5G 28 GHz high efficiency integrated phased array transceivers

Thibaut Despoisse

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

SPÉCIALITÉ ÉLECTRONIQUE

Par Thibaut DESPOISSE

5G 28 GHz High Efficiency Integrated Phased Array Transceivers

From System-Level Specifications to CMOS SOI Front-End Module Circuits

Sous la direction de : Nathalie DELTIMPLE
(Co-directeur : Anthony GHIOTTO, co-encadrant : Pierre BUSSON)

Soutenue le 23 octobre 2020

Membres du jury :

Mme	Patricia DESGREYS	Professeur	Telecom Paris	Rapporteur
M.	Serge VERDEYME	Professeur	Université de Limoges	Rapporteur
M.	Eric KERHERVE	Professeur	Bordeaux INP	Président
M.	Bernard JARRY	Professeur	Université de Limoges	Examinateur
M.	Alain KARAS	Ingénieur	Safran DS	Examinateur
Mme.	Nathalie DELTIMPLE	Maître de conférence	Bordeaux INP	Directrice de thèse
M.	Anthony GHIOTTO	Maître de conférence	Bordeaux INP	Co-directeur
M.	Pierre BUSSON	Ingénieur	STMicroelectronics	Co-encadrant

« *Faites que votre tableau soit toujours une ouverture au monde* »

Léonard de Vinci

Remerciements

Avant toute chose, je remercie sincèrement Madame Patricia Desgreys, professeur à Telecom Paris, et Monsieur Serge Verdeyme, professeur à l'Université de Limoges, d'avoir accepté d'être les rapporteurs de mes travaux de thèse. Je remercie également Messieurs Bernard Jarry et Alain Karas, respectivement professeur à l'Université de Limoges et ingénieur chez Safran Data Systems, d'avoir examiné mon travail et pris part au jury de la soutenance. Enfin, Monsieur Eric Kerherve, professeur à Bordeaux INP et directeur du laboratoire commun « ST – IMS », d'avoir présidé de ce jury.

Je tiens à exprimer toute ma gratitude à la directrice de cette thèse, Nathalie Deltimple, pour m'avoir convaincu de prendre part à cette aventure et pour son aide tout au long de ces presque quatre années. A Anthony Ghiotto, co-directeur de cette thèse, pour la confiance qu'il m'a accordé dès ma première année à l'ENSEIRB MATMECA, où il était mon enseignant. A Pierre Busson, co-encadrant industriel, pour avoir su trouver le sujet qui m'a séduit, et pour sa disponibilité tout au long de cette thèse.

Cette thèse s'est déroulée au sein du laboratoire commun « ST – IMS ». Je remercie donc chaleureusement Monsieur Yann Deval, directeur du laboratoire IMS, de m'avoir accueilli dans ses locaux. Je remercie également tous les membres du groupe Conception et de l'équipe Circuits et Systèmes Hyperfréquence (CSH) pour leur aide, tant technique qu'humaine ; et particulièrement Magali sans qui nos circuits ne pourraient pas être mesurés, Manuel qui a donné beaucoup de son temps pour que lesdits circuits fonctionnent, et Simone pour son aide précieuse. Enfin, je remercie les doctorants du laboratoire qui ont fait de ces années une expérience très enrichissante. Je pense notamment à John et à mes compagnons d'organisation de conférences : David et Samuel (Dr. Redois).

Je remercie également mes collègues de STMicroelectronics Grenoble et Crolles, qui m'ont toujours apporté leur aide lors de mes séjours Isérois : Pascal, Joël, Jérôme, Patrice, Jocelyn, Andreia, Fréderic ; et bien sûr mes voisins de box : Jean-Yves, Vincent et Jérémie qui m'ont beaucoup soutenu lors de la conception des différents circuits.

Enfin, un très grand merci à ma famille et à mes amis qui me soutiennent depuis toujours, et qui ont continué à le faire pendant ces années de travail. Ils m'ont aidé à prendre du recul lors des moments difficiles. J'ai également une pensée émue pour toi, papi, dont la curiosité, la culture et la rigueur ont toujours été des exemples pour moi. Merci.

Je terminerai par remercier ma compagne, Camille, pour son aide quotidienne précieuse et pour la patience exemplaire dont elle a fait preuve. Merci de m'avoir épaulé chaque jour. Je t'aime.

Thibaut Despoisse

Contents

Glossary

1G = First generation of mobile communications

2G = Second generation of mobile communications

3G = Third generation of mobile communications

4G = Fourth generation of mobile communications

3GPP = 3G partnership project

5G = Fifth mobile communication generation

AMPS = Advanced mobile phone system

AP = Access point

BB = Base band

BC = Body Contact

BEOL = Back end of line

BOX = Buried Oxide

BSIMSOI = Berkley short-channel IGFET model SOI

CAGR = Compound annual growth rate

CAPEX = Capital expenditure

CDMA-One = Code division multiple access one

CG = Common gate

CMOS = Complementary metal oxide semiconductor

CRAN = Cloud radio access network

DAC = Digital to analog converter

D-AMPS = Digital – advanced mobile phone system

DCS = Digital cellular system

DIBL = Drain-induced barrier lowering

DMBA = Digital multi-beam antenna

EDGE = Enhanced data rates for GSM evolution

eMBB = Enhanced mobile broadband

ESAR = Electronically steered array radar

ETSI = European telecommunication standards institute

E-UTRAN = Evolved universal terrestrial radio access network

EVM = Error vector magnitude

FB = Floating body

FD = Fully depleted

FDD = Frequency division duplex

FDMA = Frequency division multiple access

FEM = Front-end module

FET = Field effect transistor

FoM = Figure of merit

FSPL = Free space path loss

GND = Ground

GPRS = General packet radio service

GSG = Ground-signal-ground

GSM = Groupe speciale mobile

HBT = Heterojunction bipolar transistor

HLC = Home location register

HR = High resistivity

HSDPA = High-speed downlink packet access

HSPA = High-speed packet access

HSUPA = High-speed uplink packet access

IC = Integrated circuit

IF = Intermediate frequency

IGFET = Insulated gate FET

IL = Insertion loss

IMS = IP multimedia system

IoT = Internet of things

IP = Internet protocol

IS-95 = Interim standard 95

IMT = International mobile telecommunications

ISD = Inter site distance

ITU = International telecommunication union

ITU-R = ITU radio-communication sector

KPI = Key performance indicator

LNA = Low noise amplifier

LO = Local oscillator

LOS = Line of sight

LP = Low power

LTE = Long term evolution

LTE-A = Long term evolution – advanced

MBA = Multi-beam antenna

MBPAA = Multi-beam phased array antenna

MCS = Modulation and coding scheme

MEMS = Micro electro-mechanical systems

MHC = Measurement to hardware correlation

MIMO = Multiple-input multiple-output

mMTC = Massive machine type communications

MOM = Metal-oxide-metal

NF = Noise figure

NLOS = Non-line of sight

NMT = Nordic mobile telephone

NYU = New York university

O2I = Outdoor to indoor

OFDMA = Orthogonal frequency division multiple access

OPEX = Operational expenditure

OS = Open-short

PA = Power amplifier

PAE = Power added efficiency

PAPR = Peak to average power ratio

PBO = Power back-off

PD = Partially depleted

QAM = quadrature amplitude modulation

QoS = Quality of service

QPSK = Quadrature phase shift keying

RFFE = Radio frequency front-end

RTPS = Reflection-type phase shifter

RX = Reception / Receiver

SDMA = Spatial division multiple access

SMS = Short message service

SNR = Signal to noise ratio

SOI = Substrate on insulator

SOS = Substrate on sapphire

SPDT = Single pole double throw

TACS = Total access communication system

TDD = Time division duplex

TDMA = Time division multiple access

TL = Thru-line

TR = Trap-rich

TRL = Thru-reflect-line

TRx = Transceiver

TX = Transmission / Transmitter

UE = User equipment

ULA = Uniform linear array

UMi = Urban micro

UMTS = Universal mobile communication system

URA = Uniform rectangular array

URLLTC = Ultra-reliable and low latency communications

UTBB = Ultra-thin body and BOX

WCDMA = Wideband code division multiple access

WRC-19 = 2019 world radio conference

Introduction

The limitations of the current mobile communication generation (4G) are being reached. Indeed, the available bandwidth in the sub 6 GHz spectrum is no longer sufficient to cope with the increasing needs of higher data rates and more simultaneous users. Thus, a new generation is being developed: 5G. It will target different use cases from internet of things to high data rate communications, including critical applications such as autonomous vehicles and remote surgery. To cope with these highly demanding needs, innovative solutions are being studied. This thesis work focuses on the high data rate use case. The expected improvements to meet the specifications are the use of the millimeter-wave spectrum, network densification, and spatial multiplexing. This leads to degraded circuit performances, and to the need of many more components. Indeed, the number of required front-end modules may be a hundred times higher due to network densification and up to a thousand times higher due to the implementation of phased array transceivers. Consequently, this leads to drastic increases of the network deployment cost and the global power consumption. However, the use of phased arrays enables to relax the system requirements, and therefore its power consumption. Besides, it also enables to move from III-V circuits to less expensive advanced CMOS technologies. Hence, the aim of this thesis is twofold: determine the best suitable system architecture, meeting high data rate 5G requirement while minimizing the power consumption; and design millimeter-wave circuits using an advanced CMOS technology. It is targeted to implement RF switches and low noise amplifiers for 5G base station front-end modules

The first chapter is dedicated to the system-level study. Firstly, a use case is defined among the different 5G scenarios: high data rate communications in crowded urban environment. Then, a methodology to derive the system-level specifications from the standard early requirements is implemented. It is applied to the selected use case. Thus, several base station architectures meeting the specifications have been identified. Finally, a new methodology has been developed in order to estimate the power consumption of the different solutions. This enables to determine the best suitable topology in terms of performances and costs for a defined use case. Finally, it is applied to the selected scenario. However, both developed methodologies may be extended to other 5G scenarios, since they are easily scalable. The following chapters detail the design of millimeter-wave switches and low noise amplifiers for 5G front-end modules. The selected advanced CMOS technology is also presented. Besides, several runs have been achieved during this thesis work. They are depicted in Figure 0.

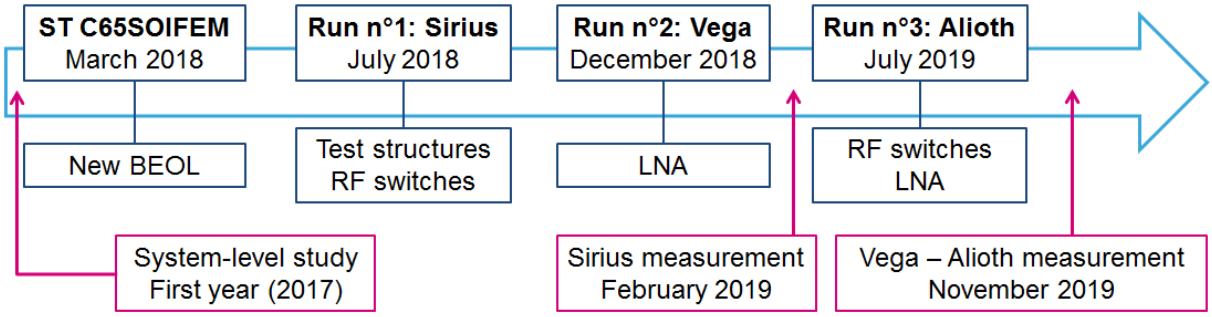

Figure 0: Overview of the runs achieved during the thesis work

The second chapter presents the circuits implemented in the first two runs (Sirius and Vega). The first aims are to characterize the new *ST C65SOIFEM* technology and to validate its suitability for the design of millimeter-wave front-end modules. Thus, several test structures and series-shunt RF switches are implemented and measured. A Ka-band low noise amplifier is also designed. These first steps enable to test the technology model accuracy and to validate the design methodology.

In the third chapter, improved versions of the RF switches and the Ka-band low noise amplifier are presented. Indeed, several drawbacks of the first designed circuits have been highlighted by accurate retro-simulations. A new asymmetrical RF switch is therefore designed, offering both low insertion loss and high power handling. The low noise amplifier is also improved to provide lower noise figure and higher gain, with consistent power consumption. These circuits are implemented in the third run, known as Alioth. After the analysis of the measurement results, some co-integration perspectives are discussed.

Chapter 1: From 5G Expectations to FEM Specifications

Daily life is increasingly a connected life. From the smartphone massive adoption to the rapid emergence of the internet of things (IoT), people have ever more connected objects, which require ever more data rate. In [1], Cisco expects 47% compound annual growth rate (CAGR) for mobile data traffic between 2016 and 2021. It is assumed that more than 80% of these data will be spent through smartphones. Moreover, 23% CAGR is expected for connected wearable devices as shown in Figure 1.

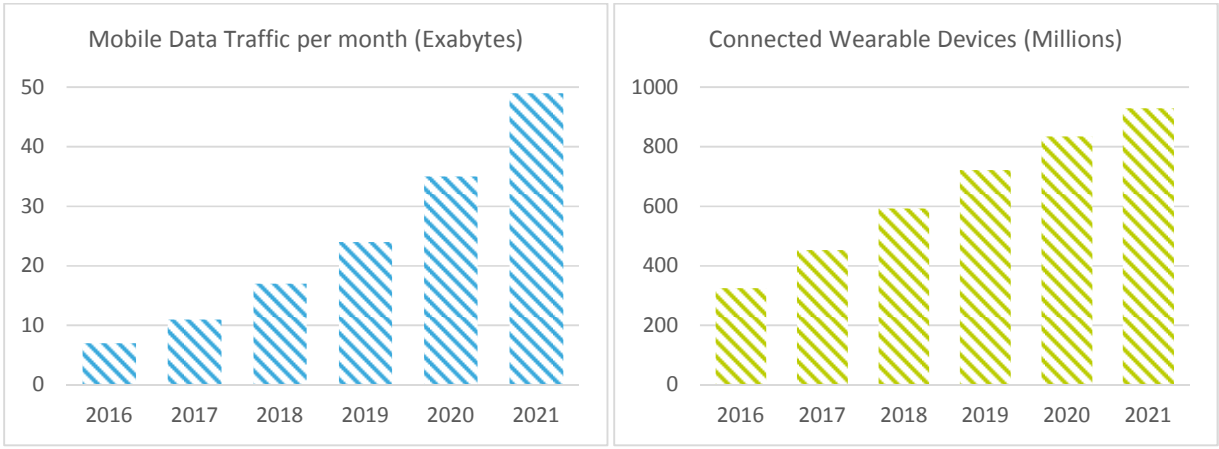

Figure 1: Mobile data traffic per month and global connected wearable devices, from [1]

To cope with the exponential number of connections and the rise of many new applications, an innovative solution is needed. That's why the fifth mobile communication generation (5G) is being developed. As detailed in [2], this new standard is willing to unify the different medium of wireless communication. From connected health, to virtual reality, through autonomous vehicles, the challenges are diverse and complex. In order to clearly explain why 5G is a paradigm shift in the telecommunication world, this chapter starts with an overview of the previous generations from the beginning of mobile communications to the current long-term evolution advanced (LTE-A) network. Then, the different 5G key performance indicators (KPI's) and the main expected solutions are presented and discussed. The deployment timeline and the main market actors are also introduced.

Secondly, the scope of the thesis is defined choosing practical use cases among prospective scenarios. The main challenges for system design are highlighted, and the system-level specifications are determined depending on the chosen use case KPI's. A new flexible methodology is set up to size and compare topologies which fill the defined requirements. A key metric is introduced: system power consumption.

This chapter ends with a state of the art of 5G Ka-band transceivers for high data rate wireless communications. Then, the system power consumption model is detailed, and the best suitable architecture for the given scenarios is highlighted. Finally, some perspectives are given as the power consumption estimation refining, and the Ka-band integrated circuit design.

I. On the Road towards the Wireless Communication Paradigm Shift

1. Legacy and Current Telecommunication Networks

a. 1G – The Foundation of Mobile Communications

The first generation of mobile communications (1G) was launched in the early 80's. It consisted in the analog modulation of user's voice, transmitted thanks to base stations, as presented in [3]. The major innovation was the ability to move while communicating. Even if 1G has disappeared today, the first mobile communication network laid the foundations for all the next generations.

The main feature in 1G is the emergence of a cellular network. It is composed of base stations which defines the cells, and the backhaul network which coordinates the communications. The base stations are spread out over the territory. Each one operates at a defined frequency to avoid interferences between neighboring cells. The number of different frequencies used to mesh the area is called the cluster size. It is used to choose a cluster size about 7 to limit the inter-cell interferences with a minor number of reused frequencies. A cellular network with a cluster size of 7 is presented in Figure 2. Moreover, a base station usually operates at 3 different frequencies. It enables to cover 3 cells simultaneously.

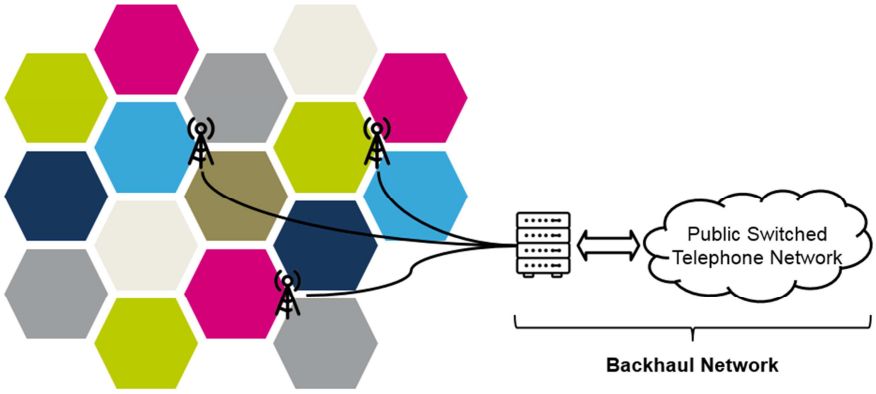

Figure 2: Sectorized cellular network with a cluster size of 7 and connection to the backhaul network

In a given cell, the frequency band is divided in 30 kHz sub-channels. Each one could only handle one mobile communication. As users are accessing the spectrum, at the same time, with different frequencies, it is frequency division multiple access (FDMA). Moreover, large guard bands are necessary between sub channels to avoid user interferences. Consequently, the spectrum efficiency is very poor. Moreover, the mobility is limited due to the difficulty to manage the move from a cell to another, as known as hard handoff. The lack of security is also a major concern for 1G. Indeed, the analog voice modulation makes the communications easy to be intercepted and deciphered.

Besides, it is the emergence mobile communication standards. It defines the frequency bands dedicated to mobile communications, the sub-channel and guard band sizes, and the inter site distances. However, each world region has set up its rules without any compatibility between them. The best know are the Nordic mobile telephone (NMT) for Nordic countries, the advanced mobile phone system (AMPS) for North America and Australia, and the total access communication system (TACS) for the United Kingdom [**3**]. Nowadays, this standard has disappeared. It has been replaced by digital mobile technologies.

b. 2G – The Beginning of Digital Telecommunications

As it was the case for 1G, the second generation of mobile communications (2G) is composed of different standards, including, digital AMPS (D-AMPS) and *groupe speciale mobile* (GSM), which are improvements of AMPS and NMT, respectively. Their working principles are very similar. Both are based on time division multiple access (TDMA). With this new spectrum access technique, several users could share a sub-channel using different time slots. As shown in Figure 3, GSM sub-channels are split in 8 time slots. It could be shared by 7 different users. The last slot is dedicated to time synchronization, and short message service (SMS), whereas, D-AMPS slots could only be shared by 3 users. That's why the GSM standard, established by the European Telecommunications Standards Institute (ETSI) in 1990, has been quickly adopted in a great part of the world. Actually, it is made by 3 standards: GSM 900, GSM 1800, also known as digital cellular system (DCS), and GSM 1900. They respectively operate at 900, 1800, and 1900 MHz.

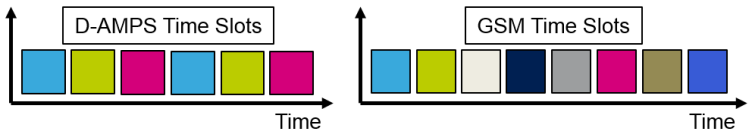

Figure 3: Sub-channel splitting in time slots for D-AMPS and GSM standards

The main change compared with 1G is the shift from analog to digital modulation. It provides more security and flexibility. However, the transmitter and receiver architectures are more complex since an analog to digital conversion stage is needed as shown in Figure 4 for transmitters. It may be highlighted that 2G is the generation of mass mobile adoption thanks to technology scalability which reduces user terminal cost.

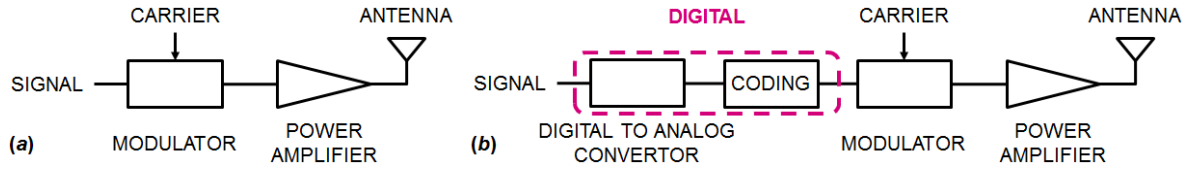

Figure 4: (a) Analog transmitter architecture; (b) Digital transmitter architecture

Moreover, up to [4], 2G has been improved twice to enable data transmissions. Firstly, with the general packet radio service (GPRS). The data rate is only between 9, and 21 kbps per time slot, and is limited to 40 kbps per user. GPRS is also called 2.5G as it is the first step towards the third generation (3G). Secondly, there is the enhanced data rates for GSM evolution (EDGE) which offers a data rate up to 60 kbps per time slot. This standard, sometimes known as 2.75G, could provide up to 200 kbps per user.

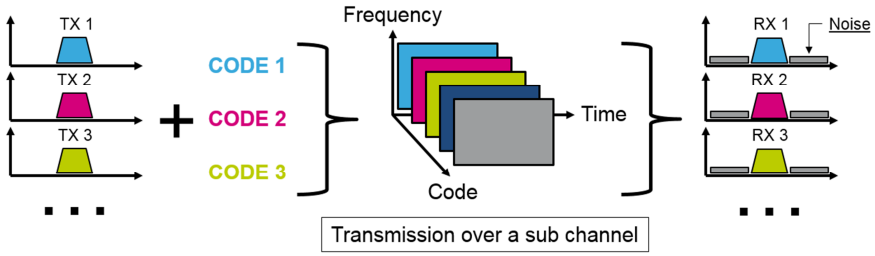

Figure 5: Transmission and reception of CDMA telecommunications

Finally, 2G laid another foundation for next generations with a less well-known standard: interim standard 95 (IS-95) also called code division multiple access one (CDMA One). First of all, this option was left behind due to GSM standardization. Then, its capacity, more than 4 times higher than GSM, makes it the favorite candidate for 3G. Actually, CDMA enables users to share the same time and spectrum resource thanks to a code assigned to each user in a sub channel. Consequently, the spectrum efficiency is considerably improved as shown in [3]. Figure 5 depicts the principle of CDMA standard. First of all, data for each user are coded independently before the transmission (TX). The signals are transmitted over a unique sub channel at the same time. Finally, each receiver (RX) deciphers the signal with its corresponding code. Thus, desired data are received and the others are dissolved in the noise.

c. 3G – Towards Broadband Mobile Communications

Due to its popularity, 2G quickly reached its limitations. So, a new standard has been launched to cope with the significant demand: the universal mobile communication system (UMTS). Firstly, 3G has been standardized by ETSI. Then the *IMT-2000* project defined the standard rules. It was led by the 3G partnership project (3GPP) forum, and the international telecommunication union (ITU). Initially, two candidates, both inspired by CDMA One, were competing. The most popular is wideband code division multiple access (WCDMA), which was then replaced by high-speed packet access (HSPA) as mentioned in [3]. It is also known as “Release 5”. It is important to notice that 3G is complementary with the previous generation. The new base stations (known as Node B) are deployed in crowded areas to provide broadband communications and enhanced internet access. Whereas 2G still covers large part of the territory with reliable communications. Even if new base stations are deployed, the core network is shared by the two co-existing standards, as presented in Figure 6. Voice and data from both 2G and 3G base stations are respectively split towards circuit and packet switching. This kind of network implies the deployment of an expensive backhaul, based on optical fiber. Finally, the home location register (HLC) manages mobile authentication on the network, and route signals towards the proper nodes. It is the emergence of the signal control plane.

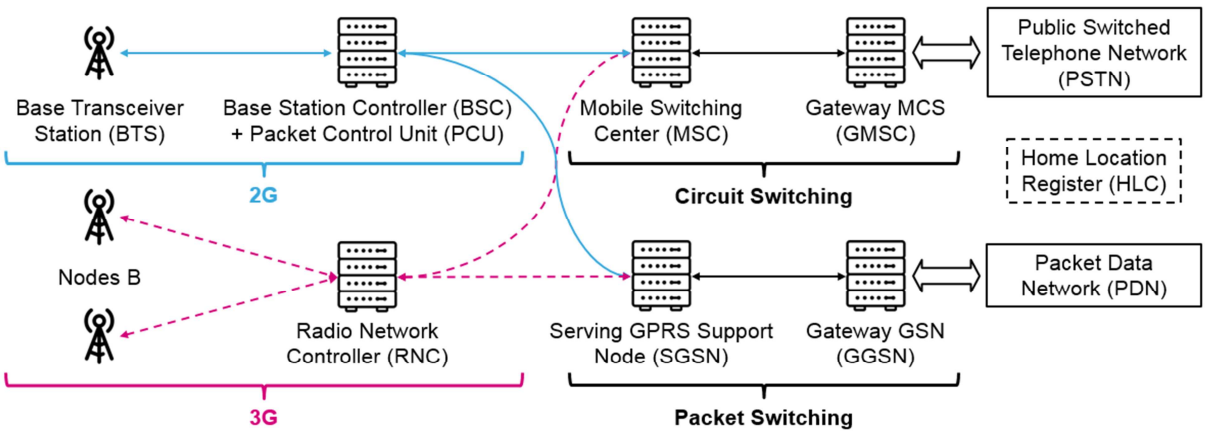

Figure 6: Presentation of the co-existing 2G and 3G networks

HSPA has evolved twice. Firstly, with high speed downlink packet access (HSDPA) to improve data rates from the base station to the user terminals up to 63 Mbps. Secondly, with high-speed uplink packet access (HSUPA) to enhance uploading from mobile phones [4]. The merge of these two standards is known as HSPA+.

d. 4G – Towards a Unique Mobile Communication Standard

Today, 3G is widely deployed and provides a reliable data rate. However, the use of 5 MHz channels is mandatory with HSPA+. Consequently, flexibility is limited [5]. Moreover, the theoretical channel capacity limit, defined by C. Shannon in 1948, has been reached. It is given by *Equation (1)*, where *C* (bit/s) is the channel capacity limit, *B* (Hz) is the available spectrum, and *S/N* is the signal to noise ratio before the demodulation, i.e. at the end of the receiving chain.

$$C = B.\log_2\left(1 + \frac{S}{N}\right) \quad (1)$$

The emergence of the fourth generation of mobile communications (4G), also known as Long Term Evolution (LTE), is therefore motivated by the needs of higher capacity and flexibility. To face these challenges, several innovations are implemented. Firstly, a new spectrum access technique is used: orthogonal frequency division multiple access (OFDMA). It consists in splitting the signal over different sub-carriers, which are orthogonal to avoid interferences. It is combined with carrier aggregation, that is, several distinct sub-carriers could be used for the same user. Consequently, the sharing of time and frequency domains is very flexible. Secondly, LTE has introduced the use of several antennas, whereas previous generations had only one antenna broadcasting the signal for its whole cell. These kinds of systems are called multiple-input multiple-output (MIMO). It could operate in different modes, providing higher reliability, or improving the channel capacity. The Shannon's law is updated as in Equation (2), with M the number of transmitting antennas [4].

$$C = M.B.\log_2\left(1 + \frac{S}{N}\right) \quad (2)$$

Furthermore, LTE introduces major changes in network architectures. As presented in Figure 7, the base stations (named eNode B) are now connected between them thanks to X2 links. This new radio interface is called evolved universal terrestrial radio access network (E-UTRAN). The core network has also evolved. Data and control planes are split in order to highly reduce network latency. Finally, the IP (internet protocol) multimedia system (IMS) enables to merge voice and data. This unification into a single communication protocol makes a break, since previous generations started with different standards. However, LTE and its enhanced version (LTE-A) have to face spectrum congestion [6], as shown in Figure 8.

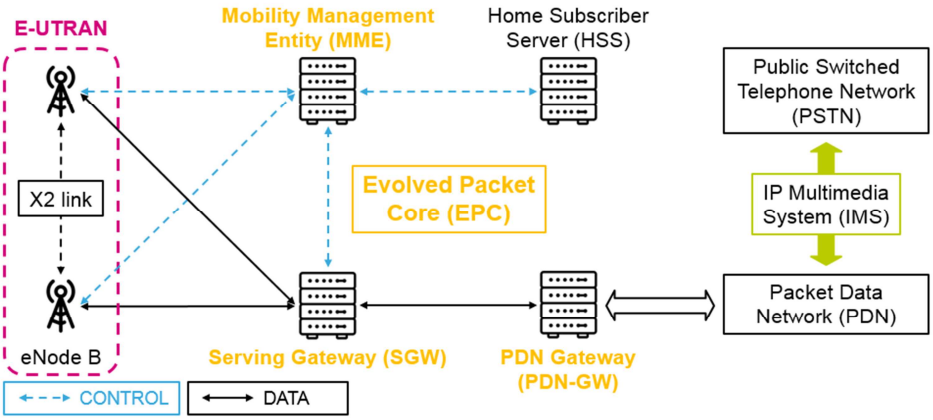

Figure 7: Presentation of the LTE network

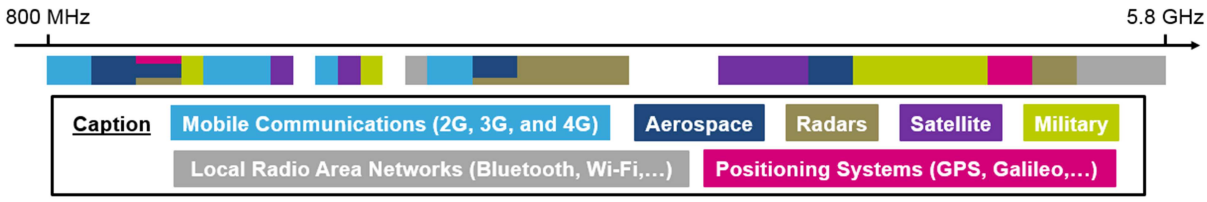

Figure 8: Simplified view of frequency spectrum users in France, from 800 MHz to 5.8 GHz

Finally, LTE-A provides high data rates, and improves network capacity. Network latency is also highly reduced. However, the mobile communications are still concentrated in the crowded sub-6 GHz spectrum. Moreover, the current standard is, as for previous generations, cell-phone focused. Thus, it is no longer appropriated regarding to the emerging applications such as IoT. LTE-A is also requiring more energy than the previous standards. In [7], it is shown that LTE-A power efficiency is 23 times worse than Wi-Fi, and even worse than 3G. To conclude, reducing the network power consumption while including the emerging mobile uses will be the major concern of the next mobile communication generations.

2. 5G Overview

a. The Emergence of 3 Standards

The upcoming 5G standard aims to reach unprecedented performance levels, as 20 Gbps peak data rate, or 10 million connections per km square. These targets have been defined by the radio-communication sector of international telecommunication union (ITU-R) in [8]. They launched the international mobile telecommunications (IMT) 2020 project, capitalizing on IMT 2000 (i.e. 3G), and IMT-Advanced (i.e. 4G), to set up 8 KPI's for 5G. Figure 9 presents these indicators, and gives a comparison with current levels. As these targets are too high to be reached simultaneously, 3 usage scenarios are defined.

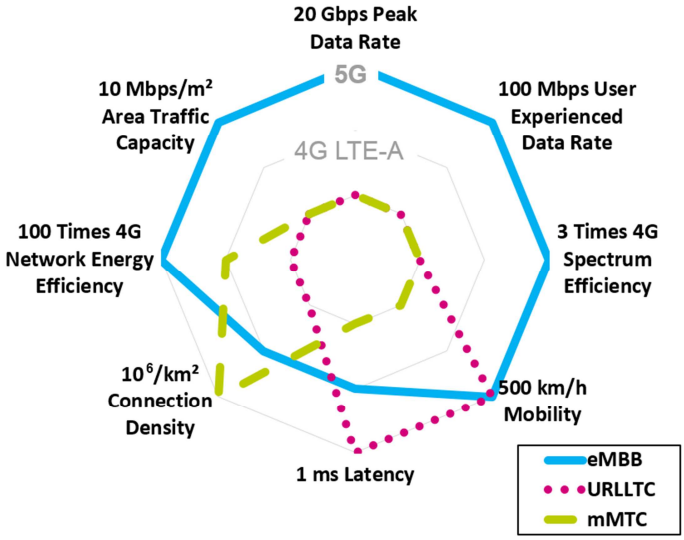

Figure 9: 5G KPI for 3 Different Scenarios

- Enhanced mobile broadband (eMBB): This first scenario is focused on high data rate communications for mobile user terminals. Consequently, high requirements are set on peak data rate, and area capacity. To reach the targets, spectrum efficiency has to be improved 3 times compared with LTE-A.

The implementation of a new network architecture is also scheduled. It will improve the user experience, including at cell edges. Besides, it is expected to severely reduce the global power consumption and increase the user mobility.

- Ultra-reliable and low latency communications (URLLTC): This second scenario is designed for critical applications such as health robotics, energy infrastructures, or factory processes. High reliability and extremely low latency are required. Up to [2], some use cases may also require important mobility (autonomous cars), or high data rates (collaborative gaming).
- Massive machine type communications (mMTC): Finally, the last case is tailored for large IoT networks, including wearable and e-health devices, or home and factory sensors. Consequently, massive connection density is needed in addition to good energy efficiency to ensure adequate operating time.

In a nutshell, 5G is pursuing the telecommunication standard unification as shown in [4]. On the one hand, it tends to build a whole connected world, but on the other hand, the demanding KPI's imply a splitting into three less demanding scenarios, which are closer to the usual mobile standard vision.

b. EMBB Spectrum and Network

This thesis work focuses on the EMBB scenario. Particularly, it is about providing extreme data rates into highly populated areas, while reducing the network latency, and the global power consumption. Furthermore, these researches are concentrated around the base station, i.e. on the radio interface (communications with the user), and on the backhaul (communication between base stations, and the core network). Consequently, mobility is not an EMBB targeted KPI here.

The first way to obtain higher data rates and network capacity is to use larger frequency bands, as shown by *Equation (2)*. But, as shown previously in Figure 8, the sub-6 GHz spectrum is very crowded. Even if a wide frequency band has been booked between 3.4 and 3.8 GHz for 5G, it is not enough to cope with the demanding requirements. Consequently, it is expected to move towards millimeter-wave [6]. Several bands, presented in Figure 10, have been identified during the 2019 world radio conference (WRC-19).

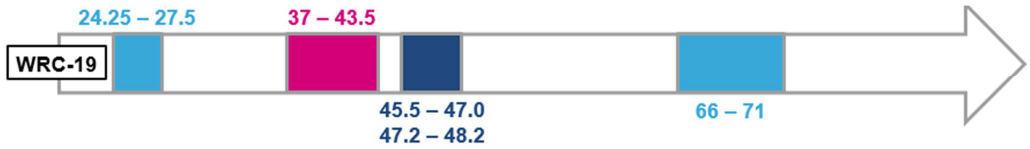

Figure 10: Millimeter-wave frequency bands (in GHz) selected at WRC-19

However, the use of higher frequencies for mobile communications also implies serious challenges. Firstly, as presented by *Equation (3)*, the wireless channel free space path loss (FSPL) increases with the signal frequency and the link length. In this equation, also known as Friis formula, *FSPL* is in dB, and the propagation distance (R) and signal wave length (λ) are both given in meters. The additional FSPL between usual LTE (f_1 = 2600 MHz) and 5G (f_2 = 28 GHz) communications is computed in *Equation (4)*. Considering the same range for both links, the additional losses are 20.6 dB. It should be noticed that the path loss given in *Equation (3)* is a negative value in dB, since it accounts for power attenuation.

$$FSPL = 20.\log\left(\frac{\lambda}{4\pi R}\right) \qquad (3)$$

$$FSPL_1 - FSPL_2 = 20.\log\left(\frac{\lambda_1}{\lambda_2}\right) = 20.6\ \text{dB} \qquad (4)$$

Secondly, as shown in Figure 11, the losses due to the atmospheric absorption are also stronger at higher frequencies. Actually, signal absorption increases with the frequency for the usual materials (air, water, glass, concrete, and wood) [9]. The additional atmospheric losses are not significant, but they are much stronger for other materials. For example, glass and concrete penetration losses are respectively 5.1 dB and 101.6 dB higher at 28 GHz compared with LTE (2600 MHz), for a typical building wall. Channel modeling is also more challenging at higher frequencies.

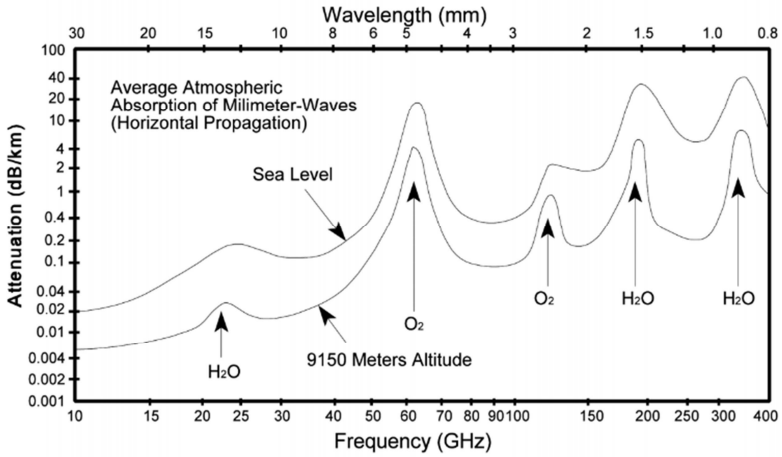

Figure 11: Average atmospheric absorption, from everythingrf.com

An accurate model for 5G communication channels is given in [9]. It includes the distinction between line of sight (LOS) and non-line of sight (NLOS) propagation within different environments. For the urban use case; the probability to have direct LOS between the access point (AP) and the user equipment (UE) is about 60% at 50 m, and decreases below 15% from 200 m. The associated NLOS propagation penalties are respectively 39 dB, and 53 dB. Consequently, it has a major impact on the communication reliability.

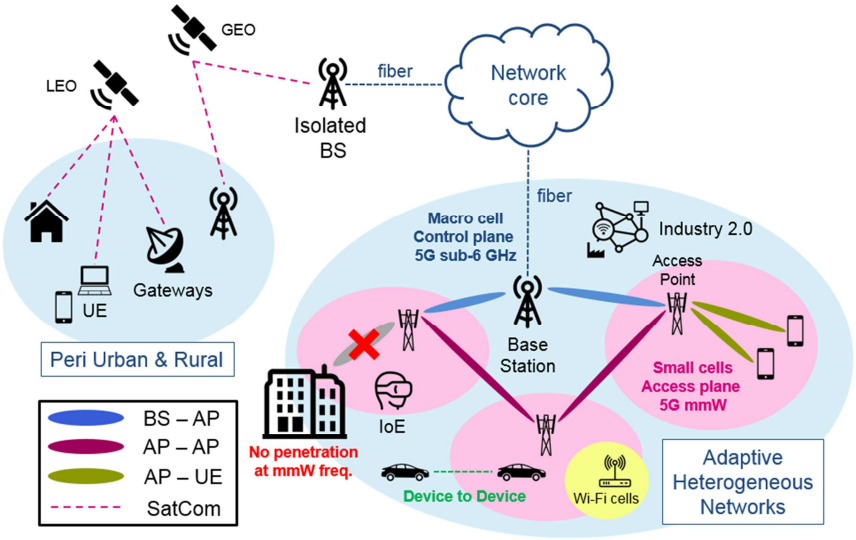

Figure 12: 5G Network, including small cells, and control and user plane splitting

To cope with the limited range of mm-wave communications, and to deal with the significant rise of user density, a new network architecture is needed. Several expected innovations are introduced in [10]. The first improvement is the implementation of many mm-wave small cells which provide high data rates in very crowded areas. This enables to shift from a base station centric to a user-centric network. Actually, the adaptive cells could change their configuration to fit properly with user density and to offer high quality of service (QoS) even at cell edges. Secondly, the split between control and user planes is stronger than previously. It is expected to route mm-wave signals from base stations to the core network. Consequently, the baseband processing is pushed out of the access point. This reduces both deployment complexity and network latency. The centralization of processing steps into virtual baseband unit pools is called cloud radio access network (CRAN). Finally, 5G also expects to benefit from the existing Wi-Fi network, particularly indoor, to offload outdoor facilities. The expected 5G network is shown in Figure 12.

To conclude, small cell densification enables to reduce the inter site distance to cope with high path loss but it also implies higher costs for mobile operators. In fact, more base stations have to be deployed and connected with the others so the capital expenditure (CAPEX) is increased. However, this cost is reduced thanks to CRAN and mm-wave backhauling. Actually, it replaces the usual fiber links which are very expensive. The global power consumption is also more important. Thus, the operational expenditure (OPEX) rises. Low-power small cell transceivers are required to minimize the operating costs. Finally, high interoperability with current LTE-A, and Wi-Fi standards is needed to optimize user quality of experience.

c. Beamforming: a Key Enabler

The use of mm-wave frequencies unlocks vast amounts of bandwidth but it also brings major drawbacks as propagation issues and lower circuit performances. However, the smaller wavelength is, the smaller the antennas are. Thus, it is possible to simultaneously use many antennas to improve the transceiver ranges, and support the network densification. This is the so-called beamforming. As proven in [11] with *Equation (5)*, the maximum antenna array directivity (D_{max}) depends on the signal wave length (λ) and array area (A). Thus, for a given transmitting antenna area, the frequency increase offers more advantages than drawbacks.

$$(D_{max})_{dB} = 20.\log\left(4.\pi.\frac{A}{\lambda^2}\right) \qquad (5)$$

In addition to a significant directivity increase, beamforming enables to dynamically modify the direction of the beam. As shown by Figure 13, the signal is divided among several paths. Then, it is delayed, and emitted through the antennas. Finally, the transmitted signals are combined over the air to form a focused beam. In uniform linear arrays (ULA), as illustrated in Figure 13, elements are equally spaced. Thus, the delays vary proportionally, from a line to the nearby one. For a ULA composed of N_{el} half wavelength spaced elements, the beam steering (ϑ) is given by *Equation (6)*. It depends on the element spacing (d), the wave speed (c), and the set delay (τ). Moreover, the 3 dB beam width (ϑ_3) is defined in *Equation (7)* for this kind of array. Consequently, larger arrays allow establishing more directive communications. This introduces spatial division multiple access (SDMA), i.e. the ability to spatially separate different users. Finally, for narrow band signals, i.e. if bandwidth is much lower than the carrier frequency, the time delay could be approximate by phase shifting (φ) as in *Equation (8)*.

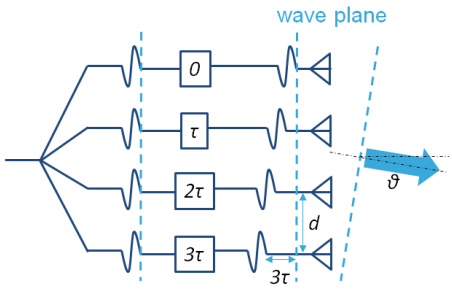

Figure 13: Beamforming operating principle

$$\vartheta = \sin^{-1}\left(\frac{c.\tau}{d}\right) \quad (6)$$

$$\vartheta_3 = \frac{1.772}{N_{el} - 1} \quad (7)$$

$$\varphi = \tau.\left(\frac{2\pi.c}{\lambda}\right) \quad (8)$$

The beamforming principle is applicable for both the transmitting and receiving arrays. *Equations (9)* and *(10)*, respectively from [12] and [13], demonstrate the benefits of these topologies. The received signal on each antenna is delayed and then combined to increase the signal-to-noise ratio (SNR) and filter uncorrelated noise. As a result, the receiving array noise figure (NF_{array}) is improved compared with the noise figure of a single element (NF_{el}). Likewise, the array effective isotropic radiated power ($EIRP_{array}$) is improved regarding to the elementary antenna radiated power ($EIRP_{el}$).

$$NF_{array}(dB) = NF_{el}(dB) - 10.\log(N_{el}) \quad (9)$$

$$EIRP_{array}(dB) = EIRP_{el}(dB) + 20.\log(N_{el}) \quad (10)$$

In a nutshell, larger arrays enable to overcome mm-wave communication drawbacks. Also, it introduces a new spectrum access technique. However, larger systems imply many more circuits. Depending on the chosen topology, overall power consumption and system complexity may increase in varying degrees. Indeed, several architectures are able to achieve beamforming. Up to [14], the first ones had been introduced in the early 1900's by K. F. Braun, for radar purpose. It consisted in a fixed beam, generated by a bulky and lossy passive phase shifter network. Then, the beam could be steered mechanically.

Since the 1940's, the beam could be steered changing the phase shift electronically. But the beamforming network is always behind the power amplification stage. Consequently, much power is lost causing power dissipation. A famous example is the electronically steered array radar (ESAR) designed in 1958. It has only been in the 1980's that phased array technology shifts from passive to active systems, thanks to circuit scaling. Actually, it was possible to integrate phase shifters or time delays before the amplifier to design more compact systems with better energy efficiency. This topology is illustrated in Figure 14 (*a*). Another consequence on phase shifter integration is the opportunity to move it towards lower frequencies, i.e. before the transmitter up conversion. It is also possible to operate the phase shifting in the local oscillator path. By the way, it is the case for the first fully integrated CMOS phased array receiver and transmitter introduced in [12] and [13], respectively. By now, phased arrays are going to become more digital thanks to the scaling of silicon technology nodes, significantly increasing the processing capabilities, as illustrated in Figure 14 (*b*).

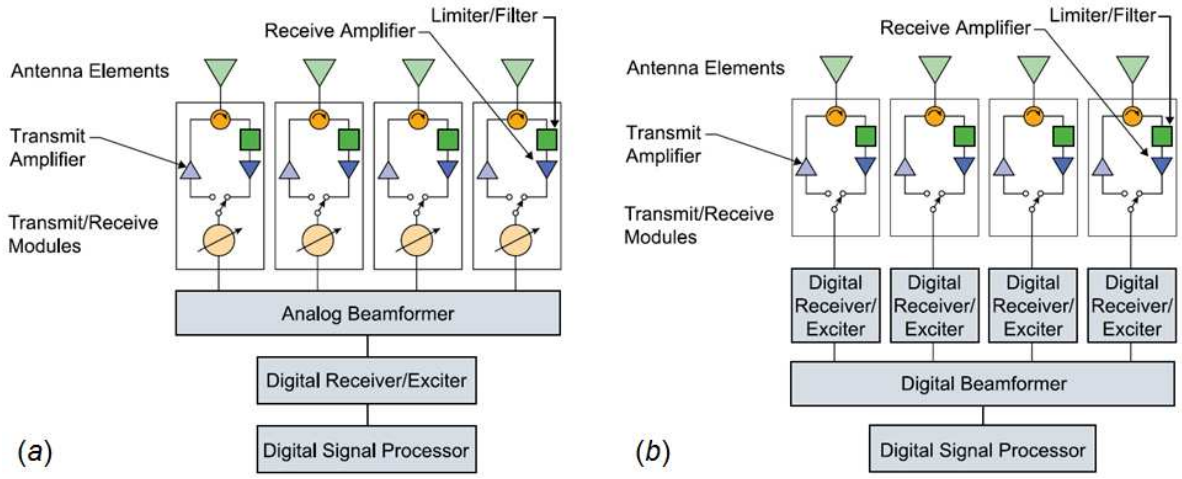

Figure 14: (a) Modern analog transceiver architecture; (b) Digital phased array transceiver architecture

d. The 5G Standardization Timeline

Following the general mobile standardization trend, 5G is expected to be released around 2020. Whereas previous generations have led the market, the 5G advent is triggered by the emergence of new uses. It could explain the large gap between current mobile networks and the next generation requirements. As shown in Figure 15, the deployment is expected to be very progressive, to cope with this large step. Since WRC-15, which has set the first 5G requirements, directions for below 6 GHz standardization, and recommendations for mm-wave frequency bands, many trials have been conducted. Up to [15], 40% of these trials are performed below 6 GHz, and 31% in the 24 – 29.5 GHz frequency band.

