont la variation explique la variation des performances du CUT.
- Concevoir des capteurs non intrusifs qui monitorent les variations de ces paramètres duprocédé.
- En outre, les capteurs non intrusifs doivent être mesurés en continu ou à basse fréquence et doivent limiter l'augmentation de la surface du circuit, afin de parvenir à un faible coût.

Il y a plusieurs lignes directrices qui peuvent être suivies pour relever ces défis:

- Une approche directe qui a montré son bon fonctionnement est de décomposer le CUT en étages analogiques et en composants individuels et les utiliser comme capteurs non intrusifs.
- Les kits de conception de processus pour les nœuds technologiques avancés ont des centaines de paramètres de procédé et il n'est pas facile de bien comprendre les relations complexes entre les paramètres du procédé et les performances d'un circuit. Pour cette raison, nous pouvons compter sur les principaux paramètres de conception (de la géométrie des transistors, de la transconductance des transistors, de la valeur des résistances et des condensateurs, des courants et des tensions de polarisation, etc.), qui, à leur tour, sont des fonctions complexes des paramètres du procédé, et choisir de concevoir des capteurs non intrusifs qui capturent efficacement ces paramètres de conception. Les idées et l'expérience du concepteur peuvent être très utiles pour effectuer cette tâche au mieux.
- Inclure un ensemble initial de capteurs non intrusifs qui est plus simple d'un point de vue conceptuel et d'effectuer une vaste campagne de simulation pour examiner si cet ensemble est largement suffisant. Sinon, le problème de la sélection des capteurs non intrusifs peut être revu et étudié de manière plus approfondie.
- Inclure un grand nombre de capteurs non intrusifs et d'effectuer une vaste campagne de simulation pour étudier la redondance et de garder un ensemble optimal.

10.3 RÉSULTATS ET DISCUSSION

10.3.1 *Cas d'étude 1 : LNA avec dégénération inductive à 2.4GHz en technologie 65nm*

La première cas d'étude de ce travail de cas comprennent un LNA avec dégénération inductive à 2.4GHz en technologie 65nm. Les Figures 70 et 71 représentent respectivement les schémas du LNA 2.4 GHz et les capteurs non-intrusifs sélectionnés.

Le LNA et les capteurs non intrusifs de variations de procédé ont été conçus en utilisant la technologie CMOS065 bulk 65nm fournie par ST Microelectronics. Le schéma du LNA est représenté sur la Figure 70. Les capteurs que nous avons utilisées dans l'analyse comprennent (a) un miroir de courant dummy, (b) un transistor MOS PCM connecté en diode, (c) un condensateur MOM PCM, et (d) une inductance PCM, comme représenté sur la Figure 71, ainsi que (e) des lignes de métal et un PCM de résistances alucap. L'ensemble des mesures

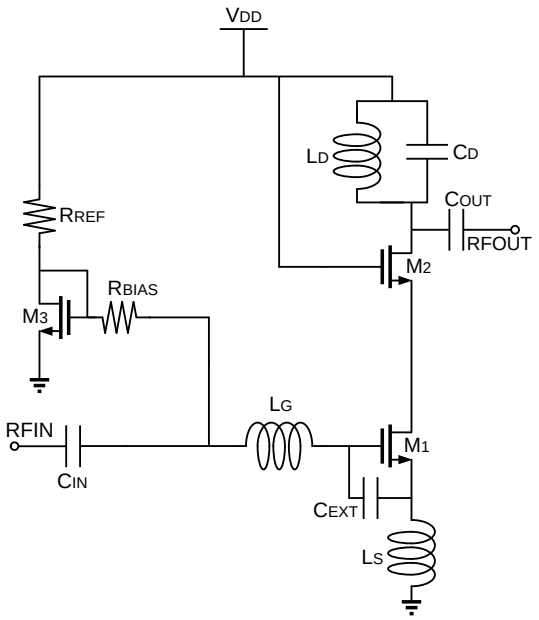


Figure 70: Schéma du LNA avec dégénération inductive à 2.4GHz en technologie 65nm.

