N° d'ordre : 4284

THÈSE

PRÉSENTÉE A

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGENIEUR

Par Raffaele Roberto SEVERINO

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : ELECTRONIQUE

Design Methodology for Millimeter Wave Integrated Circuits: Application to SiGe BiCMOS LNAs

Soutenue le : 24 juin 2011

Après avis de :

MM.	D. ZITO Professeur, University College Cork, Tyndall National Institute	**Rapporteur**
	S. MIRABBASI Professeur, University of British Columbia, Vancouver	**Rapporteur**

Devant la commission d'examen formée de :

MM.	D. ZITO Professeur, University College Cork, Tyndall National Institute	**Rapporteur**
	S. MIRABBASI Professeur, University of British Columbia, Vancouver	**Rapporteur**
	D. BELOT Ingénieur, STMicroelectronics, Minatec, Grenoble	**Examinateur**
	D. DUBUC Maitre de Conférences, LAAS, Toulouse	**Examinateur**
	Y. DEVAL Professeur, ENSEIRB, Bordeaux	**Président**
	E. KERHERVE Professeur, ENSEIRB, Bordeaux	**Examinateur**
	J-B. BEGUERET Professeur, Université de Bordeaux 1	**Directeur de thèse**
	T. TARIS Maître de Conférences, Université de Bordeaux 1	**Co-Directeur de thèse**

Table of Contents

Introduction

The development of microelectronic industry involves an incessant search for new technological applications capable to intensify the economic exploitation of existing market segments and to establish new business areas.

For this reason, the interest towards millimeter waves has rapidly increased during the last few years, leading to the development of a large number of potential applications in the millimeter wave band, such as WPANs and high data rate wireless communications at 60GHz, short and long range radar at 77-79GHz, and imaging systems at 94GHz.

Historically, the millimeter-waves market segment has been dominated since decades by the III-V semiconductors, thus relegating these applications into some niche market, because of the high manufacturing cost and reduced integration scale of the III-V solutions. Recently, improvements in silicon-based technologies have led to a significant extension of the upper frequency limit of silicon integrated devices. The very high frequency performances of silicon active devices (bipolar and CMOS) have dramatically increased, featuring both f_T and fmax close to or even higher than 200GHz. As a consequence, modern silicon technologies can now address the demand of low-cost and high-volume production of systems and circuits operating within the millimeter wave range.

However, even if the latest CMOS and BiCMOS technology nodes can support millimeter wave operation, nevertheless the augmentation of the operative frequencies and bandwidths entails considerable difficulties. The implementation of active and passive devices becomes more complex because the frequency augmentation unavoidably implies the degradation of device performances. In addition, the equivalent electromagnetic wavelength on silicon substrates at millimeter wave is comparable with the physical dimension of integrated circuits. Therefore, the influence of parasitics on the performance of intrinsic devices cannot be neglected, and a special effort is demanded in developing devices' models. Consequently, millimeter wave design still requires special techniques and methodologies to overcome a large number of constraints which appear along with the augmentation of the operative frequency.

The aim of this thesis is to define a design methodology for integrated circuits operating at millimeter wave and to provide an experimental validation of the methodology, as exhaustive as possible, focusing on the design of low-noise amplifiers (LNAs) as a case study.

The first chapter is dedicated to a brief description of the most important millimeter wave applications and standards. Also, several state-of-the-art architectures of millimeter wave building blocks and transceivers are briefly reviewed.

Chapter 2 reports a detailed study of all the key elements needed for successful design of millimeter wave integrated circuits based on the BiCMOS9MW technology by STMicroelectronics. Starting from a detailed insight on the design facilities provided by the technology, a complete design methodology is presented. Both the lumped and distributed design approaches are considered.

In chapter 3, the proposed methodology is applied to the practical case of LNAs design. Examples concerning the most common application standards are presented, at 60GHz, 80GHz and 94GHz. In particular, the design and the experimental characterization of a two-stage cascode 60GHz LNA dedicated to WPAN applications, of a two-stage LNA operating at 80GHz, and of a single-stage 94GHz amplifier for imaging applications are reported. Then a comparative analysis of lumped and

distributed approach is made, based on two different implementations of a single-stage 80GHz LNA using, respectively, distributed transmission lines and lumped inductors.

Finally, in the last part of the chapter, the design of an 80GHz front-end and the co-integration of an LNA with an integrated antenna are also considered, opening the way to the implementation a fully integrated receiver.