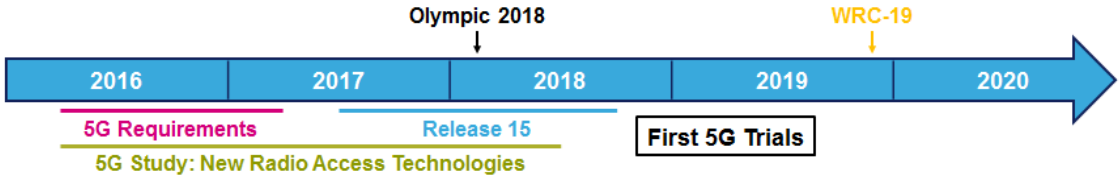

Figure 15: 5G Deployment Timeline, based on ARCEP Expectations [2]

As summarized in [10], market players have different targets. For example, *Qualcomm*'s vision is focused on enabling new uses and improving user experience, while *Samsung* is focused on multimedia experience and IoT, and *Huawei* on massive capacity and network deployment. Moreover, the mobile terminal market is split between sub 6 GHz, and mm-wave system vendors. There are few actors able to provide mm-wave circuits. Up to [16], *Qualcomm* leads the market with the *QTM525* radio frequency front-end (RFFE), and the *Snapdragon X55* modem. More recently, *Qualcomm* have introduced the *Snapdragon X60* [17]. This 5G modem-RF system based on the 5 nm process node enables to aggregate the sub 6 GHz and mm-wave spectrums. Then, some players are focused on modems, as *Samsung* with its *Exynos 5100*, and *Intel* with its *XMM 8160*. Others provide RFFE's as *Qorvo*, or *SkyWorks*.

However, the big 5G challenge is to ensure high network capacity. In this context, *Ericsson* has presented the first integrated mm-wave phased array antenna module in [18]. *Anokiwave* is also an early actor in mm-wave phased arrays [19]. For now, they exhibit a complete offer including beamforming ICs and

front-ends, for macro cells and small cells operating in the 28 GHz and 39 GHz frequency bands. They have developed advanced features such as the *zero-cal* technology and the *kinetic-green* functionality which reduce respectively the system complexity and power consumption [20]. Besides, *Movandi* has designed the *BeamX* module, including amplifiers and frequency converters. It is used with a repeater (*BeamXR*) to optimize NLOS communications. Finally, *SatixFy* has developed the first fully digital commercial beamformer, which confirms the digitalization of such systems [20]. For now, their 256 element array is only dedicated to satellite communications in the Ku-band (12 – 18 GHz). These first 5G developments are advertised by major public events, as 2018 Winter Olympics in PyeongChang and 2021 Olympics in Tokyo.

To summarize this overview, 5G is expected to be a paradigm shift in the wireless communication world, led by the ever more connected society. To fill the demanding requirements, the network will firstly be deployed in the current sub 6 GHz spectrum. At the same time, mm-wave trials are performed to prepare the move to higher frequencies.

II. From 5G EMBB Expectations to System-Level Specifications

1. Thesis Use Case: Definition & Challenges

a. Chosen Scenarios

Since 5G is a broad standard, with a lot of applications and as many frequency bands, it is necessary to bound the study with clear use cases. This work focuses on mm-wave frequency bands, as the sub 6 GHz network is currently being deployed. Consequently, the main target is providing high data rates in very crowded areas. As presented in Figure 9, based on ITU-R recommendations [8], it is expected to provide a 10 Tbps/km² network capacity. Moreover, the data rate offered to each user has to be at least 10 Mbps, and should reach 20 Gbps in optimum cases. These requirements are consistent with current population densities. For cities like Paris or West New York, it is about 20 000 people per km² [21].

One of the key principles of 5G radio access network is significant access point densification. Consequently, it is important to determine how many users will be related to each small cell. The cell different radiuses are detailed in [22]. Results are reported in Table 1.

Cell Type	Radius	Number of Users	Location
Femto Cell	10 to 100 m	Up to 30	Indoor
Pico Cell	100 to 200 m	From 30 to 100	Indoor & Outdoor
Micro Cell	0.2 to 2 km	100 to 2000	

Table 1: Mobile cell features for different kind of cells

To conduct an accurate study, the following work is limited to pico cells. However, the system sizing methodology is suitable for other scenarios. Considering an inter site distance (ISD) of 100 m, the cell density is about 100 cells/km². It implies a 100 Gbps/cell capacity, i.e. simultaneously providing 1 Gbps to 100 users in a single cell. Likewise, cells with ISD of 200 m have to deal with 400 concurrent users. These use cases are reported in Table 2, considering 6 sectors per access point.

Inter Site Distance	Cell Density	Users / Cell	User Data Rate	User / Sector
100 m	100 /km²	100	1 Gbps	17
200 m	25 /km²	400	0.3 Gbps	67

Table 2: Identified 5G scenarios

Finally, the chosen operating frequency, among available mm-wave bands, is 28 GHz. This allows taking benefit from the many recent developments. Based on previous parts, the mm-wave access point is mainly composed of the phased array transceiver. It both provides data for users, and ensures low-cost backhaul with fast deployment. Consequently, this work also takes into account AP to AP links.

b. First 5G Standard

The 5G standardization has been started in 2017 by the ETSI, with Release 15 [23]. The aim of this first release is to meet the crucial market needs thanks to existing technologies. Thus, it includes detailed guidelines for the sub 6 GHz spectrum and sparse directions for mm-wave frequencies (24.25 – 27.5, 26.5 – 29.5, and 37 – 40 GHz). Indeed, mm-wave system requirements are only given over the air, i.e. after the transmitting antenna array. These partial guidelines include the maximum channel bandwidth (400 MHz), the receiver sensitivity (from -86 to -109 dBm), and the maximum error vector magnitude (EVM) depending on the signal modulation (17.5% for QPSK, 12.5 % for 16-QAM, and 8% for 64-QAM). But some requirements are missing, as the maximum EIRP for the different kinds of small cells. This lack of information is expected to be bridged with Releases 16 (June 2020) and 17 (September 2021). They should target respectively the mm-wave use cases and the mMTC applications.

c. Challenges

The design of mm-wave systems for 5G small cells implies several challenges. Firstly, the requirements are very demanding, whereas the standardization is only at the beginning. It is therefore necessary to set up a sizing methodology to determine the system-level specifications. Then, the many studied solutions have to be compared. It enables to select the best suited solution for 5G small cell deployment. Finally, the significant densification implies high CAPEX and OPEX which have to be reduced. Thus, the design of 5G systems require a particular care about the overall power consumption.

2. Sizing Methodology for mm-Wave 5G Small Cells

a. Step 1 – Introduction

5G requirements are expected to be as broad as demanding. Thus, a new sizing methodology has been developed in this work. It enables to derive the system-level specifications from the early requirements. This methodology is summed up in Figure 16. It starts with the use case definition. Here, the previously presented use case is chosen. However, the sizing method may be applied to different requirements. The next steps are detailed in the following sections: the analysis of signal properties (modulation order, coding rate, and bandwidth) and propagation channels. Then, a link budget is performed. It is based on the performances of recently reported circuits, and it leads to the array sizing. Finally, the overall power consumption of the different phased array architectures is estimated to select the best suitable.

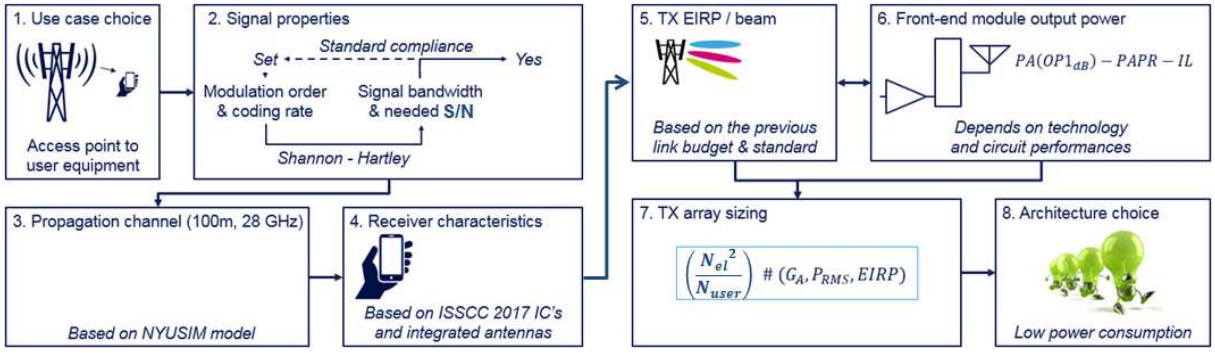

Figure 16: Main steps of the 5G access point sizing methodology

b. Step 2 – Signal Properties

This step consists in determining the optimal modulation and coding scheme (MCS), i.e. the modulation order, and the coding rate, which are respectively represented by *M*, and C_R in the following equations. It is important to notice that antenna arrays are supposed to operate in beamforming mode. So, the network capacity is only increased due to higher SNR. Actually, this kind of transceiver does not bring diversity as it is the case with previously presented MIMO systems. Consequently, the Shannon's law is applied using *Equation (1),* not *Equation (2)*. It is reminded that these equations are valid for additive white Gaussian noise channels. Finally, the communication channels are supposed to be independent for each user, as it is the case in [24] with the use of orthogonal beamforming. Consequently, the maximum data rate per user, termed R_{USER}, is equivalent to the channel capacity. This is expressed in *Equation (11)*. Moreover, *Equation (12)* gives the relation between the signal bandwidth, the data rate, and the MCS.

$$R_{USER} = B * log_2\left(1 + \frac{S}{N}\right) \quad (11)$$

$$B = \frac{R}{C_R * log_2(M)} \quad (12)$$

These equations enable to compute the required signal to noise ratio (*S/N*) and bandwidth (*B*) to ensure the targeted user data rate. The results for different MCS are reported in Table 3. This includes communications up to 2 Gbps for backhauling. The first remarkable result is that *S/N* only depends on the MCS. This could be shown thanks to *Equation (13)*, which is a combination of *Equation (11)*, and *Equation (12)*. Moreover, a lower coding rate implies a higher reliability but also a higher bandwidth. Consequently, the *S/N* requirements are relaxed. It stresses the tradeoff between signal bandwidth and required *S/N*, which both impact the system power consumption.

$$log_2\left(1 + \frac{S}{N}\right) = C_R * log_2(M) \quad (13)$$

Based on prior 5G specifications, which expected channel bandwidths about 100, 200, and 400 MHz, the 16-QAM modulation has been chosen for 300 Mbps communications, with 7/8 coding rate. It enables to minimize the occupied bandwidth, with relaxed EVM requirements regarding to 64-QAM. Likewise, higher data rates could be reached using the same coding rate, with higher modulation orders. However, the latest specifications suggest using lower *CR* to improve reliability [23]. Consequently, the coding rate is reduced to 3/4 for 300 Mbps communications, using 16-QAM. For higher data rates, i.e. higher

modulation orders, it is chosen to decrease C_R to 2/3. Thus, the maximum throughput per beam is 1.6 Gbps to respect the 400 MHz bandwidth limitation. Based on this knowledge, several beams may be required for backhaul. As these links should often be in LOS, the coding rate may be increased to reduce the occupied bandwidth.

R_{USER} = 300 Mbps		**QPSK**	**16-QAM**	**64-QAM**
C_R = 7/8	*S/N* (dB)	3.75	10.1	15.7
	B (MHz)	171	85.7	57.1
C_R = 3/4	*S/N* (dB)	2.62	8.5	13.4
	B (MHz)	200	100	66.7
C_R = 2/3	*S/N* (dB)	1.82	7.24	11.8
	B (MHz)	225	113	75
R_{USER} = 1 Gbps		**QPSK**	**16-QAM**	**64-QAM**
C_R = 7/8	*S/N* (dB)	3.75	10.1	15.7
	B (MHz)	570	286	190
C_R = 3/4	*S/N* (dB)	2.62	8.5	13.4
	B (MHz)	667	333	222
C_R = 2/3	*S/N* (dB)	1.82	7.24	11.8
	B (MHz)	750	375	250
R_{USER} = 2 Gbps		**QPSK**	**16-QAM**	**64-QAM**
C_R = 7/8	*S/N* (dB)	3.75	10.1	15.7
	B (GHz)	1.14	0.57	0.38
C_R = 3/4	*S/N* (dB)	2.62	8.5	13.4
	B (GHz)	1.33	0.67	0.44
C_R = 2/3	*S/N* (dB)	1.82	7.24	11.8
	B (GHz)	1.50	0.75	0.50

Table 3: Signal to noise ratio and signal bandwidth for different MCS's and data rates

c. Step 3 – Propagation Channel Modeling

As discussed previously, the selected operating frequency is 28 GHz. This implies many new challenges, as rain and atmospheric attenuation, specular reflection on buildings, and channel modeling. A lot of channel modeling studies have been gathered in [9]. This is supported by the New York University (NYU) Wireless mm-wave channel simulator: NYUSIM 2.01 [25]. It enables accurate modeling for wireless mm-wave links, as it is based on many field studies.

Input parameters are divided in 4 sections. The main objective of this part is to determine the channel path loss, in order to perform the link budget. Thus the *channel parameters* section is of particular interest. The simulator may take into account atmospheric features (pressure, humidity, temperature), environmental constraints (rain, foliage, building penetration), and the transceiver localization. Moreover, different scenarios could be tested. The chosen use case corresponds to the urban micro (UMi) environment. Finally, the simulator also includes LOS probability. The other sections are related to *antenna properties*, *signal consistency*, and *human blockage*. These features are optional for path loss computing. Thus, they are not considered in this work.

Range (m)	**Pressure (hPa)**	**Freq. (GHz)**	**Humidity (%)**	**Bandwidth (MHz)**
10 – 500	1013	28	50	400
Temperature (°C)	**Scenario**	**Polarization**	**TX Power (dBm)**	**Number of RX**
20	UMi	Co-Polarization	30	1

Table 4: NYUSIM default channel parameters

All the channel parameters reported in Table 4 are set to their default values, and the base station height is set to 10 m. The *Signal consistency* and *Human blockage* sections are disabled. The path loss has been computed using these settings. The effects of environmental constraints and NLOS scenario have been highlighted. Results are plotted in Figure 17. It is shown that heavy rains have low impact on the path loss. The main issues are NLOS scenario and outdoor to indoor (O2I) penetration losses. They respectively lead to 27 dB and 37 dB penalties. It is reminded that the LOS probability is between 15% and 60 % in the 50 to 200 m range. Finally, as shown in [9], the base station height also impacts the path loss from a certain distance. Indeed, it modifies the limit between the LOS and NLOS scenario.

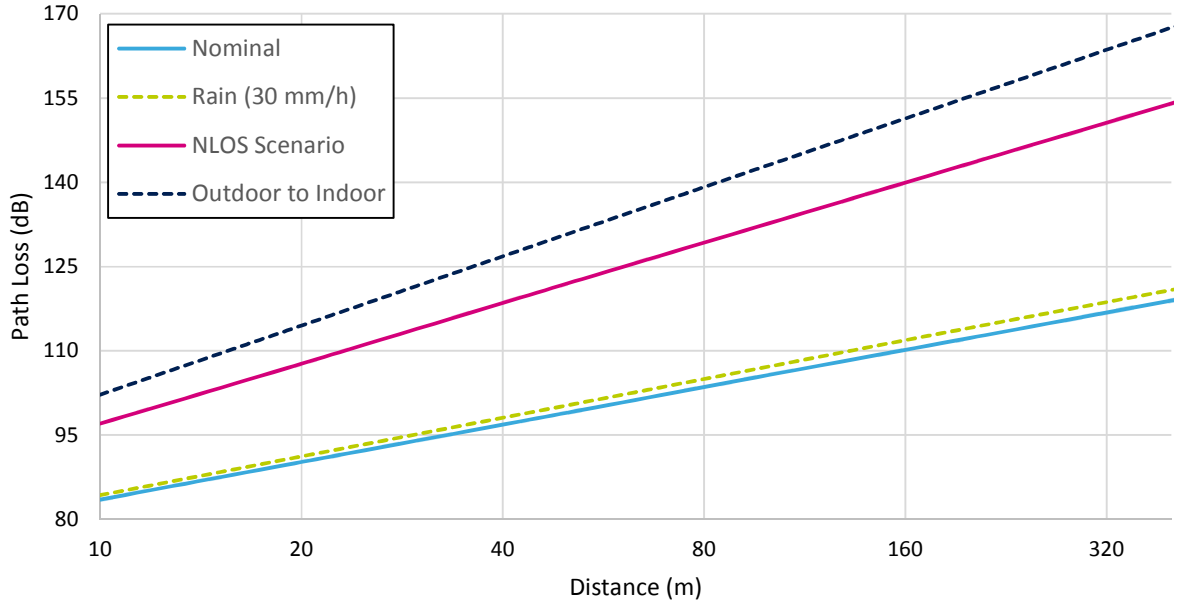

Figure 17: UMi path loss at 28 GHz, for different scenarios

To conclude the path loss estimation, some relevant values are reported in Table 5.

Distance (m)	50	100	200
Nominal path loss (dB)	99	106	112
NLOS path loss (dB)	122	133	158
O2I path loss (dB)	131	143	172

Table 5: Path loss relevant values at 28 GHz

d. Step 4 – Receiver Characteristics

To establish the link budget of the whole mobile communication system, it is necessary to determine the receiver (i.e. the user equipment) characteristics. Based on recent state of the art, mobile receivers may be composed with small antenna arrays. As it is highly integrated, its gain per element is expected to be about 0 dBi, as shown in [26]. Moreover, circuit performances depend on the chosen technology, the operating frequency and the integration level. For example, the CMOS integrated receiver designed in [27] exhibits a noise figure (NF) lower than 4 dB and a maximum output power about 14.5 dBm. For cost and integration issues, SiGe or CMOS technologies may be preferred to GaAs and GaN. The expected user equipment characteristics are summed up in Table 6.

Freq. (GHz)	NF (dB)	Antenna Gain (dBi)	Number of Antenna	Technology
28	4	0	4	SiGe / CMOS

Table 6: Expected user equipment features

Finally, the base station receiver will be sized depending on the transmitter properties (front-end module features and array size). The current methodology may be used, after the base station transmitter sizing.

e. Step 5 – Link Budget

Thanks to the previous parts, transmitted signal, mm-wave channel, and receiver characteristics have been determined. Thus, the transmitter EIRP could be derived from a link budget. The wireless link is pictured in Figure 18.

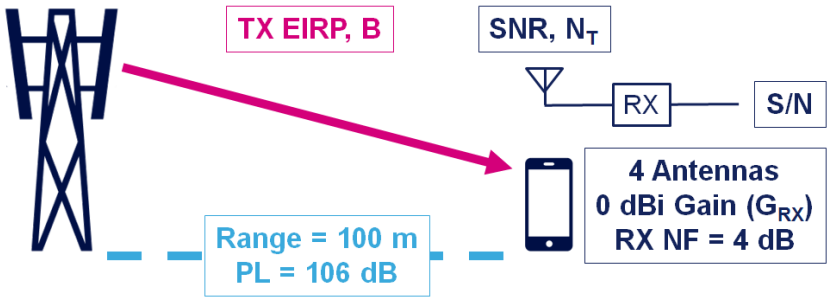

Figure 18: Wireless link in the chosen use case

Firstly, the SNR is determined thanks to *Equation (14)*. It gives the relation between the required S/N, the receiver *NF*, and the *SNR* (all expressed in dB). Then, the link budget is completed with *Equation (15)*.

$$SNR = \frac{S}{N} + NF - 10.\log_{10}(N_{el}) \quad (14)$$

$$EIRP = SNR + PL - G_{RX} + N_T \quad (15)$$

$$N_T = 10.\log_{10}(k_B * T * B) \quad (16)$$

Finally, *Equation (16)* enables to compute the thermal noise, N_T. It depends on the Boltzmann constant k_B, which value is $1.3806 * 10^{-23}$ J.K^{-1} and the temperature T. It also depends on the signal bandwidth. As it is much higher than it was for the previous standards, wireless systems are no longer interference limited, but noise limited. Results for the targeted use cases are reported in Table 7.

Range (m)	PL (dB)	LOS (%)	S/N (dB)	B (MHz)	T (K)	N_T (dB)	EIRP (dBm)
100	106 – 143	40	11.8	250	298	- 119.9	25.9 – 62.9
200	112 – 172	15	8.5	100		- 123.9	24.6 – 84.6

Table 7: Link budget results for the chosen use cases

Up to these results, O2I and NLOS scenarios seem highly demanding due to their high path loss. Actually, the current specifications for sub 6 GHz pico-cells limit the base station EIRP to 33 or 47 dBm [23]. Fortunately, the expected network densification should increase the LOS probability. The use of signal repeaters is also expected [20]. It enables to cope with the large path loss penalty. As well, indoor coverage will be provided thanks to dedicated networks and offloading.

f. Step 6 – Front-End Module Features for mm-Wave Access Point

This part focuses on the base station front-end module (FEM) performances in the TX mode, i.e. on the power amplifier (PA) properties and the way it is connected to each antenna. As the EIRP has been determined, FEM features will enable to size the antenna array. The choice of the FEM topology mainly depends on the standard requirements, and the chosen technology.

Firstly, 5G PA's will probably operate at an important power back-off (PBO). This means that the input power is much lower than the one corresponding to the PA saturation. Actually, the use of the OFDM modulation implies high peak to average power ratio (PAPR). Thus, the back-off is necessary to ensure adequate linearity. As discussed in [28], the input PBO is about 9.6 dB. Consequently, the power added efficiency (PAE) of the reported PA drops from 35.5 to 10 %. It is possible to save efficiency thanks to dedicated PA architectures [29], [30], or using PAPR reduction techniques, which lower the required PBO. The technology choice has also a significant impact on the PAE, as shown by the results from [31], reported in Table 8. It is shown that CMOS may achieve slightly higher PAE than SiGe, but it drops faster with the increase of P_{SAT}. Moreover, GaAs and GaN exhibit PAE up to 40% for P_{SAT} above 30 dBm.

P_{SAT} (dBm)	PAE_{SiGe} (%)	PAE_{CMOS} (%)
20	43	47
30	20	10

Table 8: Reported PAE depending on the technology and P_{SAT}, from [31]

Then, in order to choose the best suitable technology, several recently designed Ka-band PA's have been reported in Table 9. The PAE at PBO is reported with a 6 dB output PBO which is often equivalent to the required 9.6 dB input PBO [26]. The first remarkable result is that GaAs and GaN technologies enable to reach high P_{SAT} with good PAE performances. However, the use of such technologies leads to high area penalties and challenging digital integration. Thus, CMOS and SiGe seem to be more suitable for the design of energy and cost efficient FEM's. For now, these technologies exhibit comparable features.

Reference	[**32**]	[**29**]	[**33**]	[**34**]	[**30**]	[**35**]
Technology	90 nm CMOS	45 nm SOI CMOS	250 nm SiGe	65 nm CMOS	GaN-Si	150 nm GaAs
Freq. (GHz)	28	26	27.5	38	28	28
P_{SAT} (dBm)	26	25	17.1	24.8	32	26.5
PAE_{peak} (%)	34.1	31	26.5	24.3	30	42
PAE_{PBO} (%)	≈ 10 *	24 – 27	11.6	≈ 10 *	30	31
Area (mm²)	0.401	0.63	0.36	0.146	6	2.86
P_{DC} (mW)	1083 **	286 **	83.3 **	1035 **	1350 **	375 **

*: graphically estimated; **: graphically estimated with a 6 dB output PBO

Table 9: Ka-band power amplifier state of the art

Moreover, MIMO and beamforming systems require an accurate and complex channel coding. To reduce the processing cost, time division duplex (TDD) is preferred to frequency division duplex (FDD). Actually, TDD enables to use the same frequency for both transmitting and receiving operations. Consequently, the channel coding only has to be realized once. This step may also be completed at the base station to reduce the constraints on user equipment. TDD implies the use of antenna switches. Several circuits designed in different technologies are reported in [36]. The comparison highlights the benefits of SOI CMOS in regards with bulk CMOS and SiGe. Actually, SOI CMOS switches may exhibit insertion losses lower than 1 dB, high power handling, and high integration level. Bulk CMOS and SiGe circuits have higher insertion losses, about 1.5 dB. Moreover they are often larger due to the implementation of bulky improvement techniques.

Finally, multi-beam phased array architectures have two effects on power amplifiers. On the one hand, the transmitted signal is divided among the antenna elements. Thus the power is divided as many paths. Consequently, the PA output power specification is relaxed. Moreover, as shown by *Equation (10)*, an N element array radiates N² times more power than a single antenna. Thus the PA requirements are reduced again. On the other hand, the signals for each user are combined before the amplification stage. So, the PA output power is increased. This contrast is shown by *Equation (17)*. P_{SAT} is expressed depending on the required power per user P_{USER}, the power back-off *PBO*, and the number of users N_{USER}. The powers are given in dBm, and the PBO is in dB. Then, the power transmitted by a single element P_{TX} is expressed in dBm by Equation (18). It depends on the switch insertion losses *IL*, and on the antenna element gain G_{TX}. Both are in dB. Besides, other losses between the PA and the antenna are neglected.

$$P_{SAT} = P_{USER} + PBO + 10.log_{10}(N_{USER}) \quad (17)$$

$$P_{TX} = P_{SAT} - PBO - IL + G_{TX} \quad (18)$$

The antenna designed in [37] gives an estimation of the antenna gain. A 64-element array is presented with about 4 dBi gain per element at 28 GHz. Moreover, the dual-polarized elements show gain higher than 0 dBi over a large scanning range. More directive antennas could relax the circuit requirements, but the scanning range would be degraded. Lastly, the relation between the transmitter power capabilities and the array radiated power requirements will be determined thanks to the array sizing, in the following part. Then, the receiver characteristics may be derived from the technology features and the array geometry.

g. Step 7 – Array Sizing Issues

Previously, the available FEM output power and the required array EIRP have been determined, thanks to the technology features, and the system link budget. Replacing $EIRP_{el}$ in *Equation (10)* with P_{TX} from *Equation (18)* enables to relate these results into *Equation (19)*.

$$EIRP_{array} = P_{TX} + 20.log_{10}(N_{el}) \quad (19)$$

Moreover, based on *Equation (17)*, the needed EIRP per user, $EIRP_{user}$, is expressed in *Equation (20)*. Finally, the desired relation is given in *Equation (21)*, from the combination of *Equations (18)*, *(19)*, and *(20)*. It is applied to the two identified use cases, using different FEM technologies. The parameters are summed up in Table 10. Results are reported in Table 11, considering 4 dBi antenna gain and the required 6 dB output PBO.

$$EIRP_{user} = EIRP_{array} - 10.log_{10}(N_{USER}) \quad (20)$$

$$EIRP_{user} = P_{SAT} - PBO - IL + G_{TX} + 20.log_{10}(N_{el}) - 10.log_{10}(N_{USER}) \quad (21)$$

	Range	N_{USER}	$EIRP_{user}$		GaAs	SOI	CMOS / SiGe
Case 1	100 m	17	19.9 dBm	**P_{SAT} (dBm)**	30	25	20
Case 2	200 m	67	18.6 dBm	**IL (dB)**	0.8		1.5

Table 10: Use case (left) and technology parameters (right) for array sizing

Array Size (*N*)	GaAs	SiGe / CMOS	SOI
Case 1 (LOS)	2	7	4
Case 1 (NLOS)	61	209	109
Case 2 (LOS)	4	11	6
Case 2 (NLOS)	742	2541	1319

Table 11: Array sizing, for different LOS and NLOS scenarios

It could be noticed that for the considered use cases and technologies, the required array size is not so high. Unfortunately, it drastically increases for the NLOS and the O2I scenarios. Then, the use of GaAs technology enables to reduce the array size thanks to its high-power capabilities. Consequently, there is a major tradeoff on the system power consumption: increasing *N* leads to reduced PA requirements, but also increases the overall complexity. However, the array size not only influences the link budget.

As shown in *Equation (7)*, the larger the antenna array is, the sharper the beams are. Thus, the SDMA is more efficient. Besides, it is illustrated in [38] that larger arrays generate more radiation pattern nulls, providing better spatial filtering.

Based on the 100 m range use case, the array has to form 17 beams. Which corresponds to 4.1 beams in both azimuth and elevation planes. Considering a uniform rectangular array (URA) covering a 60°-wide sector, it leads to 14.6° beam width. Likewise, the 200 m range use case leads to beam width about 7.3°. Using *Equation (7)*, the corresponding array sizes are respectively 8 x 8 and 15 x 15 elements. However, users are not equally distributed in beam areas. Actually, the further from the access point they are, the more they are in a sector. Likewise the further the beam target is, the larger the beam area is as shown in Figure 19 (*b*), where HPBW is the Half Power Beam Width, i.e. ϑ_3 in *Equation (7)*. Finally, this phenomenon mainly depends on the access point height. Thus, it is harder to discriminate users far from the access point than closer ones. To determine the optimum access point elevation range R_E and height H, *Equation (22)* is used. It is derived from trigonometry formulas applied to Figure 19 (*a*).

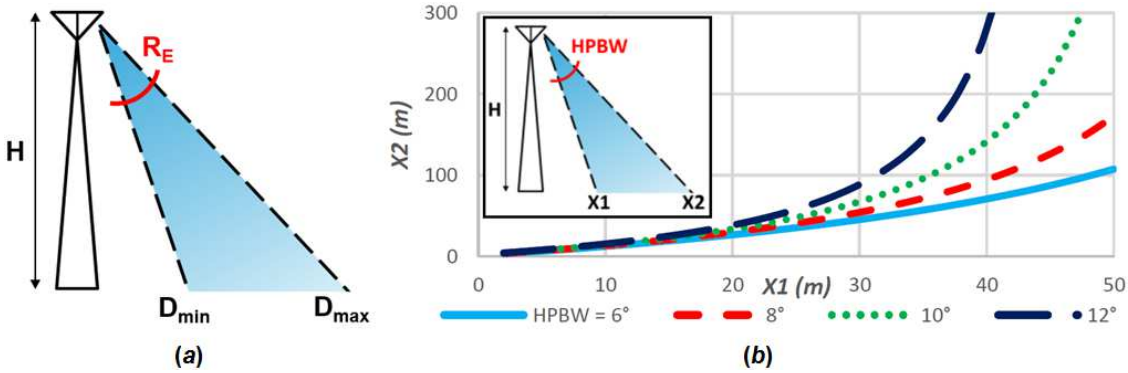

Figure 19: (a) Parameters of the whole access point array; (b) Beam spreading of a single beam, with H = 10m

$$\frac{\frac{D_{min}}{H} + tan(R_E)}{1 - \frac{D_{min}}{H}.tan(R_E)} = \frac{D_{max}}{H} \qquad (22)$$

Typically, the elevation range R_E is close to 60° as antenna arrays are square, and the minimum distance between user and access point, D_{min}, is a few meters. The chosen access point heights approximately correspond to street lights (4m), or 3-floor buildings (10m). Results are reported in Table 12. It could be noticed that for small heights, D_{max} could be infinite which leads to much energy loss. Consequently, it is necessary to properly set the access point deployment parameters to efficiently cover the targeted sectors. Finally, considering a 100 element array, different conformations are available. Several are reported in Table 13.

R_E (°)	55				60			
H (m)	4		10		4		10	
D_{min} (m)	2	5	2	5	2	5	2	5
D_{max} (m)	27	∞	23	67	67	∞	30	167

Table 12: D_{max} for Different access point configurations

Antennas	Antennas per Row	HPBW in Azimuth	Antennas per Column	HPBW in Elevation
100	10	11.3°	10	11.3°
99	9	12.7°	11	10.2°
96	8	14.5°	12	9.2°

Table 13: Possible 100-element array conformations

To conclude about beam spreading, the beam projections on the ground are illustrated in Figure 20, for the worst case: H = 4 m. Each line represents the limits of a single beam area, i.e. the minus 3 dB boundaries. It shows that it is harder to discriminate users in areas far from the access point. Moreover, access points at higher heights have both higher ranges and better spatial filtering.

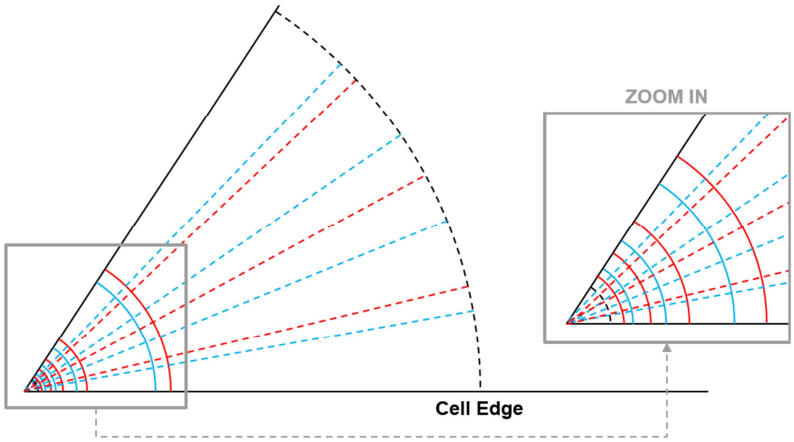

Figure 20: Beam projection on the ground for 10 x 10 (red), and 12 x 8 (blue) arrays

Finally, the array sizing and positioning impacts the link budget and the spatial filtering ability. Obviously, it also has a significant effect on the overall complexity and power consumption. As these points are as well related to beamforming topology and chip partitioning, it will be discussed in the following section.

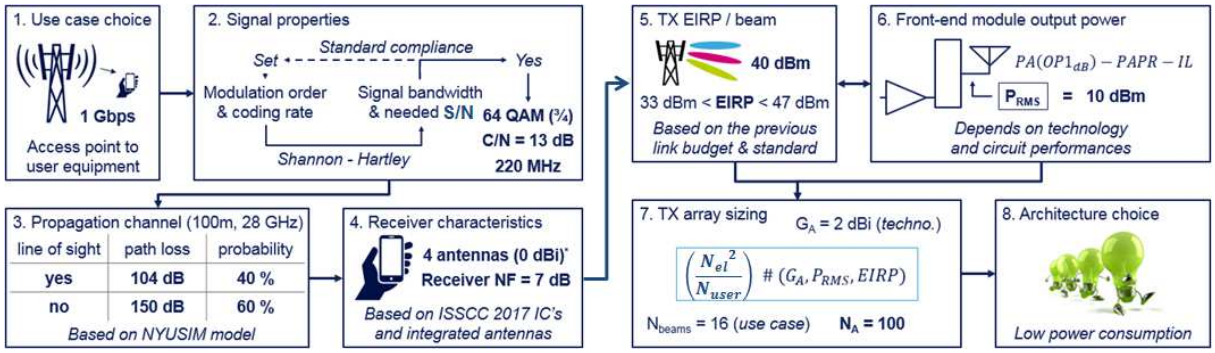

Figure 21: Main steps of the 5G access point sizing methodology

Figure 21 reminds the different steps achieved in the previous sections. From the early 5G requirements, the system-level specifications have been determined. In conclusion, a suitable phased array size has been selected, reflecting the choice of the technology. The last step aims to determine the best optimal phased array architecture, meeting the defined system-level specifications with low power consumption.

Indeed, the next section gives an overview of the existing beamforming architectures. Then, recently reported systems are compared, based on their main advantages and drawbacks. Finally, the last step will be achieved, introducing a new method to estimate the power consumption of such systems.

h. Preliminary Step 8 – Beamforming Topologies Overview

Previously, beamforming history and operating principle have been introduced. This section details the different modern phased array architectures able to concurrently transmit and receive several beams. The consequences on the overall system are also discussed. These architectures are divided in two distinct families. On the one hand, there are passive multi-beam antennas (MBA's) which directly comes from the historical topology. Such systems could be based on reflectors, lenses, or beamforming feeding circuits [39]. The different beams are generated in the RF domain, and no additional power consumption is needed. However, the directions of the beams are often settled by construction. Thus, it is not suitable for 5G mobile communications. Moreover, in passive beamforming, the signal is amplified before it is split, as illustrated in Figure 22 (*a*) from [39]. Thus the PA constrains are higher. It also implies high losses between the antenna and the FEM. If it could be offset in TX mode by increasing the PA output power, but it still degrades the overall NF in RX mode.

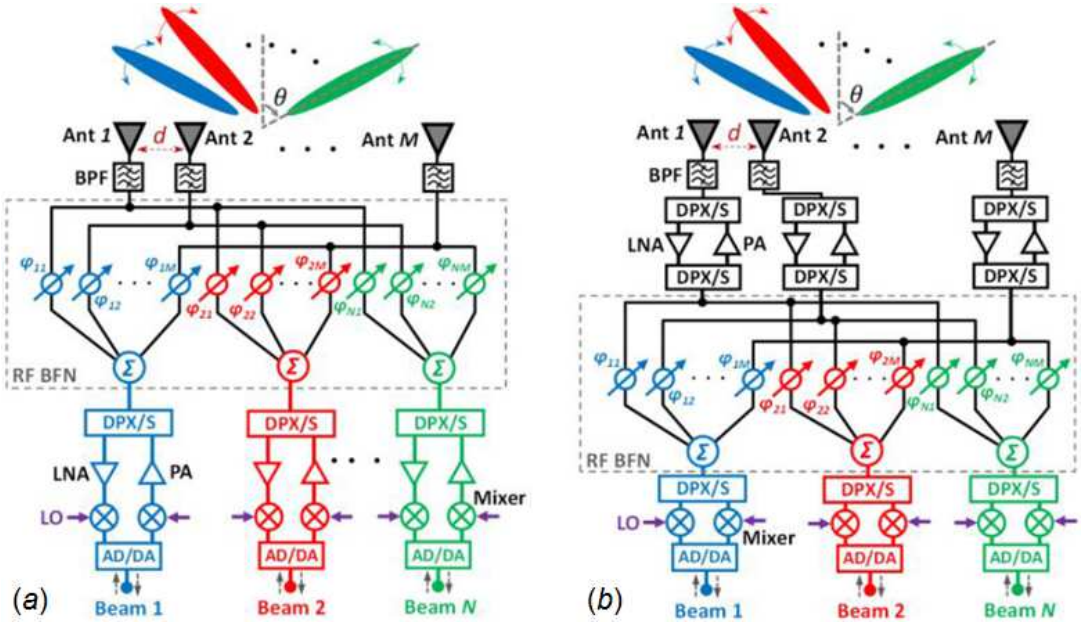

Figure 22: (a) System architecture of a passive MBA; (b) System architecture of an active RF MBPAA, from [39]

On the other hand, there are multi-beam phased array antennas (MBPAA's) which generate the beams thanks to active circuits. Consequently, it provides higher flexibility. This family is split into 2 sub-families depending on which circuit generates the beams. The first one is RF MBPAA, compared with passive MBA in Figure 22 from [39]. The beamforming is performed by phase shifters. These architectures are very popular. As presented previously, the PA requirements are relaxed, so the overall efficiency is improved. Besides, there are fewer losses between the antenna and the low noise amplifier. Both solutions benefit from the array effect, which improves SNR. Finally, the phase shifters may be replaced by true time delays to avoid beam squinting in wideband systems. The second one gathers architectures which generate the beams at intermediate frequency (IF), local oscillator (LO) frequency, or at base band

(BB). As the phase shifting is realized at lower frequency, these systems are more accurate and scalable [38]. However, the LO distribution is very challenging. Also, wideband beamforming is not supported because of prohibitive low frequency true time delays. Figure 23 (*a*) presents a MBPAA receiver with LO phase shifting. MBPAA's are more flexible than MBA's, but implies RF phase shifting or LO distribution issues. They also have a number of beams settled by construction. So the system adaptability is limited. Besides, the beams may also be generated in the digital domain. These architectures known as digital multi-beam antennas (DMBA's) are illustrated in Figure 23 (*b*), from [39]. They provide more flexibility and accuracy than previously presented systems. But the processing required to simultaneously generate the beams is very challenging. Finally, RF and digital phase shifting may be combined into hybrid beamforming. This enables to reduce both processing requirements, and number of circuits. Different partitioning is possible as discussed in [40].

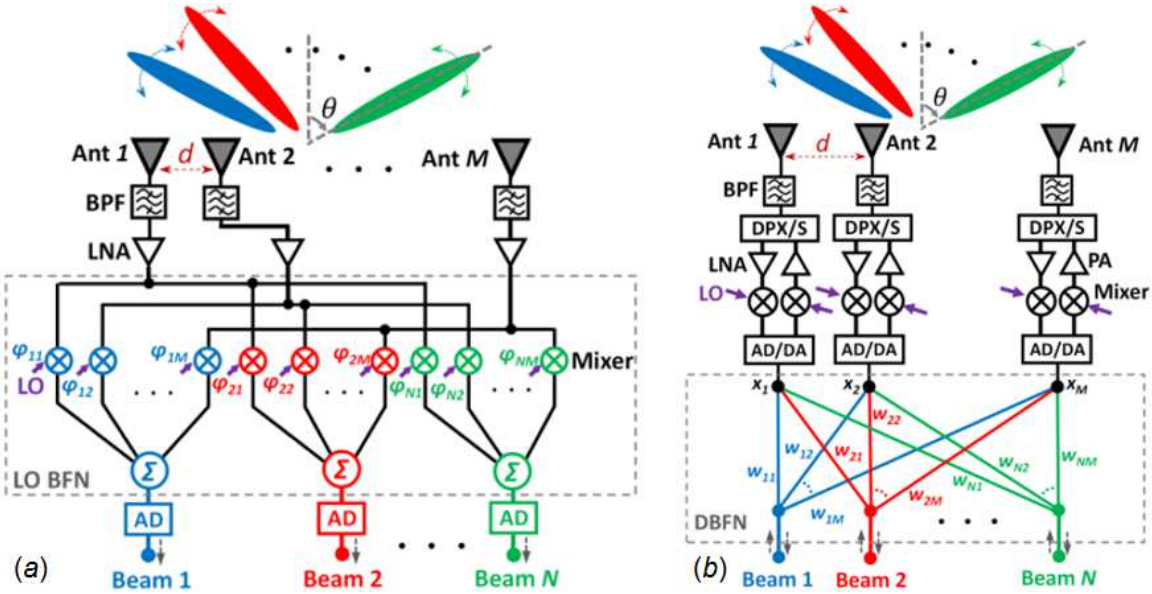

Figure 23: System architecture of: (a) an MBPAA receiver with LO phase shifting and (b) a full DMBA, from [39]

Whether beamforming is realized thanks to passive MBA, MBPAA, DMBA, or a hybrid combination of RF and digital, the design of mm-wave circuit remains challenging. Moreover, a qualitative comparison of these architectures is not sufficient to determine the best suited system for 5G access points. Thus, a more accurate analysis based on power consumption will be completed in the next section.

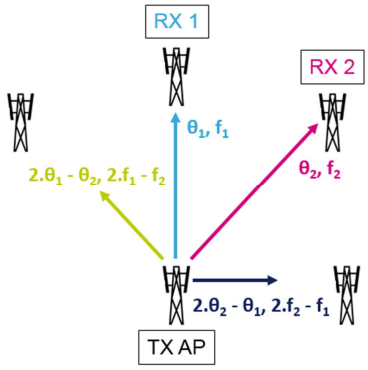

Figure 24: Undesired beams due to the intermodulation products

Finally, the transmission of several concurrent beams thanks to a single array implies the generation of undesired intermodulation products [41]. Two beams simultaneously transmitted at f_1, and f_2 frequencies, respectively in ϑ_1 and ϑ_2 azimuths, generate spurious at $2.f_1 - f_2$, and $2.f_2 - f_1$ frequencies, respectively in $2.\vartheta_1 - \vartheta_2$, and $2.\vartheta_2 - \vartheta_1$ azimuths. It is illustrated in Figure 24. Their power depends on PA features.

III. Phased Array Antenna Transmitter Comparison

1. Ka-band 5G Phased Array Transceiver State of the Art

a. RF Phased Arrays

This overview of 5G transceivers starts with the most popular architectures: RF phased arrays. In [42], a 16 dual-polarized element transceiver (TRx) is presented. Each integrated circuit (IC), designed in 130 nm SiGe BiCMOS technology, is composed of 16 TRx per polarization. As 4 IC's are used, the system could handle 2 concurrent beams. Each beam takes advantage from the 64 antenna array. Besides, it could also operate in multi-beam mode. 8 beams are simultaneously generated, i.e. 2 per 16 antenna array. Finally, the asymmetrical switch enables to reduce the PA output power requirement, which leads to 20 % power consumption gain. Each IC spends 4.6 W per polarization in TX state, and 3.3 W in RX state. The IC block level schematic is shown in Figure 25. Another phased array transceiver is introduced in [43]. It is based on 4-element TRx's, made in 180 nm SiGe technology. 16 IC's are gathered, and combined with a 64 antenna PCB array. This transceiver can only handle a single beam. However, 8 to 12 Gbps communication has been achieved using one array as TX, and another as RX. The whole system consumes 11.9 and 7.8 W in TX and RX states. The system block diagram is shown by Figure 25.