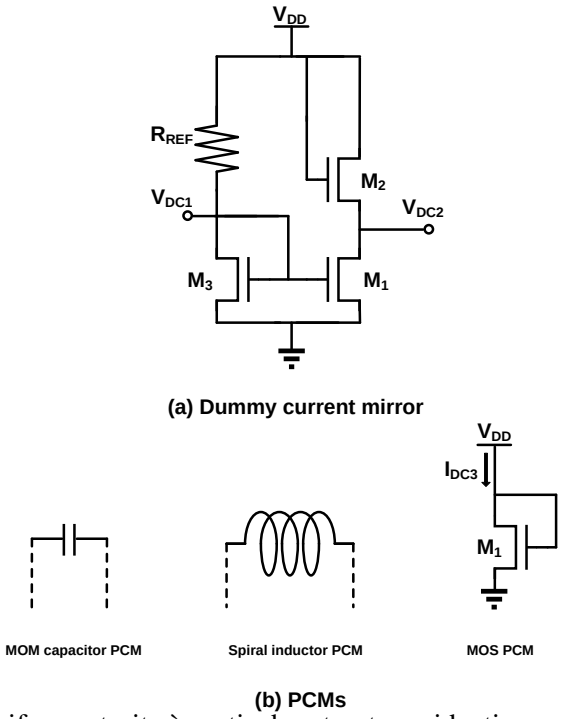


Figure 71: Capteurs non-intrusifs construits à partir des structures identiques issus de la topologie initiale du LNA.

alternatives comprennent (a) les tensions continues V_{DC1}, V_{DC2} du miroir de courant fictif, (b) le courant continu I_{DC3} de la PCM MOS connecté en diode (c) la capacité du condensateur PCM MOM, appelée C_{PCM} (d) l'inductance de l'inductance PCM (par exemple la partie réelle R_s et immaginaire L_s de l'impédance de l'inductance), et (e) la résistivité de la ligne de métal et de la résustance de la PCM alucap, dénoté par ρ_{metal} et ρ_{alucap} respectivement obtenues en divisant la tension continue appliquée à leurs bornes avec le courant continu qui circule à travers elles. La capacité et l'inductance sont extraites par des paramètres S à basse fréquence (de l'ordre de quelques dizaines de MHz).

Dans les expériences avec le test alternatif suivant, nous utilisons des différents groupes de capteurs de variation de procédé non-intrusifs et leurs mesures correspondantes, comme indiqué dans le Tableau 45.

	Dummy current mirror	MOS PCM	MOM PCM	Inductor PCM	Inductor PCM	Resistor PCMs
	V_{DC1}, V_{DC2}	I_{DC3}	C_{PCM}	R_s	L_s	$\rho_{metal}, \rho_{alucap}$
Group 1	✓	✓				
Group 2	✓	✓	✓			
Group 3	✓	✓	✓	✓	✓	
Group 4	✓	✓	✓	✓		
Group 5	✓	✓	✓		✓	
Group 6	✓	✓	✓		✓	✓
Group 7	✓	✓	✓			✓

Table 45: Groupes de capteurs non-intrusifs ainsi que les groupes correspondants de mesuresutilisées dans l'analyse.

Le LNA et les capteurs non intrusifs ont été simulées dans l'environnement Cadence Virtuoso en utilisant SpectreRF. Les distributions réelles des paramètres de processus définis dans le Process Design Kit (PDK) ont été utilisés lors de l'analyse de Monte Carlo pour modéliser les variations D2D et WID. Le Tableau 46 énumère les performances du LNA, les valeurs nominales, et l'écart type observé dans un échantillon issu de 1000 simulations Monte Carlo considérant (a) uniquement les variations D2D et (b) les variations D2D et WID (WID an anglais).

Performance	Standard Deviation D2D	Standard Deviation D2D and WID
S_{21} (dB)	0.34	0.34
S_{11} (dB)	0.3	0.3
S_{22} (dB)	2.53	2.54
NF (dB)	0.07	0.07
1-dB CP (dBm)	0.44	0.42
IIP_3 (dBm)	0.28	0.27

Table 46: Liste des performances du LNA qui montrent la valeur nominale des performances ainsi que l'écart-type issues de simulations Monte Carlo qui considèrent (a) seulement les variations D2D et (b) les variations D2D et WID.

Le même échantillon a été utilisé pour apprendre les fonctions de régression en utilisant le test alternatif, en utilisant différents groupes de capteurs non-intrusifs, comme indiqué dans le Tableau 45. Sur les 1000 instances, nous avons utilisé 800 instances de circuit pour l'entraînement et 200 instances pour la validation, pour fournir une estimation non biaisée de l'erreur de prédiction lors du test alternatif. L'erreur de prédiction est exprimée avec différentes métriques d'erreur, y compris le (a) l'erreur Root Mean Square (RMS) moyenne exprimée sur l'ensemble des instances de validation en %, (b) l'erreur RMS moyenne absolue sur l'ensemble des instances de validation exprimée en unité de chaque performance, (c) l'erreur maximale observée sur l'ensemble des instances de validation exprimée unités de chaque performance, et (c) le coefficient de corrélation entre les performances simulées (appelées «real») et prédites. Les métriques (a) et (b) montrent la précision globale du test alternatif. La métrique (c) montre le pire cas observé dans l'ensemble des instances de validation, qui doit être inférieur ou comparable à l'erreur de mesure de répétabilité sur l'ATE. Enfin, la métrique (d) est une mesure qualitative qui montre l'ajustement linéaire entre les performances réelles et prédites. Plus le coefficient de corrélation s'approche de 100%, au mieux les performances sont prédites. Toutefois, il convient de souligner que le coefficient de corrélation nettement inférieur à 100% ne conduit pas nécessairement à des résultats de prédiction insatisfaisants.