Chapter 1

Millimeter Wave Integrated Circuits: Applications and Architectures

1.0 Introduction

The expression *millimeter waves (mm-Waves or mmW)* designates a radio frequency band having a wavelength of ten to one millimeter.

According to the International Telecommunications Union (ITU) radio regulations resumed in Table 1, it corresponds to the region of extremely high frequency (EHF: 30GHz to 300GHz).

Table 1. IUT Radio Bands.

Nr	Symbols	Designation	Frequency Range	Wavelength Range
1	ELF	extremely low frequency	3Hz to 30Hz	100000km to 10000 km
2	SLF	super low frequency	30Hz to 300Hz	10000km to 1000km
3	ULF	ultralow frequency	300Hz to 3000Hz	1000km to 100km
4	VLF	very low frequency	3kHz to 30kHz	100km to 10km
5	LF	low frequency	30kHz to 300kHz	10km to 1km
6	MF	medium frequency	300kHz to 3000kHz	1km to 100m
7	HF	high frequency	3MHz to 30MHz	100m to 10m
8	VHF	very high frequency	30MHz to 300MHz	10m to 1m
9	UHF	ultrahigh frequency	300MHz to 3000MHz	1m to 10cm
10	SHF	super high frequency	3GHz to 30GHz	10cm to 1cm
11	EHF	extremely high frequency	30GHz to 300GHz	1cm to 1mm

In the context of microelectronics it is more useful to consider the wavelength in a silicon oxide layer that is roughly half the wavelength in open space, according to the well-known relation:

$$\lambda_{Ox} = \frac{1}{f\sqrt{\varepsilon_{Ox}\varepsilon_0\mu_0}} = \frac{1}{f\sqrt{\varepsilon_{Ox}}\sqrt{\varepsilon_0\mu_0}} = \frac{c}{f\sqrt{\varepsilon_{Ox}}} = \frac{\lambda_0}{\sqrt{\varepsilon_{Ox}}} \cong \frac{\lambda_0}{1.9}$$

where the relative permittivity of silicon oxide ε_{Ox} is assumed to be equal to 3.6 and frequency-independent.

A more appropriate placement of the millimeter wave band is therefore in the 15-158 GHz range. In particular, we will consider as millimeter waves the electromagnetic spectrum between 60 and 110GHz.

As a matter of fact, electronics at lower frequencies (at least up to 20-30GHz) can rely over the traditional radiofrequency design methods and techniques. On the contrary, as it will be explained in the next chapter, design of silicon-based integrated circuits operating in the 60-110GHz range, requires the development of a dedicated methodology capable to address the numerous difficulties emerging along with the frequency increase.

Very recently the same design methodology has been gradually extended to higher frequency known as sub-millimeter wave or Terahertz radiation, opening the way to the so-called Terahertz-electronics that however will not be considered in this thesis.

During the last few years, the interest towards millimeter waves has rapidly grown, leading to the development of a large number of potential applications. As well, the recent improvements in silicon-based technologies like CMOS and BiCMOS processes have made it possible to realize low-cost implementations of microelectronic systems operating in the millimeter wave band.

Without any pretention of being exhaustive, the next three sections give an overview of the most common applications in a millimeter wave technology such as WPANs and high data rate wireless communications at 60GHz, short and long range radar at 77-79GHz, and imaging systems at 94GHz. Then, in section 1.4, the main features and the most common parameters concerning the radio-frequency building blocks are briefly reviewed. Finally, some examples of integrated transceivers taken from the present-day state-of-the-art are briefly illustrated in section 1.5.

1.1 Millimeter wave Applications at 60GHz

1.1.1 Overview

The 60GHz band [1]-[2] has attracted the interest of researchers in the past several years, as it offers ample and license-free bandwidth. This advantage becomes more evident when considering that the availability of about 7GHz of unlicensed bandwidth is guaranteed in many countries worldwide, as shown in Fig. 1. For example, the range from 57 to 64 GHz is currently available in the US, while 59 to 66 GHz are available in Japan. Furthermore 60GHz band is less restrict in terms of power limits when compared to other concurrent wideband systems, like ultra-wideband (UWB), for instance. These characteristics make 60GHz technology particularly suited for gigabit wireless applications that currently are technically constrained at lower frequency.