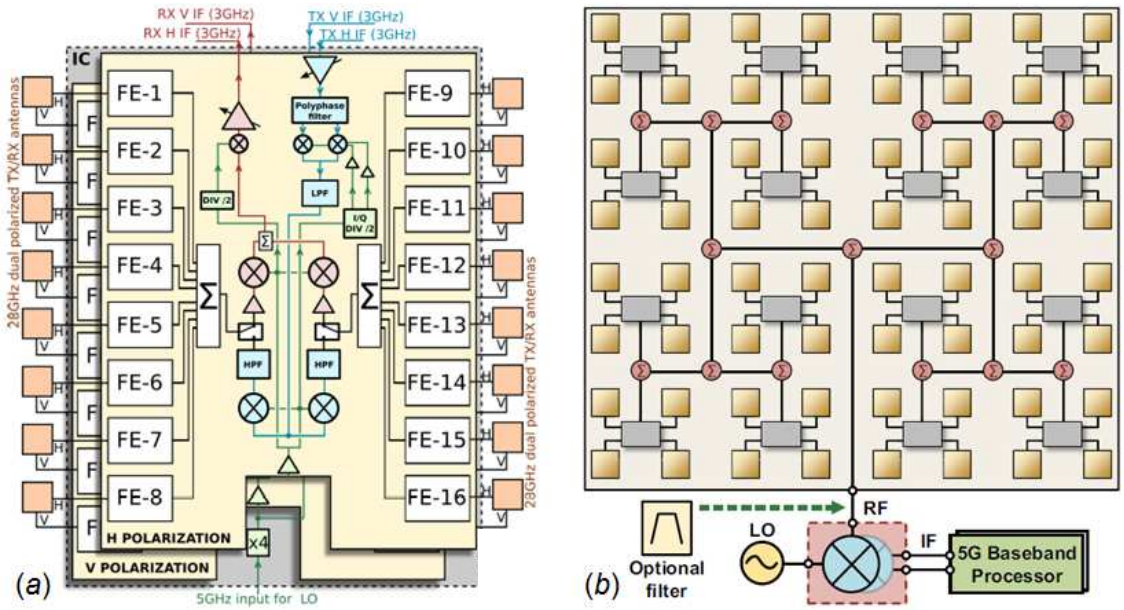

Figure 25: (a) IC architecture, from [42]; (b) Phased array block diagram, from [43]

Finally, a 24 element transceiver is presented in [44]. The generation of 2 concurrent beams is supported thanks to dual-polarization. It is manufactured with 28 nm CMOS technology, which leads to a significant power consumption saving. The system consumes only 2.85 W when operating at maximum output

power. These 3 architectures show that RF beamforming is easy to implement. They exhibit about ±50° scanning range, without any calibration. However, only few concurrent beams may be handled. Moreover, there is a tradeoff between the number of beam and the array size highlighted in [42]. Finally, it is shown in [38] that, for now, the number of signal chains per IC does not exceed 64 due to yield considerations.

b. Other Phased Array Topologies

Even if RF phased arrays are very popular, other architectures may also show interesting features. In [45], an LO phase shifting transceiver, operating at 39 GHz, is introduced. The 64 element system is composed of 4-element chips, designed in 65 nm CMOS technology. This architecture provides high phase and gain accuracy. However, calibration is needed to reach such performances. The power consumption is also higher than previously presented systems, about 24 W in TX state. As shown in Figure 26, the calibration circuits are included in the chip.

Besides, beamforming systems tend to become more and more digital. Thus, a reconfigurable front-end module for hybrid and digital beamformers is presented in [46]. It is able to operate at 28, 37, and 39 GHz. The single 65 nm CMOS chip has a low power consumption, 140 mW for the transmission of 2 QPSK beams. It could be extrapolated that the power consumption of a 64 element array would be 4.48 W per beam. Finally, a fully digital 64 element transceiver is illustrated in Figure 26, from [47]. It enables to generate 8 concurrent beams, which provide up to 50.7 Gbps data rate. Each beam takes benefit from the whole 64 element array. However, it has significant power consumption, higher than 980 W when transmitting. It is due to the use of substrate integrated waveguides which lead to higher losses, i.e. PA requirements. Then, it is not an integrated circuit, but the combination of FPGA, and daughter boards.

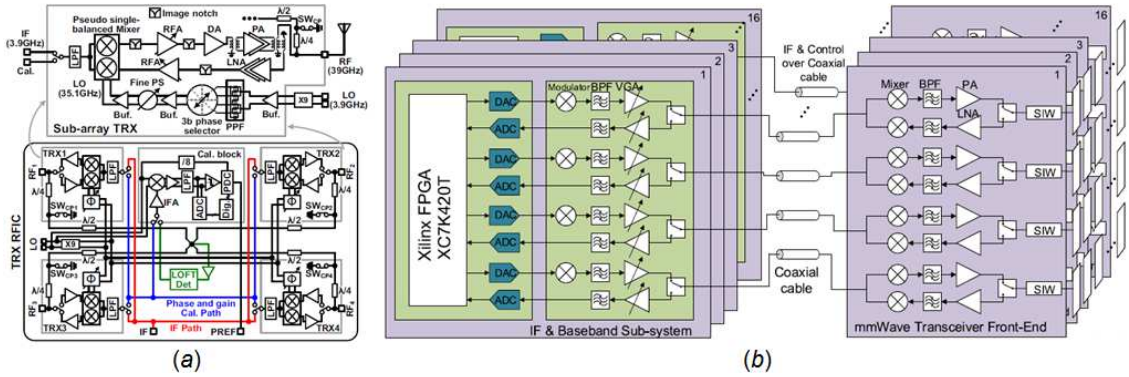

Figure 26: (a) LO phase shifting TRx, from [45]; (b) Block diagram of the mm-wave DBF TRx, from [47]

c. Phased Array Figure of Merit

In order to compare the different phased array architectures, a Figure of Merit (FoM) has to be defined. For 5G small cell applications, the most significant parameter is the system power consumption, P_{DC}. Then, these systems have to provide high data rates to concurrent users. So the first idea was to base the FoM on the previously defined R_{USER}, and N_{USER}. However, the data rate depends on many parameters, and is not often studied. So, it seems better to replace the data rate by the PA output power P_{OUT} (in W), and to add the array size N. This FoM, shown in *Equation (23)* is dimensionless. Results for the 6 previously presented systems are reported in the next section. It also sums up the main

beamformer features. Finally, as most of the reported circuits operate at 28 GHz, frequency is not taken into account in the FoM. The chip size is not considered either, as it may be increased to optimize chip partitioning [43].

$$FoM = \frac{N_{el} * P_{OUT} * N_{USER}}{P_{DC}} \quad (23)$$

d. Phased Array Comparative Table

Reference	[42]	[43]	[44]	[45]	[46]	[47]
Freq. (GHz)	28	28	28	39	28 / 37 / 39	28
Technology	130 nm SiGe	180 nm SiGe	28 nm CMOS	65 nm CMOS	65 nm CMOS	FPGA
Phase shift	RF	RF	RF	LO	Hybrid	Digital
Element /IC	2 * 16	4	2 * 12	4	1	4
Number of beams	8 / 2	1	2	1	2	8
Number of antennas	64	64	12	64	1	64
TX OIP1 (dBm)	14.0	10.4	12.0	13.4	14.0	25.0
TX P_{DC} (W)	36.8 [1]	11.9 [2]	2.9 [2] / 2.1 [3]	24	0.14 [3]	982 [1]
RX NF (dB) / channel	6	4.6	3.8 – 4.4	7	6.2	5.3
RX P_{DC} (W)	26.4	7.8	0.84	8	0.038	729
Size (mm²)	661	188	29.0	192	0.48	-
Data Rate (Gbps)	-	8 – 12	-	-	1.5	50.7
FoM	0.087	0.059	0.13 / 0.18	0.059	0.36	0.17

[1] at linear output power, [2] at maximum output power, [3] at 6 dB PBO

Table 14: Phased array state of the art

Table 14 gathers several recent works, which try to connect with 5G base station requirements. RF phased arrays based on SiGe technology seem very popular. Even if it enables to reach high data rate [43] and have good scalability [42], it is outperformed by 28 nm CMOS for the power consumption [44], and by digital architectures for the number of antennas / number of users tradeoff [47]. The FoM highlights that hybrid and digital systems are competitive with RF ones in terms of power consumption. Moreover, the latest achievements could be more integrated to reduce the power consumption. To support this assumption, a power consumption estimation method is introduced in the next section.

2. Transceiver Power Consumption Estimation

a. Power Consumption Model

Up to Table 14, the TX state consumes between 30 and 270 % more power than RX state. Consequently, the following method is focusing on the transmitter power estimation. But it could be employed for RX.

As illustrated in Figure 22, and Figure 23, the transmitter architecture, so the number of circuits needed, depends on which beamforming topology is selected. Firstly, high losses between the antenna and the FEM imply power consumption increase [42]. It also leads to higher NF. To avoid this, the FEM is placed as close as possible to the antenna. Thus, whatever the chosen topology, it is needed to have as many PA as antennas. The PA power consumption may be estimated thanks to the definition of its power added efficiency (*PAE*), which is given by *Equation (24)*. It is expressed in %, and it depends on the output power P_{OUT}, the power consumption P_{DC}, and the PA gain G_{PA}.

$$PAE = \frac{P_{OUT} * \left(1 - {}^{1}/_{G_{PA}}\right)}{P_{DC}} \qquad (24)$$

Equation (21) enables to compute the required P_{OUT}, depending on N and N_{USER}. The best suited technology is chosen regarding to this result. This brings to G_{PA} and *PAE* values, which are reminded in Table 15. An output PBO of 6 dB is considered, i.e. about 9.6 dB input PBO [26].

References	[**35**]	[**28**]	[**33**]	[**29**]
Technology	150 nm GaAs	28 nm CMOS	250 nm SiGe	45 nm SOI
P_{SAT} (dBm)	26.5	15	17.1	25
Gain (dB)	11.8	14	21.7	10
PAE_{PBO} (%)	31	11	11.6	25.5

Table 15: PA gain and PAE depending on the technology

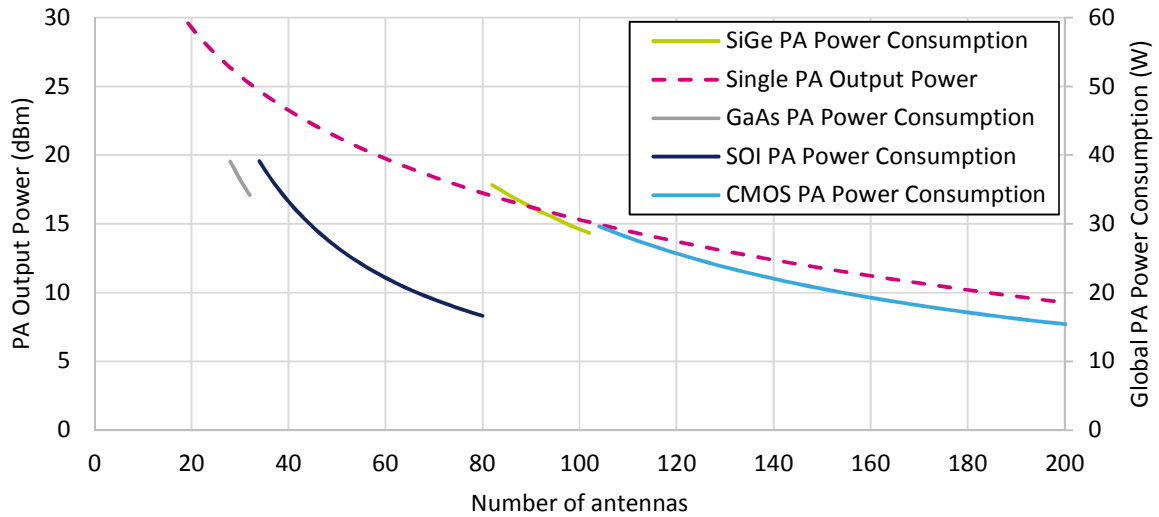

Figure 27: Required output power per PA and global PA power consumption, both depending on the array size

To simplify this step, switch losses are supposed to be 1 dB whatever the technology. It is also assumed that there are no other losses between the antenna and the FEM. Finally, the EIRP is set between the LOS and NLOS requirement, about 40 dBm. The needed P_{OUT} for each PA and the global power consumption of PA's are plotted in Figure 27 for use case 1 (100 m range, 17 users).

The first remarkable result is that PA output power requirement decrease with the increase of the array size. This leads to a reduction of their power consumption. Moreover, the results for SiGe and CMOS highlight that *PAE* has a significant impact on the overall power consumption. Thus, the reported SOI PA exhibits interesting features for energy savings.

Circuits before the PA depend on the beamforming architecture. For active RF MBPAA's, shown in Figure 22, each beam is split, phase shifted, and combined. This part also includes gain control. The beams may also not be combined in the case of sub-array architectures. In both cases, $N * N_{USER}$ variable gain amplifiers and phase shifters are needed. Their resolution could be relaxed in the case of hybrid beamforming. But the complexity is reported to the digital part. Finally, it is supposed that splitting and combining loss are offset by the gain control stage. For full DMBA's, shown in Figure 23, this section is not necessary. In [27], the power consumption for a highly accurate 6 bit phase shifter and gain control stage, designed in 65 nm CMOS, is about 48 mW per channel. Another 65 nm CMOS realization is presented in [48]. As it is based on a passive phase shifter and a variable gain amplifier which compensates the losses, its power consumption is only 10 mW. Finally, the 45 nm SOI IC presented in [49] exhibits a 42 mW power consumption, including 12 mW for the LNA which compensate the passive phase shifter loss. It is assumed that medium phase shifter power consumption is about 20 mW. So, the global power consumption of this section for an active RF MBPAA is 34 W, in use case 1.

The next component of interest is the mixer, which achieves the up and down frequency conversions. Their number depends on the beamforming architecture, but also on chip partitioning. For DMBA's, 1 mixer per antenna is needed, plus LO distribution among all RF chains. While for active RF MBPAA's, at least N_{USER} mixers are required. In this case, high frequency connections are needed between the FEM and the single frequency conversion circuit. To relax these constraints, several mixers may be used, each one for a reduced number of antennas $N_{USER} < N_C < N$, as discussed in [38]. The number of mixers is then $N_C * N_{USER}$. Finally, the power consumption of this section depends on the required conversion gain and NF. Several IC's, designed in different technologies, are reported in [50], [51], and [52]. Their power consumption are respectively 90 mW for 130 nm SiGe, 15 mW for 65 nm CMOS, and 0 mW for resistive 90 nm SOI mixer. It is supposed that the mixer power consumption is about 40 mW, including LO distribution. As previously, a 100 element array is considered, fulfilling use case 1 requirements. This leads to power consumption about 40 W for DMBA's, at least 680 mW for RF active MBPAA's, and up to 2.72 W using $N_C = 4$. Even if DMBA's power consumption is much higher, it does not depend on the number of users, unlike MBPAA's.

The only major circuit not estimated yet is the first step of TX systems: digital to analog converter (DAC). Based on the study achieved in [53], the CMOS DAC power consumption P_{DAC} is given by *Equation (25)*. It depends on two technological parameters: the supply voltage V_{DD}, and the minimum transistor channel length L_{min}. It also depends on sampling and signal frequencies, f_{sample} and f_{signal}, and on the effective number of bits, *ENOB*. As technology scaling have a significant impact on P_{DAC}, it supports the use of more integrated technologies for the other system IC's. Besides, the significant increase of data rates implies the use of higher f_{sample}. It is related to f_{signal} thanks to Shannon's law, given by *Equation (26)*.

$$P_{DAC} = \frac{{V_{DD}}^2 * L_{min} * \left(f_{sample} + f_{signal}\right)}{10^{4.838-(0.1525*ENOB)}} \quad (25)$$

$$f_{sample} > 2 * f_{signal} \quad (26)$$

Consequently, there is a major tradeoff between f_{sample} and *ENOB* for P_{DAC}. Besides, DAC requirements are relaxed in DMBA's thanks to array effect. And they are increased as each DAC has to handle N_{USER} users. Finally, based on [54], the required *ENOB* for 5G wireless communications is between 6 and 11 bits. Likewise, the reported f_{sample} is about 1 GHz. And the analog to digital converter power consumption is 40 mW. Assuming a medium required *ENOB* of 8 bits, the power consumption is 20 mW.

Besides, there is additional power consumption in DMBA's due to digital processing. Up to [55], it depends on the number of operations needed to form a beam, on the beam bandwidth, and on the frequency reuse, i.e. the number of users sharing the same frequency channel. Digital processing power consumption is supposed to be 62.5 mW per antenna for 16 users, divided among two 250 MHz channels.

b. Phased Array Transmitter Architecture Comparison

To estimate the overall power consumption of MBPAA and DMBA transmitters, the number of circuits for each architecture is reminded in Table 16. Then, it is combined with the estimated power consumptions of the previous section. Results are plotted in Figure 28. Moreover, as hybrid topology may be close to DMBA or MBPAA depending on the way it is realized, its power consumption is supposed to be comparable.

	Digital	DAC	Mixer & LO	Σ	Phase & Gain	Σ	PA	Switch	Antenna & Filter
MBPAA	ø	N_{USER}	$4 * N_{USER}$	N_{USER} 1 to N	$N * N_{USER}$	N N_{USER} to 1	N	N	N
DMBA	N_{USER}	N	N	ø	ø	ø	N	N	N

Table 16: Number of IC's for different beamforming transmitter topologies

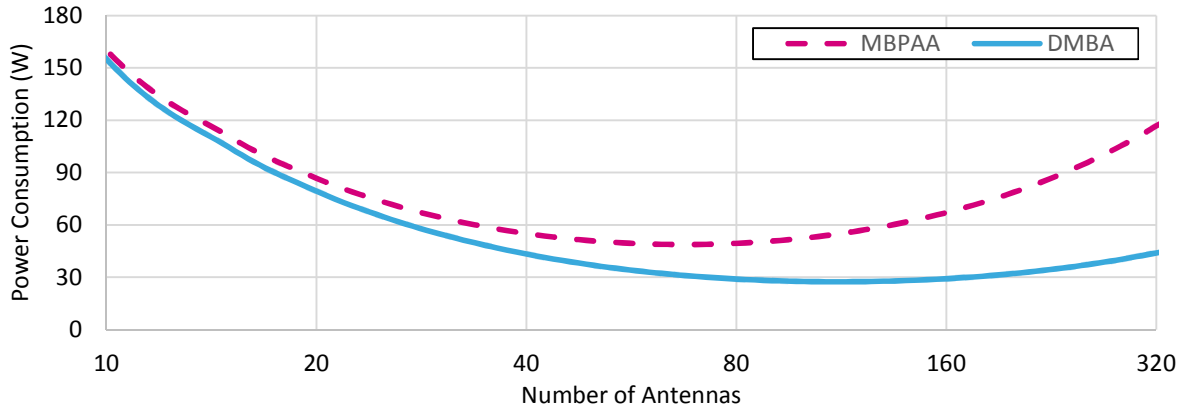

Figure 28: Comparison of DMBA and MBPAA power consumptions, depending on the number of antennas, for use case 1

Figure 28 shows that there is an optimal power consumption for MBPAA's and DMBA. It is due to reduced PA requirements, and higher number of phase shifters, or digital processing complexity for large arrays. This point is also highlighted in [56]. The optimal system power consumptions are obtained with 65-element MBPAA, and 110-element DMBA. It respectively leads to 48.8 W, and 27.5 W.

Consequently, DMBA's seem most suitable for 5G small cells than MBPAA's. First of all, its lower power consumption enables to reduce network OPEX. Moreover, its high scalability makes it better for quick deployment, which also decreases small cell CAPEX. Finally, the choice of DMBA topology is supported by a higher accuracy, as discussed in [57]. However, smaller arrays are required for MBPAA architectures. So, the overall cost and complexity could be reduced. Besides, the results also highly depend on the chosen technology.

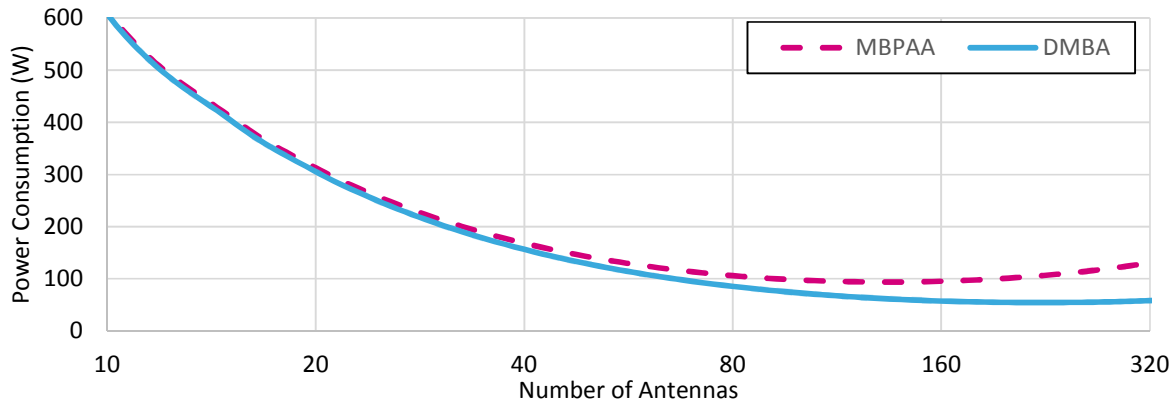

Figure 29: Comparison of DMBA and MBPAA power consumptions, depending on the number of antennas, for use case 2

For use case 2, the results are plotted in Figure 29. The optimal sizes are 135 elements for MBPAA's and 220 elements for DMBA's, leading respectively to power consumptions of 93.9 W and 54.5 W. Digital beamforming is therefore the most suitable topology for both use cases.

c. Prospects of Improvement

As a conclusion of this section, the above results are compared with the DMBA introduced in [47]. It is reported in Table 17. Firstly, the power consumption repartition among the different sections is consistent between use cases 1 and 2. However, there are differences between the proposed model and the current state of the art. Actually, a way to improve the model may be to take into account more accurately the beamforming algorithm effects. Besides, in [47], the DMBA may be improved, since it is not fully integrated.

	Total P_{DC} (W)	PA	Phase Shifters and Mixers	DAC and Digital Processing
MBPAA 1 – 2	48.8 – 93.9	48.6 – 47.8 %	50.9 – 51.8 %	0.5 – 0.4 %
DMBA 1 – 2	27.5 – 54.5	53.5 – 50.6 %	15.3 – 16.1 %	31.2 – 33.3 %
[**47**]	729	30 %	29 %	41 %

Table 17: Power Consumption Comparison between Model and State of the Art

IV. Conclusion

In this chapter, the evolution of mobile networks has firstly been presented. It introduced the main mobile network features, as cellular network and spectrum access techniques. In particular, the description of the current 4G LTE-A standard enabled to discuss about modern network architectures and MIMO wireless systems. Then, the limitations of current network have been presented to introduce the upcoming mobile communication generation. The 5G KPI's have been discussed, and the expected solutions enabling to cope with spectral congestion and emerging applications have been introduced. It is based on 3 pillars: the use of high-frequency spectrum, access point densification, and phased array antennas. The deployment timeline and several market actors have been presented.

Secondly, the targeted use case is defined. This work focuses on mm-wave access points for 5G small cells. It is of particular interest, as it takes benefits from the 3 pillars of 5G. Starting from the standard KPI's, the use case KPI's and system-level specifications are derived. In this part, signal properties, mm-wave propagation issues, and integrated circuit technologies have been discussed. The array sizing, including the impact on beam resolution, has also been studied.

Based on the system-level study, the circuit specifications have been discussed. Then, the impact of the choice of beamforming architecture has been presented. Finally, as power consumption has significant effect on the overall network cost, a methodology has been developed to estimate it. This enables to compare the different topologies and choose the best suitable for a given use case. It has been highlighted that DMBA's are of particular interest due to their low power consumption and high flexibility. However, there still is a gap between the proposed power consumption model and current state of the art systems. Possible sources of these differences may be: the model is based on the Shannon channel capacity, which is an optimal use case; digital power consumption is very complex to estimate accurately; the state of the art is still limited, the best reported system is not integrated so it has a high potential for improvements.

At the end, it has been shown that there is an optimal transmitter power consumption depending on the array size and on the chosen use case. However, this phenomenon is less pronounced with phased array receivers. Actually, the SNR is not as improved as EIRP when the array size increases. Moreover, as the signal combination is performed in the digital domain, DMBA's signal chains are more sensitive to receiver NF. That is why in the following chapters, this work focuses on the design of FEM receiver circuits: switches and low noise amplifiers.

V. References

[1] Cisco. (2017, February) Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021. [Online]. https://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html

[2] ARCEP. (2017, March) 5G: Issues & Challenges. [Online]. https://archives.arcep.fr/uploads/tx_gspublication/Report-5G-issues-challenges-march2017.pdf

[3] Qualcomm. (2014, July) The Evolution of Mobile Technologies: 1G to 2G to 3G to 4G LTE. [Online]. https://www.qualcomm.com/documents/evolution-mobile-technologies-1g-2g-3g-4g-lte

[4] C. Gruet. (2019) 5G Overview. [Online]. https://www.telecom-evolution.fr/

[5] Y. Bouguen, E. Hardoin, and F-X. Wolff, *LTE et les réseaux 4G.*: EYROLLES, 2012.

[6] ANFR. (2018) Agence Nationale des Fréquences (Plaquette 2018). [Online]. https://www.anfr.fr/fileadmin/mediatheque/documents/ANFR/ANFR_Plaquette-2018.pdf

[7] J. Huang et al. (2012, June) A Close Examination of Performance and Power Characteristics of 4G LTE Networks. [Online]. http://www.cs.columbia.edu/~lierranli/coms6998-7Spring2014/papers/rrclte_mobisys2012.pdf

[8] Radiocommunication Sector of ITU (ITU-R). (2015, September) IMT Vision - Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond. [Online]. https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!PDF-E.pdf

[9] T. S. Rappaport et al., "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks—with a focus on Propagation Models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213-6230, 2017.

[10] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617-1655, Third Quarter 2016.

[11] R. J. Mailloux, *Phased Array Antenna Handbook*. Norwood: Artech House, Inc, 2005.

[12] X. Guan, H. Hashemi, and A. Hajimiri, "A Fully Integrated 24-GHz Eight-Element Phased-Array Receiver in Silicon," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2311-2320, December 2004.

[13] A. Natarajan, A. Komijani, and A. Hajimiri, "A Fully Integrated 24-GHz Phased-Array Transmitter in CMOS," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2502-2514, 2005.

[14] J. S. Herd and M. D. Conway, "The Evolution to Modern Phased Array Architectures," *Proceedings of the IEEE*, vol. 104, no. 3, pp. 519-529, March 2016.

[15] YOLE Développement. (2018, June) 5G Impact on RF Front-End Modules & Connectivity for Cellphones. [Online]. https://www.i-micronews.com/produit/5gs-impact-on-rf-front-end-module-and-connectivity-for-cell-phones-2018/

[16] P. Moorhead. (2019, June) Who is 'Really' Leading In Mobile 5G, Part 2: 5G Mobile Chipsets. [Online]. https://www.forbes.com/sites/patrickmoorhead/2019/06/12/who-is-really-leading-in-mobile-5g-part-2-5g-mobile-chipsets/#5b3db7b844d4

[17] Qualcomm. (2020, February) Snapdragon X60 5G Modem-RF System. [Online]. https://www.qualcomm.com/products/snapdragon-x60-5g-modem

[18] IBM and Ericsson, "IBM and Ericsson announce 5G mmWave Phased Array Antenna Module," *Microwave Journal*, February 2017.

[19] Anokiwave. Products. [Online]. anokiwave.com/products

[20] P. Hindle, "Comprehensive Survey of Commercial mmWave Phased Array Companies," *Microwave Journal*, January 2020.

[21] Wikipedia. List of cities by population density. [Online]. https://en.wikipedia.org/wiki/List_of_cities_by_population_density

[22] Qorvo. (2017, May) Small Cell Networks and the Evolution of 5G (Part 1). [Online]. https://www.qorvo.com/design-hub/blog/small-cell-networks-and-the-evolution-of-5g.

[23] European Telecommunications Standards Institute (ETSI), "5G; NR; Base Station (BS) Radio Transmission and Reception (3GPP TS 38.104 version 15.5.0 Release 15)," 2019.

[24] P. Wongchampa and M. Uthansakul, "Orthogonal Beamforming for Multiuser Wireless Communications: Achieving higher received signal strength and throughput than with conventional beamforming," *IEEE Antennas and Propagation Magazine*, vol. 59, no. 4, pp. 38-49, 2017.

[25] NYU Wireless. (2019, November) NYUSIM Version 2.01 Now Available. [Online]. https://wireless.engineering.nyu.edu/nyusim/

[26] J. Curtis, H. Zhou, and F. Aryanfar, "A Fully Integrated Ka-Band Front End for 5G Transceiver," in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, CA, 2016.

[27] H. A. Ameen et al., "A 28 GHz four-channel phased-array transceiver in 65-nm CMOS technology for 5G applications," in *2017 29th International Conference on Microelectronics (ICM)*, Beirut, 2017.

[28] S. Shakib, H. Park, J. Dunworth, V. Aparin, and K. Entesari, "A 28GHz Efficient Linear Power Amplifier for 5G Phased Arrays in 28nm Bulk CMOS," in *2016 IEEE International Solid-State Circuit Conference (ISSCC)*, San Francisco, CA, 2016.

[29] N. Ozen, M. Rostomyan and P. Asbeck, "A Ka-band asymmetric dual input CMOS SOI Doherty power amplifier with 25 dBm output power and high back-off efficiency," in *IEEE topical Conference on*

RF/microwave power amplifiers for radio and wireless applications (PAWR), 2019.

[30] R. Giofrè, A. Del Gaudio, and E. Limiti, "A 28 GHz MMIC Doherty power amplifier in GaN on Si technology for 5G applications," in *IEEE/MTT-S International Microwave Symposium*, 2019, pp. 611-613.

[31] H. Wang et al. (2020, February) Power amplifiers performance survey 2000-present. [Online]. https://gems.ece.gatech.edu/PA_survey.html

[32] W-C. Huang and H. Wang, "An inductive-neutralized 26-dBm K-/Ka-band power amplifier with 34% PAE in 90-nm CMOS," *IEEE Transactions on Microwave Theory and Techniques (TMTT)*, vol. 67, no. 11, pp. 4427-4440, November 2019.

[33] D. Wang, W. Chen, L. Chen, and Z. Feng, "A broadband linear millimeter-wave power pmplifier with an adaptive bias circuit," in *IEEE MTT-S International Wireless Symposium (IWS)*, Guangzhou, China, 2019, pp. 1-3.

[34] Y. Chang, B-Z. Lu, Y. Wang, and H. Wang, "A Ka-band stacked power amplifier with 24.8-dBm output power and 24.3% PAE in 65-nm CMOS technology," in *IEEE/MTT-S International Microwave Symposium*, 2019, pp. 316-319.

[35] D. P. Nguyen, B. L. Pham, and A-V. Pham, "A compact Ka-band integrated Doherty amplifier with reconfigurable input network," *IEEE transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 205-215, January 2019.

[36] C. Li et al., "Ka Band FEM Design Comparison with 45nm RFSOI CMOS and High Performance SiGe BiCMOS," in *14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Qingdao, 2018.

[37] X. Gu et al., "A Multilayer Organic Package with 64 Dual-Polarized Antennas for 28GHz 5G Communication," in *2017 IEEE MTT-S International Microwave Symposium (IMS)*, Honololu, HI, 2017.

[38] X. Gu, A. Valdes-Garcia B. Sadhu, "The More (Antennas), the Merrier," *IEEE Microwave Magazine*, vol. 20, no. 12, pp. 32-50, 2019.

[39] W. Hong et al., "Multi-Beam Antenna Technologies for 5G Wireless Communications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6231-6249, 2017.

[40] A. F. Molisch et al., "Hybrid Beamforming for Massive MIMO: A Survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134-141, 2017.

[41] W. A. Sandrin, "Spatial Distribution of Intermodulation Products in Active Phased Array Antennas," *IEEE Transactions on Antennas and Propagation*, vol. 21, no. 6, pp. 864-868, 1973.

[42] B. Sadhu et al., "A 28GHz 32-Element Phased-Array Transceiver IC with Concurrent Dual Polarized Beams and 1.4 Degree Beam-Steering Resolutionfor 5G Communication," in *2017 IEEE International*

Solid-State Circuit Conference (ISSCC), San Francisco, CA, 2017.

[43] M. Sayginer, G. M. Rebeiz K. Kibaroglu, "A scalable 64-element 28 ghz phased-array transceiver with 50 dbm eirp and 8-12 gbps 5g link at 300 meters without any calibration," in *2018 IEEE/MTT-S International Microwave Symposium - IMS*, Philadelphia, PA, 2018.

[44] J. D. Dunworth et al., "A 28GHz Bulk-CMOS dual-polarization phased-array transceiver with 24 channels for 5G user and basestation equipment," in *IEEE International Solid - State Circuits Conference - (ISSCC)*, San Francisco, CA, 2018.

[45] Y. Wang et al., "A 39GHz 64-Element Phased-Array CMOS Transceiver with Built-in Calibration for Large-Array 5G NR," in *2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Boston, MA, 2019.

[46] S. Mondal, R. Singh, and J. Paramesh, "A Reconfigurable Bidirectional 28/37/39GHz Front-End Supporting MIMO-TDD, Carrier Aggregation TDD and FDD/Full-Duplex with Self-Interferance Cancellation in Digital and Fully Conneted Hybrid Beamformers," in *2019 IEEE International Solid-State Circuits Conference - (ISSCC)*, San Francisco, CA, 2019.

[47] B. Yang et al., "Digital Beamforming-Based Massive MIMO Transceiver for 5G Millimeter-Wave Communications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 7, pp. 3403-3418, 2018.

[48] R. Garg and Natarajan A. S., "A 28-GHz Low-Power Phased-Array Receiver Front-End With 360° RTPS Phase Shift Range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 11, pp. 4703-47-14, 2017.

[49] U. Kodak and G. M. Rebeiz, "A 42mW 26-28 GHz Phased-Array Receive Channel with 12 dB Gain, 4 dB NF and 0 dBm IIP3 in 45nm CMOS SOI," in *2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, San Francisco, CA, 2016.

[50] J. A. Qayyum, J. Albrecht, and Ulusoy A. C., "A Compact 24-32 GHz Linear Upconverting Mixer with -1.5 dBm OP1dB using 0.13-um SiGe BiCMOS Process," in *IEEE 19th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Orlando, FL, 2019.

[51] C. W. Byeon, J. H. Lee, D. Y. Lee, M-R. Kim, and J-H Son, "A High Linearity, Image/LO-Rejection I/Q Up-Conversion Mixer for 5G Cellular Communications," in *2015 10th European Microwave Integrated Circuits Conference (EuMIC)*, Paris, 2015.

[52] F. Ellinger, "26.5-30-GHz Resistive Mixer in 90-nm VLSI SOI CMOS Technology With High Linearity for WLAN," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 8, pp. 2559-2565, 2005.

[53] E. Lauwers and G. Gielen, "Power estimation methods for analog circuits for architectural exploration of integrated systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 10, no. 2, pp. 155-162, 2002.

[54] Aurangozeb, F. Aryanfar, and M. Hossain, "A Quad Channel 11-bit 1 GS/s 40 mW Collaborative ADC based Enabling Digital Beamforming for 5G Wireless," in *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Honolulu, HI, 2017.

[55] Fenech H., Roux, L. , A. Hirsch, and V. Soumpholphakdy, "Satellite Antennas and Digital Payloads for Future Communication Satellites," *IEEE Antennas and Propagation Magazine*, vol. 61, no. 5, pp. 20-28, 2019.

[56] T. Cameron. RF Technology for the 5G Millimeter Wave Radio. [Online]. https://www.analog.com/media/en/technical-documentation/white-papers/RF-Technology-for-the-5G-Millimeter-Wave-Radio.pdf

[57] D. Sikri and R. M. Jayasuriya, "Multi-Beam Phased Array with Full Digital Beamforming for SATCOM and 5G," *Microwave Journal*, April 2019.

Chapter 2: Standalone Ka-Band Front-End Module Circuit Design

By now, a system-level study has been completed in the previous chapter. It has highlighted that energy and cost efficient Ka-band circuits are needed, in order to cope with their skyrocketing number. It has also introduced prior results, which tend to favor substrate on insulator (SOI) CMOS technologies, compared with usual III-V technologies, which are more expensive. Thus, this chapter firstly presents the benefits of SOI technologies. Then, the *ST C65SOIFEM* technology is introduced, and its main features are measured.

Secondly, two standalone circuits are designed using this technology. As the receive path seems to be more critical, it is chosen to focus on low noise amplifier (LNA) and duplexer designs. Their specifications are discussed, based on the previous system-level study. Finally, the circuits have been measured, and enhancement perspectives are given.

I. The *ST C65SOIFEM* Technology

1. Substrate on Insulator Technologies

a. From Bulk CMOS to SOI Technologies

In 1965, Gordon Moore has observed that the number of transistors in an integrated circuit doubled every year [1]. Then, it has been adjusted: “the number of transistors in an integrated circuit doubled every 2 years.” It is the famous Moore’s law. Since then, circuit manufacturers have tried to follow this rule. It has triggered a transistor scaling race, illustrated by Figure 30, from [2]. However, device scaling implies the emergence of many parasitic effects [3]. It requires complex manufacturing to be reduced, which results in reaching the Moore’s law limits. Moreover, the current concerns about power saving favor advanced technologies, over the race to ever more transistors. Consequently, new solutions as FinFET and SOI have been promoted to cope with the device scaling issues

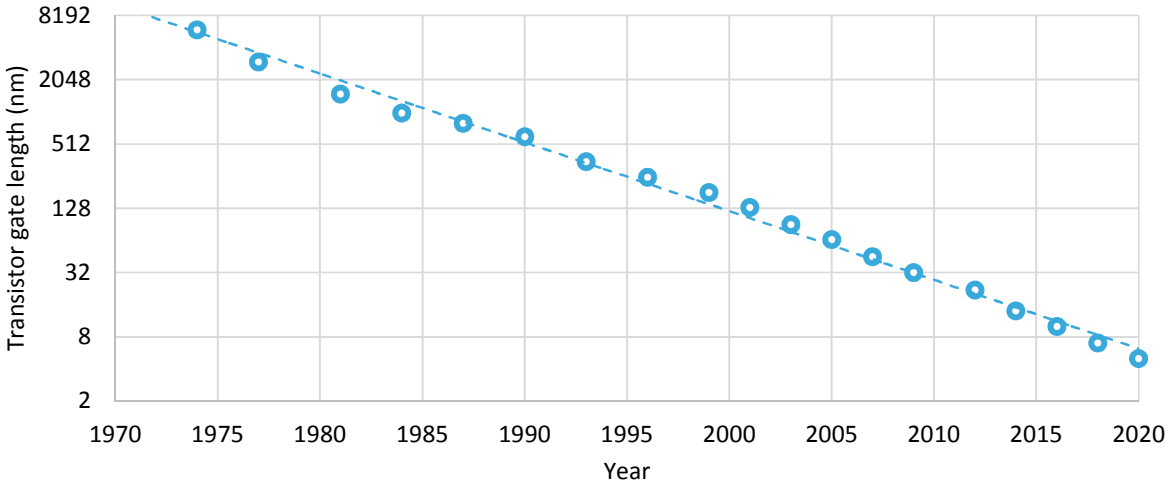

Figure 30: Transistor gate length scaling since 1970, from [2]

Generally, SOI refers to circuits in which active silicon is separated from the substrate by a buried oxide layer (BOX), as depicted by Figure 31. The first available devices have been developed in the early 1970's. They were based on sapphire substrate, also known as silicon on sapphire (SOS). The buried oxide provided high robustness against radiation. So, the technology was chosen for space and defense uses. As the process was very expensive, these devices were dedicated to these niche applications. Later, the use of cheaper substrates has reduced SOI technology cost. Moreover, it enables to enhance the device performances, as it suppresses some parasitic effects due to scaling. Thus, expensive technology scaling is delayed. These devices have been highly used in digital circuits, since the 1990's. For example, *IBM* has developed the *PowerPC* processors, based on SOI [3]. Finally, recent works have introduced hybrid GaN on SOI circuits [4]. They take benefit from the power handling and energy efficiency of III-V technologies, and the integration density of SOI. Therefore, it seems very suitable for the design of certain 5G FEM's.

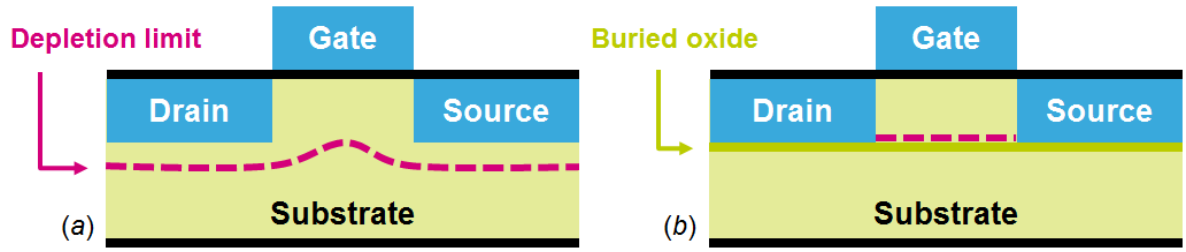

Figure 31: (a) Bulk CMOS transistor; (b) SOI CMOS transistor

The different kinds of substrates are depicted by Figure 32. The trap-rich (TR) SOI, also known as RF SOI, is composed of high resistivity (HR) substrate and a trap rich layer which improves its RF performances. The comparison completed in [5] has highlighted that RF SOI provides higher quality factor, and higher linearity compared with usual HR SOI. Even if SOS and quartz substrates exhibit better performances, their cost is often prohibitive.

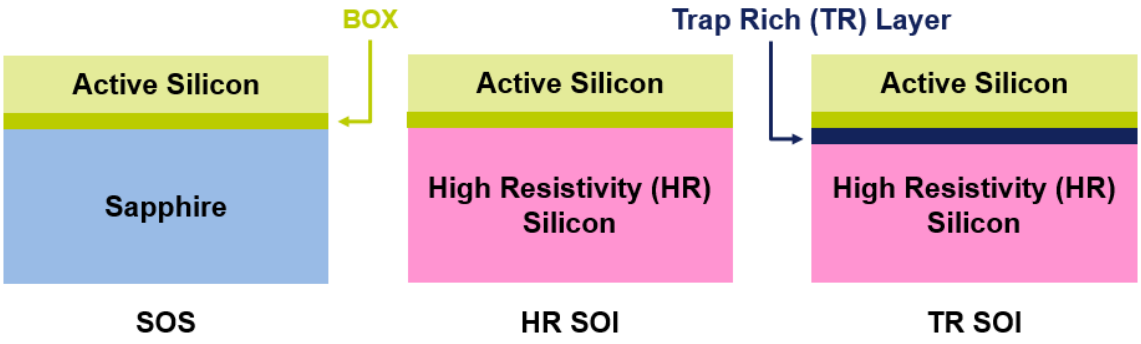

Figure 32: Cross section of SOI technologies using different substrates

The benefits of SOI technologies are not limited to radiation robustness. Actually, the BOX significantly reduces the drain to substrate and source to substrate parasitic capacitances [6]. It provides a speed improvement in digital circuits. Conversely, it leads to power consumption reduction at a given level of performance. Moreover, the BOX offers inherent isolation between transistors, which immunizes the circuits against the latch-up effect [3]. It may occur when the PNPN structure, formed by two bulk parasitic bipolar transistors, is activated. In a CMOS inverter, this could short V_{DD} to the ground (GND), as depicted in Figure 33. Thus, the circuit will be damaged.

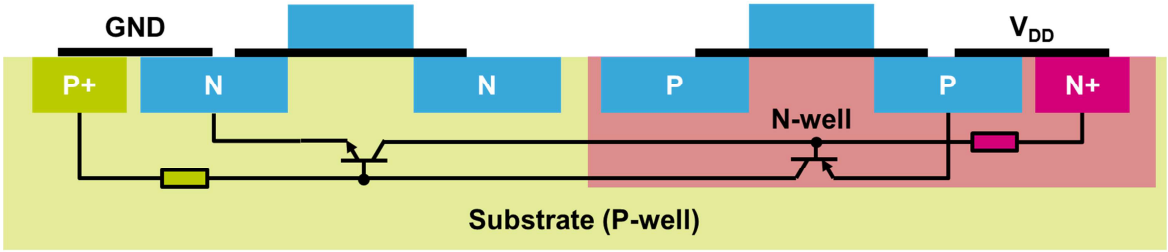

Figure 33: Schematic of a CMOS inverter and its parasitic bipolar transistors

Transistor scaling also results in channel lengths closer to the depletion layer width. Consequently, the device threshold voltage is degraded, leading to increased leakages [6]. In SOI circuits, the depletion layer is confined by the BOX, as illustrated by Figure 31. The short channel effects are therefore reduced. Finally, the device threshold is reduced thanks to SOI. This makes this technology suitable for low-power designs.

To sum up, SOI provides higher integration density, power consumption savings, and robustness against short channel effects. It is also compatible with high resistivity substrates, which are unusable with bulk due to latch-up effects. This offers good RF perspectives, since the passive quality factors are improved [3]. However, the SOI technologies have also significant drawbacks. As the parasitic capacitances may be used in certain designs to reduce the supply noise, its reduction is not only advantageous. The active region shrinking also induces more self-heating than in bulk. Finally, some SOI topologies could lead to higher drain to source leakages, and unwanted history or kink effects. These are described in the next section.

b. Comparison between PD SOI and FD SOI

The previous part has introduced some SOI drawbacks. They depend on the BOX implementation. In fact, there are two kinds of SOI transistors. As depicted in Figure 34, the different kinds are defined by the maximum depletion layer depth D_{max}. This value is given by *Equation (27)*, from [3]. It depends on the Fermi potential φ_F, the silicon permittivity ε_{Si}, the charge of an electron q, and the density of acceptor atoms N_a.

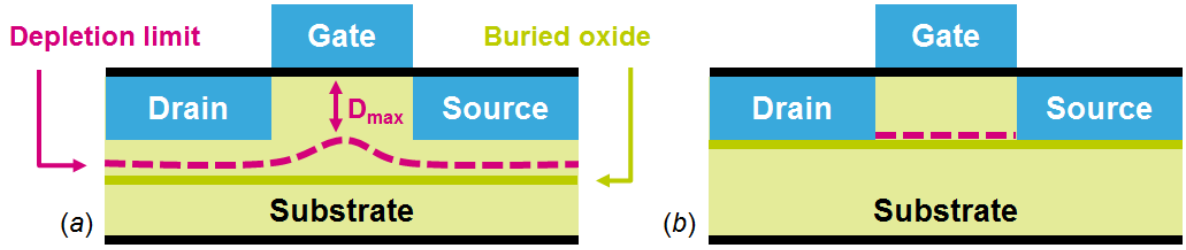

Figure 34: Comparison between (a) partially depleted and (b) fully depleted SOI transistors

$$D_{max} = \sqrt{\frac{4 * \varepsilon_{Si} * \varphi_F}{q * N_a}} \qquad (27)$$

When D_{max} is shorter than the BOX depth, the transistor is called partially depleted (PD) SOI. Conversely, if the depletion layer is larger than the active zone, it is called fully depleted (FD). In PD SOI devices, it is

possible to connect the substrate part above the BOX, i.e. the body, to the ground. It results in body contact (BC) MOS. If it is left electrically floating, it is called floating body (FB). Due to the migration of holes towards the floating body, the diode between the body and the source may become forward biased. This increases the potential of the body, leading to threshold voltage reduction. This is the so-called kink effect, which increases the drain current [3]. It is illustrated in Figure 35. Moreover, the body potential variation makes the circuit dependent of the previous state. Thus, history effects may occur. As these parasitic effects are related to FB, the use of BC MOS enables to cope with them. The body may also be tied to the source or the ground, in order to eliminate the parasitic effects in FB transistors. It implies a small area penalty.

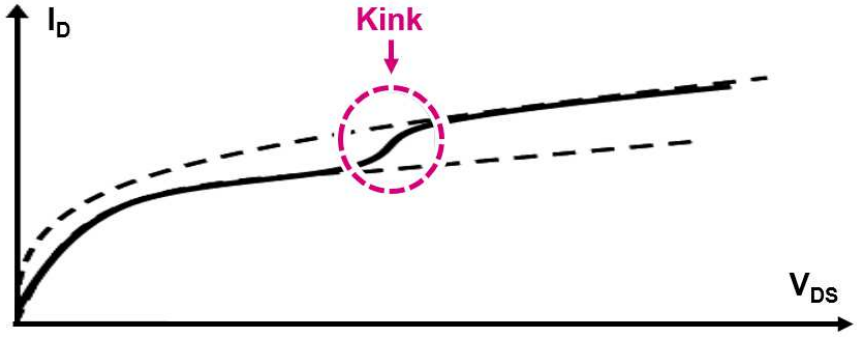

Figure 35: The I(V) characteristic of a PD-SOI transistor with kink effect, from [3]

As FD-SOI's active layer is thinner, these unwanted effects do not occur [6]. However, there is a parasitic open-base NPN transistor between the NMOS source and drain. This leads to leakage drain to source current when activated. Finally, the FD-SOI thin depletion layer could be combined with an ultra-thin BOX. These ultra-thin body and BOX (UTBB) transistors have an additional gate, known as back gate, which enables to control the threshold voltage. This new way to control the channel enables to design tunable circuits. However, it implies a cost penalty, since a more accurate manufacturing is needed.