Le Tableau 47 présente les résultats de prédiction en utilisant le Groupe de capteurs 1 en ne considérant que les variations D2D. Comme on peut le constater, les résultats de prédiction ne sont pas satisfaisants, ce qui implique qu'une inductance PCM de surveillance est nécessaire pour augmenter la précision de la prédiction.

Le Tableau 48 présente les résultats de prédiction en utilisant le Groupe de capteurs 2, où l'inductance PCM est ajoutée au Groupe de capteurs 1 en utilisant les R_s et L_s dans la liste des mesures. Comme on peut le voir, les résultats de prévision sont presque parfaits en examinant toutes les métriques. Ce résultat démontre que, lorsque l'on considère que des variations de D2D, un test alternatif à faible coût basé sur le Groupe 2 de capteurs est capable de remplacer les tests de spécification pour la mesure des performances tout en conservant la même précision de test.

L'étape suivante consiste à déterminer si nous pouvons éviter d'utiliser une inductance PCM car elle induit à une trop forte augmentation de la surface de la puce. Le Tableau 49 présente les résultats de prédiction en utilisant le Groupe 3 de capteurs où l'inductance PCM est ajoutée au Groupe de capteurs 1 en utilisant seulement R_s dans la liste de mesures, tandis que le Tableau 50 montre les résultats de prédiction en utilisant le Groupe 4 de capteurs où l'inductance PCM est ajoutée au Groupe 1 des capteurs en utilisant seulement Ls dans la liste des mesures. En comparant les Tableaux 49 et 50 par rapport au Tableau 48, nous observons que les résultats se détériorent quand nous n'incluons pas R_s dans la liste des mesures, tandis que, si nous n'incluons pas Ls dans la liste des mesures, les résultats montrent la même précision de test dans la pratique à l'exception de la performance $S22$ pour laquelle le coefficient de corrélation diminue jusqu'à 54,89%. Toutefois, le test de la performance $S22$ est connu pour être redondant, car si un circuit ne satisfait pas à la spécification du $S22$, il échouera aussi la spécification d'une autre performance. Dans tous les cas, malgré un coefficient de corrélation assez faible, l'erreur quadratique moyenne absolue n'est pas trop importante et est comparable à l'erreur de mesure de la reproductibilité sur l'ATE. Nous concluons que nous pouvons éviter d'utiliserLs de mesure sans affecter de manière significative la précision du test.

Une autre étape est d'essayer de trouver d'autres capteurs pour remplacer la mesure de R_s. L'idée est de placer des lignes des métaux de plus haut niveau ainsi que la résistance la PCM alucappour contourner la nécessité d'inclure une inductance PCM. Le Tableau 51 présente les résultats de prédiction en utilisant le Groupe 5 de capteurs où en plus de Groupe de capteurs 1, nous avons inclus la mesure de L_s et des lignes des métaux de plus haut niveau ainsi que la résistance la PCM alucap. En comparant le Tableau 51 au Tableau 48, on constate que la précision de la prédiction est pratiquement le même, donc, nous pouvons remplacer la mesure R_s avec la résistivité de lignes métalliques et la PCM alucap. Le Tableau 52 présente les résultats de prédiction à l'aide du Groupe de capteurs 6 où, par rapport au Groupe 2, l'inductance PCM est éliminée et des lignes des métaux de plus haut niveau ainsi que la résistance la PCM alucapsont ajoutés. En comparant le Tableau 52 au Tableau 49, on constate que la précision de la prédiction est pratiquement la même.

Enfin, nous avons répété l'analyse pour les Groupes 5 et 6 de capteurs qui envisagent cette fois à la fois les variations D2D et WID. Les résultats de prévision sont présentés dans les Tableaux 53 et 54. En comparant les Tableaux 53 et 54 aux Tableaux 51 et 52, respectivement, où seules variations D2D sont prises en compte, on constate que les variations WID dégradent légèrement les corrélations, mais les corrélations demeurent solides et le test alternatif offre encore d'excellents résultats.