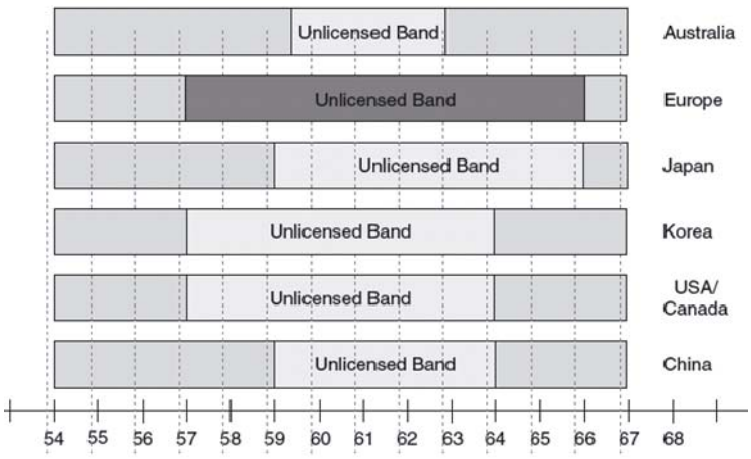


Fig. 1. Worldwide frequency allocation for 60 GHz band and operation.

As depicted in Fig. 2, the atmospheric absorption presents a peak at 60GHz due to oxygen absorption that puts a severe limit on the distance that 60GHz transmissions can handle – at least in the case of low-power applications – limiting it to close-range communications. In home-space, however, oxygen absorption is not a big problem, so 60GHz is beginning to look like a useful way to provide dramatic speed-ups over traditional wireless local area networks (WLANs) and is particularly indicated for

wireless personal area networks (WPANs). A WPAN is designed to provide short-range (<10m), very-high-speed (>2Gb/s) multi-media data services to computer terminals and consumer appliances located in rooms, office space, "hot spots" and kiosks. Wireless PANs will provide higher data rates, and shorter range, than comparable WLANs but sufficient to cover the size of most offices, medium-size conference rooms, and rooms in personal home. Various electronic devices could be interconnected, including laptops, cameras, and monitors. Potential applications include wireless display, wireless docking station, and wireless streaming of data from one device to the other. A 60GHz link could be used to replace various cables used today in the office or home, including gigabit Ethernet (1Gbps), USB 2.0 (480Mbps), or IEEE 1394 (~800Mbps). Currently, the data rates of these connections have precluded wireless links, since they require so much bandwidth. While other wireless standards are evolving to address this market (802.11n and UWB), 60GHz is still a viable candidate.

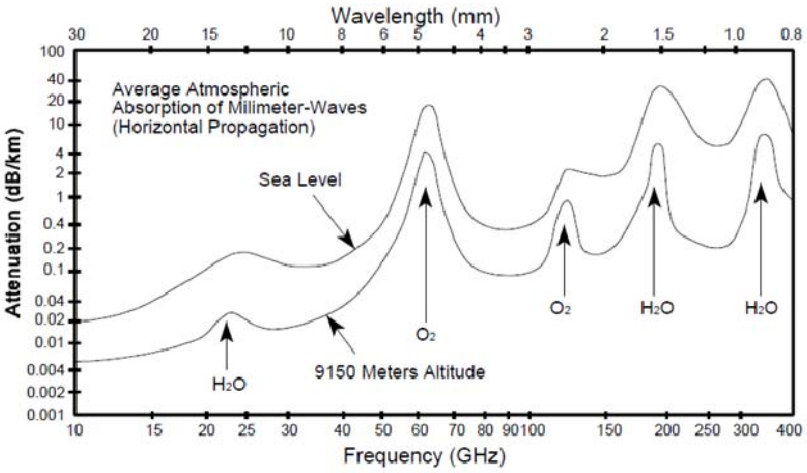


Fig. 2. Atmospheric absorption of millimeter waves.

Besides oxygen absorption in open-space, different materials affect the propagation of 60GHz signals in different ways and, whereas plasterboard or drywall absorption is not much greater than in the case of 2.4GHz signals, other materials, on the contrary, cause a big fall-off in signal strength. Also human skin, for instance, absorbs 60GHz radiation pretty efficiently and crossing a link between a computer and a media player synchronized over a 60GHz system would break it.

To reduce the effect of absorption loss, many solutions such as directional antennas or beamforming techniques are currently investigated. If sufficient directionality is guaranteed in transmission, then atmospheric and material properties could also prove more of a benefit than a limitation, since they can prevent the signal from leaking into adjacent environment, reducing the risk of interference and enhancing security.

1.1.2 Standards and Regulations for 60GHz WPAN applications

Several attempts of standardization have been recently proposed or are still in course of definition. For example, the IEEE 802.15 Task Group 3c is working since 2005 on the wireless personal area network (WPAN), leading to the definition of the IEEE 802.15.3c standard in September 2009 [3]-[4].