The main SOI features are summarized in Table 18. FD-SOI devices exhibit less parasitic effects, lower power consumption, and higher integration level. It makes them highly favored for digital circuits. Conversely, PD-SOI ones are easier to manufacture, so their cost is reduced. As they could be combined with high resistivity or trap rich substrates, they seem suitable for analog and RF circuits.

	PD SOI	**FD SOI**
Manufacturing cost	Medium	High
History and kink effects	Yes	No
Self-heating	Moderate	High
Power consumption	Medium	Low

Table 18: PD and FD SOI features comparison

Finally, as BC transistors do not suffer from the kink effect, they are preferred for analog design. Up to [3], they also offer better low frequency noise, higher power handling, and better reliability compared with FB MOS. These ones seem to be more suitable for LNA, mixer, and switch design.

The previous sections have introduced the SOI technologies, and their main features have been discussed. The next one is an overview of the SOI technologies developed within *ST Microelectronics*. The section will provide keys to choose the best suitable process for the targeted Ka-band 5G applications.

c. Overview of ST SOI Technologies

In order to address a wide spectrum of applications, ST has developed a diversified portfolio of SOI technologies. Two of them are based on PD-SOI, combined with TR substrate to improve their RF performances. These are known as RF SOI. They have gate lengths of 130 nm and 65 nm. They have been designed for FEM's. The third technology is a 28 nm channel length FD-SOI technology [7]. It is therefore more suitable for digital circuits. The following Table 19 reports their main features.

	130 nm SOIFEM	**65 nm SOIFEM**	**28 nm FD SOI**
Metal layers (BEOL)	3 thin + 1 thick	4 thin + 1 thick	6 thin + 2 thick
Noise figure @ 6 GHz	0.74 dB	0.16 dB	0.30 dB
Noise figure @ 20 GHz	4.3 dB	1.1 dB	1.9 dB
V_T (mV)	480	640	600
g_{ds} (mS)	1.5	5.7	2.8
f_T (GHz)	49	118	344

Table 19: ST SOI technologies main feature comparison, from STMicroelectronics measurements

Firstly, it highlights that the 130 nm SOI FEM technology is not suitable for the targeted applications. Its cut-off frequency (f_T) is too low. Thus, it results in high noise figure in the Ka-band. The 28 nm FD-SOI technology is also discarded, as its back end of line (BEOL) makes it very expensive. However, its low V_{GS} is highly suitable for low-power digital designs. Moreover, it exhibits good RF performances. Thus, it could be chosen for mixed-signal circuits. Finally, thanks to its better RF features and lower cost, the 65 nm SOI FEM technology is chosen. It exhibits sufficient f_T to be used in the Ka-band, and low noise figure for LNA design. Finally, Figure 36 gives an overview of the SOI technologies developed at *ST*, and their scope of applications. The details about 65 nm SOIFEM technology are presented in the next section, in order to introduce the design of the Ka-band LNA's and duplexers.

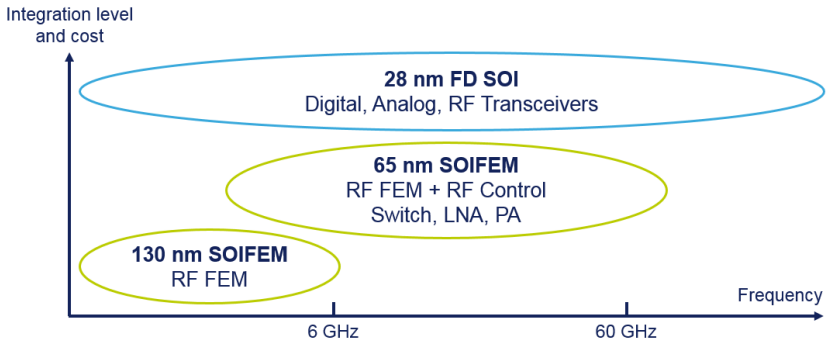

Figure 36: ST SOI technologies and their applications

2. The *ST C65SOIFEM* Technology

a. *Overview of the Technology Features*

The *ST C65SOIFEM* technology derives from a previous 65 nm SOI technology. The complexity of its BEOL has been reduced in order to make its manufacturing less expensive. As depicted in Figure 37, there is a thick top metal (M5) which enables to reach higher passive quality factors.

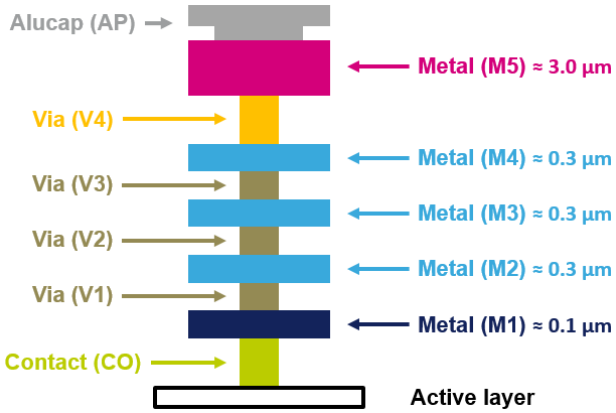

Figure 37: The ST C65SOIFEM BEOL

This technology offers low-power (LP) MOS transistors, known as GO1 due to their thin gate oxide. They have been enhanced to reduce their noise figure. Since *ST C65SOIFEM* is an RF SOI technology, FB and BC transistors are available. Both offer a standard threshold voltage (V_T) version, and a low V_T alternative. It has to be noticed that LP FB MOS are only available in NMOS, while BC exists in NMOS, and PMOS. Finally, there are standard V_T BC transistors which have been improved for switch design. These devices have a thicker gate oxide compared with GO1. Thus, the nominal drain to source voltage (V_{DS}) they can handle is increased to 2.5 V. However, they are only available in NMOS. The *ST C65SOIFEM* technology also offers standard resistors between 14 and 710 Ω /sq, high-resistance resistors about 6 kΩ /sq, and Metal-oxide-metal (MOM) capacitors with 2.5 fF /µm².

b. *Device Models*

The GO1 MOS simulations are all based on the Berkley short-channel IGFET model SOI (BSIMSOI). It is a standard model for SOI circuit design. It is called physical model, since its parameters are close to process and device parameters. It has been developed for DC simulations. Then, it has been enhanced for RF applications [8]. BSIMSOI RF is composed of the BSIM core, and added parasitic elements, as depicted in Figure 38. All the available BC MOS are simulated using usual BSIM3SOI RF. For FB MOS, BSIM4SOI RF is used. It has new features such as stress effect model, improved noise model, and more accurate V_T model [9].

Conversely, the thick gate oxide MOS (GO2) are not simulated thanks to a physical model like BSIM. They use the PSP SOI model, which is a compact model based on surface potential [10]. This model is widely used for RF SOI designs, since it accurately takes into account large signal behaviors as second and third order harmonics [11]. Moreover, it includes significant features for switch design, such as drain-induced barrier lowering (DIBL). It is a short channel effect which lowers V_T, depending on the drain voltage.

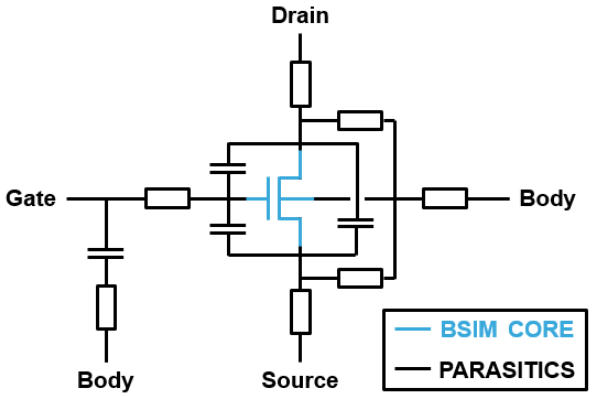

Figure 38: BSIMSOI RF schematic, from [8]

c. First Simulation Results

The *ST C65SOIFEM* technology, with its cost reduced BEOL is fairly recent. Actually, it has been released in early 2018. This thesis work is among the first uses of this technology for mm-wave applications. So, the following sections aim to extract the main transistor features, to compare the available devices, and to investigate the agreement between simulations and measurements. The passive elements such as pads for measurements, and inductors are also simulated. Firstly, it is necessary to define a test bench for the device parameter extraction. It is presented in Figure 39. The use of DC block (capacitor) and DC feed (inductor) elements enables to simultaneously complete the DC and small signal (SP) simulations.

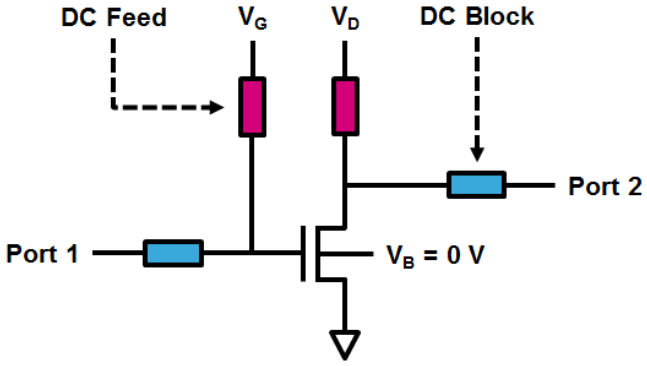

Figure 39: DC and small-signal simulation bench for BC MOS

The DC simulation is used to extract the MOS transconductance (g_m), drain to source resistance (r_{ds}), and threshold voltage (V_T), from the $I_D(V_{GS})$, and $I_D(V_{DS})$ characteristics. The simulation also provides the parasitic capacitances, including the gate to source (C_{gs}), and gate to drain (C_{gd}) capacitances. These give an approximation of f_T, through *Equation (28)*, from [12]. Then, the SP simulation is processed to extract more precisely the value of f_T, thanks to *Equation (29)*. It also enables to extract C_{gd}, C_{gs}, the drain to source capacitance (C_{ds}), and the gate resistance (R_G) through *Equations (30)*, *(31)*, *(32)*, and *(33)*, from [12]. These parameters give an approximation of the maximum frequency (f_{max}), through *Equation (34)*, also from [12].

$$f_T \approx {g_m}/{2 * \pi * (C_{gs} + C_{gd})} \quad (28)$$

$$f_T = {f_m}/{\mathrm{Im}\left({H_{21}}^{-1}\right)} \tag{29}$$

$$C_{gd} = \frac{-\mathrm{Im}(Y_{12})}{2 * \pi * f} \tag{30}$$

$$C_{gs} = \frac{\mathrm{Im}(Y_{11} + Y_{12})}{2 * \pi * f} \tag{31}$$

$$C_{ds} = \frac{\mathrm{Im}(Y_{22} + Y_{21})}{2 * \pi * f} \tag{32}$$

$$R_G = {\mathrm{R}(Y_{11})}/{[\mathrm{Im}(Y_{11})]^2} \tag{33}$$

$$f_{max} \approx \sqrt{\frac{f_T}{8 * \pi * R_G * C_{gd}}} \tag{34}$$

The SP simulation also provides the device minimum noise figure (NF_{min}), which is related to R_G, as detailed in [13]. Finally, the transistor f_{max} may be extracted more precisely, thanks to the Mason's gain (U), defined in *Equation (35)*. The relation between f_{max}, and U is given by *Equation (36)*. The device parameter extraction is detailed for the GO1 low V_T BC NMOS, and then applied to the other available devices.

$$U = \frac{|Y_{21} - Y_{12}|^2}{4.[\mathrm{R}(Y_{11}).\mathrm{R}(Y_{22}) - \mathrm{R}(Y_{12}).\mathrm{R}(Y_{21})]} \tag{35}$$

$$f_{max} = f * \sqrt{U} \tag{36}$$

Firstly, the DC simulation is processed. The GO1 low V_T BC NMOS $I_D(V_{GS})$, and $I_D(V_{DS})$ curves are respectively plotted in Figure 40, and Figure 41, with V_{DS} from 0.4 to 2 V. The studied device has two 0.5 µm fingers, a 60 nm channel, one gate and one body accesses. The transconductance of the transistor (g_m) is plotted, which highlights that the optimal operating zone is reached when V_{GS} is about 0.8 V. C_{gd}, and C_{gs} are extracted in this operating zone for the estimation of f_T.

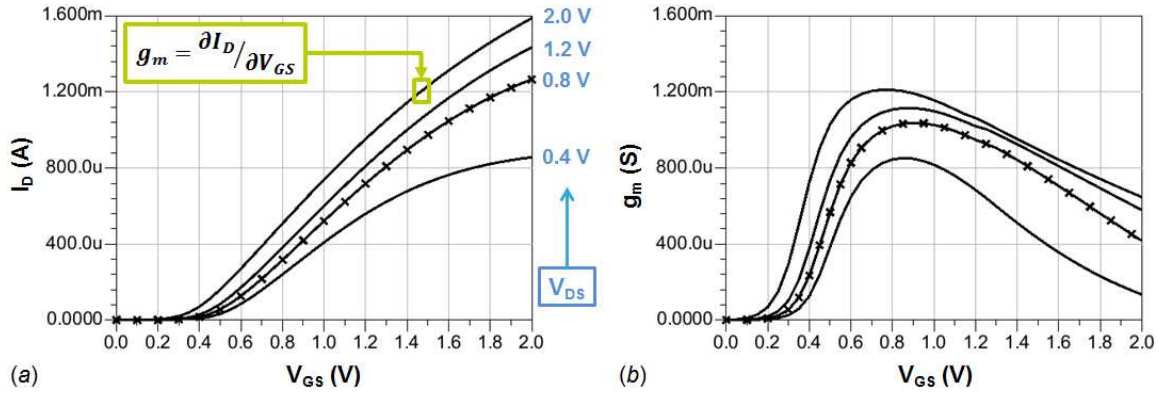

Figure 40: (a) I_D and (b) g_m for a GO1 low V_T BC NMOS transistor, both depending on V_{GS} and for different V_{DS}

The previously introduced DIBL effect is clearly visible in Figure 40 (*a*). The transistor starts to conduct at lower V_{GS} when V_{DS} is higher. It is also shown in Figure 41 (*b*): when V_{GS} is 0.4 V, and V_{DS} is 1 V, V_T is reduced from 456 mV to 392 mV. Thus, the transistor becomes open. Besides, the $I_D(V_{DS})$ curves show that the saturation is not ideal at high V_{DS}', the device has therefore poor r_{ds}. Finally, all the DC simulation results for this transistor are reported in Table 20, including the potential process variations.

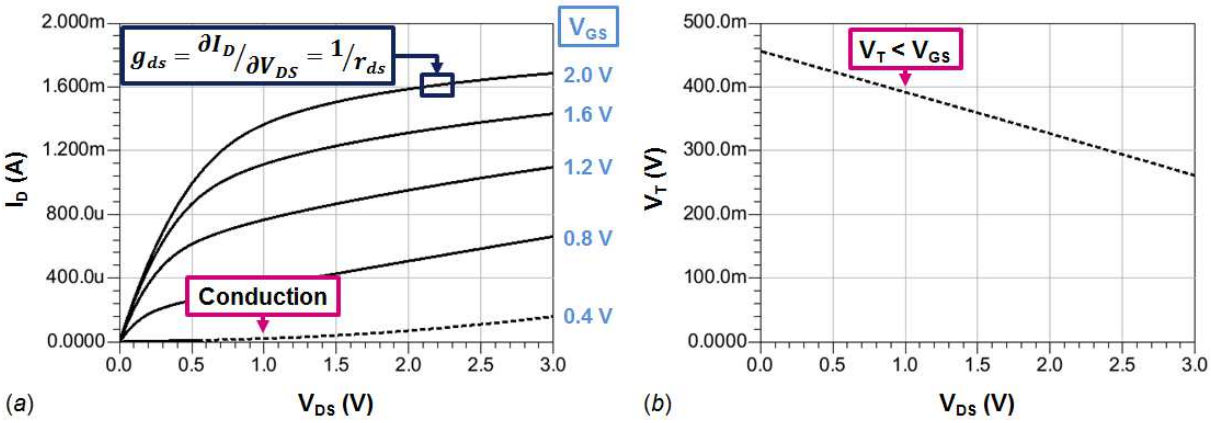

Figure 41: GO1 Low V_T BC NMOS (a) I_D for different V_{GS} and (b) V_T for V_{GS} = 0.4 V, both depending on V_{DS}

Parameter	Comment	Slow - Slow	Typical	Fast – Fast
V_T (mV)	V_{DS} = 0 V	484	456	424
	V_{DS} = 1 V	428	392	351
I_D (mA)	$V_{DS} = V_{GS}$ = 1 V	0.49	0.56	0.65
g_m (mS)	V_{DS} = 1.2 V	1.03	1.11	1.18
g_{ds} (mS)		0.14	0.16	0.19
g_m / g_{ds}		7.52	6.79	6.10
C_{gd} (fF)	V_{GS} = 0.8 V	0.32	0.32	0.32
C_{gs} (fF)		0.85	0.84	0.83
f_T (GHz)		141	152	163

Table 20: GO1 Low V_T BC NMOS DC parameter extraction, for typical and extreme processes

The DC parameters of the available GO1 NMOS devices are reported in Table 21. The device operating points and size are the same as in the study above. Only typical process is considered now. Besides, the PMOS devices are not in the scope of this comparison. As expected, the low V_T devices exhibit higher drain currents than the standard ones, for the same V_{GS} value. Besides, FB transistors have higher f_T than BC, thanks to lower gate capacitances. Their g_m/g_{ds} (self-gain), which is related to the circuit linearity, are also better, thanks to low g_{ds}. However, the floating body effects highly impact the high power performances.

Finally, the DC parameters of GO2 BC NMOS have not been extracted. Indeed, these devices are not based on a physical model, but on PSP SOI. Thus, some parameters are not extracted by the simulator. Besides, this device model has not been optimized for DC simulations, since it is dedicated to RF switch design.

Device	V_T^0 (mV)	I_D (mA)	g_m / g_{ds}	C_{gd} + C_{gs} (fF)	f_T (GHz)
GO1 Low V_T BC NMOS	456	0.56	6.79	1.16	152
GO1 Std V_T BC NMOS	538	0.43	9.38	1.07	145
GO1 Low V_T FB NMOS	449	0.51	8.22	0.53	261
GO1 Std V_T FB NMOS	515	0.47	9.75	0.47	303

Table 21: DC parameter comparison of the available GO1 NMOS devices, for typical process

To conclude the DC parameter extraction, the $I_D(V_{DS})$ curves of GO1 Low V_T NMOS BC, and GO1 Low V_T NMOS FB are compared in Figure 42. The first relevant result is that these devices have different behaviors at high V_{DS}, i.e. V_{DS} higher than its nominal maximum value (1.2 V). It is due to the parasitic bipolar effect in the FB NMOS. Besides, the kink effect seems to be not taken into account for the FB's.

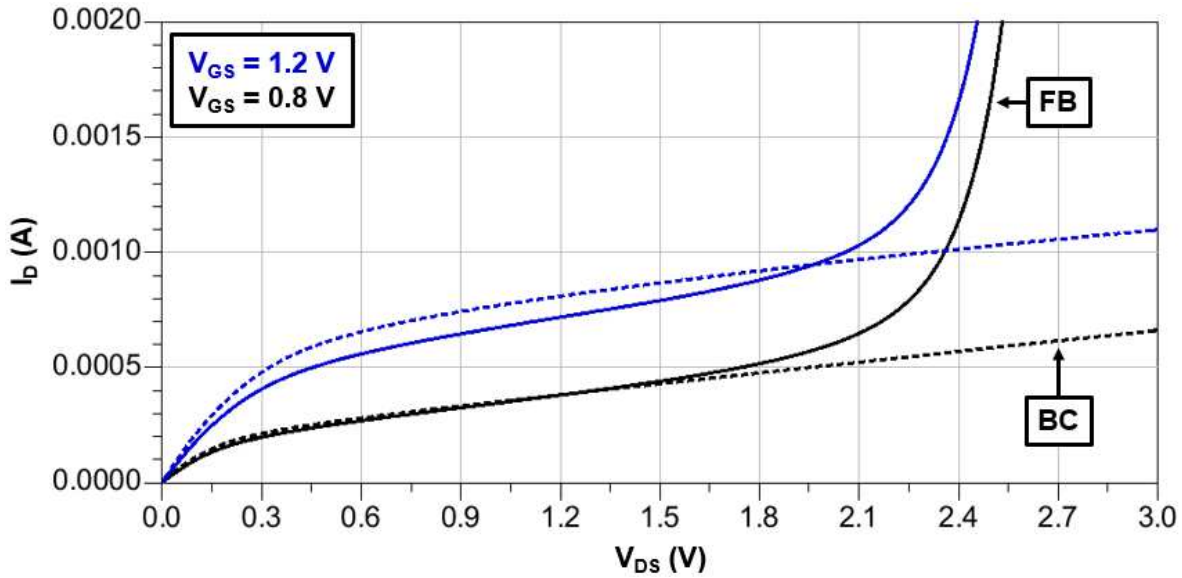

Figure 42: Comparison of $I_D(V_{DS})$ curves of GO1 low V_T NMOS BC and FB, for V_{GS} = 0.8 V and V_{GS} = 1.2 V

By now, the SP simulations are processed to extract the other parameters. As previously, the extraction is achieved for the GO1 Low V_T BC NMOS transistor. Then, it is applied to the other devices. Firstly, f_T and f_{max} are computed thanks to *Equation (29)* and *(36)*, as depicted in Figure 43. Secondly, the capacitances, and gate resistance are determined thanks to *Equations (30)-(33)*. These results are compared with the values obtained in the DC simulations. Lastly, the NF_{min} is observed at 5.86 GHzand 28.0 GHz.

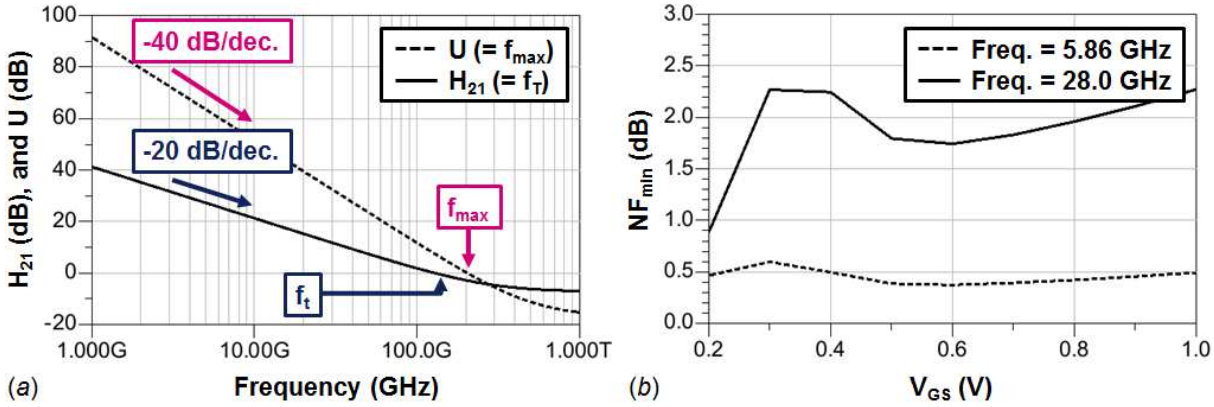

Figure 43: (a) H_{21} (dB) and U (dB), depending on frequency (GHz); (b) NF_{min} (dB), depending on V_{GS} (V)

The device parameters are extracted at V_{DS} = 1.2 V, and V_{GS} = 0.8 V. Since the capacitances and the gate resistance are computed from Y parameters, they also depend on frequency. Thus, they are extracted at 28 GHz. Besides, the minimum NF_{min} is obtained at V_{GS} = 0.6 V. The results are reported in Table 22.

Parameter	Comment	Slow - Slow	Typical	Fast - Fast
f_T (GHz)	DC Simulation Equation (28)	141	152	163
	SP Simulation Equation (29)	118	128	139
f_{max} (GHz)	SP Simulation Equation (36)	171	201	258
C_{gd} (fF)	V_{DS} = 1.2 V V_{GS} = 0.8 V @ 28 GHz	0.28	0.27	0.27
C_{gs} (fF)		1.23	1.24	1.25
C_{ds} (fF)		0.07	0.12	0.18
R_G (Ω)		329	252	177
f_{max} (GHz)	Equation (34)	226	271	341
NF_{min} (dB)	@ 5.86 GHz	0.44	0.37	0.31
	@ 28.0 GHz	2.04	1.74	1.46

Table 22: GO1 low V_T BC NMOS SP parameter extraction, for typical and extreme processes

The first relevant result is that there are discrepancies between the f_T approximation from *Equation (28)*, and the graphically determined f_T from *Equation (29)*. Moreover, the extracted gate capacitances are also different from the DC to the SP simulations. This impacts the computation of the f_T approximation. Likewise, there are mismatches between the approximated and the graphically obtained f_{max}'. Thus, it is chosen to estimate f_T and f_{max} graphically, since their approximations from the DC simulations are not accurate enough.

Device	f_T (GHz)	f_{max} (GHz)	R_G (Ω)	$C_{gd} + C_{gs}$ (fF)	NFmin (dB) @ 5.96 GHz	NFmin (dB) @ 28.2 GHz
GO1 Low V_T BC NMOS	128	201	252	1.51	0.37	1.74
GO1 Std V_T BC NMOS	113	209	257	1.50	0.67	2.09
GO1 Low V_T FB NMOS	145	209	284	1.06	0.23	1.08
GO1 Std V_T FB NMOS	157	227	282	1.01	0.19	0.91
GO2 Std V_T BC NMOS	31.6	200	87.2	1.67	0.26	1.19

Table 23: SP parameter comparison of the available NMOS devices, for typical process

The extracted SP parameters of the available devices are reported in Table 23. It exhibits that FB NMOS have higher f_T than BC transistors, thanks to their lower gate capacitances. This result was previously shown with the DC simulation approximations. In spite of high gate resistance, FB NMOS also provide good NF_{min}. Finally, since GO2 gate length is thrice the GO1 one, their R_G is roughly divided by three. This leads to good f_{max}, and NF_{min}. However, their poor f_T makes it not suitable for mm-wave design. As a conclusion, low V_T FB NMOS seems highly suitable for low-power low noise amplifier design whereas the BC NMOS will be preferred for high power circuits. These results will be compared with measurements, in Table 27 and Table 28. Comparatively, the 45 nm RF-SOI technology from *Global Foundries*, released in 2008, exhibits f_T/f_{max} about 290/410 GHz [14].

d. Electromagnetic Simulations of the Passives

In addition to active devices, the *ST C65SOIFEM* technology provides high quality factor passives. A 250 µm long inductance is simulated thanks to *Momentum 3D Planar EM Simulator*, from *Keysight*. This lumped element is made with the thick top metal level; it is depicted in Figure 44. The performances are observed for different widths: 2 µm and 5 µm. Besides, the designed circuits will be measured on wafer. Thus, it is necessary to characterize the ground – signal – ground (GSG) pads, also presented in Figure 44.

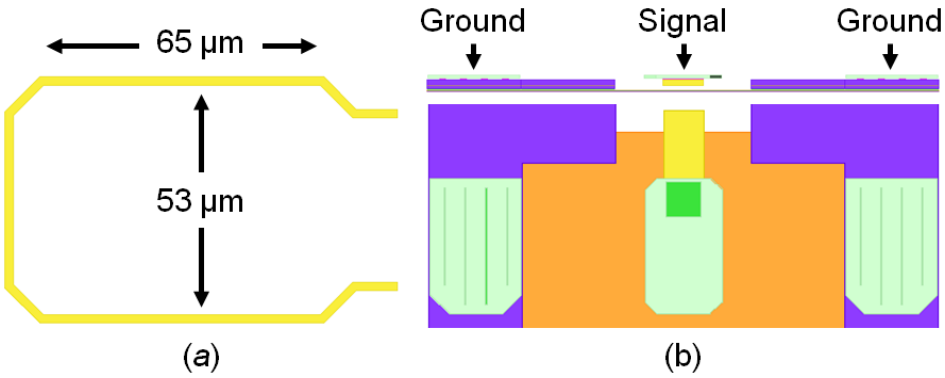

Figure 44: (a) A 250 µm Inductance; (b) Ground – signal – ground pads for on-wafer measurements

The simulation results (inductance value and quality factor, *Q*) are plotted in Figure 45. Wider path leads to lower inductance and enhanced the quality factor. The resonance frequency (f_R), defined by a quality factor equal to zero, is the same for both widths. Finally, the GSG pads have been simulated, and measured from 1 to 60 GHz. The simulated pads are equivalent to a capacitance about 40 fF, while the measured pads have a capacitance about 50 fF. This difference comes from the metal tilling, i.e. the addition of dummies for the manufacturing, which are not taken into account by the EM simulator.

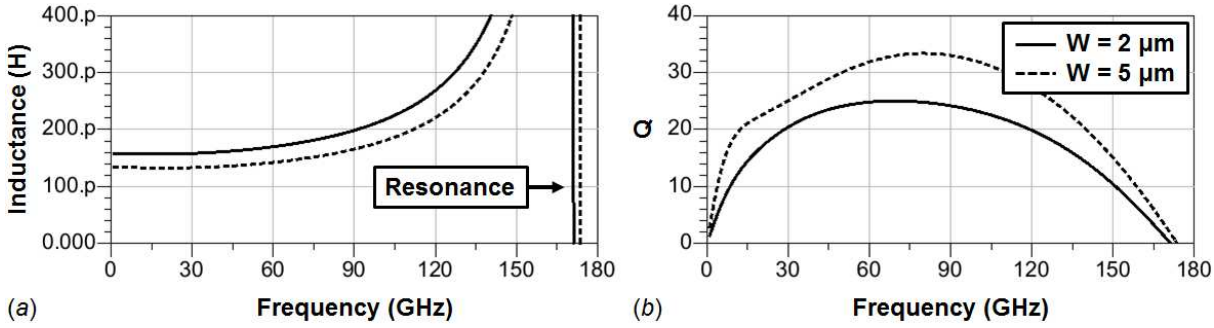

Figure 45: (a) Inductance (H) and (b) Inductor quality factor, both depending on frequency (GHz)

	Inductance (pH) F = 28 GHz	Bandwidth (GHz) Q > 20	Q_{max}	f_R (GHz)
W = 2 µm	158	28.7 – 119	25.0	171
W = 5 µm	133	12.5 – 140	33.3	174

Table 24: EM simulation results for 250 µm long inductors, with different widths (2 µm and 5 µm)

e. De-embedding Method

The on-wafer measurement bench is shown in Figure 46. A calibration is made in the S_1-G_1, and S_2-G_2 planes to remove the probe contributions. However, the effect of the pads is still present. Moreover, as demonstrated in the previous section, their behavior is not the same in simulations and in measurements. Thus, it is necessary to use de-embedding, to recover the device RF performances. Three de-embedding techniques are detailed in [15]: thru-reflect-line (TRL), thru-line (TL), and open-short (OS). Their names correspond to the structures needed to achieve the de-embedding.

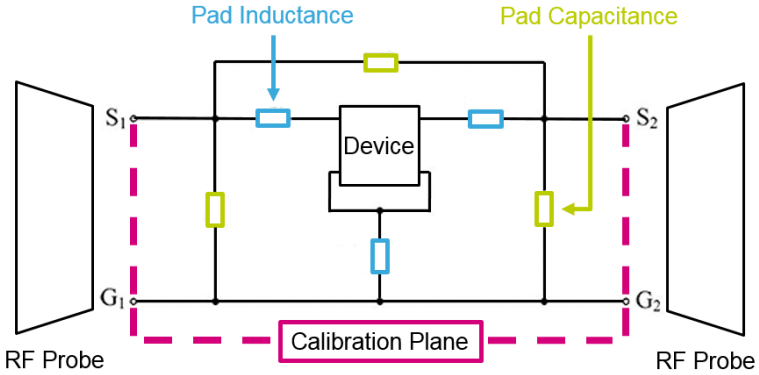

Figure 46: On-wafer measurement setup, including pad equivalent model

Inherently, TRL de-embedding needs more measurement structures. It is therefore not selected, since a less demanding method is preferred. This constraint may be relaxed using TL de-embedding. However, it is necessary to assume that both pads have exactly the same characteristics. Consequently, it is chosen to achieve OS de-embedding. Indeed, it is easier to implement, and the measurement setup is less limited. Firstly, the three measurement steps are performed (open, short, and device). The open structure has been depicted in Figure 44. For the short, both pads are connected between them and to the ground. The short and the equivalent models of each measured structure are shown in Figure 47.

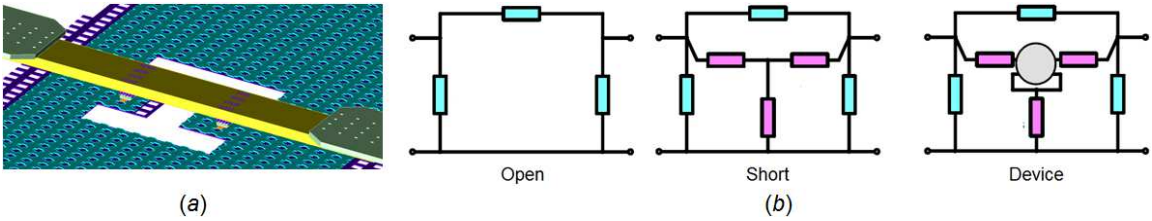

Figure 47: (a) 3D view of the short; (b) Equivalent models of the 3 measured structures

Secondly, three calculation steps are needed to recover the device performances. The pad capacitances are removed from the short, and the measured circuit Y matrixes, respectively Y_{SHORT}, and Y_{DUT}. It is done thanks to *Equation (37)* and *(38)*. The results are translated into Z parameters, and the new short matrix ($Z_{S\text{-}O}$) is deducted from the new measured circuit matrix ($Z_{D\text{-}O}$). This step is achieved with *Equation (39)*; it removes the remaining pad inductances from $Z_{D\text{-}O}$. The device performances (Z_{DEVICE}) are recovered.

$$Y_{D-O} = Y_{DUT} - Y_{OPEN} \quad (37)$$

$$Y_{S-O} = Y_{SHORT} - Y_{OPEN} \quad (38)$$

$$Z_{DEVICE} = Z_{D-O} - Z_{S-O} \quad (39)$$

Finally, a comparison of different de-embedding methods has been completed in [16]. It has highlighted that the accuracy of OS de-embedding is limited at high frequencies. Indeed, measurement errors higher than 5% have been sown above 40 GHz. However, this range of confidence is sufficient in this work.

f. Measurement Results

In order to check the device models, 5 transistors have been implemented on the first wafer: two GO1 low V_T FB NMOS dedicated to LNA design, and two GO2 standard V_T BC NMOS optimized for switch design. These ones are based on standard cells from the ST C65SOIFEM design kit. An additional custom device is also implemented, with double source and drain contact rows. These circuits are shown in Figure 48. The 3D view of the common source implementation is also presented in Figure 48. The source of the transistor is directly connected to the ground thanks to the bottom metal layer. Staircase vias are used at the drain to reduce the parasitic capacitances. Finally, a major part of the top metal line contribution is removed with de-embedding. This first run is known as *Sirius*. Its process has been reported as slow-slow. Firstly, a V_T extraction of the 40 µm GO1 low V_T FB NMOS, with V_{DS} = 0.8 V and V_{DS} = 1.2 V, is achieved in order to confirm the process corner. Results are reported in Table 25. The measured circuit seems to match with the expectations. However, it is shown that the DIBL effect is more significant in measurements than in simulations. The GO1 FB NMOS parameters are extracted for 40 and 160 µm devices. The results are reported in Table 26 and compared with the model expectations.

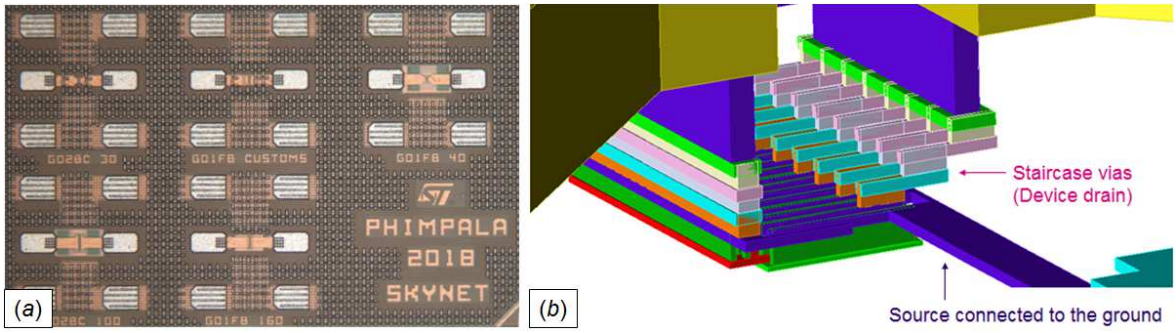

Figure 48: (a) Die micrograph; (b) Transistor's layout 3D view

40 µm GO1 low V_T FB NMOS	Simulations			Measurements		
	Fast-Fast	Typical	Slow-Slow	#1	#2	#3
$V_T^{0.8\ V}$ (mV)	349	397	444	442	445	442
$V_T^{1.2\ V}$ (mV)	282	332	381	329	336	326

Table 25: V_T extraction of 40 µm GO1 low V_T FB NMOS, simulations versus measurements

	Comments	40 µm GO1 FB NMOS		160 µm GO1 FB NMOS	
		Simulation	Measurement	Simulation	Measurement
V_T (mV)	V_{DS} = 0.8 V	444	442	436	405
I_D (mA)	V_{DS} = 1.2 V V_{GS} = 1.2 V	26.1	22.4	84.2	66.1
g_{ds} (mS)	V_{DS} = 1 V V_{GS} = 1 V	7.1	5.1	22	15
g_m (mS)		31.4	26.7	101	75.8
g_m / g_{ds}		4.42	5.24	4.59	5.05

Table 26: GO1 FB NMOS simulated and measured parameters

The measured devices exhibit I_D, g_m, and g_{ds} lower than expected. Actually, measurements are in good agreement with the models for low V_{DS}, i.e. V_{DS} lower than 0.4 V. Since the discrepancies occur for higher biasing points, it may come from self-heating, or from electro migration issues. Indeed, the vias may be too small to carry the expected current. Thus the two hypotheses are related. Moreover, the results for the 160 µm device are not 4 times higher than for the 40 µm one. It is due to the layout effects, which are more significant for a larger device. It is reminded that there is no de-embedding for DC measurements. Finally, the I_D (V_{DS}) curve of the 40 µm GO1 FB NMOS, plotted in Figure 49, highlight that the kink effect is not taken into account by the simulator. This effect also reduces the drain current. Besides, the first manufacturing of the GO2 devices was not successful. Indeed, the gate length is 20% shorter than expected. Measurements exhibit a significant V_T shift: it is two times higher than the model value. However, for the same V_{GS} - V_T, I_D and g_m measurements are good in agreement with simulations.

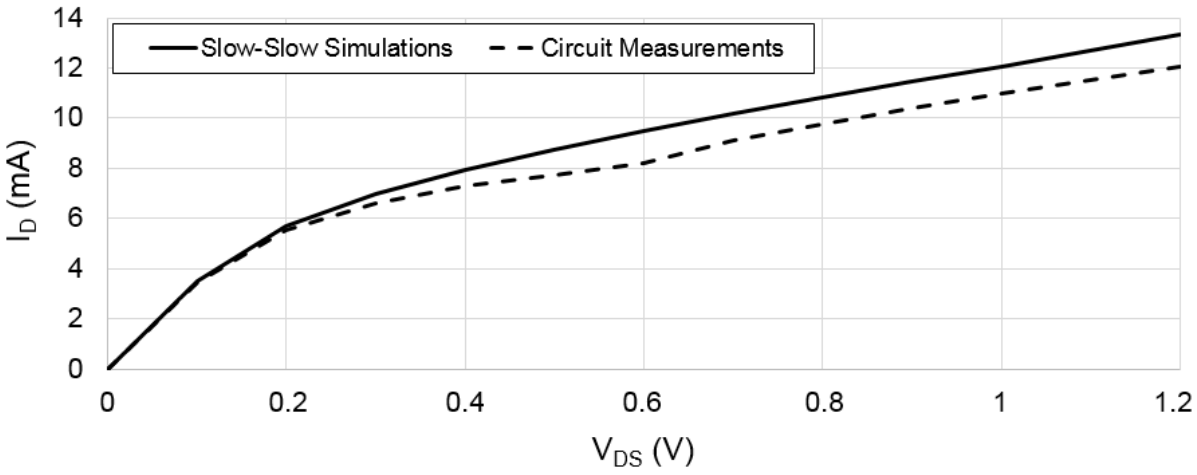

Figure 49: Simulated and measured I_D (V_{DS}) of a 40 µm GO1 FB NMOS device, at V_{GS} = 0.8 V

The small signal de-embedded measurement results are reported in Table 27, for GO1 FB NMOS. They are compared to the post layout simulations of the device, without pads and top metal lines. This exhibits a good agreement between the measurement results and the post layout simulation results. Besides, the custom device was designed to reduce the parasitic capacitances. This tentative is unsuccessful, since the custom device behavior is not predicted accurately by the model. Thus, only standard cells will be used.

		40 µm GO1 FB NMOS	**Custom GO1 FB NMOS**	**160 µm GO1 FB NMOS**
f_T (GHz)	Simulated	145.3	137.4	129.3
	Measured	135.7	128.3	120.8
f_{max} (GHz)	Simulated	136.1	157.2	134.6
	Measured	131.6	153.5	134.7
C_{gd} (fF)	Simulated	10.9	11.7	45.9
	Measured	12.1	12.7	51.3
C_{gs} (fF)	Simulated	25.3	27.1	83.6
	Measured	26.1	30.7	78.7
C_{ds} (fF)	Simulated	14.1	14.5	46.3
	Measured	14.9	18.9	42.5

Table 27: GO1 FB small signal measurement results and comparison with post layout simulation results

Finally, the small signal measurement results are reported in Table 28, for GO2 BC NMOS. Since the device gate is shorter than in simulations, the measured f_T / f_{max} are increased. Likewise, the device capacitances are also reduced. In spite of these discrepancies, there is a decent agreement between the measurement results and the post layout simulations, with harmonized V_{GS} - V_T.

		30 µm GO2 BC NMOS	100 µm GO2 BC NMOS
f_T (GHz)	Simulated	44.4	46.1
	Measured	41.9	48.8
f_{max} (GHz)	Simulated	99.6	102.7
	Measured	102.5	108.8
C_{gd} (fF)	Simulated	13.6	45.8
	Measured	10.9	38.6
C_{gs} (fF)	Simulated	32.1	93.9
	Measured	36.4	87.3
C_{ds} (fF)	Simulated	8.3	20.2
	Measured	13.4	14.7

Table 28: GO2 BC small signal measurement results and comparison with post layout simulation results

To summarize the results of this section, de-embedding enables to recover the core device performances with good accuracy. Moreover, it has been shown that post layout simulations and measurement results are in good agreement. This confirms the validity of the design kit models. However, there are still some discrepancies. For example, all the floating body effects are not taken into account yet.

II. Single Pole Double Throw RF Switches

1. Introduction

a. The Need of RF Switches

As presented in Chapter I, it is preferred to implement TDD systems for 5G EMBB. Indeed, it enables to achieve signal processing only once, at the base stations. Thus, the constraints on the user equipment are relaxed. Besides, it provides more scalability for the uplink/downlink ratio. It is easier to change the time interval lengths than the sub-channel frequency widths. However, TDD systems have larger delays than FDD ones, and they require accurate time synchronization [17]. Consequently, FDD will be preferred for applications requiring low latency, or low power consumption, in the sub 6 GHz spectrum.

TDD systems may be implemented in different ways. Firstly, separate arrays may be used for transmission and reception. This solution is quick to deploy, and has low loss between the PA/LNA and the antenna. However, it implies a high deployment cost, as two arrays are needed. Moreover, split TX and RX systems prevent the possibility to merge certain functions, as phase shifting. It is also possible to use different polarizations through a single antenna; the implementation cost is therefore reduced. However, it requires a specific care about TX to RX isolation. The other way to implement TDD systems is through the use of RF switches. There are various ways to realize such circuits. Mechanical switches have

good performances (insertion loss, isolation, and power handling), but are bulky, slow and expensive. Conversely, semiconductor switches have poorer performances, but they are smaller, faster, and cheaper. They may be implemented either thanks to PIN diodes, or FET's. The first solution provides high speed at a significant DC consumption cost, whereas the other one offers power consumption savings, but it is a bit slower. Finally, there is a tradeoff between mechanical and semiconductor switches: micro electro-mechanical systems (MEMS). They have good performances, with low power consumption, small area, and are virtually perfectly linear. However, they have poor power handling, and are slower than semiconductor circuits [18]. For the targeted applications, mechanical switches are discarded due to their weight, size and cost. Likewise, MEMS are not selected, as they have poor power handling. Finally, PIN diode switches have significant power consumption. Moreover, they are larger than switches based on FET's, since they require a choke inductor to block the DC current. Thus, FET RF switches are chosen for the targeted applications. Besides, they are easy to co-integrate in a single chip with the LNA and the PA. Switches with a single input and two outputs are known as single pole double throws (SPDT's).

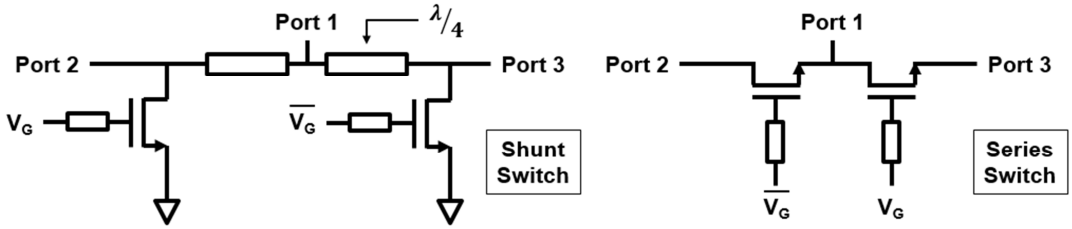

Figure 50: Shunt switch and series switch schematics

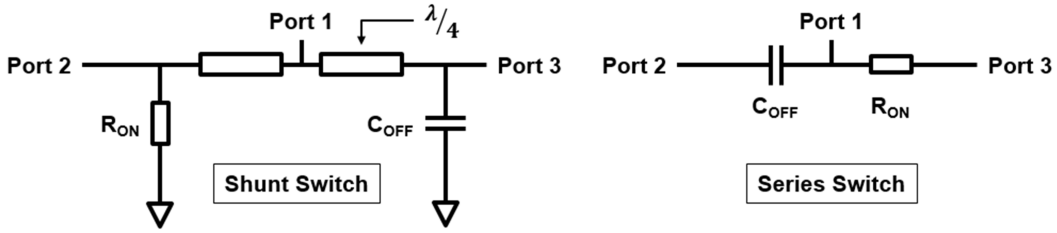

Figure 51: Shunt switch and series switch equivalent models ($V_G > V_T$)

Basically, switches based on FET may be composed of two shunt or series devices. They are inversely controlled, as illustrated in Figure 50. Shunt switches require additional quarter-wave lines, which prevent signal flowing through the undesired path. In series switches, the OFF state device degrades the insertion loss, since it is equivalent to a capacitance (C_{OFF}). Furthermore, switches are not only used as duplexers. They are useful building blocks for adaptable circuits, as variable phase shifters, variable gain amplifiers, and multi-standard circuits (PA's, LNA's…). Several mm-wave switches are reported in the next section.

b. Ka-Band Semiconductor Switches

RF semiconductor switches are commonly based on III-V transistors or diodes, since they provide low insertion loss, and high power handling. The commercial circuit presented in [19] exhibits such features. However, it includes inductors, which limit the bandwidth and increase the chip size. Besides, the CMOS technology scaling has stimulated the interest for switches based on this technology, notably below 6 GHz. In [20], the major limitations of millimeter-wave CMOS switches are detailed. It is emphasized

that the drain to substrate and source to substrate junction diode parasitic capacitances (respectively C_{jd} and C_{js}) increase the signal loss and limit the voltage swing. The model of an NMOS switch and its ON-state and OFF-state equivalent models are presented in

Figure 52, including the substrate resistance, and the capacitances of the junction diodes. It is detailed in [20] that higher substrate resistance reduces the loss, but degrades the isolation.