La Figure 72 représente les diagrammes de dispersion entre les valeurs réelles et prédites des performances pour le Groupe 6 de capteurscomptetenu à la fois D2D et variations WID. La figure 72 offre une visualisation des résultats de prévision du Tableau 55. Chaque point correspond à une instance de circuit utilisé pour la validation. Comme on peut le voir, les points sont répartis sur la diagonale pour toutes les performances à l'exception du S_{22} pour le quelle coefficient de corrélation à partir du Tableau 55 est 55.96%.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.95 %	0.28 dB	56.65 %	0.69 dB
S_{11}	3.19 %	0.28 dB	34.95 %	0.66 dB
S_{22}	13.48 %	2.13 dB	30.03 %	6.05 dB
NF	3.09 %	0.06 dB	63.14 %	0.14 dB
1-dB CP	2.01 %	0.26 dBm	81.12 %	0.93 dBm
IIP_3	12.15 %	0.18 dBm	74.92 %	0.64 dBm

Table 47: Résultats de prédiction du test alternatif en utilisant le Groupe 1 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.61 %	0.09 dB	96.66 %	0.34 dB
S_{11}	1.11 %	0.1 dB	94.37 %	0.28 dB
S_{22}	3.14 %	0.5 dB	97.67 %	1.3 dB
NF	1.29 %	0.02 dB	94.59 %	0.06 dB
1-dB CP	0.72 %	0.09 dBm	97.76 %	0.23 dBm
IIP_3	3.67 %	0.06 dBm	97.97 %	0.13 dBm

Table 48: Résultats de prédiction du test alternatif en utilisant le Groupe 2 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.06 %	0.15 dB	89.3 %	0.41 dB
S_{11}	1.39 %	0.12 dB	91.32 %	0.3 dB
S_{22}	12.23 %	1.93 dB	54.89 %	4.81 dB
NF	1.29 %	0.02 dB	94.54 %	0.05 dB
1-dB CP	1.23 %	0.16 dBm	93.3 %	0.47 dBm
IIP_3	8.22 %	0.12 dBm	89.3 %	0.38 dBm

Table 49: Résultats de prédiction du test alternatif en utilisant le Groupe 3 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.79 %	0.26 dB	65.37 %	0.66 dB
S_{11}	3.07 %	0.27 dB	43.84 %	0.65 dB
S_{22}	7.16 %	1.13 dB	87.23 %	3.8 dB
NF	3.07 %	0.06 dB	63.41 %	0.14 dB
1-dB CP	1.8 %	0.23 dBm	84.89 %	0.68 dBm
IIP_3	9.93 %	0.15 dBm	83.92 %	0.45 dBm

Table 50: Résultats de prédiction du test alternatif en utilisant le Groupe 4 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.53 %	0.08 dB	97.42 %	0.1 dB
S_{11}	1.11 %	0.1 dB	94.39 %	0.26 dB
S_{22}	2.62 %	0.41 dB	98.38 %	1.43 dB
NF	1.31 %	0.02 dB	94.41 %	0.06 dB
1-dB CP	0.68 %	0.09 dBm	98.06 %	0.22 dBm
IIP_3	3.31 %	0.05 dBm	98.33 %	0.15 dBm

Table 51: Résultats de prédiction du test alternatif en utilisant le Groupe 5 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.09 %	0.16 dB	88.62 %	0.44 dB
S_{11}	1.37 %	0.12 dB	91.4 %	0.31 dB
S_{22}	12.41 %	1.96 dB	53.59 %	4.86 dB
NF	1.3 %	0.02 dB	94.49 %	0.05 dB
1-dB CP	1.22 %	0.16 dBm	93.39 %	0.49 dBm
IIP_3	8.38 %	0.13 dBm	88.83 %	0.38 dBm

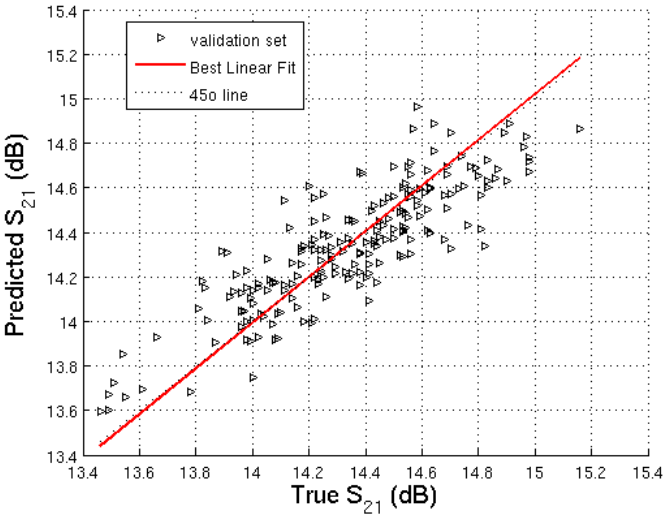
Table 52: Résultats de prédiction du test alternatif en utilisant le Groupe 6 de captuers non-intrusifs en considérant uniquement les variations D2D.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	0.72 %	0.1 dB	94.98 %	0.31 dB
S_{11}	1.35 %	0.12 dB	92.78 %	0.34 dB
S_{22}	3.03 %	0.48 dB	98.13 %	2.01 dB
NF	1.18 %	0.02 dB	95.14 %	0.05 dB
1-dB CP	1.09 %	0.14 dBm	94.14 %	0.39 dB
IIP_3	5.62 %	0.08 dBm	94.79 %	0.25 dB