Besides, Wireless HD [5], the Wireless Gigabit Alliance (WiGig) [6], the IEEE 802.11ad working group [7] and the ECMA-387 [8] have been also making standardization efforts on the 60 GHz frequency band. A detailed description of these documents is beyond the interest of this text; however the most important features of the IEEE and the ECMA standards are briefly described hereafter.

IEEE Std 802.15.3c is an amendment to IEEE Std 802.15.3 that defines an alternative physical layer (PHY) operating in the millimeter wave band and the necessary modifications to the medium access control (MAC) changes to support this PHY.

The PHY defines three operational modes as follows:

- Single carrier (SC) mode optimized for low power and low complexity.
- High-speed interface (HSI) mode optimized for low-latency bidirectional data transfer.
- Audio/video (AV) mode optimized for the delivery of uncompressed, high-definition video and audio.

For devices that implement the mm-Wave PHY, at least one of the three PHY modes is required. In addition, to promote coexistence and interoperability, a common mode signaling (CMS) is defined based on a low data rate SC PHY mode.

The single carrier mode in mm-Wave PHY (SC PHY) provides three classes of modulation and coding schemes targeting different wireless connectivity applications. As summarized in Table 2, Class 1 is specified to address the low-power low-cost mobile market while maintaining a relatively high data rate of up to 1.5Gb/s; Class 2 is specified to achieve data rates up to 3Gb/s; Class 3 is specified to support high performance applications with data rates in excess of 3Gb/s.

Table 2. Modulation and Coding Schemes Classes in the SC PHY.

Class	Categorization
1	Data rates < 1.5Gb/s
2	1.5Gb/s < Data rates < 3Gb/s
3	Data rates > 3Gb/s

The SC PHY supports a wide range of modulations, π/2 BPSK, π/2 QPSK, π/2 8-PSK, π/2 16-QAM, pre-coded MSK, pre-coded GMSK, on-off keying (OOK), and dual alternate mark inversion (DAMI). It operates on four RF channels, as defined in Table 3. A compliant implementation shall support at least 1 channel from the channels allocated for operation by its corresponding regulatory body.

Table 3. RF Channels for millimeter wave PHY.

Channel ID	Start Freq [GHz]	Center Freq [GHz]	Stop Freq [GHz]
1	57.24	58.32	59.40
2	59.40	60.48	61.56
3	61.56	62.64	63.72
4	63.72	64.80	65.88

The high speed interface mode of mm-Wave PHY (HSI PHY) is designed for devices with low-latency, bidirectional high-speed data and uses orthogonal frequency division multiplexing (OFDM). HSI PHY supports a variety of modulation and coding schemes (MCSs) using different frequency-domain spreading factors, modulations, and LDPC block codes.

The Audio/Visual (AV) PHY is implemented with two PHY modes, the high-rate PHY (HRP) and low-rate PHY (LRP), both of which use orthogonal frequency domain multiplexing (OFDM). The data rates supported by the HRP and the LRP are defined in Table 4 and Table 5, respectively.

Table 4. AV HRP data rates.

HRP mode index	Modulation	Data rate [Gb/s]
0	QPSK	0.952
1	QPSK	1.904
2	16-QAM	3.807
3	QPSK	1.904
4	16-QAM	3.807
5	QPSK	0.952
6	QPSK	1.904

Table 5. AV LRP data rates.

LRP mode index	Modulation	Data rate [Mb/s]
0	BPSK	2.5
1		3.8
2		5.1
3		10.2

Different PHYs are a result of demands of different market segments. For example, one usage model is for kiosk applications. This usage model requires 1.5Gb/s at a 1m range. The SC-PHY can provide such a data rate at that short range with less complexity thus lower cost than an OFDM PHY. Another usage model is required by the streaming of uncompressed video. Due to the nature of uncompressed video signals, a special PHY, the AV PHY, was selected to provide high throughput. A third usage model involves an ad-hoc system to connect computers and devices around a conference table. In this usage model, all of the devices in the WPAN will have bidirectional, NLOS high speed, low-latency communication, which is provided for by the HSI PHY.

Besides the suggestions for MAC and physical level, IEEE standard specifies also an optional beam forming protocol that can support a multitude of antenna configurations such as single antenna element, sectored antennas, switched antennas, and one-dimensional (1-D) and two-dimensional (2-D) beam forming antenna arrays.