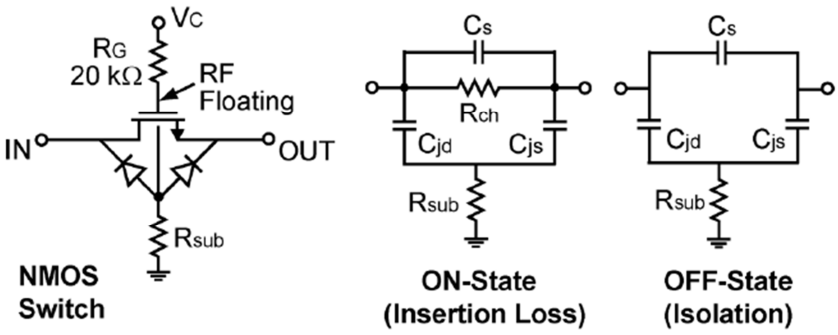

Figure 52: NMOS switch and its ON-state and OFF-state equivalent models, from [20]

Thus, it is possible to reduce the loss by adding inductors at the transistor's input and output, i.e. at its drain and source. Besides, an inductor may be connected from the transistor's drain to its source, in order to resonate with C_{OFF}, which improves the switch performances. Either solution highly increases the circuit area, and therefore the overall cost. In [21], a compact solution to reduce the loss through C_{OFF} is proposed. It is based on a single switched inductor, as depicted in Figure 53. Thus, the extra area is limited. However, this circuit suffers from a medium power handling, since it is implemented using a bulk CMOS technology.

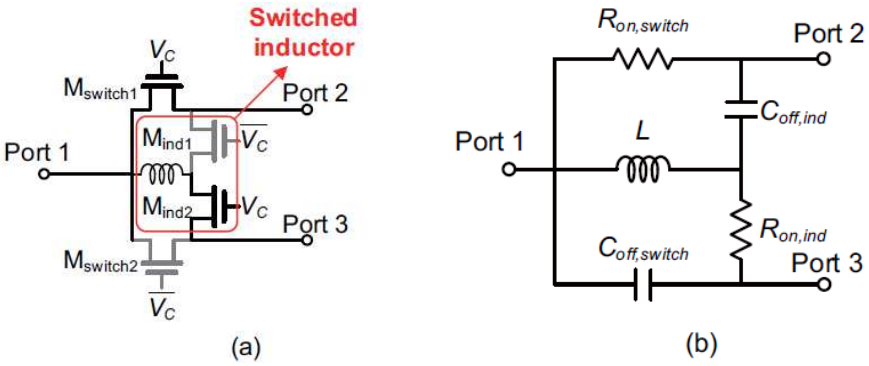

Figure 53: (a) SPT switch with switched inductor and (b) its equivalent circuit with Port 1 – Port 2 path turned on, from [21]

On the other hand, it is shown in [22] that the use of SOI technologies improves the switch performances. Indeed, the device is isolated from the substrate. The previously mentioned drawbacks are therefore avoided. This results in a very compact switch, with good performances. Moreover, layout techniques are introduced to minimize the parasitic capacitances brought by the BEOL. It includes staircased back end layout, and large width between polysilicon gate contacts. SiGe heterojunction bipolar transistors (HBT's) are sometimes chosen to overcome CMOS power handling limitations [23]. In [24], reverse saturation is used to reduce the insertion loss of the proposed 94 GHz HBT switch. However, this circuit requires several λ/4 lines, which lead to a significant area penalty at lower

frequencies. It is also possible to realize the duplexer without physical switches [25]. As illustrated in Figure 54, the PA output is connected to the LNA input. The corresponding impedances are close to ensure a proper matching based on a single network. Both ports are connected to a transformer, which is controlled by V_{CT}. Moreover, the PA and LNA supplies are turned on or off depending on the desired duplexer state. The use of a transformer results in a bulky system. With a careful design, it is possible to reduce the effect of the OFF state circuit on the activated path. Thus, the duplexer losses are reduced to the transformer losses. Finally, asymmetrical switches may be implemented in order to boost the performances of a selected path [26], [27]. In both cases, the TX gain is enhanced, at the cost of a degraded RX NF. The second asymmetrical switch is illustrated in Figure 54. There is a tradeoff between the TX linearity and the RX NF, depending on C_{OUT}. As a conclusion, several Ka-band switches are reported in Table 29. Their advantages and drawbacks are therefore compared.

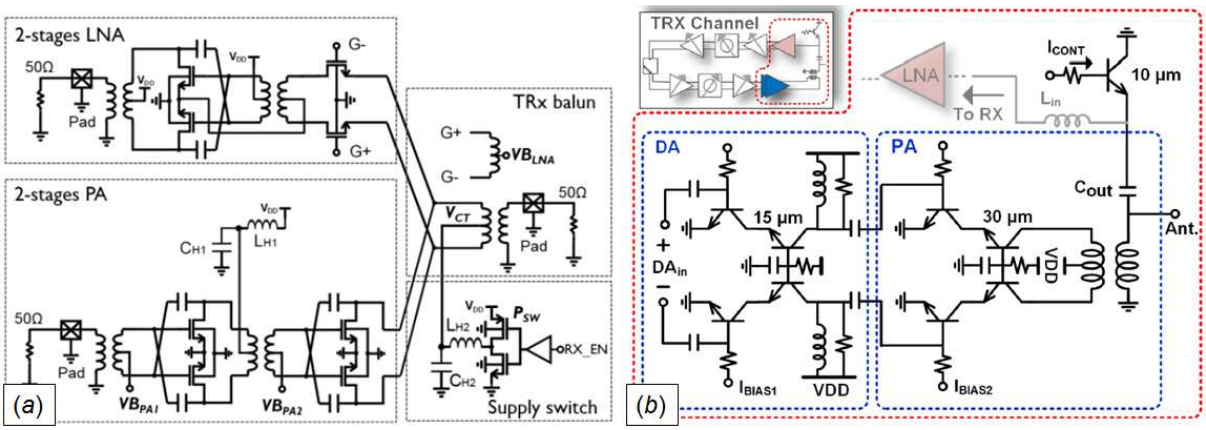

Figure 54: (a) Ka-Band transceiver without physical switch, from [25]; (b) Asymmetrical SPDT switch, from [27]

Reference	[**19**]	[**21**]		[**22**]	[**23**]
Process	GaAs	65 nm CMOS		45 nm SOI	130 nm SiGe
IL (dB)	0.9	1.06	0.89	0.76	1.5
Isolation (dB)	> 25	39	18	23	21
Bandwidth (GHz)	27 – 46	25 – 39.5		DC – 50	20 – 32
ICP1 (dBm)	33	7.25	12.55	29	17
Size (mm²)	1.19	1×10^{-2}	9×10^{-3}	4×10^{-3}	1.19×10^{-1}
Comments	PIN Diodes	Switched Inductor		Series-Shunt	Shunt

Table 29: Ka-band integrated SPDT switches for FEM state of the art

In the light of the above elements, series-shunt switches based on SOI technology seems highly suitable for the targeted applications. Indeed, it exhibits low insertion loss, small area, and high bandwidth. It

also provides enough isolation to protect the LNA in the TX state. Finally, a proper implementation may result in a high power handling compared with the other CMOS technologies [22].

2. Series-Shunt SPDT Switches

a. Operating Principle

As its name suggests, the series-shunt switch is a combination of the series and the shunt switches, which have been previously presented. The addition of a shunt switch improves the series switch isolation, while degrading its insertion loss, and limiting its power handling. Besides, the series switches replace the λ/4 lines. Thus, the circuit size is reduced. The series-shunt switch schematic and its implementation within a FEM are presented in Figure 55. The equivalent schematic of this circuit in the TX state, i.e. when $V_C > V_T$, is also shown. As detailed in Figure 51, the turned on, and turned off transistors are respectively equivalent to a resistor (R_{ON}), and a capacitor (C_{OFF}). Their values may be calculated thanks to *Equation (40)*, and *(41)*, from [28]. They depend on the drain to source resistance (r_{ds}), and on the transistor's capacitances.

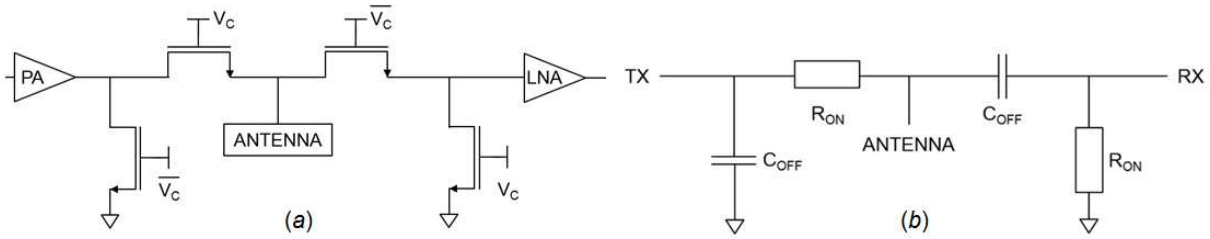

Figure 55: (a) Series-shunt CMOS SPDT implemented within a FEM; (b) its equivalent model in TX mode

$$R_{ON} = r_{ds}{}^{ON} \tag{40}$$

$$C_{OFF} = C_{ds} + \frac{C_{gs} * C_{gd}}{C_{gs} + C_{gd}} \tag{41}$$

Consequently, larger devices will exhibit lower R_{ON}, but also higher C_{OFF}. Moreover, the above equivalent model does not take into account the ON state capacitance, and the OFF state resistance of the transistor. Indeed, they are supposed to be negligible against R_{ON}, and C_{OFF}. Since the devices are isolated for the substrate in SOI technologies, the junction capacitances are also ignored. In order to compute the series-shunt switch performances, this simplified model is combined with *Equation (42)* and *(43)*, from [18]. They respectively give the loss through a series and a shunt switch, depicted in Figure 56. Either IL or isolation may be determined, depending on the transistor's impedance (Z_d). Results are reported in Table 30.

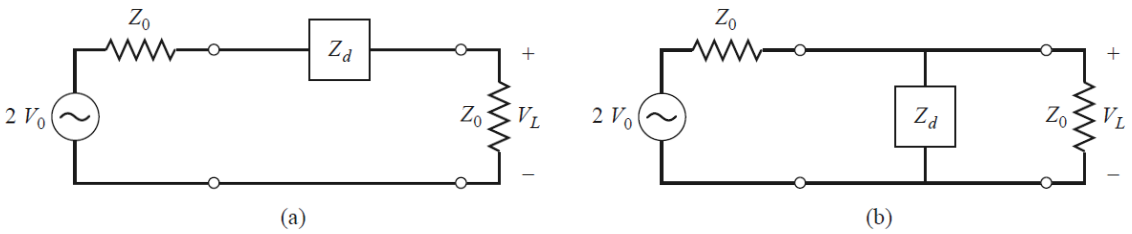

Figure 56: (a) Equivalent schematic of the series switch; (b) Equivalent schematic of the shunt switch, both from [18]

$$\left|\frac{V_L}{V_0}\right|_{series} = \left|\frac{2 * Z_0}{(2 * Z_0) + Z_d}\right| \quad (42)$$

$$\left|\frac{V_L}{V_0}\right|_{shunt} = \left|\frac{2 * Z_d}{(2 * Z_d) + Z_0}\right| \quad (43)$$

Since R_{ON} and C_{OFF} are conversely related to the transistor's width, the above results highlight that there is a significant tradeoff on this value. Indeed, improving the series ON state device reduces the shunt ON state and the series OFF state performances. Thus, there is an optimal size to maximize the performances.

Switches	Series	Shunt
ON	$2 * Z_0 \big/ (2 * Z_0) + R_{ON}$	$2 \big/ 2 + jC_{OFF}Z_0\omega$
OFF	$2 * jC_{OFF}Z_0\omega \big/ (2 * jC_{OFF}Z_0\omega) + 1$	$2 * R_{ON} \big/ (2 * R_{ON}) + Z_0$

Table 30: V_L / V_0 for series and shunt switches, in ON and OFF states

Finally, the worst case for the switch power handling is the TX state, depicted in Figure 55. Both series and shunt OFF state switches therefore have to handle a voltage swing given by *Equation (44)*. It depends on the transmitted power P_{TX}, and the PA output impedance Z_{TX}. Thus, the maximum input power of the switch is related to the maximum drain to source voltage of the selected technology. The nominal limits for *ST C65SOIFEM* GO1 and GO2 devices are respectively 1.2 V and 2.5 V, with a 10% tolerance margin. Moreover, it is assumed that Z_{TX} is equal to 50 Ω.

$$V_{TX} = \sqrt{2 * P_{TX} * Z_{TX}} \quad (44)$$

To ensure a sufficient linearity, the power handling is therefore limited to 12.4 dBm for GO1 devices, and 18.8 dBm for GO2, with a 50 Ω transmitter. However, several transistors may be stacked to improve the power handling, as in [22]. Indeed, the voltage swing is distributed among the stacked devices. Such implementation usually requires triple N-well. However, it is not necessary with SOI technology, since it provides high isolation between each transistor. The theoretical maximum power handling depending on the gate oxide width and the number of devices is reported in Table 31. For the targeted application, a single GO2, or 2 stacked GO1 are required.

Devices	Single	2 Stacked	3 Stacked	4 Stacked
GO1	12.4 dBm	18.4 dBm	22.0 dBm	24.5 dBm
GO2	18.8 dBm	24.8 dBm	28.3 dBm	30.8 dBm

Table 31: Series-shunt power handling (dBm) depending on the number of GO1 and GO2 stacked devices, from Equation (44)

As illustrated in Figure 57, stacking transistors increase R_{ON} and decrease C_{OFF}. This could be prevented using larger devices. Consequently, the switch insertion loss and isolation are preserved, while its power

handling is improved. Actually, this highlights that the global switch performances mostly depend on R_{ON} and C_{OFF}. Thus, a figure of merit is defined by *Equation (45)*. It depends only on the technology features.

$$FOM_{SW.} = R_{ON} * C_{OFF} \quad (45)$$

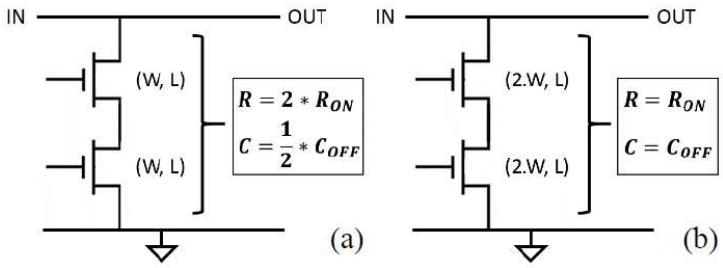

Figure 57: Stacked devices equivalent models

b. Series-Shunt Switch Sizing Method

Two series-shunt SPDT switches have been designed, using GO1 FB and GO2 BC NMOS transistors. They are depicted in 78Figure 58, and Figure 59. Thus, the performances of an intrinsic floating body transistor (FB) are compared with the "virtual" floating body, i.e. a BC device with a large resistance on its body. This implementation reduces the switch insertion loss, and enables to cope with the power handling issues of the intrinsic FB device [29], [30]. As expected, the GO1 FB's are stacked to ensure sufficient linearity.

Moreover, large resistances are connected from the drain to the source of each GO1. Thus, the voltage swing is equally divided among the stacked transistors. In order to preserve the switch performances, the resistances have been chosen large enough to be negligible against R_{ON}. Finally, the floating body and floating gate resistors of both SPDT's are 36 kΩ and the GO1 voltage balancing resistors are 18 kΩ.

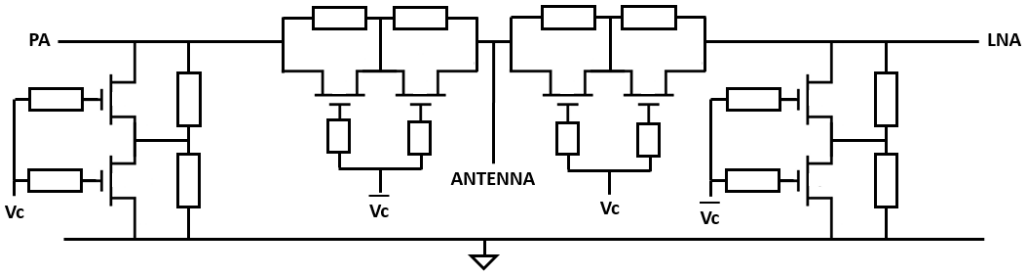

Figure 58: Schematic of the GO1 FB series-shunt SPDT

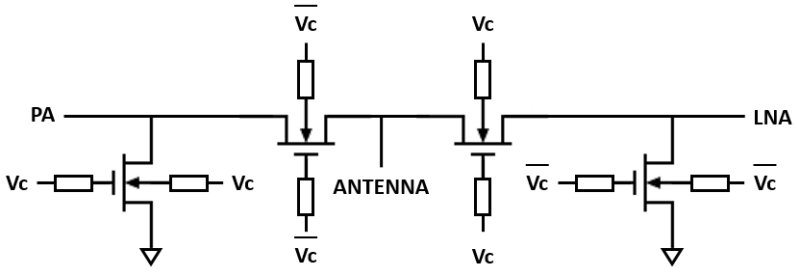

Figure 59: Schematic of the GO2 BC series-shunt SPDT

As introduced previously, the proper sizing of the switch results in optimal performances. The switch sizing method is detailed for the above presented GO1 SPDT switch, illustrated in Figure 58. The same method will be used for the GO2 BC circuit.

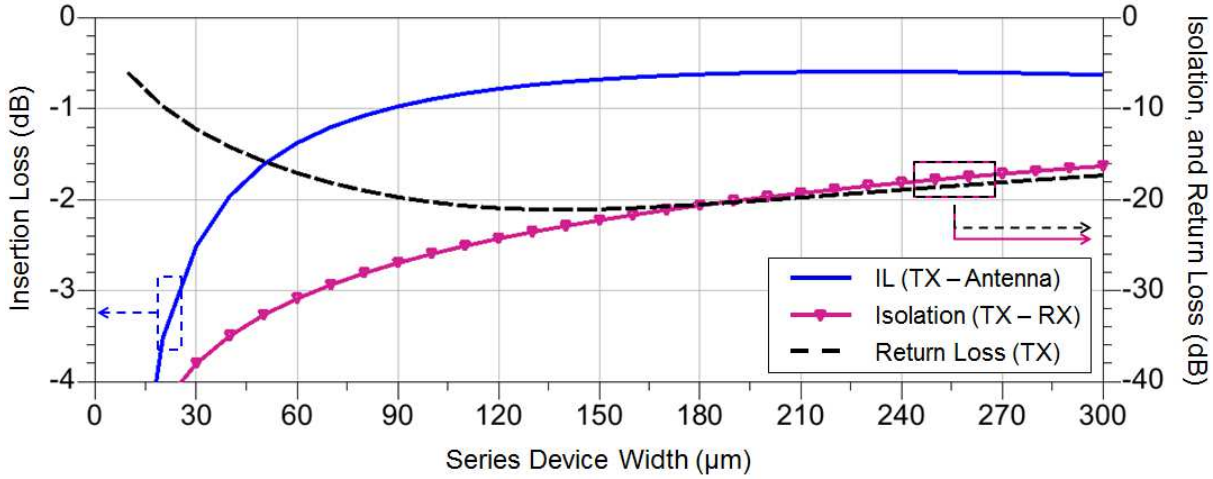

Figure 60: GO1 FB SPDT switch performances (IL, isolation, and return loss) depending on the series device width at 28 GHz

Firstly, the switch performances (IL, isolation, and return loss of the TX port) are simulated, depending on the series device width. Both series devices have the same dimension, and the shunt device width is locked. The results are plotted at 28 GHz in Figure 60. To obtain a low loss switch with sufficient isolation, the series device width is set to 160 µm. Likewise, the shunt device width is increased up to its optimal value, as shown in Figure 61. The shunt device width is therefore set to 40 µm.

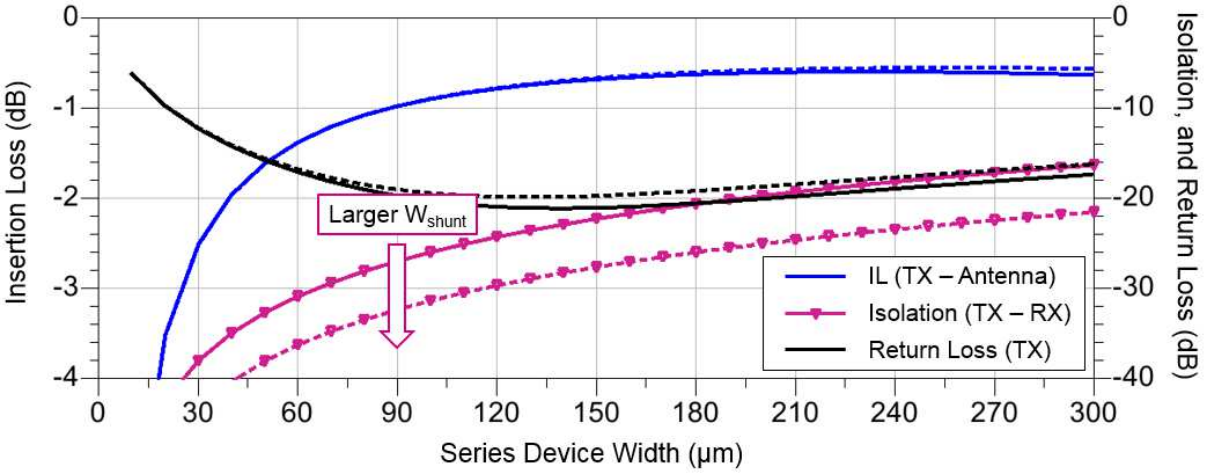

Figure 61: GO1 FB SPDT switch performances at 28 GHz depending on the series device width, for 10 µm and 50 µm shunts

The simulated results of the final circuit are plotted in Figure 62. In the TX state, it exhibits 0.6 dB insertion loss, isolation about 25.3 dB, and 19.4 dB return loss at 28 GHz. The GO2 BC switch is sized with the same method. This results in 100 µm series width, and 30 µm shunt width. The simulated performances are 0.4 dB insertion loss, isolation about 31.1 dB, and 14.9 dB TX port return loss at 28 GHz. However, it has lower bandwidth than the GO1 switch.

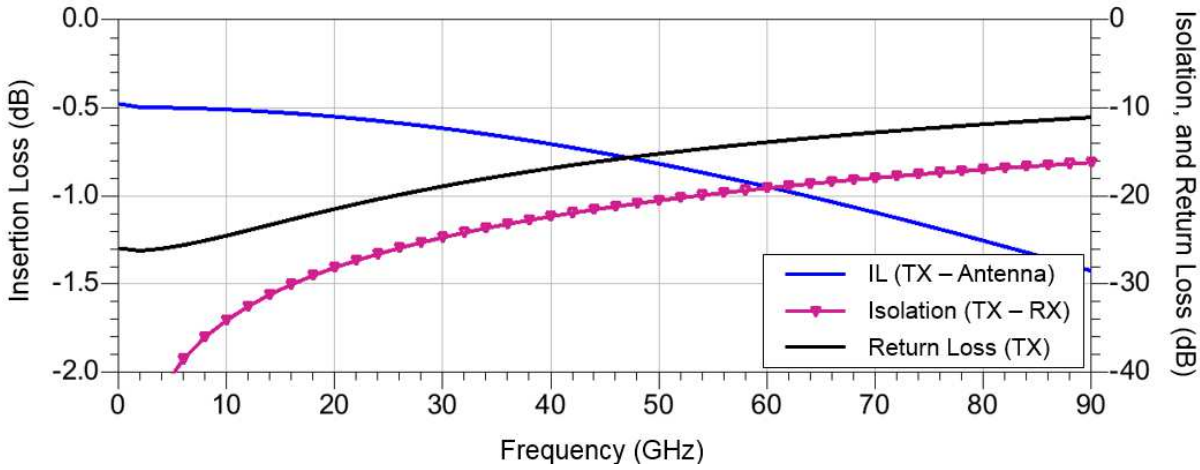

Figure 62: GO1 FB SPDT switch performances depending on frequency

c. *Circuit Layouts and Transistors' Figure of Merit Measurements*

Firstly, the GO1 SPDT switch is laid out, as depicted in Figure 63. The stacked series devices are connected to each other thanks to M1 and M2. The control signal comes from the top of the circuit. The floating gate and voltage balancing resistances are respectively above, and below the transistors. The vias from the top metal (in gold) to the transistors are staircased, in order to reduce the parasitic capacitances. The stacked shunt devices are connected to each other thanks to M1 only. The floating gate and balancing resistances are on both sides of the transistors. Then, it is connected to the ground through M2. The GO1 circuit area is $2.3 * 10^{-3}$ mm². The GO2 BC switch layout is similar to the other, with a global area of $2.4 * 10^{-3}$ mm².

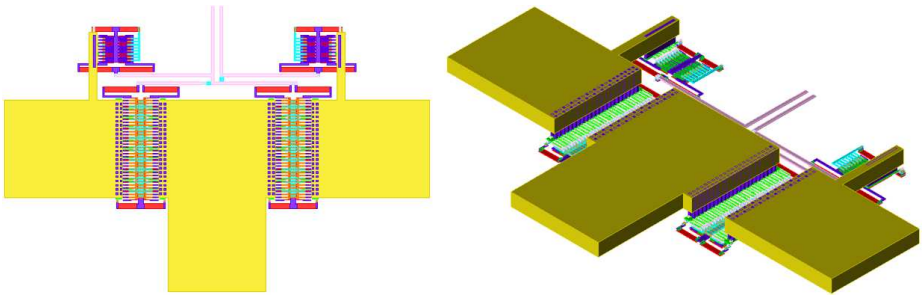

Figure 63: Layout of the GO1 FB series-shunt SPDT, from Sirius

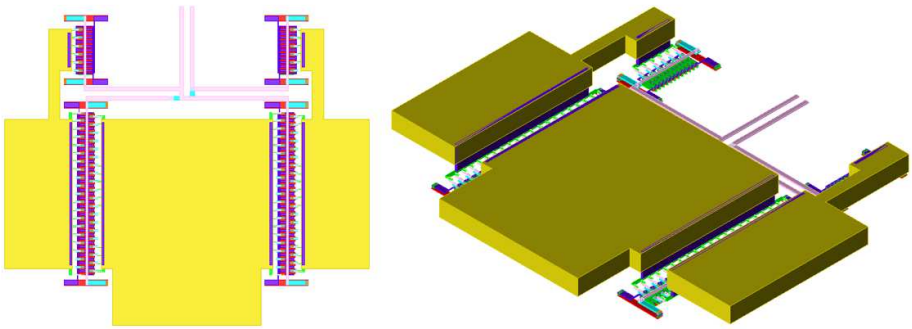

Figure 64: Layout of the GO2 BC series-shunt SPDT, from Sirius

Before measuring the switch performances, the FOM of both GO1 and GO2 devices are extracted. These measurements are realized thanks to the bench previously presented, in Figure 39. Port 1 and Port 2 are respectively used to control the transistor state, and to achieve the small signal measurement. Thus, R_{ON} and C_{OFF} are computed thanks to *Equation (44)*, and *(45)*.

$$R_{ON} = {1}/{\mathrm{Real}(Y_{22})} \tag{44}$$

$$C_{OFF} = {\mathrm{Imag}(Y_{22})}/{2 * \pi * f} \tag{45}$$

The 40 µm GO1 FB measurement results are reported in Table 33. They are compared with the post layout simulation results, with and without de-embedding. The first remarkable result is that the difference between the simulated and measured pad capacitance has an impact on C_{OFF} measurement. This mismatch is corrected by the de-embedding. After de-embedding, the measured C_{OFF} remains slightly higher than expected. Conversely, the measured R_{ON} is lower than it was expected to be. This probably reflects a small shift of the device dimensions. The results reported in Table 33 exhibit similar trends for the 160 µm GO1 FB device. Since R_{ON} and C_{OFF} are varying inversely, the device FOM are in good agreement.

		Before De-embedding		**After De-embedding**	
		Simulation	**Measurement**	**Simulation**	**Measurement**
40 µm GO1	R_{ON} (Ω)	14.9	12.9	14.1	12.1
	C_{OFF} (fF)	60.4	73.5	18.9	22.6
	FOM (fs)	900	948	266	273

Table 32: 40 µm GO1 FB post layout simulation and measurement results, with and without de-embedding

		Simulation	**Measurement**
160 µm GO1	R_{ON} (Ω)	5.0	4.6
	C_{OFF} (fF)	76.6	88.1
	FOM (fs)	383	405
30 µm GO2	R_{ON} (Ω)	17.0	21.4
	C_{OFF} (fF)	21.7	24.3
	FOM (fs)	369	520
100 µm GO2	R_{ON} (Ω)	6.8	8.5
	C_{OFF} (fF)	66.4	54.5
	FOM (fs)	425	463

Table 33: GO1 and GO2 post layout simulation and measurement results, after de-embedding

Besides, the GO2 BC measurement results confirm their unsuccessful manufacturing. Indeed, their gates are shorter than expected. Thus, they exhibit high R_{ON}, but reduced C_{OFF}. Finally, their FOM is significantly degraded. This will result in poor switch performances. However, the GO2 SPDT switch has been manufactured again, in the third run. The measurement results of the de-embedded GO1 FB series-shunt switch are presented in the next section. The results for the new GO2 BC switch are also shown. However, there were no de-embedding structures on the third run. Thus, only embedded results will be presented. Comparatively, the RF-SOI technology used in [22] exhibits a FOM of 90 fs.

d. SPDT Switch Measurement Results

Both fabricated circuits are depicted in Figure 65. Firstly, the GO1 FB SPDT is measured. The results before, and after de-embedding are plotted in Figure 66. The measurements including pads exhibit results in very good agreement. This confirms that the passives (pads, access lines…) are accurately taken into account by the simulator. The de-embedded measurements also show good agreement for IL and isolation. The simulated and measured IL's are 0.8 dB at 28 GHz. The simulated and measured isolation are respectively 22 dB and 21.3 dB, also at 28 GHz. Besides, there is a mismatch between the simulated and the measured return loss. Since the correlation is correct before de-embedding, it is supposed to be responsible for this discrepancy. To support this assumption, the embedded simulation results are de-embedded thanks to the simulated open and shorts. The results are plotted in Figure 67.

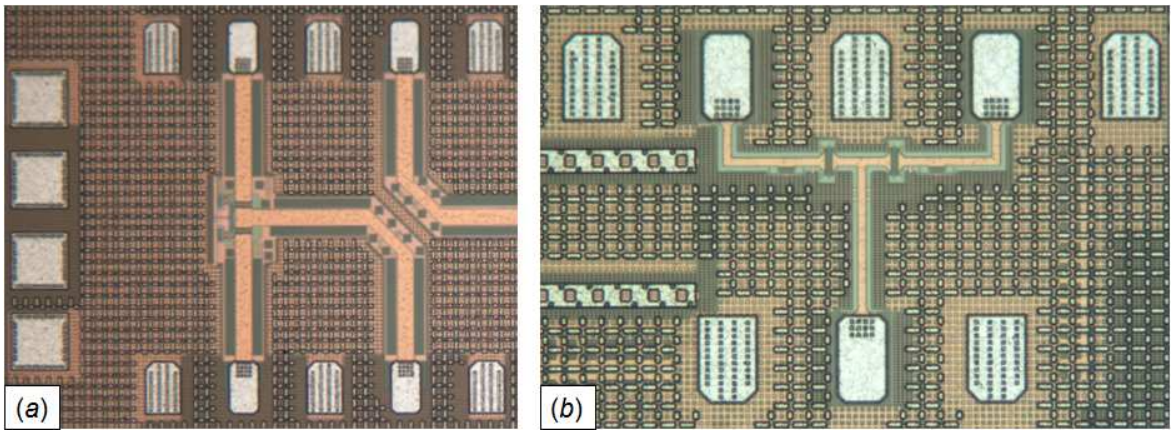

Figure 65: (a) Micrograph of the GO1 SPDT from Sirius; (b) Micrograph of the GO2 SPDT from Alioth

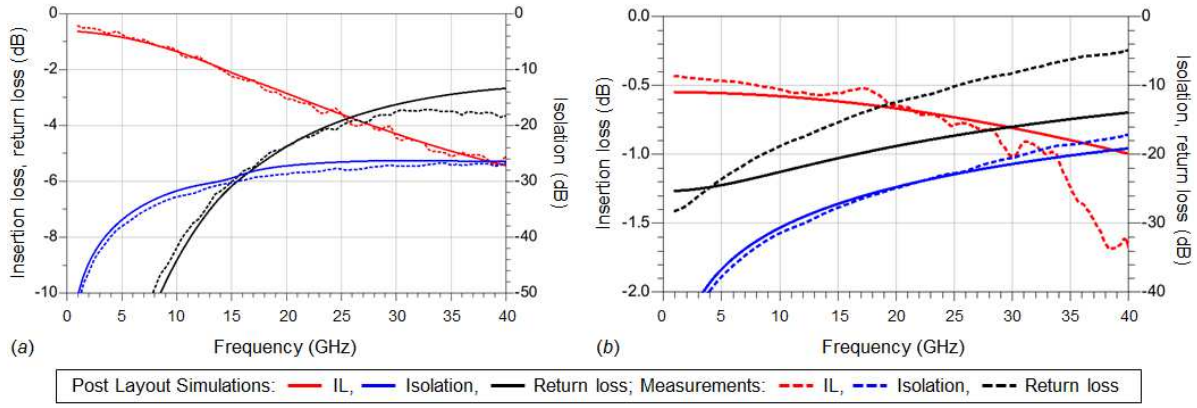

Figure 66: (a) Embedded, and (b) de-embedded simulation, and measurement results for the GO1 FB SPDT switch

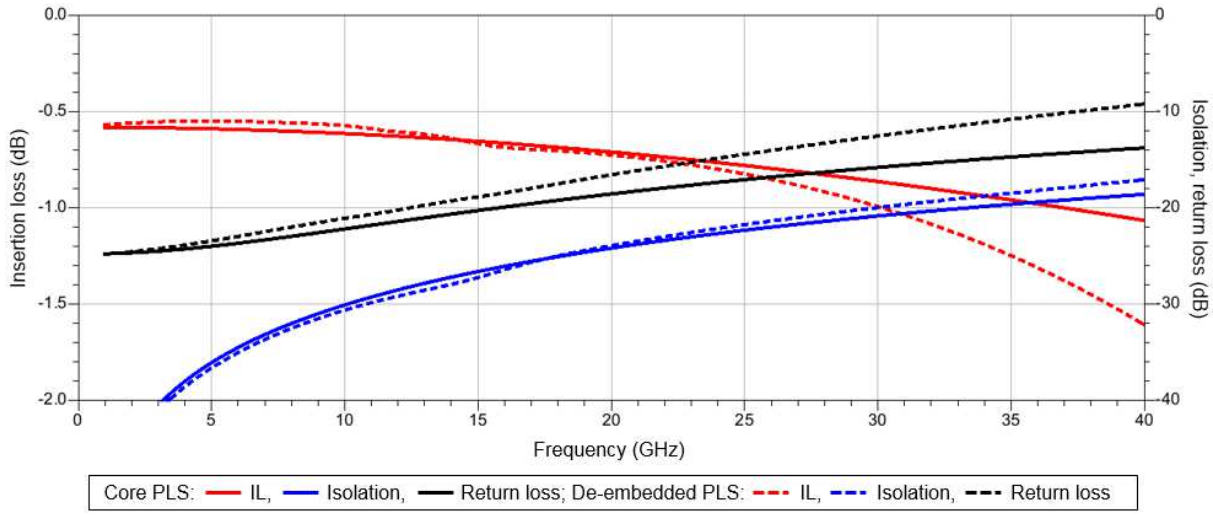

Figure 67: Core post layout simulation and de-embedded post layout simulation results for the GO1 FB SPDT switch

It should be noticed that de-embedded PLS and measurements have the same trends: high frequency IL mismatch, isolation inflexion point, and discrepancy for the TX port return loss. Thus, de-embedding is mostly responsible for these differences. Indeed, as introduced previously, the selected de-embedding method has limited accuracy after 40 GHz. Moreover, as shown in Figure 65, the de-embedding structures are too close to the SPDT. This results in signal coupling, which degrades the measurement accuracy. The poor switch matching, i.e. the measured high return loss, also impacts the de-embedding effectiveness.

Due to manufacturing issues, the GO2 BC SPDT switch of the *Sirius* run exhibits poor performances. The de-embedded measured IL's are higher than 1.6 dB at 28 GHz, while less than 0.9 dB were expected. Thus, these results are not detailed here. However, a second GO2 BC SPDT switch has been designed for the *Alioth* run. It is depicted in Figure 65. Its layout is close to the previously presented switch, as illustrated in Figure 68. Besides, this one is a bit larger, with an area about $4.48 * 10^{-3}$ mm². But the gate and body control are enhanced, thanks to double accesses. The shunt switch is connected to the ground through the control wire shields (drawn in green).

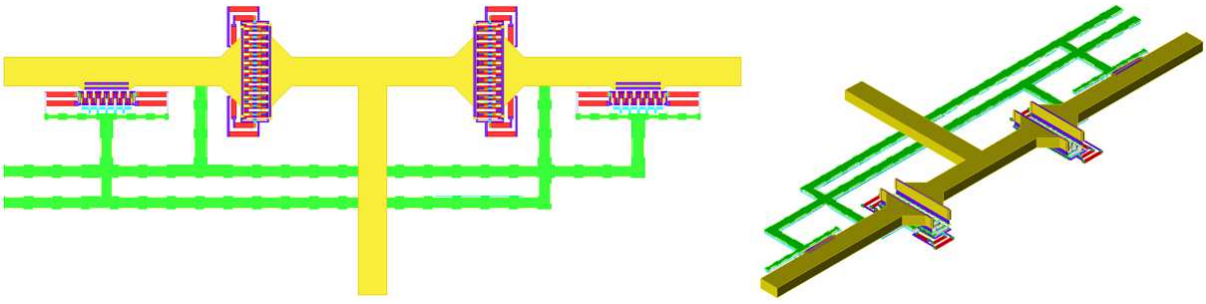

Figure 68: Layout of the GO2 BC series-shunt SPDT switch, from Alioth

The post layout simulation, and measurement results are plotted in Figure 69. These exhibit a relatively good measurement to hardware correlation (MHC). At 28 GHz, the simulated IL's are respectively 1.12 dB, and 1.16 dB. Likewise, the isolations are 26.1 dB and 31 dB. And the return losses are 10.7 dB, and 12.9 dB. The isolation discrepancy may come from impedance mismatch at the RX port, since it is

disabled. Whereas the return loss variation is supposed to come from layout elements which are not taken into account by the simulator, such as the metal dummies. Finally, the post layout simulation results without pads results in 0.73 dB IL, and 22 dB isolation, in TX mode, at 28 GHz. The symmetry of the TX and RX states is also verified for both GO1 FB, and GO2 BC SPDT switches.

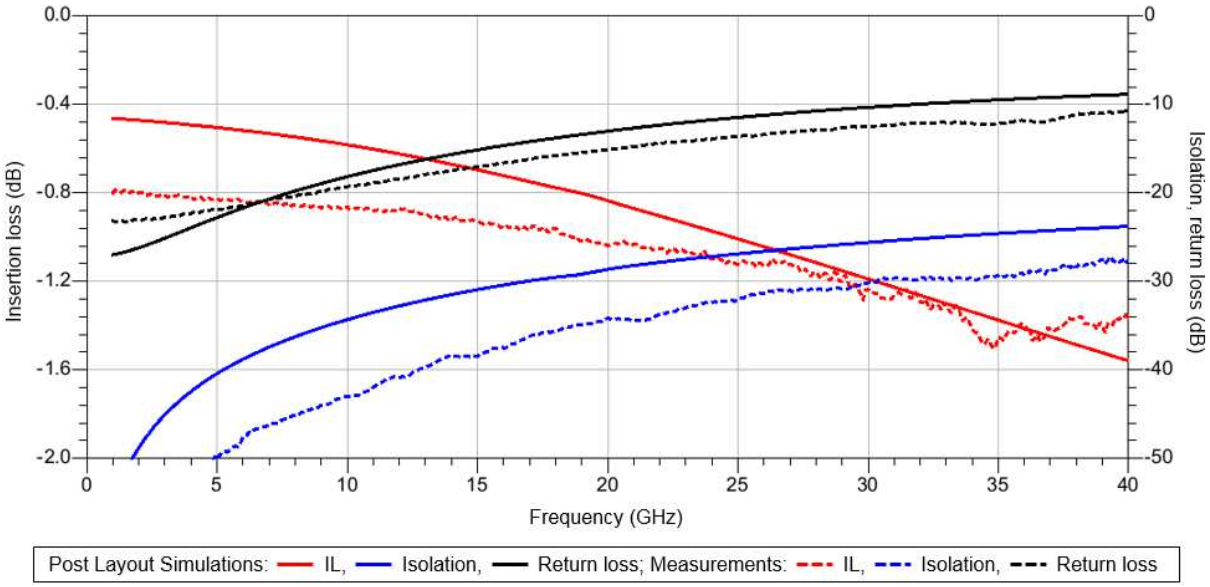

Figure 69: GO2 BC SPDT switch simulation and measurement results, including pads

Finally, continuous wave power measurements are achieved in order to determine the ICP1 of GO1 and GO2 switches. The measurement steps and the associated benches are presented in Figure 70. The first step enables to determine the effective power at the probe output, i.e. at the circuit input (P_{input}). Then, a second calibration step is completed in order to determine the output losses. They have a contribution of 1.2 dB, which comes from the output probe, and to the connector. The switch insertion losses are retrieved thanks to *Equation (46)*. The results are plotted in Figure 71.

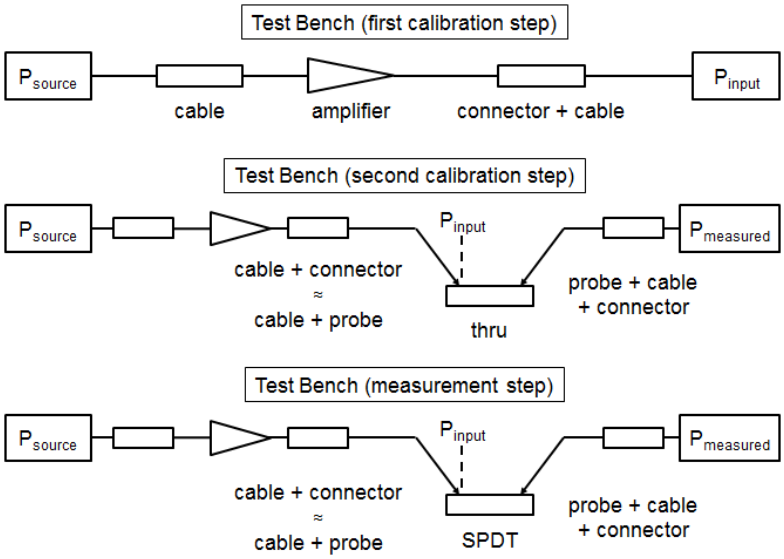

Figure 70: Power calibration and measurement benches

$$IL\ (\text{dB}) = P_{measured} - P_{input} + 1.2 \quad (46)$$

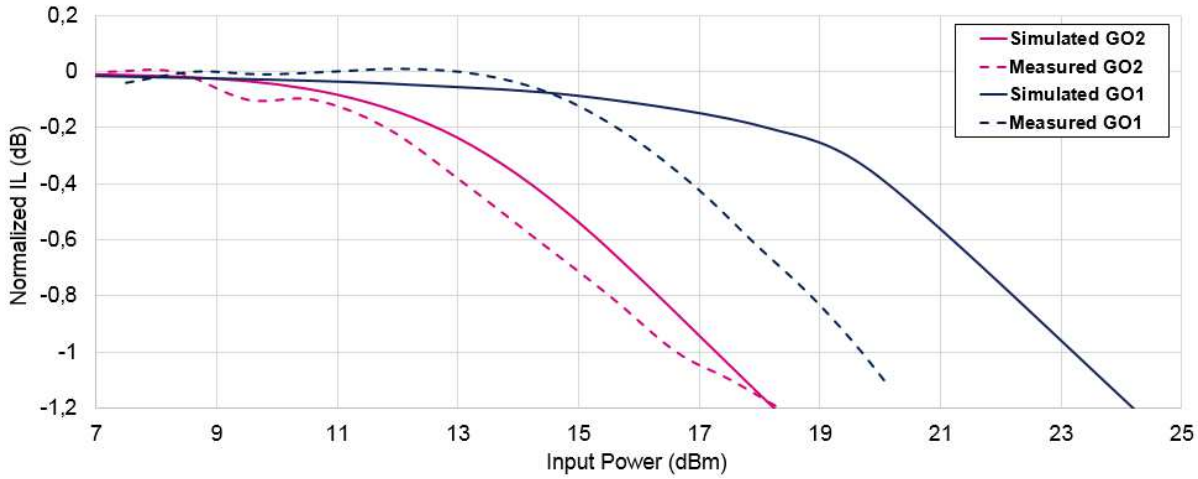

Figure 71: Simulated and measured input compression points, for GO1 and GO2 SPDT switches

The simulated ICP1 are 17.5 dBm for the GO2 switch, and 23.5 dBm for the GO1. The measured results for these circuits are 16.5 dBm and 19.5 dBm, respectively. Based on the results reported in Table 31, both ICP1 should be closer. The difference comes from the impedance of the access line, which is higher for the GO2 switch. This ensures a better impedance matching, but also implies a larger voltage swing. Besides, the large discrepancy between GO1 simulation and measurement probably comes from undesired floating body effects, not accurately taken into account by the simulator. Finally, the power meter accuracy and the limited f_T of GO2 BC may lead to discrepancies.

In this section, the need of SPDT switch for TDD systems has been introduced. Then, several topologies have been detailed. The series-shunt switch seems particularly of interest in the EMBB 5G context, since it is compact and therefore low-cost. Its performances are optimized thanks to the use of SOI technology. The operating principle and the main limitations of the selected topology are presented. Then, a circuit sizing method is introduced, and two SPDT switches are designed, based on GO1 and GO2 transistors. The circuits exhibit good performances for the targeted use case. However, the measurement results show that FB devices are not taken into account accurately by the simulator. Thus, BC switches seems more suitable for higher power handling. However, connecting the bodies to the ground, and a particular care about the access line impedances may improve the circuit performances. In the following section, different LNA circuits are presented. Then, a standalone Ka-band LNA is designed, based on ST C65SOIFEM technology.