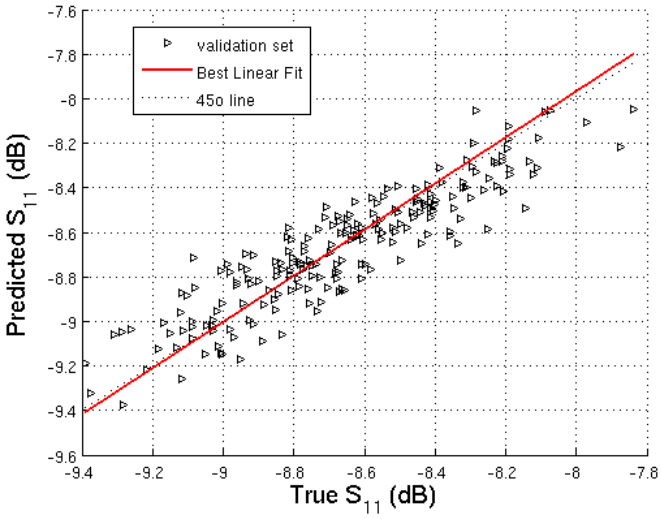
Table 53: Résultats de prédiction du test alternatif en utilisant le Groupe 5 de captuers non-intrusifs en considérant les variations D2D et WID.

Performance	Average RMS Error	Average Absolute RMS Error	Correlation Coefficient	Maximum Error
S_{21}	1.24 %	0.18 dB	84.25 %	0.5 dB
S_{11}	1.58 %	0.14 dB	89.95 %	0.37 dB
S_{22}	13.12 %	2.1 dB	55.96 %	5.26 dB
NF	1.16 %	0.02 dB	95.30 %	0.05 dB
1-dB CP	1.57 %	0.2 dBm	87.7 %	0.57 dBm
IIP_3	9.87 %	0.15 dBm	83.09 %	0.45 dBm

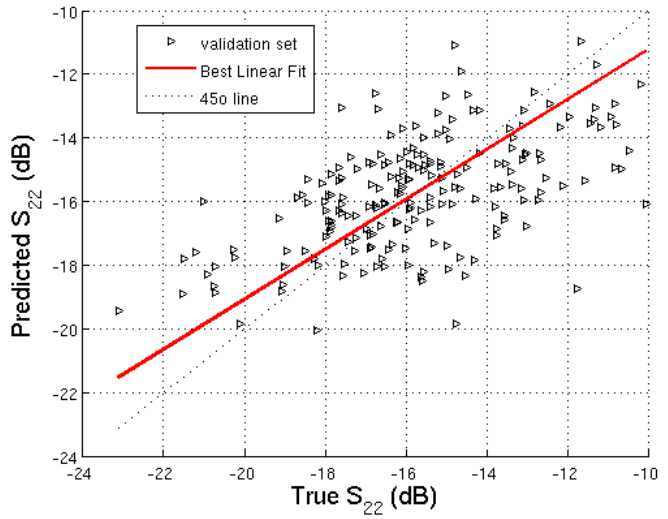
Table 54: Résultats de prédiction du test alternatif en utilisant le Groupe 6 de captuers non-intrusifs en considérant les variations D2D et WID.