The ECMA 387 standard defines three device types as follows:

- A type A device offers video streaming and WPAN applications in 10m range line of sight/not line of sight multipath environments. It uses high gain trainable antennas. This device type is considered as the 'high end' - high performance device.
- A second type, type B device offers video and data applications over shorter range (1-3 m) point to point line of sight links with non-trainable antennas. It is considered as the 'economy' device and trades off range and not line of sight performance in favor of low cost implementation and low power consumption.
- The third type, type C device is positioned to support data only applications over point to point line of sight links at less than 1m range with non-trainable antennas and no QoS guaranties. This type is considered as 'bottom end' device providing simpler implementation, lowest cost and lowest power consumption.

The A, B, and C devices can interoperate with their own types independently and can coexist and interoperate with the other types leading to the implementation of heterogeneous network solution that provides interoperability between all device types. As depicted in Fig. 3, each one of the three device types corresponds to a different organization of the PHY layer. The three PHYs converge into a single MAC level based on the ECMA-368 standard, with the necessary changes to support directional communication in 60GHz band.

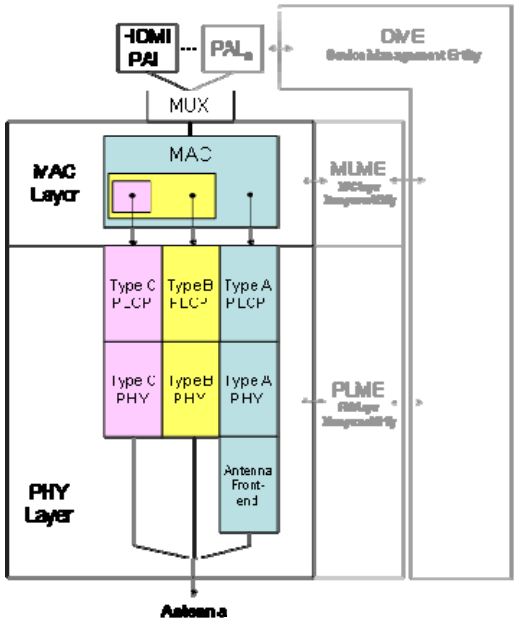


Fig. 3. Protocol Structure of ECMA standard.

The Type A PHY includes two general transmission schemes, namely Single Carrier Block Transmission (SCBT), also known as Single Carrier with Cyclic Prefix, and OFDM; the Type B PHY has been designed using a simplified single carrier transmission scheme with a common beaconing mode based on differentially encoded BPSK modulation (DBPSK), thus allowing for both simple coherent and non-coherent demodulation and minimizing the implementation overhead to support interoperability with type A devices; the Type C PHY uses the simplest single carrier transmission scheme based on the Amplitude-Shift-Keying (ASK) modulation scheme.

The standard specifies four frequency channels each with a symbol rate of 1.728Giga-Symbols/second and with a separation of 2.160GHz, as suggested in Fig. 4. All device types follow the same frequency plan.

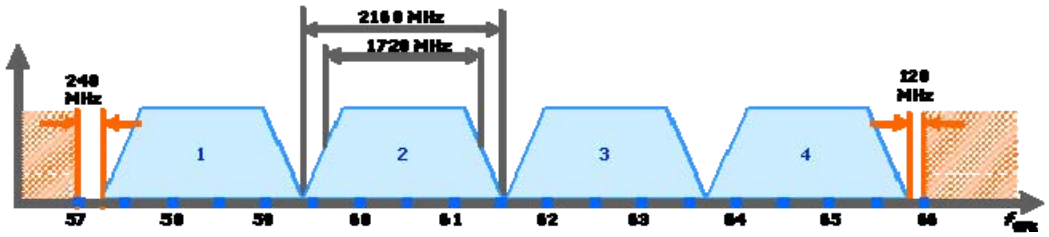


Fig. 4. Frequency plan of ECMA standard.

European regulation of WLAN and WPAN system operating in the 60GHz band is provided by the European Telecommunications Standards Institute (ETSI). The technical characteristics of these applications are described in [9] and the specific technical requirements are defined in [10]. The most important features can be summarized as follows:

- frequency band is defined from 57 to 66 GHz;
- the maximum spectral power density shall be limited to 13dBm/MHz by indoor usage and to -2dBm/MHz by indoor/outdoor;
- the maximum output power level (effective isotropic radiated power, EIRP) shall be limited to 40dBm by indoor usage and to 25dBm by indoor/outdoor;
- the limit level of unwanted emissions in the spurious domain for transmitter are reported in Table 6 (measured with 0dBi antenna gain);
- the limit level of unwanted emissions in the spurious domain for receiver are reported in Table 7 (measured with 0dBi antenna gain);
- the use of an integral antenna (that is an antenna designed as a fixed part of the equipment, without the use of an external connector and that, therefore, cannot be disconnected from the equipment by a user with the intent to connect another antenna) is required, to provide interference protection;
- a medium access protocol shall be implemented by the equipment and shall be active under all circumstances, in order to facilitate spectrum sharing with other devices in the wireless network.