III. Standalone Ka-Band Low Noise Amplifier Design

1. Ka-Band Low Noise Amplifier Overview

a. Noise Considerations in RF Microelectronics

The low noise amplifier (LNA) is often one of the first elements of an RF receiver. Thus, as given by *Equation (47)*, i.e. the Friis formula for noise [31], this circuit has a significant impact of the overall noise figure (NF_{RX}). Indeed, the noise contributions of the following components may be reduced thanks to the cascaded gain ($G_1 * \ldots * G_{m-1}$). Usually, there are passive elements before the LNA, such as SPDT

switches and filters. Their NF is equal to their IL. Consequently, most of the RX NF is due to the FEM (NF_{FEM}), as illustrated in *Equation (48)*. In order to minimize the NF, i.e. to maximize the data rate at the receiver, it is necessary to reduce the losses between the antenna and the LNA. This circuit has to exhibit low NF, as its name implies. Moreover, it also has to provide high gain, in order to minimize the following noise contributions.

$$NF_{RX} = NF_1 + {(\mathrm{NF}_2 - 1)}/{G_1} + \cdots + {(\mathrm{NF}_m - 1)}/{(G_1 * \ldots * G_{m-1})} \quad (47)$$

$$NF_{FEM} = (IL_1 * NF_{LNA}) \quad (48)$$

The main noise contributor in RF circuits is the thermal noise. In Figure 72, the resistor and MOSFET noise sources are depicted. Their noise depends on the Boltzmann constant (k), and the circuit temperature (T). The resistor's noise is also related to the resistance (R_1), whereas the transistor's drain noise ($\overline{I_n^2}$) is related to g_m and γ, which is 2/3 for long-channel devices, and may increases to 2 for short-channels. Finally, there is another noise source in MOSFET transistors: the gate resistance. It is often negligible against $\overline{I_n^2}$ [31].

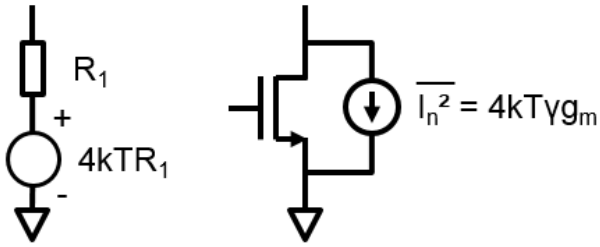

Figure 72: Thermal noise in a resistor and a MOSFET transistor

b. Low Noise Amplifier Topologies

The low noise requirement restricts the choice of the topology. Indeed, only several topologies are able to provide low noise, and sufficient gain. The impedance matching is also a major concern, which is very challenging to achieve with low noise. Usually, the targeted LNA input impedance is 50 Ω, with no imaginary part [31]. However, the source impedance may vary, due to the circuit environment, it is therefore necessary to ensure the LNA stability. For example, the coupling between adjacent array elements may modify the source impedance. Thus, the circuit performances and stability may be degraded [32]. A commonly used stability criteria is $K > 1$ and $\Delta < 1$, which ensures unconditional stability. This means the circuit will be stable, regardless of the source and load impedances. K and Δ depends on the circuit S parameters. They are respectively defined by *Equations (49)*, and *(50)*. Besides, it is suggested in [31] that maximizing the reverse isolation, i.e. $S_{12} = 0$, provides stability. Finally, the LNA power consumption is small compared with PA. Nevertheless, it has to be kept in mind that there will be many circuits in 5G EMBB systems. Thus, power savings remain a top priority.

$$K = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2 * |S_{21}| * |S_{12}|} \quad (49)$$

$$\Delta = (S_{11} * S_{22}) - (S_{12} * S_{21}) \quad (50)$$

The common LNA architectures are presented in Figure 73, from [33]. The first option uses a resistive termination to ensure an input impedance of 50 Ω. This results in a poor noise figure, due to the contribution of real resistors. The common gate (CG) topology provides a stable input impedance of $1/g_m$. Thus, the size of the input transistor is locked. Besides, its noise figure is given by *Equation (51)*. For a short-channel device, the coefficient of channel thermal noise (γ) may be greater than 1. α is the self-gain of the device, considering the zero-bias drain to source conductance ($g_{ds}{}^0$), as defined in *Equation (51)*. It may be lower than 1 for short-channel devices. Consequently, the theoretical minimum noise figure of the CG topology is greater than 3 dB.

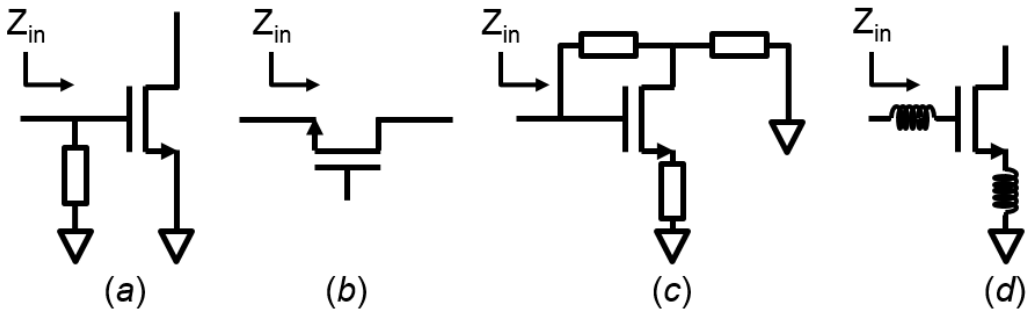

Figure 73: The common LNA architectures: (a) Resistive termination, (b) Common gate termination, (c) Shunt-series feedback, and (d) Inductive degeneration, from [33]

$$NF_{CG} = 10 * \log\left(1 + {}^{\gamma}/_{\alpha}\right) \quad (51)$$

$$\alpha = {}^{g_m}/_{g_{ds}{}^0} \quad (52)$$

The shunt-series feedback topology provides a broadband input and output matching. However, it exhibits a very high power consumption compared with the other usual LNA architectures.

Finally, the fourth topology presented in Figure 73 (*d*) is the inductive source (or emitter) degeneration. It is the most used LNA architecture, since it is able to achieve the best noise figure of the above topologies. However, the use of inductors restricts the circuit bandwidth and results in larger area. For the targeted application, i.e. mm-wave phased array, the main objective is to reduce the power consumption. Since the DMBAA topology may be selected, the NF is also a significant parameter. Indeed, the RX benefits from the array factor only in the digital domain. Thus, the inductively degenerated LNA architecture is chosen. Besides, the single-ended topology is preferred against differential structures, which usually have higher loss before the LNA, due to the input balun.

The input impedance of the selected LNA is given by *Equation (53)*, from [33]. It depends on the gate and source inductors (L_g and L_s), and on the transistor parameters (g_m and C_{gs}). When the inductors resonate with C_{gs}, i.e. at ω_0, the input impedance is given by *Equation (54)*. Since it is a first order circuit analysis, C_{gd} is neglected. Thus, Z_{in} only depends on the cut-off frequency (ω_T) and L_s, up to *Equation (28)*.

$$Z_{in} = j\omega * (L_s + L_g) + {}^{1}/_{(j\omega * C_{gs})} + \left({}^{g_m}/_{C_{gs}}\right) * L_s \quad (53)$$

$$Z_{in} = \left({}^{g_m}/_{C_{gs}}\right) * L_s \approx \omega_T * L_S \quad (54)$$

Besides, the noise factor of this circuit (F) is also studied in [33]. It is computed thanks to *Equation (55)*, from the resistance of the source (R_S), the resistance of the gate (R_G), and the parasitic resistance of L_g, i.e. R_L. It depends on the quality factor of the inductor (Q_L), as defined in *Equation (56)*. Up to the Friis formula, the first stage of the LNA is the main noise contributor. The following equation provides therefore a good approximation of the overall noise of this architecture. Moreover, it gives some directions to improve the circuit performances: reduce ${g_{ds}}^0$, increase ω_T, and improve Q_L.

$$F = 1 + \frac{R_L}{R_S} + \frac{R_G}{R_S} + \left[\gamma * {g_{ds}}^0 * R_S * \left(\frac{\omega_0}{\omega_T}\right)^2\right] \tag{55}$$

$$R_L = {(\omega * L_g)}/{Q_L} \tag{56}$$

The optimal source impedance (${R_S}^{opt}$), i.e. the source impedance offering the minimum NF, is defined as the impedance which satisfies *Equation (57)*. The result is given in *Equation (58)*. Then, the minimum noise figure (F_{min}) may be computed thanks to *Equation (58)*, replacing R_S by ${R_S}^{opt}$ in *Equation (55)*. Finally, the inductively degenerated topology is usually combined with a cascade stage. It provides higher isolation, and therefore more stability. Since C_{gd} is neglected, its noise contribution is negligible.

$$\frac{\partial F}{\partial R_S} = -\frac{(R_L + R_G)}{{R_S}^2} + \left[\gamma * {g_{ds}}^0 * \left(\frac{\omega_0}{\omega_T}\right)^2\right] = 0 \tag{57}$$

$${R_S}^{opt} = \frac{\omega_T}{\omega_0}\sqrt{\frac{(R_G + R_L)}{\gamma * {g_{ds}}^0}} = \frac{\omega_T}{\omega_0}\sqrt{\frac{\alpha * (R_G + R_L)}{\gamma * g_m}} \tag{58}$$

$$F_{min} = 1 + \frac{\omega_0}{\omega_T}\left[R_G + R_L + \sqrt{\frac{\gamma * g_m * (R_G + R_L)}{\alpha}}\right] \tag{59}$$

In the following section, several Ka-band LNAs are presented. Then, some improvement techniques are introduced, in order to cope with the inductively degenerated topology.

c. Ka-Band Low Noise Amplifier State of the Art

In [34], a CMOS SOI single-ended LNA is presented. It is based on a single inductively degenerated cascode stage, which is composed of two FB devices. It is highlighted that they provide lower NF than BC ones. The circuit schematic is depicted in Figure 74. The three inductors benefit from the use of SOI, which enables to reach high quality factors. In particular, it reaches 28 in the operating band (25-31 GHz) for the gate inductor. Thus, up to *Equation (59)*, the circuit NF is reduced. Besides, the degeneration inductor is implemented thanks to the transmission line to the ground. Then, the measured gain and NF are respectively 12.8 dB and 1.6 dB, at 28 GHz, with current density of 0.143 mA /µm. The overall LNA draws 10 mA from a 1.5 V supply. Another inductively degenerated LNA is detailed in [35]. It is based on two SiGe HBT cascode stages, as illustrated in Figure 74. The circuit operates in the 33-34 GHz frequency band. It exhibits 23.5 dB gain and 2.9 dB NF, while consuming 6 mA from a 1.8 V supply. The use of two stages provides higher gain, but degrades the overall NF. Nevertheless, as previously shown in [36], the above results confirm that CMOS SOI is more suitable for the design of Ka-band LNAs. Finally, GaAs and GaN technologies are often used to handle higher input power [37], [38].

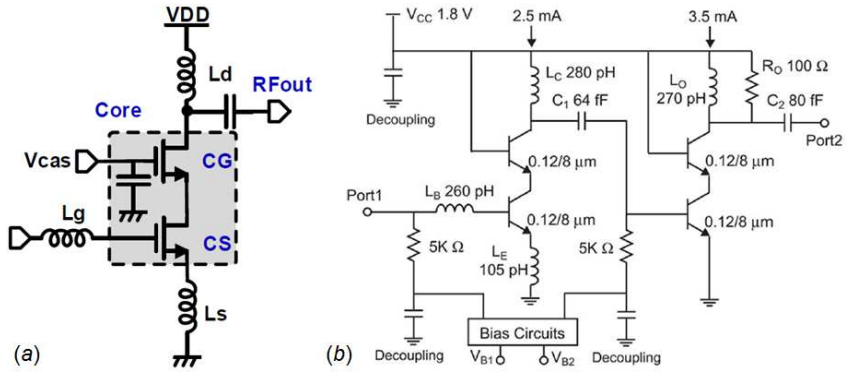

Figure 74: (a) Inductively degenerated CMOS SOI LNA, from [34]; (b) Inductively degenerated SiGe LNA, from [35]

In practice, C_{gd} has a deleterious effect on the LNA performances. It degrades the amplifier gain and reverse-isolation. Its stability is therefore also reduced. This parasitic capacitance is often neutralized thanks to cross-coupled topologies [39]. However, it usually requires the use of a differential structure, which leads to higher NF. In [40], a single-ended neutralization technique is implemented at 60 GHz. It is based on a 180° transformer and a capacitive feedback, as depicted in Figure 75. This exhibits low NF of 4 dB at 60 GHz, and wide fractional bandwidth about 15%. However, the gain is only 7.67 dB per stage due to the many transformers. They may also lead to area issues at lower frequencies.

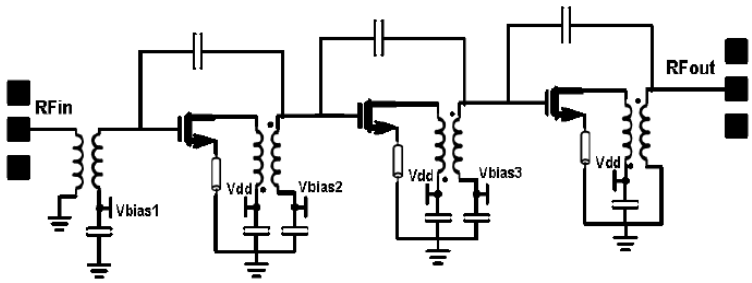

Figure 75: Schematic of the 3 Stage LNA with capacitive feedback and transformers, from [40]

In [41], another way to enhance the reverse isolation is implemented at 900 MHz. It is based on mutual-coupled degeneration, as depicted in Figure 76. The inter stage inductor improves the cascode matching, and provides an additional variable for input matching. Indeed, the coupling between L_S and L_M impacts the input impedance of the LNA, as given by *Equation (60)*. However, such implementation is challenging at mm-wave frequencies, since the degeneration inductor is very small [34].

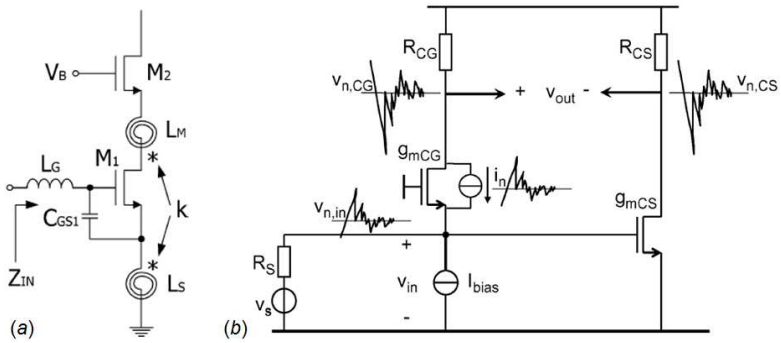

Figure 76: (a) LNA using mutual-coupled degeneration, from [41]; (b) Noise cancelling LNA, from [42]

Finally, an interesting topology is introduced in [42], and illustrated in Figure 76. It is an active balun composed of a CG and a CS stage. Thus, it benefits from both a wideband input matching and a low NF. Indeed, the noise of the CG amplifier is cancelled by the CS stage. The structure is well balanced when *Equation (61)* is verified. The gain (G_{NC}) is therefore given by *Equation (62)*. The drain resistors provide a good output matching over a wide bandwidth. However, they may cause a significant voltage drop, which degrades the overall performances. It is possible to avoid this issue replacing them by inductors. In this case, *Equation (61)* is only verified at a given frequency. The bandwidth is therefore limited.

$$Z_{in} = j\omega * (L_S + L_G) + {}^{1}/_{(j\omega * C_{GS1})} + \left({}^{g_{m1}}/_{C_{GS1}}\right) * (L_s \pm M) \quad (60)$$

$$g_{mCG} * R_{CG} = g_{mCS} * R_{CS} \quad (61)$$

$$G_{NC} = 6 + 20 * \log(g_{mCG} * R_{CG}) \quad (62)$$

In the above section, several techniques have been presented in order to improve the performances of the inductively degenerated CS LNA. However, they are challenging to implement in the Ka-band. The usual single-ended topology is therefore often selected. Moreover, the use of RF SOI technology seems to provide suitable performances for 5G applications. Consequently, the following part details the design and measurements of a conventional inductively degenerated cascode LNA based on ST C65SOIFEM. The main objective is to validate the performances of this technology for the targeted mm-wave 5G use case.

2. Standalone Ka-Band Low Noise Amplifier Design

a. Targeted Features

Depending on the selected use case (range, data rate, number of users…), 5G transceivers are expected to be either hybrid or fully digital. Consequently, the FEM has to be suitable for both architectures. Since the noise requirement is higher for DMBAs, it is expected to reach NF as low as possible. However, there is a significant tradeoff with the power consumption, since it has to be minimized. It is also necessary to provide sufficient gain, in order to neglect the noise contribution of the following stages, and low gain ripples over the 400 MHz sub-channels. Finally, the linearity is not a relevant parameter in TDD systems. Nevertheless, the LNA has to be able to support a high input power in TX state, equal to the PA output power minus the SPDT switch isolation. The targeted features for the standalone Ka-band LNA are reported in Table 34. Besides, to be competitive with recently reported circuits, the two-stage LNA should provide a gain of 20 dB at 28 GHz, while drawing a current lower than 20 mA.

Bandwidth (GHz)	Gain (dB)	NF (dB)	ICP1 (dBm)	I_{DC} (mA)
24 - 31	> 20	2.5	≈ 0	< 20

Table 34: Targeted features for the two-stage standalone Ka-band LNA

b. Design of a Standalone LNA with Inductive Degeneration

Previously, it has been highlighted that the circuit NF highly depends on the first stage. Thus, the first design steps are the choice and sizing of the input transistor. Usually, the parasitic resistance of the gate inductor is neglected against R_G. This is especially the case with the use of SOI technology, which

improves the quality factor of the inductors. Then, *Equation (59)* is simplified into *Equation (63)*. Then, ω_T is replaced by its expression from *Equation (28)*. This results in *Equation (64)*, when neglecting C_{gd}. It is therefore possible to optimize the NF thanks to the transistor size (R_G, C_{gs}), and biasing point (g_m). The other parameters are given by the application (ω_0), and by the choice of the technology (γ, α).

$$F_{min} = 1 + \frac{\omega_0}{\omega_T}\left[R_G + \sqrt{\frac{\gamma * g_m * R_G}{\alpha}}\right] \quad (63)$$

$$F_{min} = 1 + \frac{\omega_0 * C_{gs}}{g_m}\left[R_G + \sqrt{\frac{\gamma * g_m * R_G}{\alpha}}\right] \quad (64)$$

The sizing parameters are depicted in Figure 77. Firstly, it is chosen to use double gate access transistors in order to minimize R_G. Then, the device performances (I_{DC}, gain, NF) are plotted depending on the sizing parameters, for both BC and FB transistors. The results are presented in Figure 77 and Figure 78 for a BC NMOS. Besides, the performances of FB and BC devices are reported and compared in Table 35.

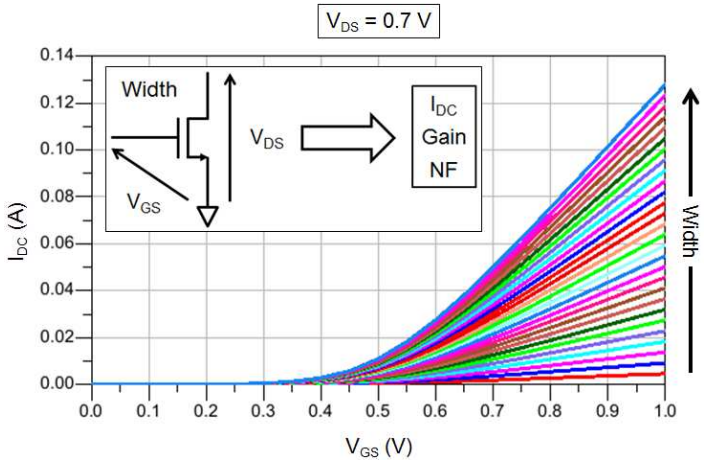

Figure 77: I_{DC} of the input transistor depending on its sizing parameters

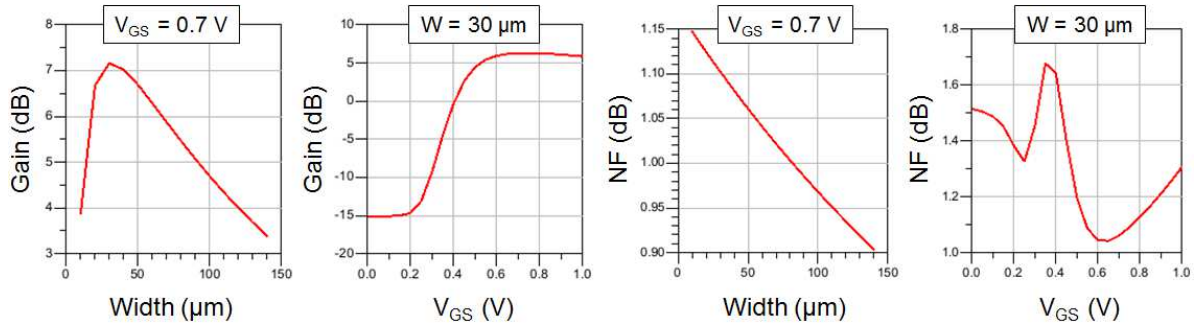

Figure 78: Device performances depending on its sizing parameters

Based on the results reported in Table 35, the BC device is preferred. Indeed, it exhibits higher gain and lower current consumption, compared with FB. Both devices have close noise performances. Secondly, the cascode device is sized in order to ensure the proper V_{DS} for the input transistor. Like previously, the performances of FB and BC cascodes are compared. Results are reported in Table 36. Here again, the BC

device is chosen, since it provides good gain, with low current consumption. Its noise performances are a bit degraded compared with FB, but the use of BC ensures a better linearity.

Device	Size (µm)	V_{GS} (V)	NF (dB)	Gain (dB)	I_{DC} (mA)
BC	120	0.5	1.5	6.6	6.7
FB	140	0.4	1.4	5.6	8.0

Table 35: Performances of the FB and BC CS devices, with V_{DS} = 1 V

Device	Size (µm)	V_{CASC} (V)	NF (dB)	Gain (dB)	I_{DC} (mA)
BC	150	1.6	1.8	9.5	7.2
FB	200	1.6	1.6	9.6	8.3

Table 36: Performances of the FB and BC CS cascodes, with a BC input stage and V_{DD} = 2 V

Since the transistor sizes are now settled, the input matching is ensured by setting L_G and L_S. As given by *Equation (60)*, L_S is used to obtain a 50 Ω input impedance, while L_G enables to cancel the imaginary part. Moreover, the input capacitor (C_{IN}) and biasing resistor respectively acts as RF and DC blockers. The shunt capacitor of the cascode transistor is used to ensure the circuit stability. Finally, L_D is both a choke inductor, and part of the output matching. Likewise, C_{OUT} is both a DC blocker, and part of the output matching. The whole schematic of the first stage is depicted in Figure 79. After several iterations, a 50 Ω input impedance is reached, with L_S = 120 pH and L_G = 210 pH. Then, the cascode cell is laid out, as presented in Figure 79. Both input and output transistors are split into two devices. The drain and source accesses are put in the middle of each device in order to reduce the parasitic capacitances, including C_{gd}. The shunt capacitor is also distributed over both sides of the cascode. The RC parasitic elements, i.e. resistors and capacitors, are extracted, and it is shown that they do not have a significant impact on the circuit matching and performances. However, this topology results in high output impedance, about 300 Ω. For measurement purpose, it is desired to have a 50 Ω output impedance. Consequently, the output stage will have to provide additional gain, while ensuring a proper output matching.

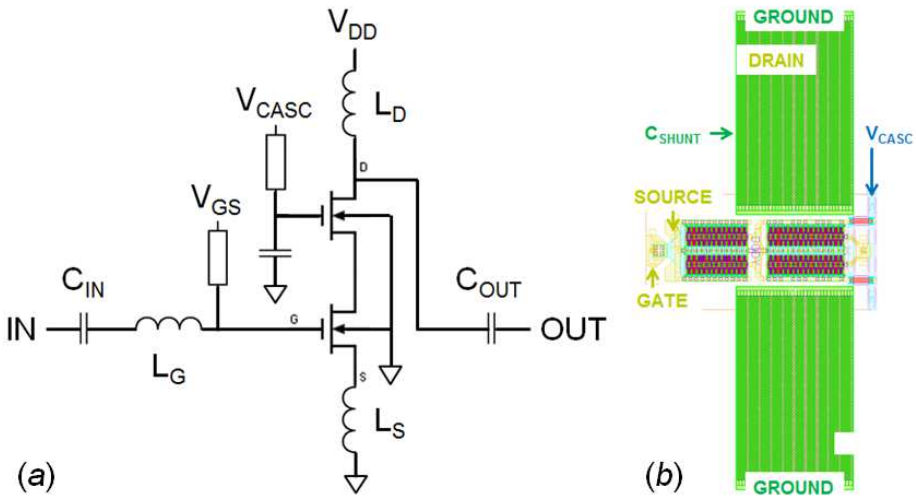

Figure 79: (a) Schematic of the inductively degenerated cascode input stage; (b) Layout of the cascode cell

Up to *Equation (47)*, the noise contribution of the second stage is reduced thanks to the gain of the first one. Thus, it is chosen to use a cascode structure without inductive degeneration. In this case, both devices are composed of two 150 µm BC transistors. The layout of this cascode cell is similar to the previously presented one. As illustrated in Figure 80, the input capacitance of the second stage is mutualized with the output capacitance of the first stage. Moreover, a gate inductor is not necessary here. Finally, the output matching is realized thanks to the output capacitor and the second drain inductor. The layout of the complete two-stage LNA is presented in Figure 80. Firstly, the two cascode cells and the DC block capacitors are placed. Then, the inductors are designed according to the available space. The major issue is to maximize their quality factor, in order to reduce their losses. The simulation of the complete circuit combines the RC extraction of the cascode cells and the 3D EM simulation of the passive elements, such as inductors and pads. The first simulation results are reported in Table 37.

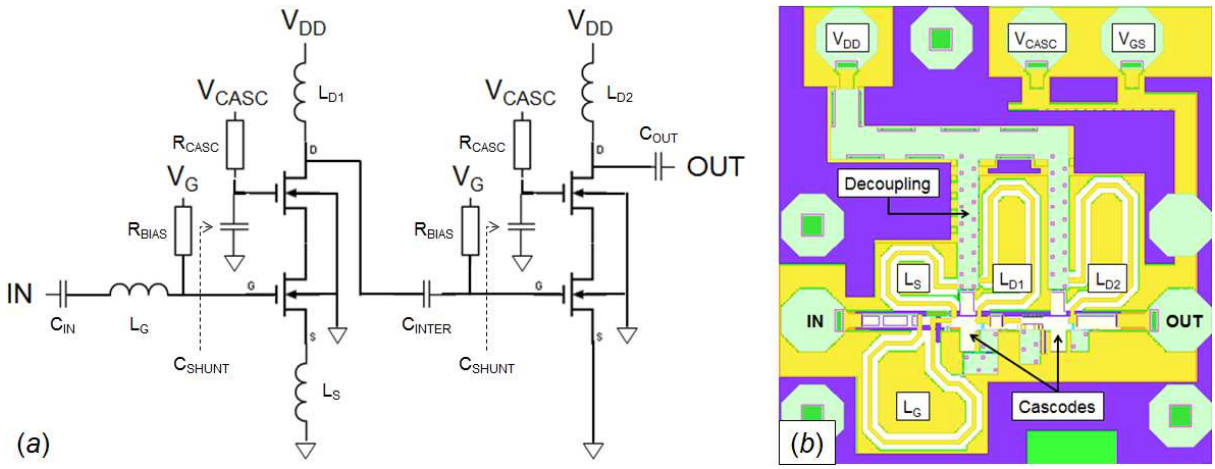

Figure 80: The complete two-stage LNA (a) Schematic and (b) Layout

L_S (pH)	L_G (pH)	L_{D1} (pH)	L_{D2} (pH)	C_{IN}, C_{SHUNT} (fF)	C_{INTER} (fF)
120	210	190	200	1000	190
C_{OUT} (fF)	**R_{BIAS}, R_{CASC} (Ω)**	**V_{DD} (V)**	**V_{CASC} (V)**	**V_{GS} (V)**	**I_{DC} (mA)**
700	1000	2.0	1.6	0.5	24
S_{11} (dB)	**Z_{IN} (Ω)**	**S_{22} (dB)**	**Z_{OUT} (Ω)**	**Gain (dB)**	**NF (dB)**
-11.3	33.1 – 15.5j	-12.9	68.6 – 19.7j	20.6	2.6

Table 37: First post-layout simulation parameters and results of the complete two-stage LNA at 28 GHz

Even if the real part of Z_{IN} and Z_{OUT} are shifted from the expected 50 Ω impedance (about ±35%), S_{11} and S_{22} are below -10 dB. Thus, the circuit matching is correct. Moreover, the overall performances are a bit lower than the previously defined targets. The circuit has been included in the second run, called *Vega*. It has been measured in order to validate the design method and the *ST C65SOIFEM* technology models and features. Measurement results are presented and discussed in the following section.

c. Circuit Measurements

The early small-signal measurements have shown poor agreement with the above post layout simulation results. In order to recover a decent agreement, a refined simulation model is implemented. Actually, the first one was only taking into account the inductors separately. Thus, the effect of their environment was neglected. The refined simulation model for retro simulations is composed of a global 3D EM view. It takes into account all the inductors simultaneously. The new results are reported in Table 38, and presented in

Figure 81. The main result is that the inductor values have been modified due to the impact of their environment. Moreover, their quality factors have also been degraded from 20 to 16, i.e. their parasitic resistances are 20% higher. This explains the degradation of the circuit matching and noise performances. However, there is still a significant difference between the simulated and measured gain. The first hypothesis is that the inter stage matching network did not operate properly. Indeed, S_{22} and gain exhibit the same frequency shift, while the measured S_{11} is in good agreement with the simulated one. Another reading may be that the second stage is not working properly, degrading the overall gain.

L_S (pH)	L_G (pH)	L_{D1} (pH)	L_{D2} (pH)	V_{GS} (V)	I_{DC} (mA)
108	231	165	176	0.4	24
S_{11} (dB)	**Z_{IN} (Ω)**	**S_{22} (dB)**	**Z_{OUT} (Ω)**	**Gain (dB)**	**NF (dB)**
≈ -9	26.0 – 13.6j	≈ -8	41.7 – 45.4j	19	3.7

Table 38: Second Post-Layout Simulation Parameters and Results of the Complete Two-Stage LNA at 28 GHz

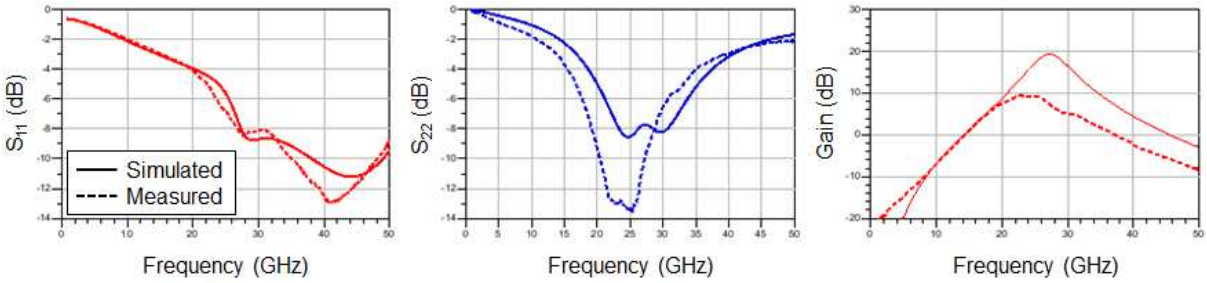

Figure 81: Comparison between the LNA post layout simulations and the measurement results (1)

It is therefore necessary to refine the model for more accurate retro-simulations. This will enable to retrieve a good agreement with measurements, and to identify the source of the gain degradation. There are still several neglected parameters in the above refined model. They are depicted in Figure 82: the ground path between the source inductor and the pad, the ground path between the second stage and the pad, and the metal dummies (tilling) in the inductors. They respectively impact the input matching, the gain of the second stage, and the inductor values. The most significant point is the second parasitic ground path, which may degenerate the second stage. Conversely, the tilling has little effect on the inductors, since a safe distance has been planned.

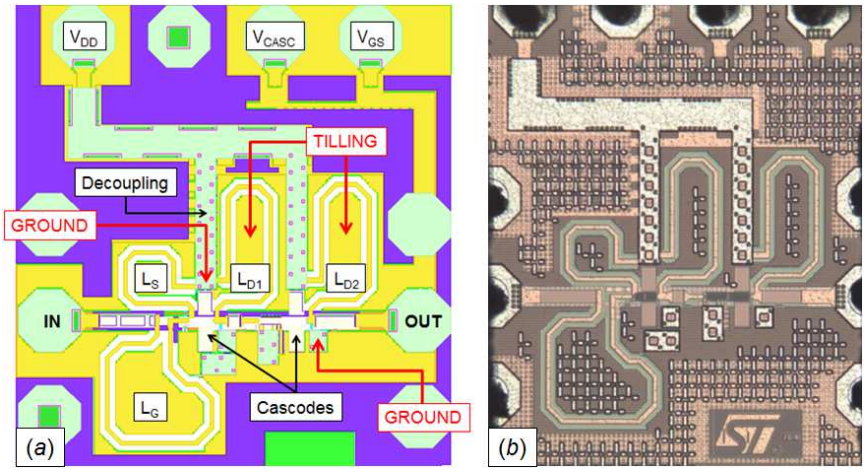

Figure 82: Complete LNA (a) Layout top view and (b) Die micrograph

Besides, a second measurement campaign has been carried out. The biasing voltages have been accurately set, ensuring the desired drain current. The results are plotted in Figure 83. They are in good agreement with the new retro-simulations, particularly for the gain and the input matching. There is still a frequency shift (10-15%) for the output matching. There are also some random measurement errors in the 5-10 GHz band. These final small signal results are reported in Table 39. They confirm that the parasitic degeneration of the second stage degrades the LNA gain. The NF is also increased, above 4 dB.

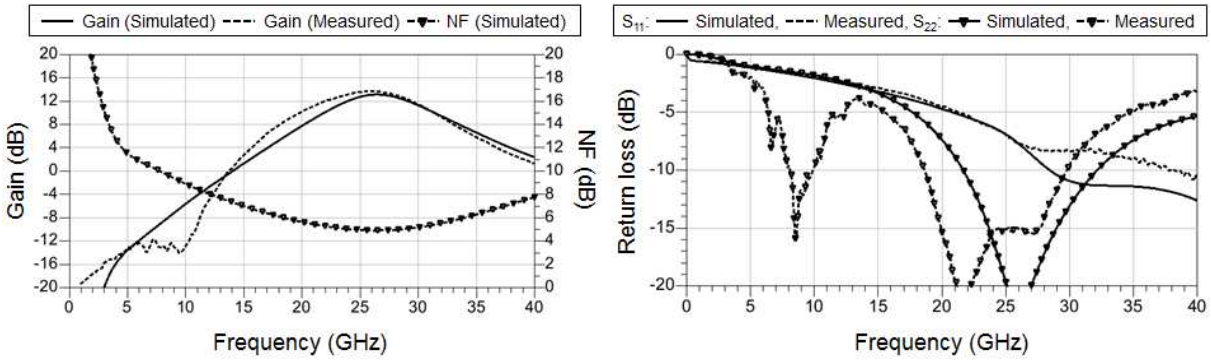

Figure 83: Comparison between the LNA post layout simulations and the measurement results (2)

	Gain (dB)			Z_{IN} (Ω)	Z_{OUT} (Ω)
Freq. (GHz)	22.0	26.4	31.3	28	
Simulated	10.1	13.1	10.1	26.7 – 10.3j	55.4 – 13.6j
Measured	12.0	13.7	10.0	25.1 – 15.8j	53.8 – 20.5j

Table 39: Comparison between the third post-layout simulation results and the second measurement campaign results

Then, the NF is measured. The bench previously presented for the SPDT switch ICP1 measurement is used. It has been detailed in Figure 70, including the calibration steps. Actually, the power source is replaced by a noise source. After calibration, the noise is measured with an accuracy of ±0.2 dB. The results are plotted in Figure 84. Indeed, the NF measurement is achieved over two different frequency

bands: 18-26.5 GHz and 25.5-30.5 GHz. They exhibit a good continuity. They are also in good agreement with the post-layout simulation results, even if there is the same frequency shift as previously.

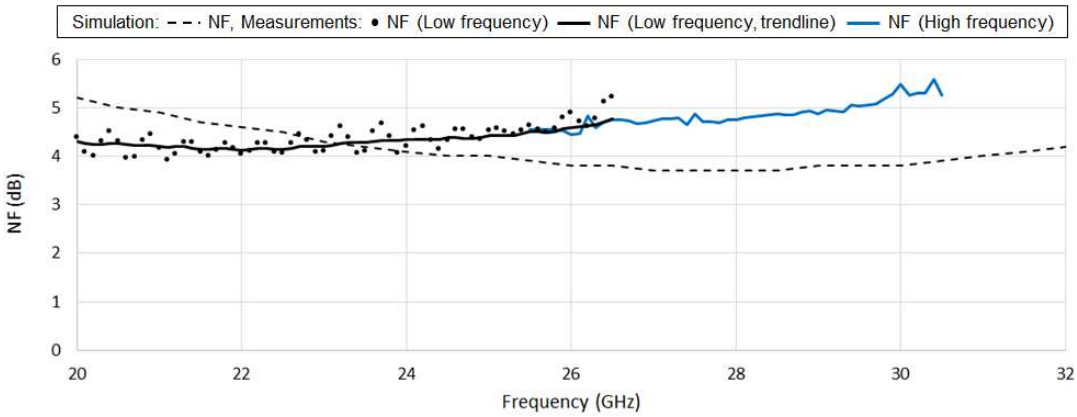

Figure 84: Post Layout Simulated and Measured NF, from 20 GHz to 32 GHz

Finally, the LNA ICP1 is measured at 28 GHz, as plotted in Figure 85. The circuit exhibits a measured ICP1 of -4 dBm, which is consistent with the post-layout simulation expectations but lower than the defined target. Overall, after several retro-simulation iterations, a good correlation is obtained with the circuit measurements. The global performances are degraded due to neglected parasitic elements: inductor environments, paths to the ground pads. The retro-simulations have also highlighted that the input stage has a small current density. Consequently, its f_T is only 60 GHz, i.e. twice the operating frequency. This limits the circuit performances. However, the technology seems to be suitable for the design of a Ka-band LNA targeting 5G applications. Moreover, the use of BC devices offers a good linearity.

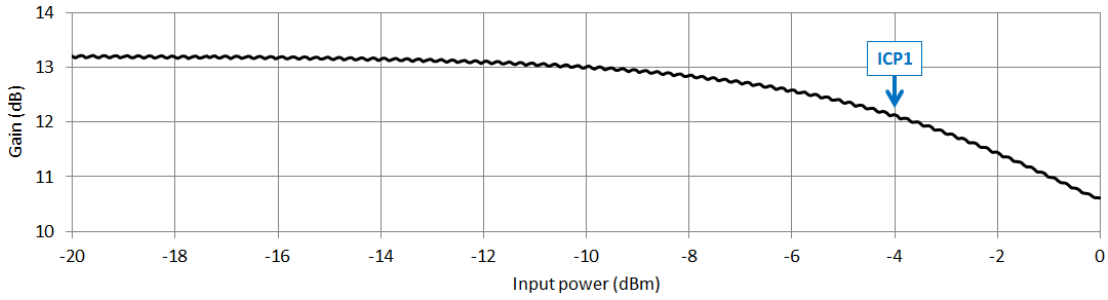

Figure 85: ICP1 measurement of the LNA, at 28 GHz

IV. Conclusion

In this chapter, the advantages of SOI technologies over traditional bulk CMOS have firstly been presented. Indeed, the use of SOI enables to cope with the drawbacks of the technology scaling. Among the different implementations, RF SOI (i.e. SOI on a TR substrate) seems to be particularly of interest for the design of 5G FEM at mm-wave frequencies. It offers sufficient power handling, good RF performances, and passives with high quality factors. The difference between PD and FD SOI has also been detailed. Then, an overview of the available SOI technologies within *STMicroelectronics* has been presented. Thus, the ST C65SOIFEM technology is selected for the targeted application. Its BEOL is detailed. Moreover, the available devices and their models are discussed. They are compared thanks to simulations, in order to determine the proper transistor for each circuit. The technology passives have also been simulated. Then, different devices are designed and measured. The open-short de-embedding method is detailed and used to retrieve the circuit performances. As a conclusion, the correlation between the device models and the simulation results is confirmed. It is also highlighted that the selected de-embedding method provides accurate results in this case. However, the floating body effects seem to be not taken into account.

Secondly, the need of SPDT switches for TDD systems is discussed. Several implementations based on different technologies have been presented. The series-shunt topology is therefore chosen, since it may offer good performances using an RF SOI technology like *ST C65SOIFEM*. The operating principle of this structure is detailed. Then, a series-shunt SPDT switch has been designed. It is targeting 5G applications in the Ka-band. The circuit measurements are presented; they exhibit a good correlation with post layout simulation results. The two designed SPDT switches, based on GO1 FB and GO2 BC transistors, offer low insertion losses and sufficient isolation. However, the large signal measurements have shown that the expected ICP1 is not reached using GO1 FB. It seems to be due to the undesired floating body effects. Besides, the GO2 BC switch has lower power handling due to lower transmission line impedances and low f_T. In this case, it is also shown that the open-short de-embedding method has limited accuracy above 40 GHz.

Finally, the basic structures enabling low noise amplification are discussed. An overview of recently reported circuits has shown that the inductively degenerated structure based on RF SOI technology is of particular interest. Several improvement techniques have been presented for this kind of circuits. Their drawbacks often make them challenging to implement at mm-wave frequencies. Then, a LNA is designed. It is composed of a first inductively degenerated cascode stage to ensure low noise, and a second one which provides additional gain and a proper output matching. After several retro-simulation iterations, the circuit measurement exhibits good correlation with simulation results. This enables to validate the technology models for SPDT switch and LNA designs. However, due to neglected parasitics, the performances are lower than targeted. Some perspectives for future designs are to take into account the parasitics more accurately and to improve performances thanks to better inter stage matching.

V. References

[1] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.," *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33-35, 2006.

[2] Wikipedia. Moore's law. [Online]. https://en.wikipedia.org/wiki/Moore%27s_law

[3] B. Martineau, Potentialités de la technologie CMOS 65nm SOI pour des applications sans fils en bande millimétrique, 2008, PhD Thesis, Université des sciences et technologies de Lille.

[4] JNM2019. (2019, May) Journée thématique : les communications de 5e génération et l'internet des objets. [Online]. https://jnm2019.sciencesconf.org/data/pages/Programme_journee_thematique_5G.pdf

[5] L. Zhu et al., "RF performance of passive components on state-of-art trap rich silicon-on-insulator substrates," in *2016 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA),* Hsinchu, 2016.

[6] A. Marshall and S. Natarajan, *SOI design: Analog, memory and digital techniques*. Dordrecht: Kluwer Academic Publishers, 2003.

[7] S. Clerc, T. Di Gilio, and A. Cathelin, *The Fourth Terminal - Benefits of Body-Biasing Techniques for FDSOI Circuits and Systems*.: Springer International Publishing, 2020.

[8] K. S. Yeo, M. Do, and C. Boon, *Design of CMOS RF integrated circuits and systems*.: World Scientific, 2010.

[9] BSIM GROUP. (2010, December) BSIMSOIv4.4 MOSFET MODEL. [Online]. http://ngspice.sourceforge.net/external-documents/models/BSIMSOIv4.4_UsersManual.pdf

[10] W. Wu et al., "PSP-SOI: A Surface Potential Based Compact Model of Partially Depleted SOI MOSFETs," in *2007 IEEE Custom Integrated Circuits Conference*, San Jose, CA, 2007.

[11] J. L. Hilbert, *Tunable RF components and circuits*.: CRC Press, 2018.

[12] O. Inac, M. Uzunkol, and G. M. Rebeiz, "45-nm CMOS SOI Technology Characterization for Millimeter-Wave Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 6, pp. 1301-1311, 2014.

[13] Gabrielle Guitton, Design methodologies for multi-mode and multi-standard low-noise amplifiers, 2017, PhD Thesis, Université de Bordeaux.

[14] Global Foundries. (2019) 45RFSOI - Advanced 45 nm RFSOI technology. [Online]. https://www.globalfoundries.com/sites/default/files/pb-45rfsoi-en.pdf

[15] T. Hirano, K. Okada, J. Hirokawa, and M. Ando, "Accuracy investigation of de-embedding techniques based on electromagnetic simulation for on-wafer RF measurements," in *Numerical Simulations:*

From theory to industry.: INTECH, 2012, pp. 233-258.

[16] X. S. Loo, K. S. Yeo, and K. W. J. Chew, "On-wafer microwave de-embedding techniques," in *Microwave systems and applications*.: INTECH, 2017, pp. 101-120.

[17] S. Chen, S. Sun, Y. Wang, G. Xiao, and R. Tamrakar, "A comprehensive survey of TDD-based mobile communication systems from TD-SCDMA 3G to TD-LTE(A) 4G and 5G directions," *China Communications*, vol. 12, no. 2, pp. 40-60, 2015.

[18] D. M. Pozar, *Microwave Engineering, fourth Edition*.: Wiley, 2012.

[19] Qorvo. (2008, February) TGS4302 27 - 46 GHz Ka band high power SPDT switch. [Online]. https://www.qorvo.com/products/p/TGS4302

[20] B-W Min and G. M. Rebeiz, "Ka-band low-loss and high isolation switch design in 0.13-um CMOS," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 6, pp. 1364-1371, 2008.

[21] W. Lee and S. Hong, "Low-loss and small-size 28 GHz CMOS SPDT switches using switched inductor," in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Philadelphia, PA, 2018.

[22] C. Li et al., "< 0.8dB IL 46dBm OIP3 Ka band SPDT for 5G communication," in *IEEE 18th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Anaheim, CA, 2018.

[23] H. Dong, J. Chen, D. Hou, Y. Xiang, and W. Hong, "A Compact Bi-directional K and Ka Band SPDT in 0.13um SiGe BiCMOS Process," in *IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)*, Melbourne, 2018.

[24] R. L. Schmid, A. C. Ulusoy, P. Song, and J. D. Cressler, "A 94 GHz, 1.4 dB insertion loss single-pole double-throw switch using reverse-saturated SiGe HBTs," *IEEE Microwave and Wireless Components Letters*, vol. 24, no. 1, pp. 56-58, 2014.

[25] Y. Liu et al., "A 28-GHz transceiver front-end with T/R switching achieving 11.2-dBm OP1dB, 33.8% PAEmax and 4-dB NF in 22-nm FD-SOI for 5G communication," in *IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, Burlingame, CA, 2018.

[26] B. Sadhu et al., "A 28GHz 32-Element Phased-Array Transceiver IC with Concurrent Dual Polarized Beams and 1.4 Degree Beam-Steering Resolutionfor 5G Communication," in *2017 IEEE International Solid-State Circuit Conference (ISSCC)*, San Francisco, CA, 2017.

[27] B. Ustundag, K. Kibaroglu, M. Sayginer, and G. M. Rebeiz, "A Wideband High-Power Multi-Standard 23–31 GHz 2×2 Quad Beamformer Chip in SiGe with>15 dBm OP1dB Per Channel," in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Philadelphia, PA, 2018.

[28] C. Nguyen, *Radio-frequency integrated-circuit engineering*.: Wiley, 2015.