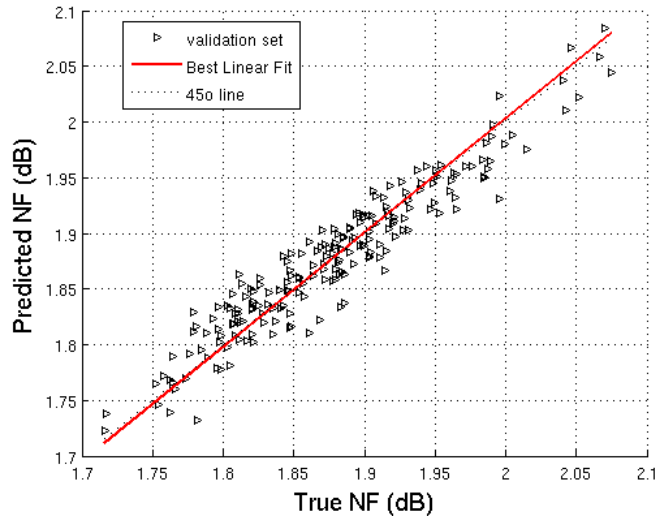
(a) S_{21}



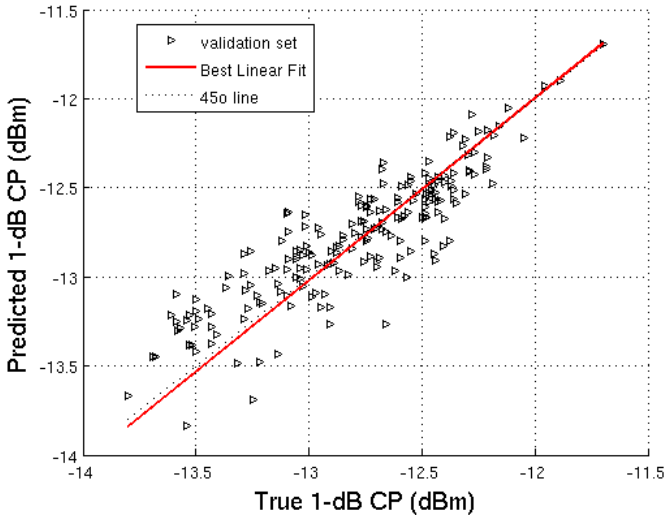
(d) S_{11}



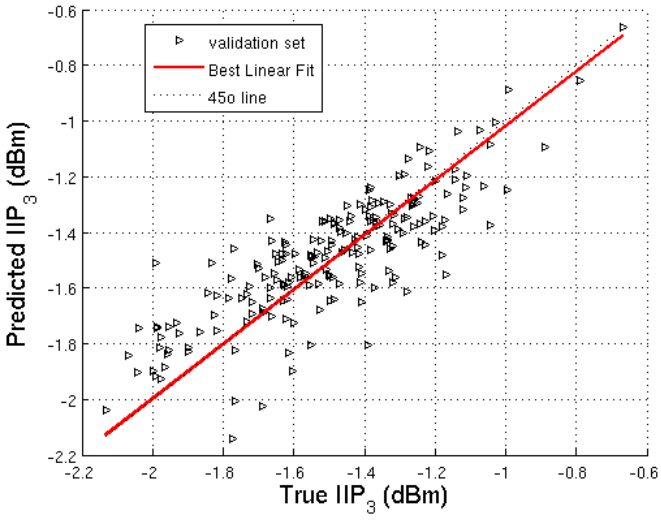
(b) S_{22}



(e) NF



(c) 1-dB CP



(f) IIP_3

Figure 72: Scatter plots of true vs. predicted values for all the performances of the LNA considering Group 7 of sensors. Both D2D and WID variations are considered.

10.3.2 *Cas d'étude 2 – LNA 3 étages millimétrique à 60GHz en technologie 65nm*

Le second cas d'étude de ce travail est un LNA à 60GHz composé de 3 étages en technologie 65nm. Les Figures 73 et 74 représentent respectivement les schémas du LNA et les capteurs non-intrusifs sélectionnés.

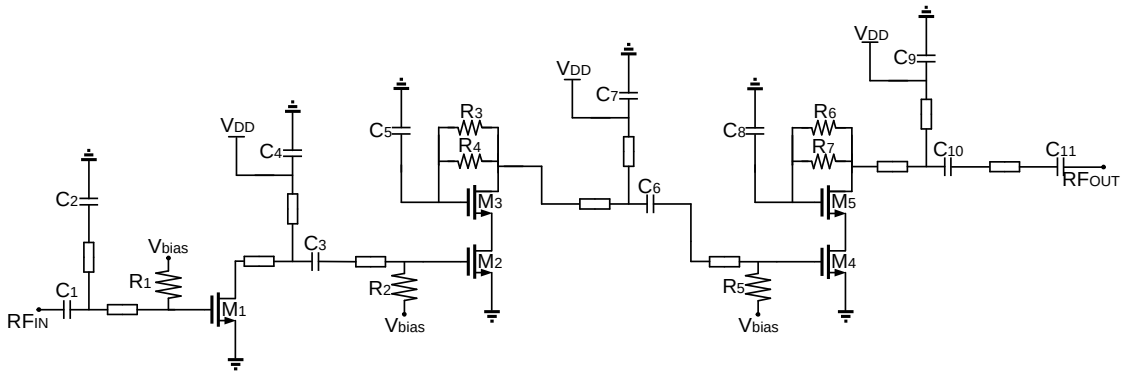


Figure 73: 60GHz mm-wave 3-stage LNA schematic implemented in 65nm CMOS bulk technology.