[29] M-C. Yeh et al., "Design and analysis for a miniature CMOS SPDT switch using body-floating technique to improve power performance," *IEEE Transactions on Microwave Theory and*

Techniques, vol. 54, no. 1, pp. 31-39, 2006.

[30] Q. Li and Y. P. Zhang, "CMOS T/R switch design: towards ultra-wideband and higher frequency," *IEEE Journal Of Solid-State Circuits (JSSC)*, vol. 42, no. 3, pp. 563-570, 2007.

[31] B. Razavi, *RF Microelectronics, second edition*.: Pearson Education, Inc., 2012.

[32] L. Belostotski, B. Veidt, K. F. Warnick, and A. Madanayake, "Low-noise amplifier design considerations for use in antenna arrays," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 6, pp. 2508-2520, 2015.

[33] D. K. Shaeffer and T. H. Lee, "A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier," *IEEE Journal of Solid-State Circuits (JSSC)*, vol. 32, no. 5, pp. 745-759, 1997.

[34] C. Li, O. El-Assar, A. Kumar, M. Boenke, and G. M. Rebeiz, "LNA design with CMOS SOI process-1.4dB NF K/Ka band LNA," in *IEEE/MTT-S International Microwave Symposium*, Philadelphia, PA, 2018.

[35] B-W. Min and G. M. Rebeiz, "Ka-band SiGe HBT low noise amplifier design for simultaneous noise and input power matching," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 12, pp. 891-893, December 2007.

[36] C. Li et al., "Ka Band FEM Design Comparison with 45nm RFSOI CMOS and High Performance SiGe BiCMOS," in *14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, Qingdao, 2018.

[37] Qorvo. (2018, June) QPF4001 - 26 - 30 GHz 1 Watt GaN Front End Module. [Online]. https://www.qorvo.com/products/p/QPF4001

[38] Qorvo. (2019, Spetember) QPF4010 - 24.25 - 27.5 GHz GaAs Front End Module. [Online]. https://www.qorvo.com/products/p/QPF4010

[39] J. Zhou et al., "A 24-30 GHz CMOS LNA with 2.05dB NF and 0.6dB in-band gain ripple for 5G-applications," 2018 IEEE MTT-S International Wireless Symposium (IWS), Chengdu, 2018, pp. 1-3.

[40] E. Cohen, O. degani, and D. Ritter, "a wideband gain-boosting 8mW LNA with 23dB gain and 4dB NF in 65nm CMOS process for 60 GHz applications," in *IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2012, pp. 207-210.

[41] C. Xin and E. Sanchez-Sinencio, "A GSM LNA using mutual-coupled degeneration," *IEEE Microwave and Wireless Components Letters*, vol. 15, no. 2, pp. 68-70, February 2005.

[42] S. C. Blaakmeer, E. A. M. Klumperink, D. M. W. Leenaerts, and B. Nauta, "Wideband balun-LNA with simultaneous output balancing, noise-cancelling and distortion-cancelling," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 6, pp. 1341-1350, June 2008.

Chapter 3: Towards an Integrated Ka-Band Front-End Module

In the previous chapter, the *ST C65SOIFEM* technology has been presented. Its main features have been extracted from post-layout simulations and compared with measurements. These results have enabled to validate the technology models. Then, several Ka-band circuits have been designed and measured: SPDT switches which are compliant with the requirements defined in the first chapter, and a standalone LNA. The switches are compact and exhibit low insertion loss. Moreover, they are suitable for the targeted use cases. The designed LNA has also confirmed the technology performances. However, these circuits have highlighted some failures. Among other things, the EM simulation model used to design the LNA was not accurate enough. This has led to degraded performances.

Thus, the limitations of the previous circuits are investigated in this chapter. Then, new versions are implemented and measured. The design methodology has been improved, which results in better MHC. In the first section, an improved SPDT switch architecture based on hybrid couplers is introduced. Secondly, a new LNA is designed to cope with the limitations that have been pointed out in the first version. This includes more accurate simulation models. Finally, perspectives for future works are discussed, including directions for a fully integrated Ka-band FEM.

I. Improved Ka-Band SPDT Switch

1. Introduction

a. Limitations of Usual SPDT Switches

In the previously designed SPDT switches, the combination of the series-shunt topology and an RF SOI technology resulted in compact circuits offering good performances over a large bandwidth. It is the case for the circuit presented in [1] as well. Indeed, the choice of an inductorless topology reduces the overall size and improves the available bandwidth. However, it has been shown that there is a major tradeoff between the insertion loss and the switch isolation. Thus, the isolation is limited. The use of SOI CMOS also limits the power handling, compared with III-V technologies. Even if it is possible to stack several devices and increase their width to maintain the FOM, the transistor's area is therefore enlarged which leads to more parasitic elements. Another drawback of such topology is its poor robustness to impedance variations. Indeed, as illustrated by *Equation (65)* deriving from Figure 86, the impedance seen at port 1 (i.e. the PA or the LNA), Z_{11}, depends on the impedance of the antenna. Thus, any impedance variation results in performance degradations. For example, impedance mismatch at the LNA input degrades the circuit NF. In this first order estimation, it is considered that C_{OFF} is small enough to neglect the path towards port 3.

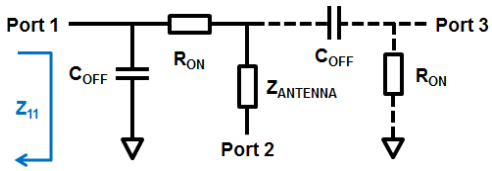

Figure 86: Series-shunt SPDT switch equivalent schematic

$$Z_{11} = (R_{ON} + Z_{ANTENNA}) \;//\; C_{OFF} \approx Z_{ANTENNA} \tag{65}$$

b. Prospects of Improvements

The main drawbacks of the selected series-shunt topology have been detailed above. In order to determine prospects of improvements, it is necessary to remind the needs of the targeted use case. Firstly, it is necessary to have low insertion loss, especially in the RX state since the SPDT switch is the first element of the receiver. Besides, high isolation and power handling are required in the TX state to cope with the high transmitted power of the PA. Thus, there are asymmetric requirements between both SPDT switch states. Usually, the TX path is favored over the RX one. It enables to reduce the PA power consumption but the overall NF is degraded [2]. Since digital beamforming is likely to be chosen, it seems necessary to keep a close eye on the NF. Indeed, the array factor is only applied in the digital domain so the whole receiving chain is dealing with the noise. The aim of this chapter is therefore to design an asymmetrical SPDT switch to reduce the RX state insertion losses while providing high isolation and power handling for TX.

Finally, an interesting perspective to improve the duplexing function is to co-design it with other parts of the FEM (i.e. with the PA, the LNA, and with potential phase or gain controls). For example, the SPDT switch presented in [3] is implemented without physical switch but integrated in the LNA input balun. Indeed, the DC supply is switched on or off instead of switching the RF signal through a path or another. Broadly speaking, baluns and couplers are often implemented within FEM's. They are useful components to realized differential or balanced circuits. Besides, they benefit from the high quality factor of the RF SOI passives. Thus, the asymmetrical duplexer will be a combination of CMOS switches and couplers in order to easily integrate it into diverse FEM topologies.

In [4], such duplexer has been implemented at 2.3 GHz using branch-line couplers at PIN diodes as shown in Figure 87. Based on this topology, a Ka-band integrated duplexer is designed. The first step is to present the operating principle of RF couplers and their main applications within 5G FEM's. Then, such coupler is implemented and measured. The complete duplexer is also introduced and designed. Measurement results are given and compared with post-layout simulation expectations. The advantages and drawbacks of this topology are also discussed, compared with the previous series-shunt SPDT switches.

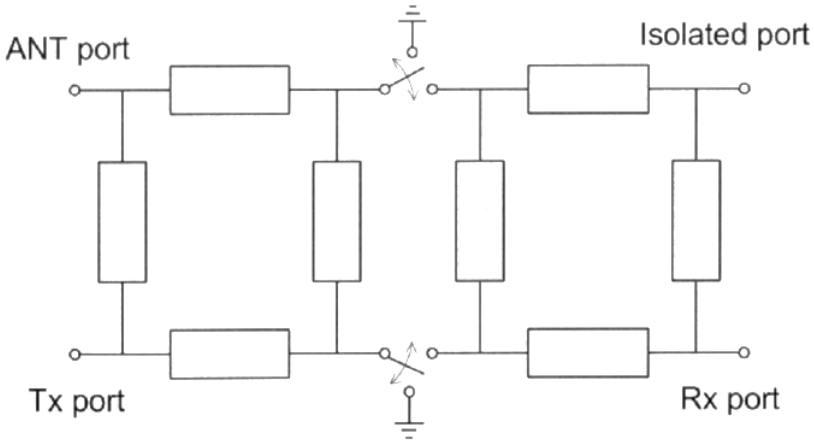

Figure 87: SPDT switch based on couplers and shunt switches, from [4]

2. Integrated Hybrid Coupler

a. Overview of RF Couplers

This section details the operating principle and the different kinds of RF couplers. Then, their use in 5G FEM's is discussed. As illustrated in Figure 88, these circuits are used for power division and combining. They may split the input signal towards two or more outputs. Likewise, two or more inputs may be combined. In Figure 88, the couplers are lossless, and α is lower than 1 since these are passive circuits.

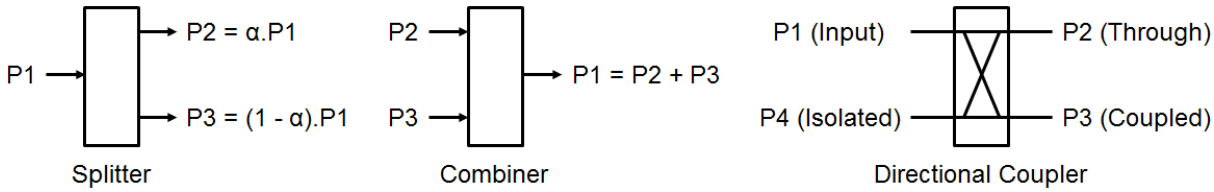

Figure 88: Overview of RF couplers (splitter, combiner, and directional coupler)

The output signals of a 3-port splitter are usually in phase [5]. Besides, 4-port couplers are known as directional couplers. The signal from the input port (P1) is transmitted to P2 and P3. The attenuations through these paths are respectively named insertion loss and coupling. In the particular case of the input signal equally split towards P2 and P3, the circuit is called hybrid coupler. Thus, there is often a 90° or 180° phase shift between the two outputs. Finally, the signal leakage from P1 to P4 defines the coupler isolation. The main directional coupler features are defined in *Equations (66)* to *(69)*, from [5].

$$Coupling = -20 * \log|S_{31}| \tag{66}$$

$$Directivity = -20 * \log\left|{}^{S_{31}}/_{S_{41}}\right| \tag{67}$$

$$Isolation = -20 * \log|S_{41}| \tag{68}$$

$$Insertion\ Loss = -20 * \log|S_{21}| \tag{69}$$

Hybrid couplers are useful components for the design of FEM's. For example, 180° hybrid couplers may be used as baluns in order to implement a differential amplifier [6]. Likewise, 90° hybrid couplers (also known as quadrature couplers) are one of the building blocks of balanced and Doherty amplifiers [7]. Both topologies provide interesting properties for FEM amplifiers.

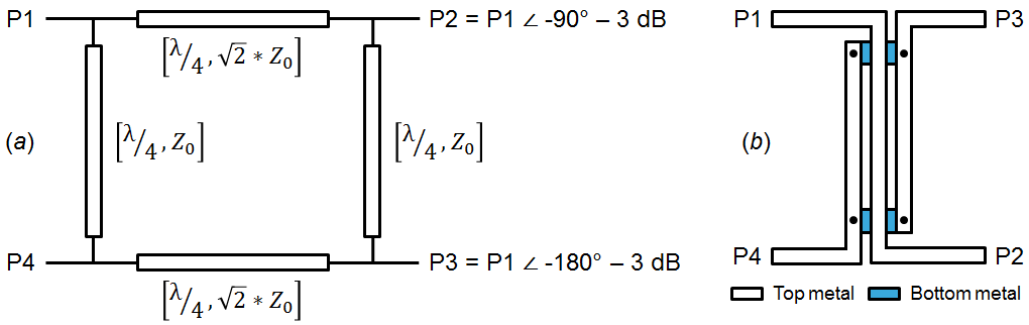

Figure 89: (a) Ideal branch-line coupler used as a divider; (b) Lange coupler top view

The simplest way to implement a quadrature coupler is the planar branchline coupler, presented in Figure 89. This topology only requires quarter wavelength lines with specific impedances. However, this kind of implementation results in a large area. In order to improve the circuit compactness, a Lange coupler may be implemented. As illustrated in Figure 89, it is based on two capacitively coupled lines and it requires the use of two metal levels. Finally, a quadrature hybrid coupler may be described thanks to lumped elements [8]. This model, depicted in Figure 90, derives from a λ/4 elementary description.

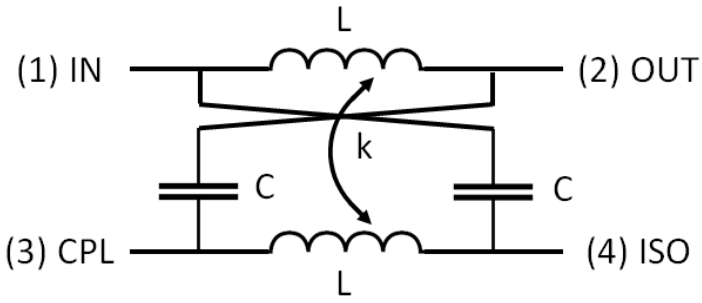

Figure 90: Lumped model of a quadrature hybrid coupler

The values of the lumped inductors and capacitors are respectively given by *Equation (70)* and *(71)*. They depend on the targeted operating frequency f_0 and coupler characteristic impedance Z_C. Moreover, the coupling factor *k* is related to several parameters (technology and implementation). It is equal to 1 for a perfect coupling; usually it is about 0.8-0.9 [8].

$$L = \frac{(2-k) * Z_C}{2 * \pi * f_0} \tag{70}$$

$$C = \frac{2-k}{Z_C * 4 * \pi * f_0} \tag{71}$$

b. Couplers in 5G Front-End Modules

Several common uses of RF couplers have been mentioned above: differential amplifiers [6], Doherty amplifiers [7], and DC switch [3]. These latter two implementations are of particular interest for 5G FEM's since they respectively improve the PAE at PBO and avoid the use of RF switches. Other examples are discussed in the current section. In [9], several reconfigurable couplers operate as power combiners in TX mode and are part of the RX matching network in the other mode. As in [3], RF switches are not required. Besides, the couplers also enable to filter out the 2nd harmonics, increasing the transmitter P_{SAT} and efficiency. Indeed, the FEM takes advantage of the limited bandwidth of the couplers.

Quadrature hybrid couplers may also be used to design phase shifters, which are major components of hybrid and analog beamformers. In [10], a reflection-type phase shifter (RTPS) is implemented this way. It is depicted in Figure 91. The output phase $\Delta\phi$ is controlled changing the impedance of the loads of the THRU and CPLD ports, Z_T. This also affects the circuit insertion losses *IL* as given in *Equation (72)* and *(73)*.

$$IL = -20 * \log|\Gamma_T| \tag{72}$$

$$\Delta\varphi = -90° + \angle\Gamma_T \tag{73}$$

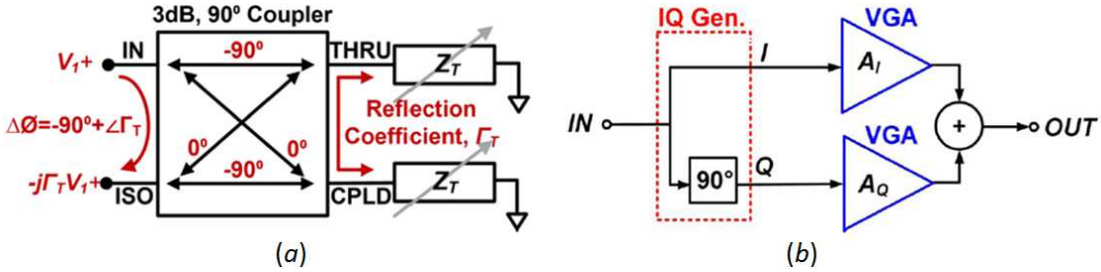

Figure 91: (a) RTPS with quadrature hybrid coupler, from [10]; (b) Vector modulator phase shifter, from [11]

Since the RTPS is a passive device, it exhibits high insertion loss. To cope with this drawback, it is possible to implement an active phase shifter as the vector modulator phase shifter presented in [11]. It is also based on a quadrature hybrid coupler. Then, the signals are amplified by variable gain amplifiers, as illustrated in Figure 91. The gain imbalance between the two paths generates the phase shifting of the recombined signal. This structure provides both amplitude and phase control.

Finally, in [12] quadrature hybrid couplers are used to implement a balanced amplifier. This topology, depicted in Figure 92, is of particular interest for 5G FEM's since it protects the PA from the output load variations, including the coupling between neighboring array elements. The variations are damped by Z_{OUT}. This circuit is protected from its environment through its construction. Thus, no additional element is required to recover the nominal performances. This topology is named self-contained PA. Indeed, such architecture is based on an RF crossover, which is depicted in Figure 93. The input quadrature hybrid coupler operates as a divider; the second one recombines the signal through the output port. Ideally, there is no leakage through the two other ports. This topology is often used to provide isolation between two RF paths crossing each other. However, the use of couplers limits the available bandwidth [13].

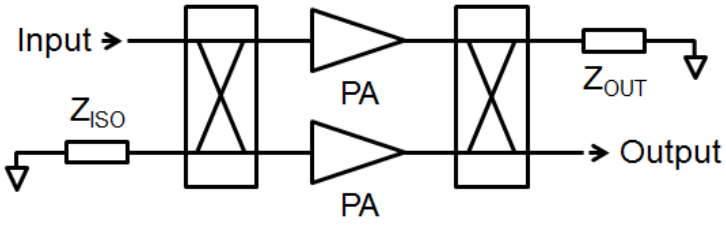

Figure 92: Block diagram of a balanced PA

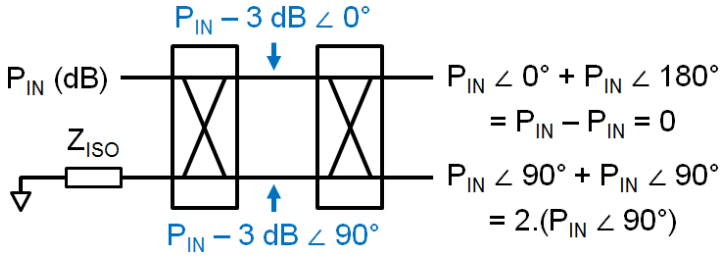

Figure 93: Block diagram of an RF crossover

c. Design of an Integrated Ka-Band Hybrid Coupler

The previous section has summed up the different uses of couplers within a FEM. In order to cope with some drawbacks of the series-shunt SPDT switch, it has been chosen to build a new version based on

hybrid couplers. The first step is therefore to implement an integrated quadrature hybrid coupler operating in the targeted frequency band and using the *ST C65SOIFEM* technology. The circuit will benefit from the thick top metal level, which enables to design passives with high quality factor. However, only one thick metal is available. Thus, the coupler design is restricted to planar topologies such as branch-line or Lange.

An interesting choice is the scalable twisted coupler described and successfully implemented in [8], at 5.85 GHz and 14 GHz, using H9SOIFEM CMOS 130 nm technology from STMicroelectronics . As depicted in Figure 94, it is composed of a combination of coupled lines and twisted elements. Besides, this topology is highly scalable since several twists may be connected. *Equations (70)* and *(71)* are used to determine the suitable values of L and C, in order to design a 50 Ω coupler operating at 28 GHz.

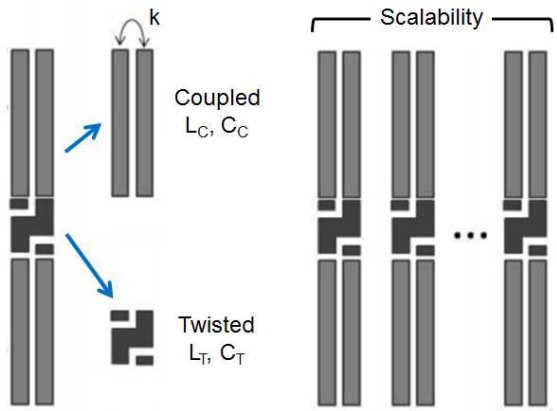

Figure 94: Scalable twisted hybrid coupler, from [8]

The results are plotted in Figure 95. The required values of L and C for a 50 Ω hybrid coupler operating at 28 GHz are respectively 313-369 pH and 62.5-73.9 fF depending on the coupling factor value. A first trial is achieved with *L* = 341 pH and *C* = 68.2 fF, which corresponds to *k* = 0.8. This results in an amplitude imbalance lower than 0.1 dB at 28 GHz and phase imbalances lower than 0.5°, as plotted in Figure 96 (*a*). This first simulation does not take into account the performance degradation due to the finite quality factor of the passives. Considering that Q = 20, the parasitic resistance of the inductors is 3 Ω. Thus, the amplitude and phase imbalance at 28 GHz are degraded. Their values are respectively 0.15 dB and 2°.

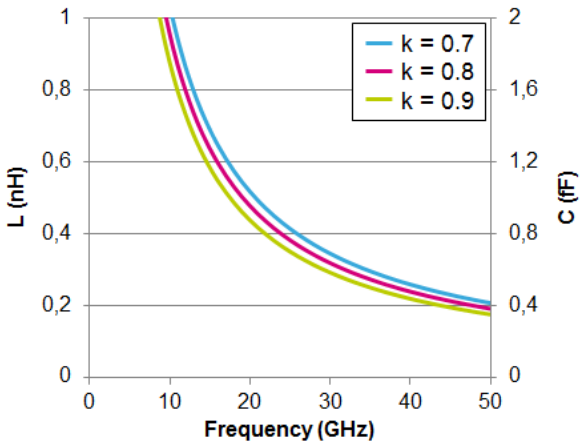

Figure 95: L and C Values for a 50 Ω Hybrid Coupler

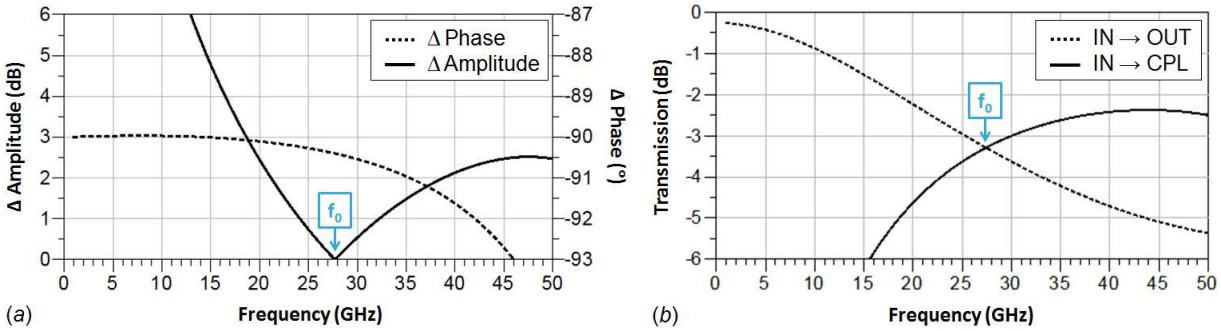

Figure 96: (a) Amplitude and phase imbalances of the 50 Ω hybrid coupler with ideal inductors; (b) Transmissions through the OUT and CPL ports with real inductors (quality factor = 20)

This first design step provides a good estimation of the required values of L and C. It also highlights that a finite quality factor decreases the operating frequency (f_0) and degrades the coupler performances. Then, the coupler is designed using the *ST C65SOIFEM* technology and simulated thanks to a 3D EM simulator. The circuit layout is shown in Figure 97. It has been chosen to distribute L and C between two twists. The post-layout simulation results, including pads, are reported in Table 40.

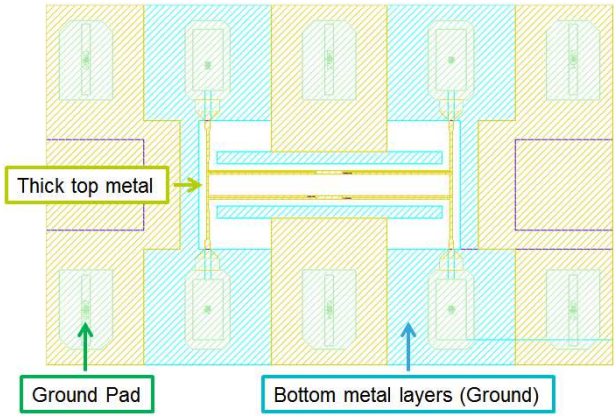

Figure 97: Hybrid coupler layout top view

Up to the post-layout simulation results, the operating frequency is 12% higher than expected. This shift is mainly due to the pad ring, as also observed in [8]. However, other parameters may also cause these discrepancies. Indeed, the quality and coupling factors have been arbitrarily estimated. Their variations could lead to differences between the first model and the post-layout simulations. The quality factor has been supposed to be 20 but it could reach 25 or 30 with the selected technology. As previously shown, increasing Q results in a higher operating frequency. A degradation of the coupling factor has the same effect, as illustrated in Figure 95. Predictably, the insertion losses are higher in post-layout simulations.

	Operating frequency (GHz)	**Insertion loss (dB)**	**Phase imbalance (°)**
First model	27.4	0.29	1.8
Post-layout simulation	31.3	0.67	1.5

Table 40: Post-layout simulation results of the designed quadrature hybrid coupler

d. Measurement Results

The above designed quadrature hybrid coupler has been manufactured during the third run of the thesis (named *Alioth*). Its small-signal performances have been measured on-wafer. The results are reported in Figure 98, with the die micrograph. The coupler exhibits insertion losses of 3.8 dB and phase imbalance lower than 1° at its operating frequency, which is 28.65 GHz. Besides, over the 26.4-30.9 GHz frequency band, the amplitude and phase imbalance are respectively lower than 1 dB and lower than 1.5°. Finally, the isolation is above 15 dB and the return losses are better than 18 dB over the whole measurement span. The discrepancies between these results and post-layout simulations may come from a degraded quality factor of the passive components. The coupling factor may also have been shifted.

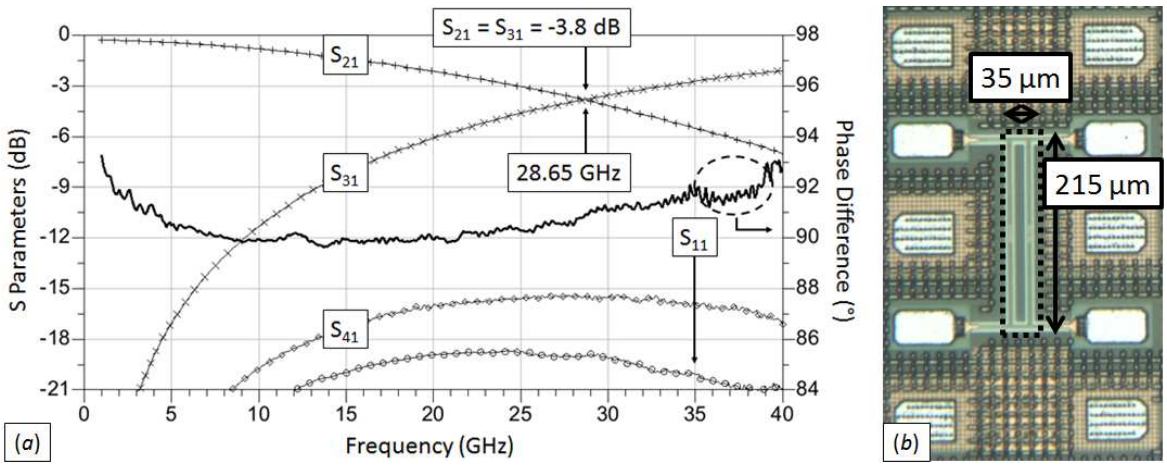

Figure 98: (a) Quadrature hybrid coupler measurement results; (b) Circuit micrograph

This first version of the quadrature hybrid coupler is intended to be used as a building block for the SPDT switch. Its design will be described in the next part. However, the coupler performances are critical for its many uses. Thus it will require future improvements. Firstly, the simulation model has to be refined to fit better with the measurement results. The coupler insertion loss has to be further reduced as well.

3. SPDT Switch Based on Hybrid Couplers

a. Principle of Operations

Firstly, an SPDT switch has been designed based on the RF crossover illustrated in Figure 93. It consists of two quadrature hybrid couplers and two couples of SPST switches, as depicted in Figure 99. The signal is routed towards OUT 1 or OUT 2, depending on the switches' states. Indeed, when SPST 1's are activated and SPST 2's are disabled, the input signal is transmitted to OUT 1. It is the same configuration as for the previously discussed balanced PA. Conversely, the signal may be routed towards OUT 2 reversing the states of the switches.

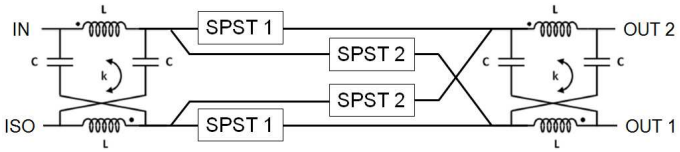

Figure 99: SPDT switch based on an RF crossover and two couples of SPST switches

The use of only series SPST switches instead of series and shunt devices results in lower IL. However, it is necessary to also take into account the quadrature hybrid coupler losses (IL_{hybrid}). Thus, the SPDT switch insertion losses (IL_{SPDT}) are given by *Equation (74)*. The series SPST losses have been detailed in Table 30.

$$IL_{SPDT} = 2 * IL_{hybrid} - 20 * \log_{10}\left|\frac{2 * Z_0}{2 * Z_0 + R_{ON}}\right| \quad (74)$$

Considering that Z_0 is 50 Ω and R_{ON} is 5 Ω, which corresponds to a 160 µm device as reported in Table 33, the series switch insertion losses are 0.4 dB. This part could be minimized using larger devices, an R_{ON} of 2 Ω results in losses of 0.17 dB. Besides, the losses of each hybrid coupler may probably be reduced by up to 0.5 dB. Consequently, the overall SPDT IL's are estimated to be lower than 1.2 dB. Finally, the first circuit simulations predict an isolation about 20 dB. The conclusion is therefore that this topology is able to provide close performances compared with the previously designed series-shunt switch. However, the use of couplers limits the available bandwidth and highly increases the circuit area.

A second version has been designed. It is still composed of two hybrid couplers but combined with shunt switches only, instead of series devices. The topology is depicted in Figure 100. It has firstly been introduced in [4], and implemented at 2.3 GHz using branch-line hybrid couplers and PIN diodes as switching components. In this section, such SPDT switch is designed to operate in the Ka-band and integrated thanks to the use of the selected RF SOI technology.

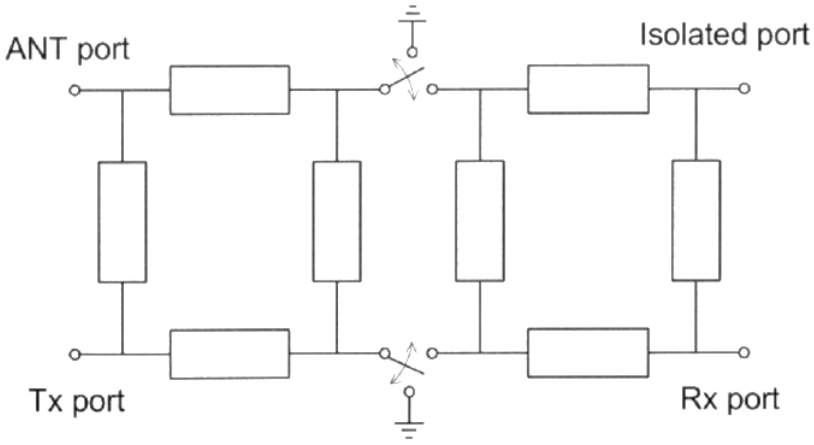

Figure 100: SPDT switch based on couplers and shunt switches, from [4]

As shown in Figure 100, this structure is asymmetrical. The operating principle of the RX and TX states are therefore different. They are detailed below, respectively in Figure 101 and Figure 102. Firstly, when the shunt switches are open, the circuit is almost equivalent to the SPDT switch in Figure 99. Its insertion losses (IL_{RX}) are computed thanks to *Equation (75)*. It derives from *Equation (74)* and Table 30, replacing the series SPST contribution by the shunt SPST one. In contrast with the previous circuit, the SPST switch contribution may be minimized by reducing the device width, i.e. reducing C_{OFF}. Moreover, the power handling is improved compared with the usual series-shunt switch, since the input signal is split by the first coupler. Thus, each shunt switch has to handle only half of the input power.

$$IL_{RX} = 2 * IL_{hybrid} - 20 * \log_{10}\left|\frac{2}{2 + j\omega * Z_0 * C_{OFF}}\right| \quad (75)$$

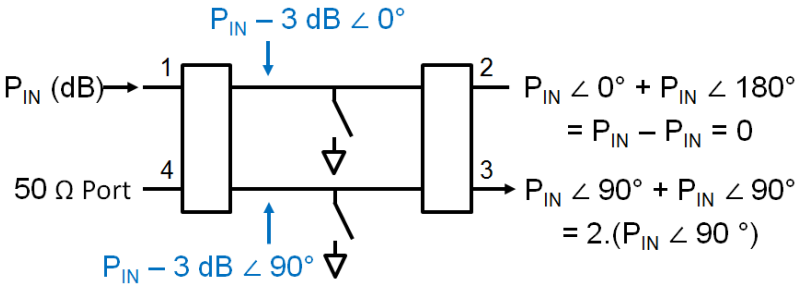

Figure 101: Operating principle of the SPDT switch in RX state

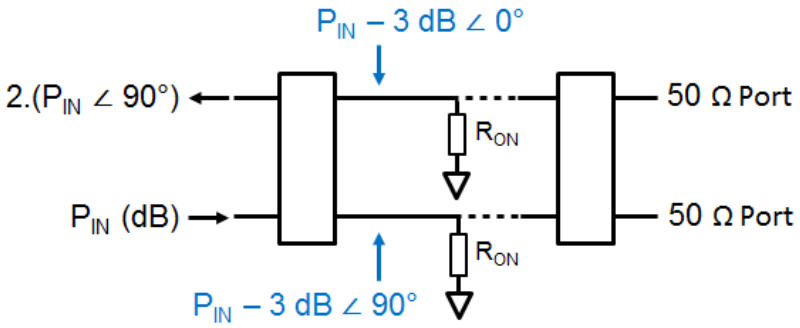

Figure 102: Operating principle of the SPDT switch in TX state

When the shunt switches are closed, the signal is reflected by the ground. Thus, the circuit is equivalent to a reflection-type phase shifter. Thanks to the destructive interferences, no signal is reflected towards the input port. Considering MOS switches, the circuit insertion losses (IL_{TX}) are given in *Equation (76)*, which derives from *Equation (72)*. Thus, it is possible to minimize the switch IL thanks to the use of larger devices, which have reduced R_{ON}. Moreover, this reflector is almost equivalent to a fully passive circuit. So, its power handling is highly enhanced compared with the usual series-shunt topology.

$$IL_{TX} = -20 * \log_{10} \left| \frac{R_{ON} - Z_0}{R_{ON} + Z_0} \right| \tag{76}$$

Finally, *Equations (75)* and *(76)* have raised a significant tradeoff on the device width, as it was the case for the previously designed series-shunt switch. Based on these equations, this topology seems to be able to exhibit low IL. However, it requires a careful sizing to minimize the loss in both states. The hybrid coupler performances have also a significant contribution. Besides, the asymmetrical structure results in a very high power handling TX state. The sizing methodology and the circuit implementation are detailed in the next part.

b. Circuit Design

The above introduced SPDT switch is implemented, as illustrated in Figure 103. The previously designed hybrid couplers are combined with GO2 BC NMOS transistors. They are preferred to their GO1 FB counterpart because they can handle a higher voltage swing. Thus the power handling is improved in the RX state. As it was the case for the series-shunt switches, the control voltage is applied at the transistors' gates and body through large resistors. In the RX and TX states V_C is 0 V and 2.5 V, respectively. This results in floating bodies and gates, which improves the switch performances. However, previous results have shown that the power handling of the GO2 series shunt switch was degraded. Further

improvements may be to connect the bodies to the ground, still through large resistors to make it floating. It avoids charge accumulation.

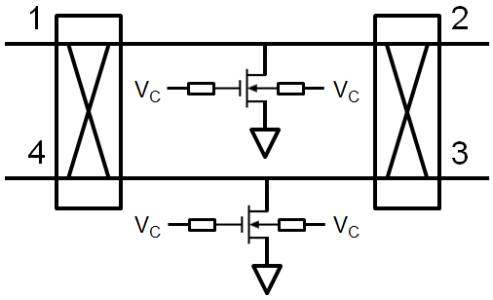

Figure 103: Asymmetrical SPDT switch schematic

As shown by *Equations (75)* and *(76)*, there is a significant tradeoff between IL_{RX} and IL_{TX}, which depends on the devices' width. Thus, the insertion losses of both states are plotted at 28 GHz, depending on the shunt switch width. It enables to determine the optimal size. The results are reported in Figure 104. As expected, the IL's in RX state are minimized with a small transistor, i.e. a reduced C_{OFF}. Conversely, larger devices improve IL_{TX}, since R_{ON} is minimized. It has been chosen to reduce IL_{RX} at the cost of higher IL_{TX} in order to lower the receiver NF. Thus, the shunt width is set to 250 µm.

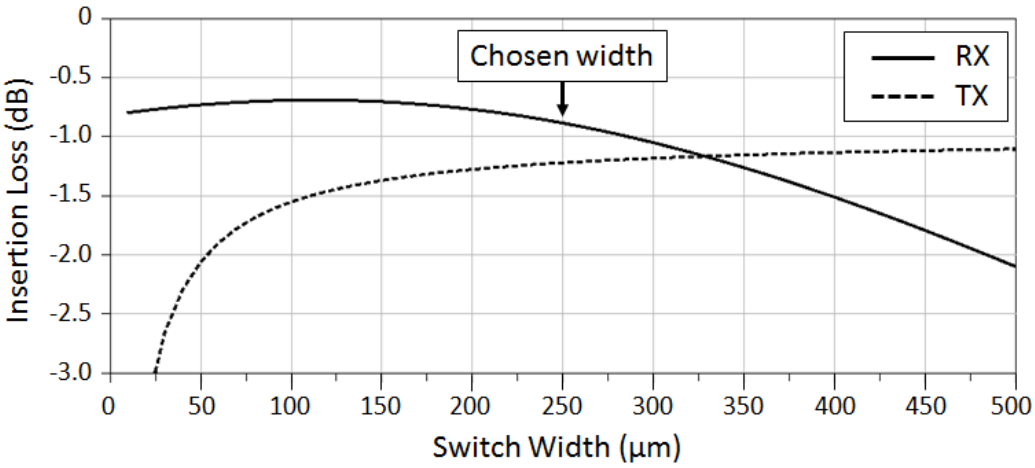

Figure 104: IL_{RX} and IL_{TX} plotted depending on the shunt switch width, at 28 GHz

Figure 105: Top view of the shunt positioning in relation to the transmission line

The two 250 µm shunt devices are split into 6 smaller transistors, as illustrated in Figure 105. Therefore, they are distributed among both sides and over the whole length of the transmission line between the two couplers. The parasitic elements (resistors and capacitors) of these elementary cells are extracted.

Then, the whole SPDT is laid out and simulated. The results are reported in Table 41. Firstly, the switch operating frequency is slightly above 28 GHz. Indeed, it comes from the coupler features. As expected, the insertion losses are lower in RX than in TX. However, they are not as good as in the first simulations. This discrepancy may come from the impedance of the transmission line between the couplers. This also degrades the isolation of the RX state. Besides, the TX state isolation is sufficient to protect the LNA.

	Frequency (GHz)	IL (dB)	Isolation (dB)	ICP1 (dBm)
RX State	30.0	1.84	15.3	15.0
TX State	29.5	2.68	22.9	> 30

Table 41: Post-layout simulation results of the asymmetrical SPDT switch

Finally, the RX state exhibits a poor power handling compared with the series-shunt results, which was 17.5 dBm in post-layout simulations. Since the body has not been properly connected to the ground, it may explain that the behavior of the transistor in large-signal simulations is not accurate, and that its performances are degraded. The impedance of the transmission line also has a significant impact. These two points will require a particular care in further designs. As expected, the TX state power handling is very high, above 30 dBm at the switch operating frequency (29.5 GHz).

c. Measurement Results

This asymmetrical SPDT switch based on quadrature hybrid couplers has been manufactured in the run called Alioth, like the standalone hybrid coupler and the second GO2 BC series-shunt SPDT switch. The die micrograph is presented in Figure 106.

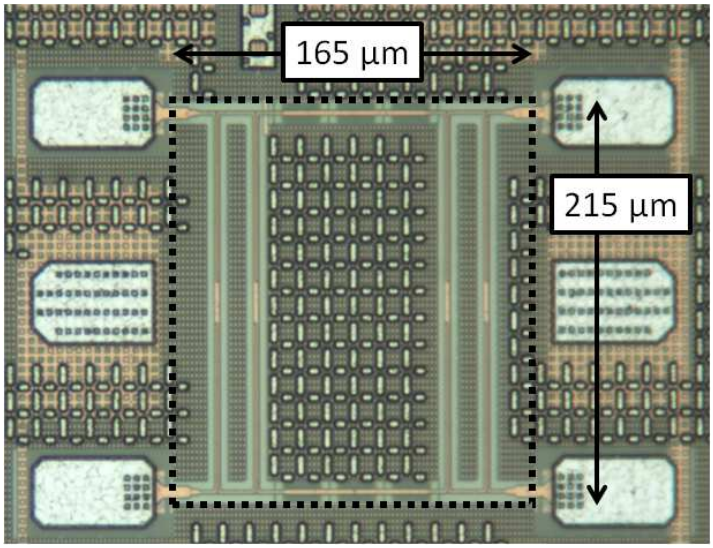

Figure 106: Micrograph of the asymmetrical SPDT switch

Firstly, small-signal measurements have been achieved in both RX and TX states. Differential probes are used; they enable to simultaneously measure the IL and the isolation. The fourth port is connected to a 50 Ω termination. Only one voltage source is required to control the 6 shunt transistors.

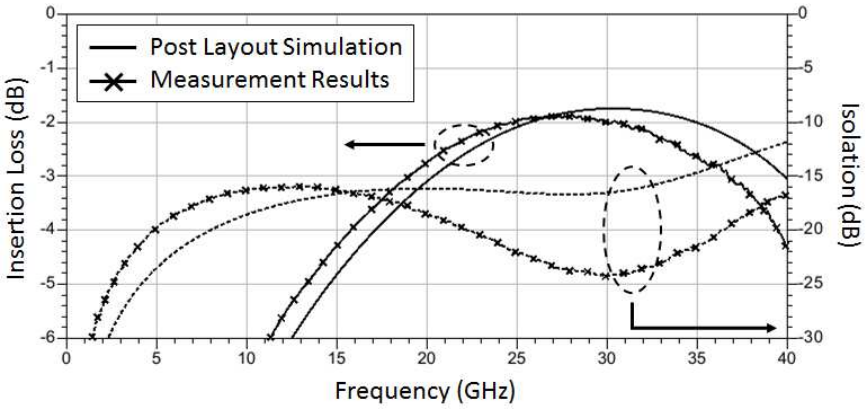

Figure 107: Small-signal measurements of the asymmetrical SPDT switch in RX state

The measurement results are plotted and compared with post-layout simulation results in Figure 107 for the RX state. Likewise, results for the TX state are plotted in Figure 109. There is the same frequency shift as the one observed for the standalone hybrid coupler. It is particularly visible in the RX state. Besides, the isolation measured in this state is far better than the simulation expectations. Indeed, as given in *Equation (76)*, the leakages through the TX port are higher as the reflection coefficient is high. Ideally, this coefficient is equal to zero and the whole signal is transmitted to the RX port. Small variations of the line impedance (Z_T in Figure 108) may highly degrade the reflection coefficient and therefore the isolation. Consequently, the line impedance is probably closer to 50 Ω practically than it is in simulations.

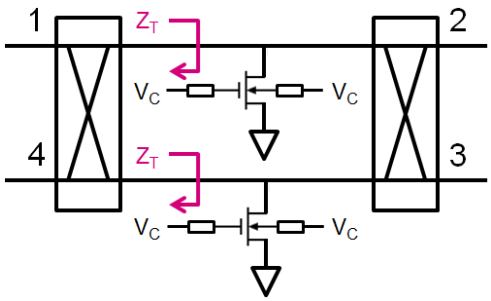

Figure 108: Asymmetrical SPDT switch schematic

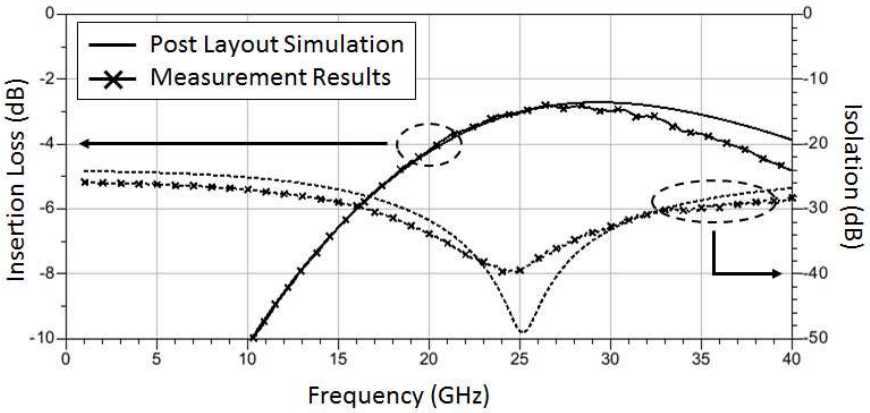

Figure 109: Small-signal measurements of the asymmetrical SPDT switch in TX state

In the TX state, the operating frequency is slightly shifted as well. However, measurement results are in good agreement with the post-layout simulation expectations, as plotted in Figure 109. Besides, the ICP1 has been measured in both states, as shown in Figure 110. It is 17.5 dBm in the RX state, which is higher than the post-layout simulation expectations. This result is not consistent with the GO2 BC series-shunt measurements. It may come from the improper body connection or to model inaccuracies. Conversely, the results for the TX state meet the expectations. The circuit gain exhibits very low compression up to an input power of 18 dBm. Due to the available laboratory instrument limitations, it was not possible to measure the IL with higher input power.

Finally, all the measurement results are reported in Table 42 and compared with post-layout simulation expectations. Overall, measurements are consistent with simulations. However, several discrepancies have been highlighted: a frequency shift due to the hybrid couplers and a better isolation in the RX state due to a shifted reflection coefficient. The large-signal simulation models also have to be refined.

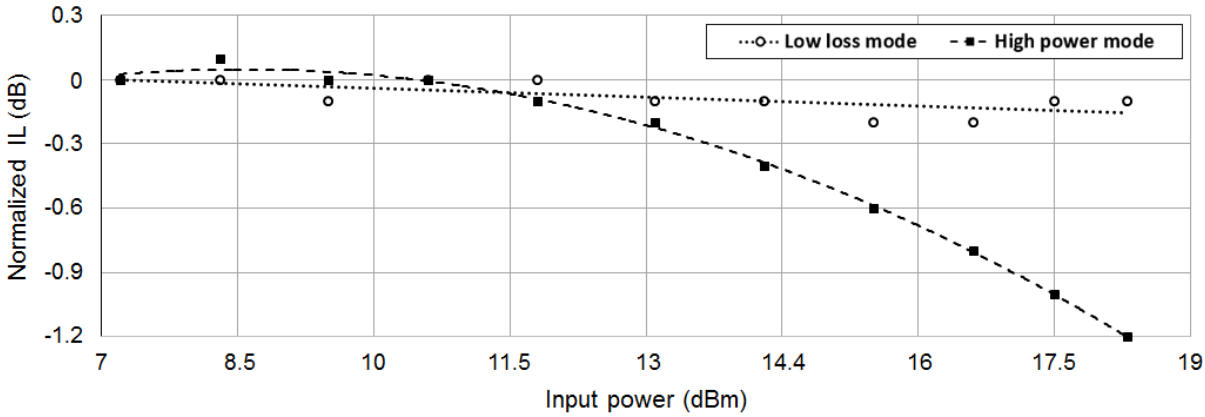

Figure 110: Large-signal measurements of the asymmetrical SPDT switch in both states, including trendlines

	Frequency (GHz)	IL (dB)	Isolation (dB)	ICP1 (dBm)
Simulated RX	30	1.84	15.3	15.0
Measured RX	28	1.88	23.8	17.5
Simulated TX	29.5	2.68	22.9	> 30
Measured TX	26.5	2.79	24.1	> 18

Table 42: Post-layout simulation and measurement results of the asymmetrical SPDT switch

d. Perspectives for Future Designs

The implemented asymmetrical SPDT switch is compared with recently reported circuits in Table 43. The first remarkable result is that its IL's are quite high compared with the other Ka-band SPDT switches. Thus, an interesting perspective of this work is therefore to improve the IL's. In the RX state, it is just necessary to reduce the losses due to the hybrid couplers while in the TX state, it requires further investigations about the ground plane and the reflection coefficient.

	[1]	[14]	Series-Shunt (GO2 BC)	This work TX State	This work RX state
Process	45 nm SOI	150 nm GaAs	65 nm SOI		
IL (dB)	0.8	0.9	0.7 [1]	2.8	1.9
Isolation (dB)	23	> 25	22 [1]	35	24
Area (mm²)	4×10^{-3}	1.2	4.5×10^{-3}	3.5×10^{-2}	
ICP1 (dBm)	29	> 33	16.5	> 30 [2]	17.5
Bandwidth (GHz)	DC-50	27-46	DC-40	21-36	20-37

[1] Core Post-layout simulations, [2] Post-layout simulation results, measured up to 18 dBm

Table 43: Ka-band SPDT switch state of the art

The reflection coefficient is also a major concern to improve the switch isolation in both states. Besides, this circuit provides a very high power handling in TX state. In the RX state, the power handling is quite good but the simulation models have to be refined in order to fit better with the measurement results. It has been shown that this kind of topology may be implemented using advanced integrated technologies. It exhibits interesting performances, which have to be enhanced thanks to further investigations. Some drawbacks have been highlighted as well: bandwidth and area penalty. However, the circuit bandwidth is sufficient for the targeted application. Moreover, implementing this topology at higher frequencies will reduce the circuit area thanks to hybrid coupler scaling. Another perspective of this work is the implementation of complex functions into the switch. For example, the TX state reflector may operate as a reflection type phase shifter at the cost of degraded IL's. Finally, based on early tests, this switch does not protect the PA and the LNA from the antenna impedance variations.

II. New Low Noise Amplifier and Co-Integration Perspectives

In the previous section, the limitations of the series-shunt SPDT switch have been discussed. Then, a new topology has been introduced and implemented. Likewise, in the following part, the main drawbacks of the previously designed LNA are reminded. Several improvement techniques are introduced and a new LNA is implemented. Measurement results will enable to verify the relevance of the selected improvements. Secondly, directions for the co-integration of the SPDT switch and the LNA into a 5G FEM are given.

1. Improved Ka-Band LNA

a. Limitations of the First LNA

The LNA designed in the previous chapter had several issues, mainly due to the design methodology and the layout parasitic elements. Firstly, BC transistors with a low biasing point have been chosen to reduce the power consumption. However, this resulted in a too low current density, i.e. to a high NF. Besides, a compact inter-stage matching network has been selected. Therefore, it had not optimal performances. Finally, the simulation model was not accurate enough. Several parasitic elements from the layout have not been taken into account, such as the coupling between inductors and the path from the source of

the second stage to the ground. Consequently, the measured performances of the LNA were below the expectations. However, the simulation model has been refined to recover these features with the retro-simulations, as reported in Table 44. In this section, design techniques to overcome the identified issues are discussed. Then, a new LNA is implemented and measured.

	Gain (dB)	NF (dB)	Z_{IN} (Ω)	Z_{OUT} (Ω)	I_{DC} (mA)	ICP1 (dBm)
Freq. (GHz)	26.4		28		-	28
Post-layout simulated	13.1	4.9	27 – 10j	55 – 14j	24	-4
Measured	13.7	5.1	25 – 16j	54 – 21j		
Targeted	20	2.5	50	50	20	0

Table 44: Comparison between post-layout simulation results and measurement results of the first LNA (Vega)

b. Design of the Second LNA

Since the above identified limitations are not due to the selected topology, it is chosen to design a two stage inductively degenerated LNA. The input stage is sized as previously: the common source transistor size is set to exhibit low noise figure while drawing a drain current lower than 10 mA. For now, however, it is chosen to combine a FB common source transistor with a BC cascode stage. This results in lower NF. Then, the input gate and source inductors are sized to provide 50 Ω input impedance at the resonance, as given in *Equation (54)*. The previous design has highlighted that it is challenging to achieve a proper matching with the subsequent stages, since the output impedance of the input stage is far from 50 Ω. Thus, a shunt resistor is used along the drain inductor to make the matching easier. It also results in larger bandwidth. However, it degrades the NF and the gain. The input stage is presented in Figure 111. Its performances are reported in Table 45 at 28 GHz, including for the extreme process corners.

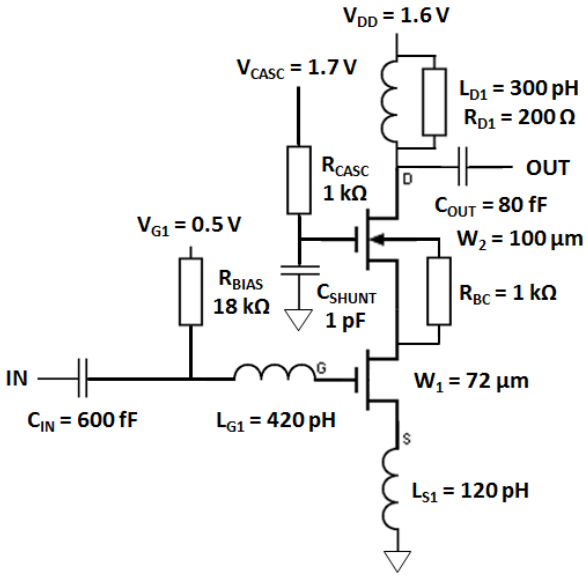

Figure 111: Input stage schematic of the second designed LNA

This new version exhibits better input and output matching. The gain and NF are decent although they are reduced due to the use of the output shunt resistor. The tradeoff with the drain current also limits the performances, in order to ensure low power consumption. Then, the input stage is simulated for different corners. This results in small impedance variations. The other features are shifted but they may be recovered with little adjustments of the biasing voltage. The typical ICP1 is -5 dBm.

	Z_{IN} (Ω)	Z_{OUT} (Ω)	S_{21} (dB)	S_{12} (dB)	NF (dB)	I_{DC} (mA)
Typical	52 – 7.8j	49 + 3.7j	8.76	- 22.6	2.13	8.9
Slow-slow	49 – 4.0j	50 + 6.0j	7.91	- 22.1	2.48	6.6
Fast-fast	55 – 11j	49 + 2.0j	9.42	- 22.9	1.86	11.6
Extracted	54 – 5.0j	48 – 8.2j	7.93	- 26.0	2.17	8.7

Table 45: Simulation Results of the Input Stage, at 28 GHz

In the first version, the second stage was degenerated by the parasitic path between the source and the ground. To cope with this issue, it is chosen to use an inductively degenerated cascode for the second stage too. Thus, the parasitic path may be neglected compared with the implemented source inductor. Since the input stage exhibits correct gain and noise performances, with a low drain current and proper 50 Ω input and output matching, it is chosen to also use this cell as output stage. This enabled to speed up the design. However, the limited ICP1 of this cell will results in a poor LNA power handling compared with the targeted features.

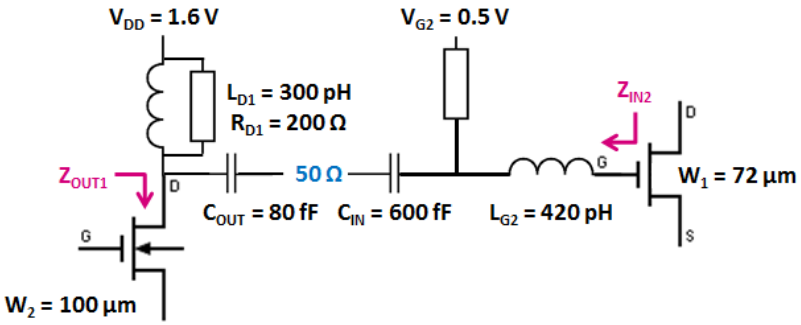

Figure 112: Inter-stage matching network

Another issue of the first LNA was the use of a compact matching network, which may have degraded the overall performances. For now, since each stage has 50 Ω input and output impedances, it is possible to directly connect them as depicted in Figure 112. However, such implementation results in two impedance transformations: Z_{OUT1} to 50 Ω and 50 Ω to Z_{IN2}. This step (reaching 50 Ω) could be avoided. Moreover, it may result in a matching network with high quality factor, limiting the available bandwidth [15]. As illustrated in Figure 113 (*a*), it is possible to design a matching network with lower quality factor to enhance the circuit bandwidth. Nevertheless, it requires the use of more components, including bulky inductors. Thus, it is necessary to find the proper tradeoff between the matching network quality factor, i.e. its bandwidth, and its complexity, i.e. its area. As shown in Figure 113 (*b*) for the first designed LNA, L_S and L_G of the first stage are located before the input cascode. Then, the space for L_{D1} and L_{D2} is limited by the decoupling capacitors. For now, it is also necessary to add L_G and L_S for the output stage.

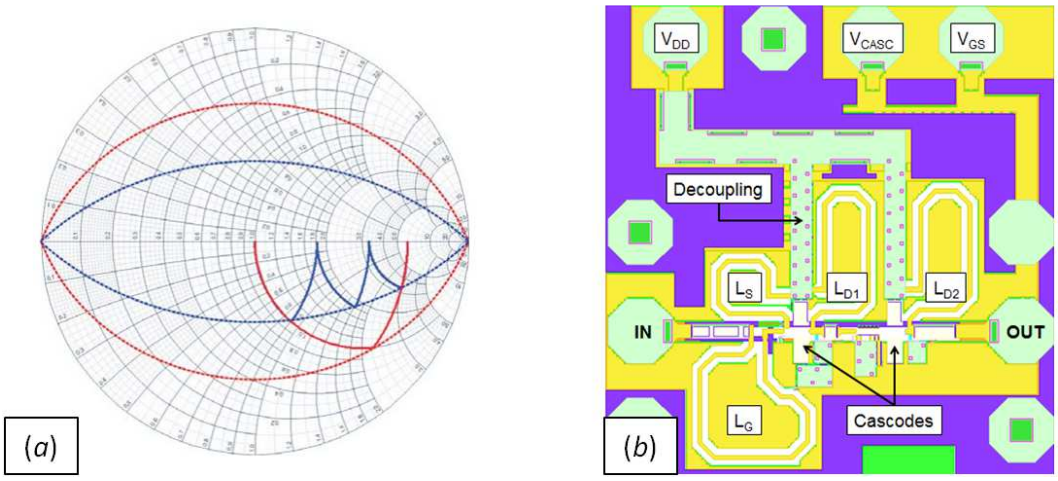

Figure 113: (a) Quality factor of generic one-element (red) and three-element (blue) matching networks, from [15]; (b) Top view of the designed first LNA layout

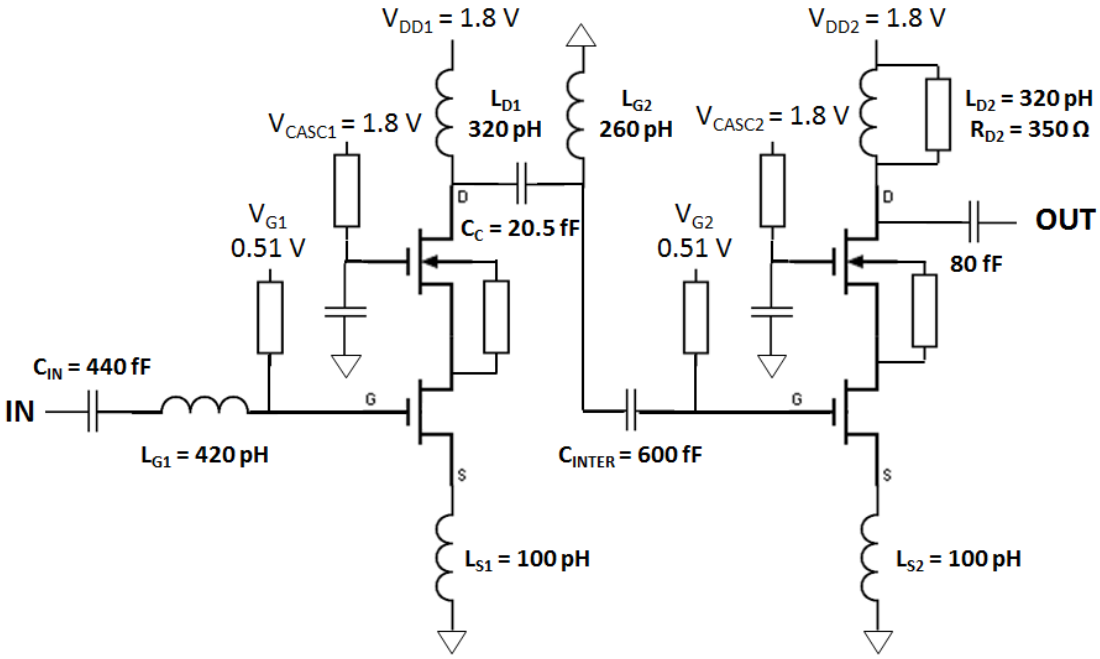

Figure 114: Schematic of the second designed LNA with improved inter-stage matching

The final schematic of the second designed LNA is depicted in Figure 114. In the input stage, the shunt resistor for a broadband output matching has been removed to provide higher gain and to improve the overall NF. This resistor has been conserved for the output stage, since it enables to achieve a better output matching. Besides, the inter-stage matching has been modified to exhibit a lower quality factor. This improves the circuit bandwidth without requiring additional inductors. The most relevant feature of this part is that L_{D1} and L_{G2} are coupled, which results in area saving. However, the co-design of these two inductors requires therefore a particular care. The designed LNA is laid out as shown in Figure 115. The parasitic elements (resistors and couplers) of the cascode cells are extracted. Then, a 3D EM simulation is achieved; it takes into account all the passives simultaneously, including the coupling between the inductors. These two parts are co-simulated. The results are reported in Table 46.

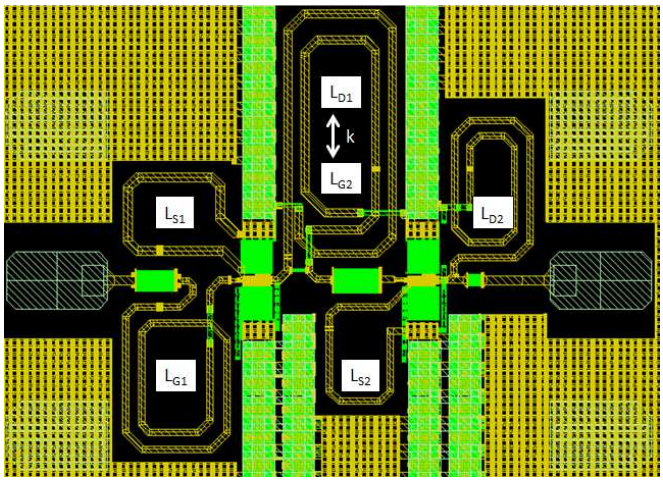

Figure 115: Top view of the second designed LNA layout

	Z_{IN} (Ω)	Z_{OUT} (Ω)	S_{21} (dB)	NF (dB)	S_{12} (dB)	I_{DC} (mA)	ICP1 (dBm)
24 GHz	24 - 18j	32 - 10j	21.2	2.2	-	21	-
28 GHz	47 + 6j	57 + 6j	20.6	2.2	-46.4		-15
31 GHz	58 + 10j	83 + 30j	17.8	2.5	-		-

Table 46: Post-layout performances of the second designed LNA over the Ka-band

The designed LNA exhibits good performances: NF lower than 2.5 dB over the whole operating band (24-31 GHz), a gain ripple of 4.4 dB over this frequency band, and a total current consumption of 21 mA. The circuit gain is slightly below 20 dB in the end of the operating band. Besides, the ICP1 is limited by the output stage; it is only -15 dBm. The circuit measurements are presented in the following section.

c. Circuit Measurements

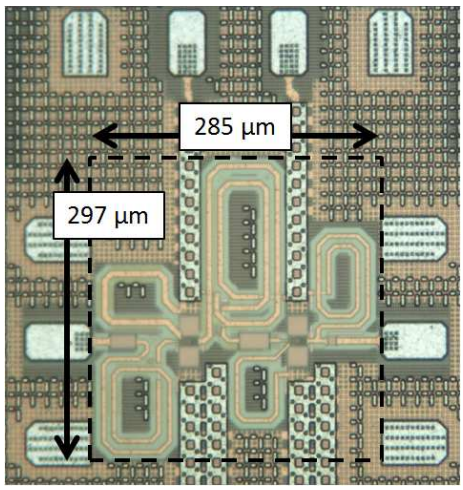

Figure 116: Micrograph of the second designed LNA

The LNA has been implemented using the selected *ST C65SOIFEM* technology. The die micrograph is shown in Figure 116. It should be noticed that more decoupling is used in this new version. Moreover, six pads are used for the DC supplies. Thus, it is possible to tune each stage separately. Firstly, small-signal measurements are achieved. The results are reported in Figure 117.

The first remarkable result is that the input return losses are above 0 dB at 16.4 GHz. Since there is also a gain spike at this frequency, it clearly reflects an oscillation. Several measurements have been made at different biasing points, they all exhibit this behavior. Thus, the circuit is not usable. However, it is necessary to identify the cause of the oscillation. Firstly, the spike is not visible on the output return losses. So this issue probably comes from the LNA input. Besides, it is a high frequency oscillation. Since the potential low frequency oscillations are avoided thanks to the decoupling, this one should be due to an internal loop. Consequently it could not be solved with additional external components.

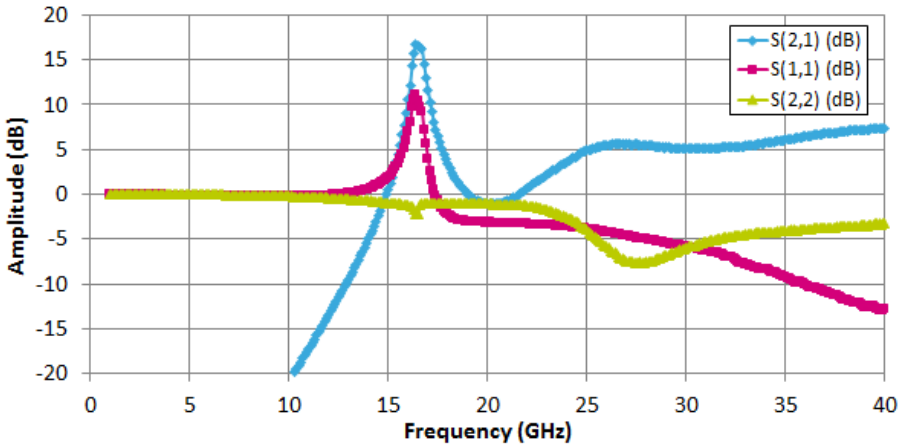

Figure 117: Small-signal measurements of the second designed LNA

After several investigations, the origin of this oscillation has been found. Actually, 8 virtual grounds have been used in the complete 3D EM simulation, as illustrated in Figure 118. This simplification was based on the assumption that the ground plane was perfect. This results in the shorting of several paths like the red one in Figure 118. Retro-simulations have been achieved with only 4 virtual grounds as shown in Figure 118. The results are plotted in Figure 119.

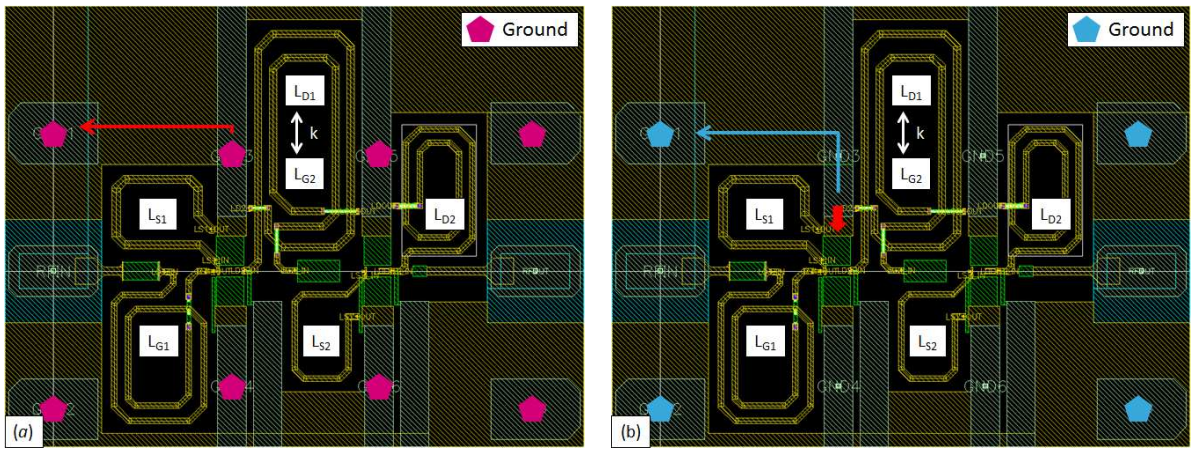

Figure 118: Layout top view including the simulation virtual ground pads: (a) First simulation; (b) Retro-simulation

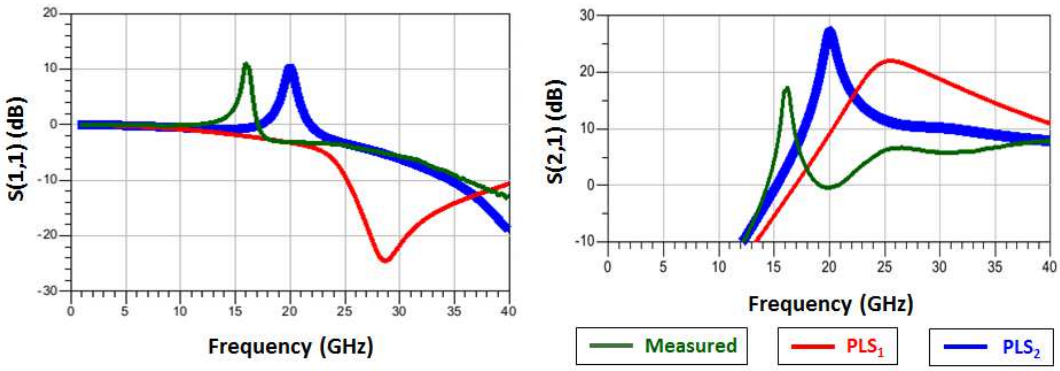

Figure 119: Retro-simulation results compared with first simulation and measurement results

The behavior observed in retro-simulations is similar to the measurement results: high frequency oscillation and significant input mismatch. The oscillation is shifted, since it is an early retro-simulation model. Two phenomena, depicted in Figure 120, occur in the circuit. The first one is that a part of L_S (shown in blue in Figure 120) is neglected. Thus the input impedance is shifted. Secondly, the output signal of the first stage is leaking into L_S through the decoupling capacitors. This creates a loop which generates the oscillation. The equivalent schematic is presented in Figure 120.

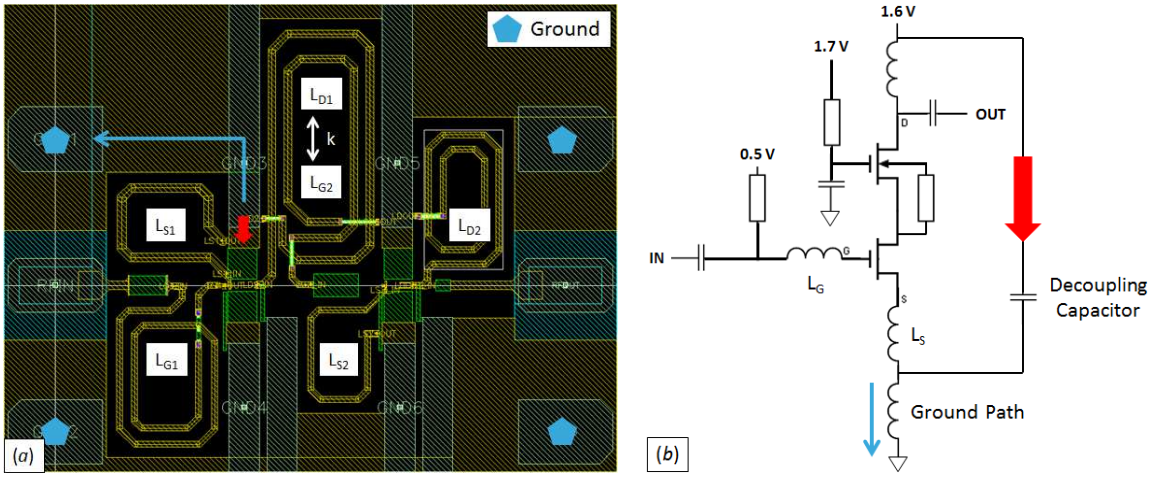

Figure 120: (a) Layout top view; (b) First stage equivalent schematic, both including the neglected effects

Two different ways to solve these issues have been identified. The first one requires significant layout modifications, but it fixes both the input mismatch and the oscillation. It consists in directly connecting the output of L_{S1} to a ground pad. Thus the path to the ground is controlled and the impedance shown at the decoupling capacitor output is too high to have signal leakages through them. The second solution is a layout patch, which is easier to implement but less effective. Indeed, the cascode cells have two ground accesses (north and south). In contrast with the first LNA, they have not been connected between each other in this new version. Adding a ring of bottom metal around the cell enables to

connect the two ground accesses, as shown in

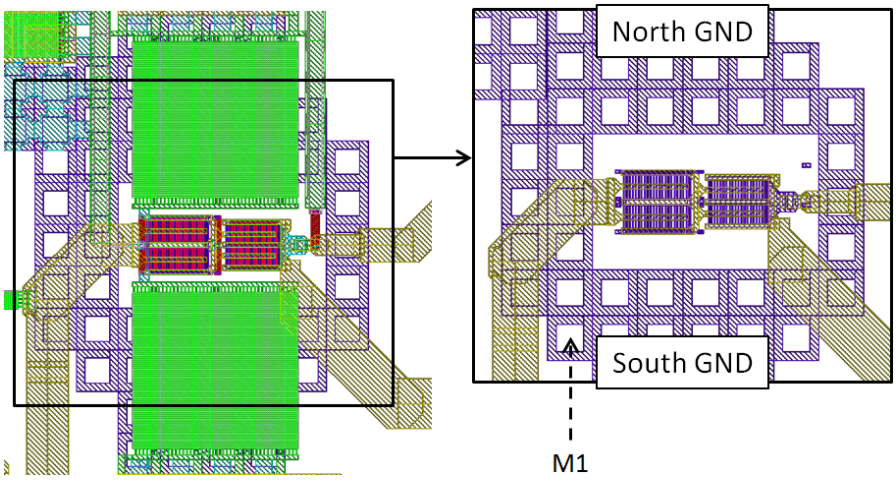

Figure 121. This enhanced ground access reduces the effect of the parasitic paths. The input matching is therefore improved and the signal loop responsible for the oscillation is broken. Even if the first solution is more effective, it is too complex to implement, in particular at the second stage. Thus, the layout patch is chosen and simulated. It enables to recover the LNA performances.

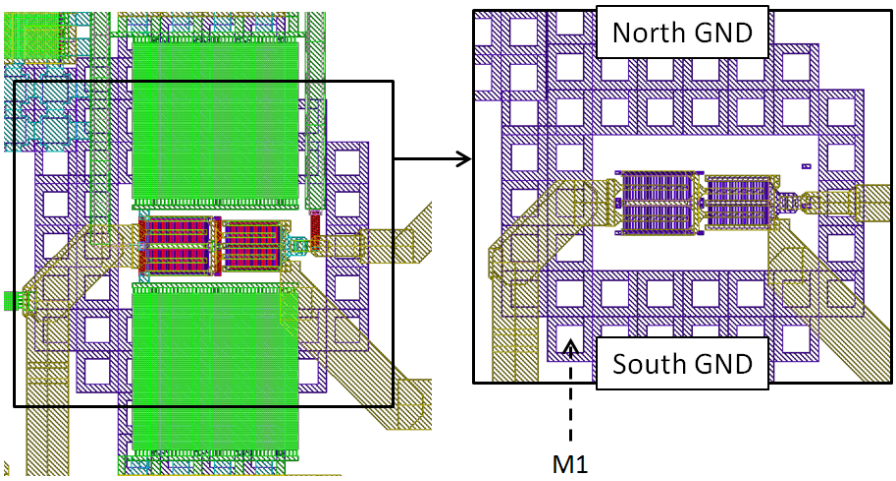

Figure 121: Top view of the layout patch

d. Conclusions

In this section, a new LNA has been designed and measured. Several improvements have been selected and implemented: higher current density thanks to smaller transistors and higher biasing points, lower NF thanks to the use of FB's, new inter-stage matching network with low quality factor in order to increase the bandwidth, and broadband output matching. These result in good performances, as shown in 1 Including SPDT Switch, [2] Post-layout Simulations

Table 47. However, the output stage limits the ICP1. Besides, measurement results have shown that there is a signal loop in the circuit generating a high frequency oscillation. The origin of this undesired behavior has been identified and a patch has been implemented. The LNA performances have been

recovered in retro-simulations. Unfortunately, it was not possible to reprocess the circuit in order to practically confirm that this solution is effective. Since the patch only consists in small layout changes, there is a high level of trust in the retro-simulation results.

Reference	[**16**]	[**17**]	[**18**]	**This work** [2]
Process	45 nm SOI	SiGe	90 nm GaAs	65 nm SOI
Bandwidth (GHz)	25 – 31	33 – 34	24.25 – 27.5	24 – 31
Gain (dB)	12.8	23.5	13.5 [1]	17.8 – 22.2
NF (dB)	1.6	2.9	4 [1]	2.2
I_{DC} (mA)	10	6	10	21
ICP1 (dBm)	-	-28.5	-8 [1]	-15

[1] Including SPDT Switch, [2] Post-layout Simulations

Table 47: Recently reported Ka-band LNA's

Finally, the above results have shown that *ST C65SOIFEM* technology seems to be suitable for the design of LNA's for Ka-band 5G. In the following section, the integration into a FEM is discussed.

2. Towards an Integrated Front-End Module

a. Overview

There are various ways to implement a transceiver: separate arrays, dual-polarization antennas [19], or with integrated duplexers [20]. Their features are summarized in Table 48. In this work, the integrated duplexer has been preferred for its low complexity and moderate size, i.e. cost.

	Separate Arrays	**Dual-Polarization**	**Integrated Duplexer**
Cost	High	Medium	Medium
Full-Duplex	Yes	Yes	No
Complexity	Low	Medium	Low

Table 48: Transceiver implementations overview

As discussed previously, the integrated duplexer may be implemented in an asymmetrical way in order to improve the performances of a state over the other [2]. Besides, it is possible to use more complex structures than the usual series-shunt switch to provide additional features. In [9], the switching circuit is based on a coupler and used as a filter to increase the linearity.

In this work, both usual symmetrical series-shunt and complex asymmetrical switches have been implemented. It has been shown that the *ST C65SOIFEM* technology is suitable for the design of such circuits for 5G in Ka-band. Besides, two LNA's have also been designed using this technology. They exhibit interesting performances for the targeted applications. The co-integration of these two parts into a single integrated FEM is the main perspective for further designs. Several directions are given below.

b. *Perspectives for Co-Integration*

One of the directions for further improvements is to co-integrate the SPDT switch and the LNA while adding advanced features. As discussed previously, the TX state of the asymmetrical switch may offer phase shifting. Then, two LNA's may be implemented between the coupler, as illustrated in Figure 122. This balanced topology offers lower gain and NF degradation compared with the usual combination of a series-shunt SPDT switch with a standalone LNA. Furthermore, such topology results in higher power consumption since two LNA's are needed and the TX gain is reduced by the reflection type phase shifter.

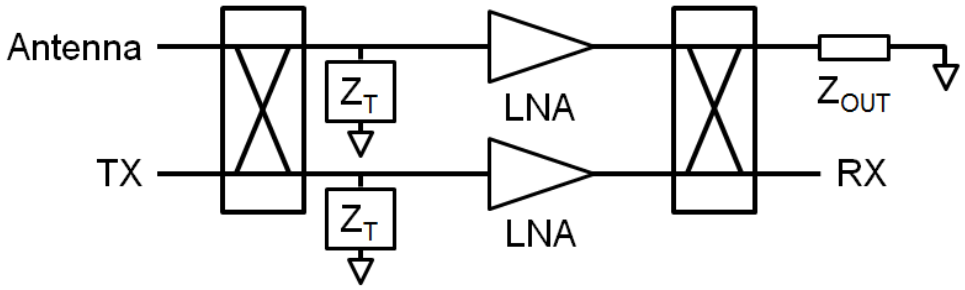

Figure 122: Schematic of the co-integration between the asymmetrical SPDT switch and the LNA

III. Conclusion

In this chapter, a new SPDT switch has firstly been designed. It enables to overcome the tradeoff between IL and isolation in usual SPDT switches. The asymmetrical topology also offers very high power handling in TX state while providing quite low IL in RX to reduce the overall NF. Besides, this architecture is favorable for the implementation of advanced features as balanced amplifiers with phase shifting.

Secondly, the limitations of the first designed LNA have been identified. Then, several improvements have been implemented such as smaller transistors to increase the current density or enhanced inter-stage matching network. This new LNA exhibits interesting post-layout simulation performances. However, the fabricated circuit has a high frequency oscillation, which makes it unusable. The issue has been identified and corrected in retro-simulations thanks to a layout patch. This gives a high level of confidence in the post-layout results. It is highlighted that the performances achieved thanks to the selected technology makes it suitable for the design of Ka-band FEM's for 5G. Finally, directions for further improvements have been discussed, including the co-integration of the LNA with the SPDT switch based on quadrature hybrid couplers.

IV. References

[1] C. Li et al., "< 0.8dB IL 46dBm OIP3 Ka band SPDT for 5G communication," in *IEEE 18th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, Anaheim, CA, 2018.

[2] B. Sadhu et al., "A 28GHz 32-Element Phased-Array Transceiver IC with Concurrent Dual Polarized Beams and 1.4 Degree Beam-Steering Resolutionfor 5G Communication," in *2017 IEEE International Solid-State Circuit Conference (ISSCC)*, San Francisco, CA, 2017.

[3] Y. Liu et al., "A 28-GHz transceiver front-end with T/R switching achieving 11.2-dBm OP1dB, 33.8% PAEmax and 4-dB NF in 22-nm FD-SOI for 5G communication," in *IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, Burlingame, CA, 2018.

[4] K. Kim and D. Ahn, "A new switching structure using branch-line hybrid couplers for time division duplex system," in *IEEE MTT-S International Microwave Symposium Digest*, Atlanta, GA, USA, 2008, pp. 999-1002.

[5] D. M. Pozar, *Microwave Engineering, fourth Edition*.: Wiley, 2012.

[6] W-C. Huang and H. Wang, "An inductive-neutralized 26-dBm K-/Ka-band power amplifier with 34% PAE in 90-nm CMOS," *IEEE Transactions on Microwave Theory and Techniques (TMTT)*, vol. 67, no. 11, pp. 4427-4440, November 2019.

[7] D. P. Nguyen, B. L. Pham, and A-V. Pham, "A compact Ka-band integrated Doherty amplifier with reconfigurable input network," *IEEE transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 205-215, January 2019.

[8] V. Knopik, B. Moret, and E. Kerherve, "Integrated scalable and tunable RF CMOS SOI quadrature hybrid coupler," in *12th European Microwave Integrated Circuits Conferance*, Nuremberg, Germany, 2017, pp. 159-162.

[9] S. Mondal, R. Singh, and J. Paramesh, "A Reconfigurable Bidirectional 28/37/39GHz Front-End Supporting MIMO-TDD, Carrier Aggregation TDD and FDD/Full-Duplex with Self-Interferance Cancellation in Digital and Fully Conneted Hybrid Beamformers," in *2019 IEEE International Solid-State Circuits Conference - (ISSCC)*, San Francisco, CA, 2019.

[10] R. Garg and Natarajan A. S., "A 28-GHz Low-Power Phased-Array Receiver Front-End With 360° RTPS Phase Shift Range," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 11, pp. 4703-47-14, 2017.

[11] D. Pepe and D. Zito, "Two mm-wave vector modulator active phase shifters with novel IQ generator in 28 nm FDSOI CMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 2, pp. 344-356, February 2017.

[12] B. Moret, V. Knopik, and E. Kerherve, "A 28GHz self-contained power amplifier for 5G applications in 28nm FD-SOI CMOS," in *IEEE 8th Latin American Symposium on Circuits & Systems (LASCAS)*,

Bariloche, 2017, pp. 1-4.

[13] Microwavse101. (2020, January) Hybrid (3 dB) couplers. [Online]. https://www.microwaves101.com/encyclopedias/hybrid-couplers

[14] Qorvo. (2008, February) TGS4302 27 - 46 GHz Ka band high power SPDT switch. [Online]. https://www.qorvo.com/products/p/TGS4302

[15] Keysight. (2006, April) Keysight EEsof EDA - Matching network Yin-Yang - Part 2. [Online]. https://www.keysight.com/us/en/assets/7018-01880/article-reprints/5989-9012.pdf

[16] C. Li, O. El-Assar, A. Kumar, M. Boenke, and G. M. Rebeiz, "LNA design with CMOS SOI process-1.4dB NF K/Ka band LNA," in *IEEE/MTT-S International Microwave Symposium*, Philadelphia, PA, 2018.

[17] B-W. Min and G. M. Rebeiz, "Ka-band SiGe HBT low noise amplifier design for simultaneous noise and input power matching," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 12, pp. 891-893, December 2007.

[18] Qorvo. (2019, Spetember) QPF4010 - 24.25 - 27.5 GHz GaAs Front End Module. [Online]. https://www.qorvo.com/products/p/QPF4010

[19] J. D. Dunworth et al., "A 28GHz Bulk-CMOS dual-polarization phased-array transceiver with 24 channels for 5G user and basestation equipment," in *IEEE International Solid - State Circuits Conference - (ISSCC)*, San Francisco, CA, 2018.

[20] J. Curtis, H. Zhou, and F. Aryanfar, "A Fully Integrated Ka-Band Front End for 5G Transceiver," in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, CA, 2016.

Conclusion and Perspectives

I. Summary

The objectives of this thesis work were twofold: system-level sizing and circuit design. In the first part, the new mobile communication generation (5G) has been introduced, and the different scenarios have been identified. Then, we have selected a use case: high data rate communications in a crowded urban environment. Based on the first standard requirements, it was aimed to reach 1 Gbps data rate per user and up to 100 users per small cell. This led to 100-element phased array transceivers. Then, the power consumption of different beamforming topologies has been estimated based on recently reported circuits. To meet the defined specifications, it appeared that digital beamforming is more suitable. Indeed, it theoretically enables to save about 40% power consumption compared with multi-beam phased array antenna. However, the only reported digital beamformer is composed of FPGA circuits. Thus, its power consumption is not comparable with the existing integrated phased array systems.

Secondly, several mm-wave circuits have been designed in order to demonstrate the performances of the advanced *ST C65SOIFEM* technology for 5G front-end modules. It has been characterized and compared with other integrated technologies. It has been shown that it is suitable for mm-wave circuit design. Then, two parts of the front-end module have been designed: standalone RF switches and low noise amplifiers. Thus, two compact low loss switches have been implemented. They exhibit insertion losses lower than 1 dB over the Ka-band. Moreover, their power handling and isolation are compliant with the defined specifications. The results for the two-stage low noise amplifier were less convincing, the measured gain and minimum noise figure are 13.7 dB and 4 dB, respectively. However, several design perspectives have been identified: improved inter-stage matching network and increased current density.

In the last chapter, improved versions of both the RF switches and the low noise amplifier have been presented. Firstly, a new Ka-band duplexer composed of hybrid couplers and shunt switches is designed. It enables to overcome the tradeoff between insertion loss and isolation, while providing high power handling in the transmitting mode. This circuit exhibits 1.9 dB insertion loss and 24 dB isolation in the receive state, and 2.8 dB insertion loss and 35 dB isolation in the transmit state. Moreover, the simulated ICP1 is higher than 30 dBm in this state. It has been shown that this topology has interesting performances for mm-wave 5G front-end modules. However, it has to be improved in order to reduce the insertion loss in both states. The low noise amplifier has also been enhanced. The new inter-stage matching network and the enhanced current density lead to a noise figure lower than 2.5 dB over the whole operating band. Circuit measurements have shown that there is an undesired high frequency oscillation. In post-layout simulations, the origin of this issue has been found. It comes from a signal loop composed by the source inductor and the decoupling capacitors. A layout patch has enabled to recover the simulated performances.

II. Perspectives

The first perspectives of this thesis work are to enhance the system-level sizing methodology. Indeed, the above results are sufficient to estimate the size of base station arrays but several features are not taken into account. For example, the consequences of the architecture choice on the adjacent leakage power ratio are not studied. Likewise, the above work focuses on the transmitter sizing, since it is the most significant contributor to the power consumption. Thus, the impact of the front-end module noise figure on the overall receiver is not estimated. The next steps of the 5G standardization will also enable to validate the sizing methodology and confirm the obtained results. Then, the power consumption model may be refined, taking into account the choice of the beamforming algorithm or using more accurate circuit models. The increasing number of 5G systems will also enable to populate the state of the art and to compare these implementations with the model expectations.

Secondly, the circuit measurements have paved the way to several short-term perspectives. Indeed, the implemented series-shunt switches have demonstrated good broadband performances. It has been shown that the combination of this topology with the use of SOI technology results in low-loss compact circuits. However, the power handling was limited by the use of floating body transistors. It is therefore expected to improve the performances thanks to body contacted devices. The standalone low noise amplifier has been enhanced; it exhibits interesting performances but also an undesired oscillation. An immediate perspective is therefore the implementation of the layout patch to validate the simulated performances thanks to measurements. The last short-term perspectives involve the asymmetrical switch and the co-integration. The design of low-loss couplers will drastically reduce the overall insertion losses, making this circuit more competitive with the recently reported switches. The co-integration of low noise amplifiers with this topology into an integrated 5G front-end module is also being studied.

Finally, there are several longer-term perspectives, as targeting other applications. Indeed, the high power handling of the asymmetrical switch makes it suitable for space and defense purposes. The move towards higher frequencies will also result in a more compact circuit.

III. Publications

Journals

T. Despoisse, N. Deltimple, A. Ghiotto, M. De Matos, J. Forest, and P. Busson, "**An integrated 65 nm CMOS SOI Ka-band asymmetrical single-pole double-throw switch based on hybrid couplers,**" in *IEEE Microwave and Wireless Components Letters*, doi: 10.1109/LMWC.2020.3028290.

International Conferences

T. Despoisse, A. Ghiotto, P. Busson and N. Deltimple, "**A Comparison of Beamforming Schemes for 5G mm-Wave Small Cell Transmitters,**" 2018 16th IEEE International New Circuits and Systems Conference (NEWCAS), Montreal, QC, 2018, pp. 6-9.

T. Despoisse, N. Deltimple, A. Ghiotto, M. De Matos, and P. Busson, "**Low-Loss Ka-band SPDT Switch Design Methodology for 5G Applications in 65 nm CMOS SOI Technology**," *2020 IEEE 20th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)*, San Antonio, TX, USA, 2020.

National Conferences

T. Despoisse, A. Ghiotto, P. Busson and N. Deltimple, "**A Comparison of Beamforming Schemes for 5G mm-Wave Small Cell Transmitters,**" 13ème colloque du GDR SoC/SiP, 13-15 Juin 2018, Paris.

T. Despoisse, P. Busson, A. Ghiotto and N. Deltimple, "**Émetteurs à Réseau Phasé pour Stations de Base 5G aux Fréquences Millimétriques en Zone Urbaine Dense,**" 21èmes Journées Nationales Micro-Ondes (JNM), 14-17 Mai 2019, Caen.

Titre : Transmetteurs intégrés à haut rendement pour les réseaux phasés 5G à 28 GHz

Résumé : Actuellement, nous atteignons les limites de la quatrième génération de télécommunications mobiles (4G), notamment en termes de débit disponible et de nombre d'utilisateurs maximal. Une nouvelle génération (5G) est en cours de développement pour répondre à ces besoins. Elle adressera différent cas d'usages : l'internet des objets, les communications à ultra-haut débit ainsi que des applications plus critiques telles que les véhicules autonomes ou la chirurgie à distance. Les exigences étant bien supérieures aux capacités actuelles du réseau, des solutions de rupture ont été proposées.

Dans ce travail, nous avons proposé une nouvelle méthode de dimensionnement des systèmes 5G adressant les communications à très haut débit en bande Ka. Différentes architectures permettant de répondre aux spécifications définies ont été étudiées. Nous avons ensuite mis au point une méthode pour comparer leurs performances par rapport à leur consommation énergétique afin de choisir le système le plus adapté. Enfin, nous avons choisi et caractérisé une technologie CMOS avancée afin de réaliser certaines parties de ces systèmes. Nous avons conçu des commutateurs RF et des amplificateurs faible bruit intégrés, adressant les fréquences millimétriques.

Mots clés : 5G, amplificateur faible bruit, bande Ka, CMOS, commutateur RF, formation de faisceau, méthode de dimensionnement, transmetteur à réseau phasé.

Title: 5G 28 GHz high efficiency integrated phased-array transceivers

Abstract: The limitations of the current mobile telecommunication generation (4G) are actually reached. Indeed, the available data rate and the maximum number of users are no longer sufficient. A new generation (5G) is being developed to cope with these needs. It will target different use cases: internet of things, ultra-high data rate communications, and some critical applications such as autonomous vehicles or remote surgery. The needs are much higher than the existing network capabilities. So, innovative solutions have been proposed.

In this thesis work, a new sizing methodology has been developed for 5G systems. It is applied to the Ka-band high data rate communication use case. Several architectures which meet the defined specifications have been studied. A methodology has been implemented to compare their performances depending to their power consumption. Thus, the best suitable architecture for the targeted use case is chosen. Finally, an advanced CMOS technology has been chosen and characterized in order to realize parts of the 5G system. RF switches and low noise amplifier operating in the Ka-band have been designed.

Key words: 5G, low-noise amplifier, Ka-band, CMOS, RF switch, beamforming, sizing methodology, phased-array transceiver.

Unité de recherche

University of Bordeaux, Bordeaux INP, UMR CNRS 5218, IMS Research Center, Talence, France