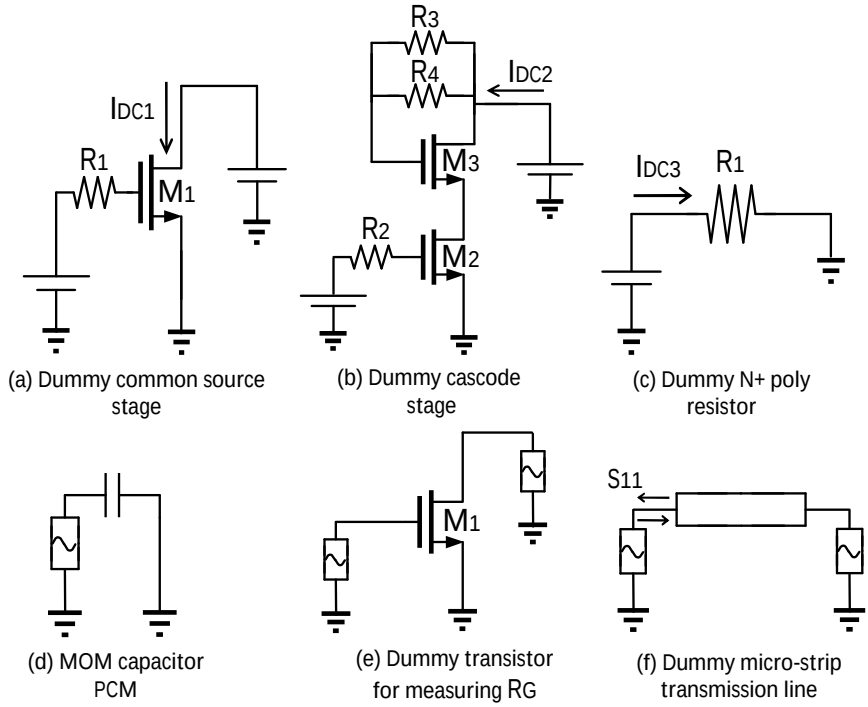


Figure 74: Variation-aware sensors built from identical components that exist in the topology of the 60GHz LNA.

Le LNA à 60GHz ainsi que les capteurs non-intrusifs ont été conçus en technologie 65nm bluk fournie par ST microelectronics. Les simulations ont été réalisées dans l'environnement Cadence Virtuoso en utilisant SpectreRF.

La Figure 75 montre la disposition montrant également la mise en place du LNA et les différents capteurs non intrusifs sur la matrice. La scène fictive de source commune, l'étage cascode factice, et la résistance factice sont placés dans la proximité physique à des structures correspondantes dans la LNA. A l'exception de leurs patins, ils peuvent être considérés comme ayant les frais généraux zéro depuis leur mise en page peut être mêlé à la disposition de la LNA et placés dans des zones qui sont laissés vides de façon à respecter les règles de conception électromagnétiques. Le condensateur factice a une valeur relativement faible de 4.3pF et est placé à proximité de ses plots GSG RF de manière à effectuer une mesure qui est aussi précis que possible. Notez que dans cette conception chacune de ces structures fictives a ses propres tampons dédiés. Cependant, étant donné qu'ils sont testés à courant continu ou basse fréquence, rien ne nous aurait empêché de multiplexage

leurs entrées et sorties de manière à utiliser seulement deux pads et de réduire la zone de tête de pad-dépendant. D'autre part, la structure factice pour mesurer Rg et la ligne de transmission fictive occupent une grande surface sur la matrice. Une grande partie de cette zone est consacrée aux structures de-enrobage qui devraient être utilisés de manière à estimer et enlever le parasite des plages RF GSG. Comme nous le verrons plus tard, alors que Rg est très utile pour prédire les variations des performances, y compris la ligne de transmission factice dans l'ensemble des capteurs non intrusifs ne propose pas un avantage important et, par conséquent, nous pouvons aussi bien choisir de l'enlever.

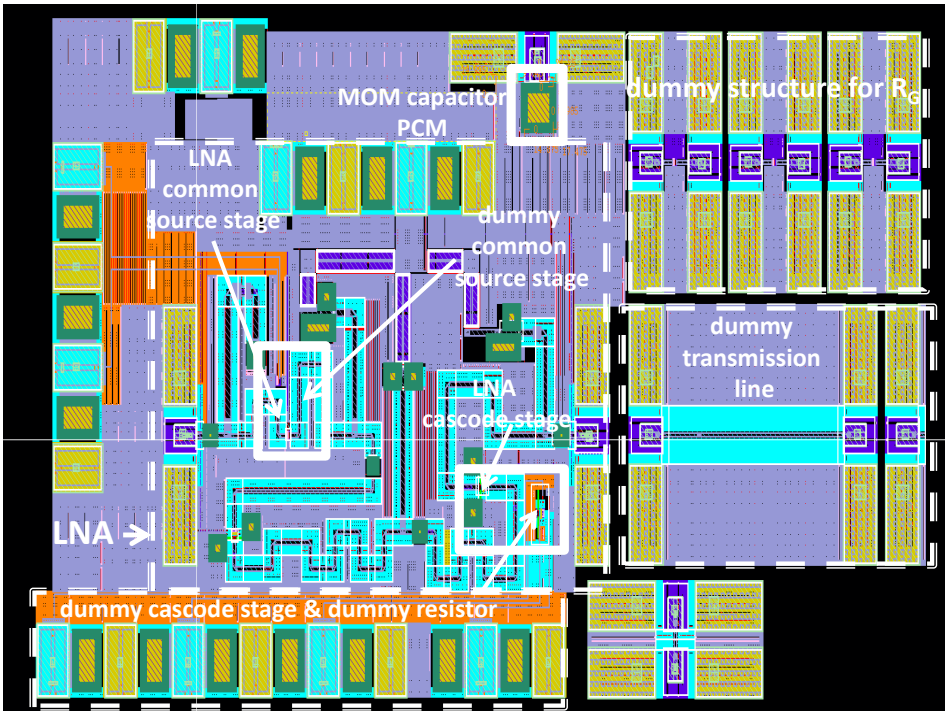


Figure 75: Layout view of the 60GHz mm-wave LNA and the embedded non-intrusive variation-aware sensors.

Performance	Standard Deviation (a)	Standard Deviation (b)
S_{21} (dB)	0.87	0.97
S_{11} (dB)	1.04	1.09
S_{22} (dB)	2.49	2.4
NF (dB)	0.23	0.27
IIP_3 (dBm)	0.64	0.69

Table 55: Standard deviation of 60GHz LNA performances when statistical variations in transmission lines are (a) disabled and (b) enabled.

Set of Sensors	Alternative Measurements
A	$I_{DC1}, I_{DC2}, I_{DC3}$
B	$I_{DC1}, I_{DC2}, I_{DC3}, C_1$
C	$I_{DC1}, I_{DC2}, I_{DC3}, C_1, R_G$
D	R_G
E	$I_{DC1}, I_{DC2}, I_{DC3}, C_1, R_G, S_{11}$

Table 56: Sets of non-intrusive sensors and corresponding groups of alternative measurements.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	4.83 %	0.79 dB	64.07 %	3.08 dB
S_{11}	6.41 %	0.98 dB	47.82 %	2.41 dB
S_{22}	9.68 %	1.74 dB	64.31 %	4.16 dB
NF	4.74 %	0.27 dB	12.55 %	0.62 dB
IIP_3	5.11 %	0.63 dBm	34.55 %	1.45 dBm

Table 57: Alternate test prediction results using Set B of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	3.77 %	0.62 dB	80.03 %	2.36 dB
S_{11}	3.42 %	0.53 dB	88.39 %	1.37 dB
S_{22}	9.68 %	1.74 dB	63.78 %	5.19 dB
NF	2.71 %	0.15 dB	81.40 %	0.41 dB
IIP_3	3.13 %	0.38 dBm	81.09 %	1.07 dBm

Table 58: Alternate test prediction results using Set C of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	5.60 %	0.92 dB	46.65 %	4.35 dB
S_{11}	5.09 %	0.78 dB	71.82 %	2.42 dB
S_{22}	12.68 %	2.28 dB	8.25 %	7.24 dB
NF	2.69 %	0.15 dB	81.73 %	0.41 dB
IIP_3	3.54 %	0.43 dBm	75.01 %	0.93 dBm

Table 59: Alternate test prediction results using Set D of non-intrusive sensors.

Performance	Average RMS Error	Absolute Average RMS Error	Correlation Coefficient	Maximum Error
S_{21}	3.16 %	0.52 dB	86.40 %	1.69 dB
S_{11}	3.14 %	0.48 dB	90.25 %	0.95 dB
S_{22}	9.40 %	1.69 dB	65.76 %	4.11 dB
NF	2.29 %	0.13 dB	87.14 %	0.38 dB
IIP_3	2.39 %	0.29 dBm	89.63 %	0.64 dBm

Table 60: Alternate test prediction results using Set E of non-intrusive sensors.

10.4 CONCLUSIONS

Nous démontrons qu'en ajoutant quelques capteurs non-intrusifs sur la puce, qui n'engendrent pratiquement pas de surcout de surface, et en obtenant de ces capteurs non-intrusifs certaines mesures dans le domaine continu et à basse fréquence, nous sommes capable de suivre les variations de toutes les performances du CUT avec une erreur de prédiction moyenne inférieure à l'écart-type de la performance, et une erreur de prédiction maxi-

mum qui est inférieure ou au moins comparable aux erreurs de mesure dans un équipement de test automatisé conventionnel.