Table 6. Transmitter spurious emissions.

Frequency band	Measurement bandwidth	Field Strenght at 3 m (dBµV/m)
30MHz to 1GHz	100kHz	59
1GHz to 132 GHz	1MHz	65

Table 7. Receiver spurious emissions.

Frequency band	Measurement bandwidth	Field Strenght at 3 m (dBµV/m)
30MHz to 1GHz	100kHz	38
1GHz to 132 GHz	1MHz	48

1.2 Millimeter wave Automotive Radar sensors

1.2.1 Overview

The first experimental applications of radar in the automotive industry date back to the 1950s and yet in 1970s automotive radars were the object of systematic investigations [11]-[12].

Years after years, safety has been one of the main concerns in the development of car industry, leading to several structural improvements capable to reduce the consequences of accidents on the driver and the passengers. In such a context, the introduction of radars represents a disruptive event, focusing on accident avoidance and prevention that, undeniably, are preferred to any system of crash protection.

Nowadays, the so-called autonomous or adaptive cruise control (ACC) is probably the most common application of automotive radar and is used to assist the driver and to augment its comfort.

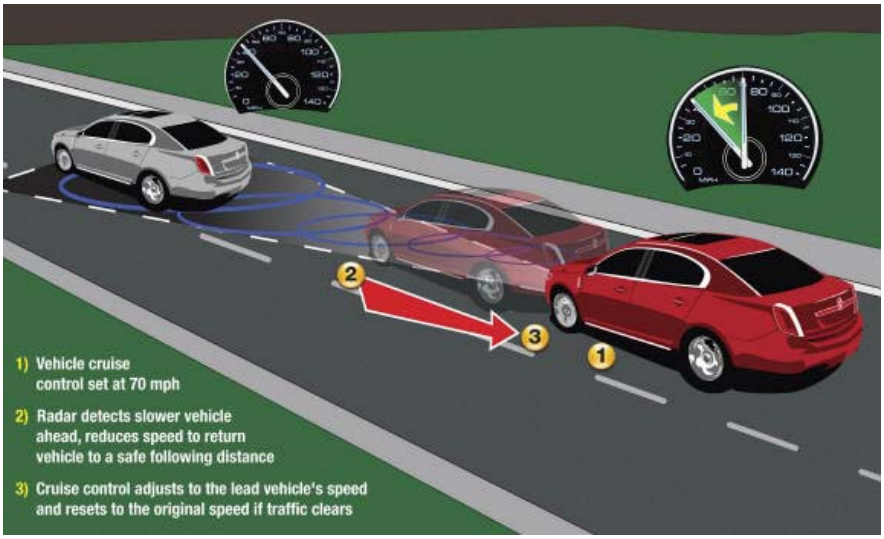


Fig. 5. Schematic example of Adaptive Cruise Control (ACC) operation.

Practically speaking, a radar sensor is mounted behind the front bumper of a vehicle at a height of less than 1m, where it is able to interrogate the road ahead and the adjacent traffic lanes forward of the vehicle's location. Using this radar, the control system within the vehicle adjusts the cruise speed in response to a slower vehicle in a merging lane, or when following a vehicle in the same lane, in order to maintain the driver's selected minimum separation distance behind the other vehicle. An ACC system can constantly provide the driver with information about traffic situation in the environment, making driving less strenuous, especially in flowing traffic and in critical context like on motorways or dual carriageways.

ACC today uses long-range radar (LRR) operating between 76GHz and 77GHz with a maximum bandwidth of 1GHz. It uses distance scanning, which requires an operating range of approximately 150m and is used at vehicle velocities not below 30km/h. One or multiple narrow lobes control or

scan the driving path in front of the car to determine the distance to the vehicle driving ahead for maintaining a constant minimum safety distance.

In 1999 Mercedes-Benz has been the first car manufacturer who introduced radar-based ACC system in its S-class vehicles [13]. Since then, radar based ACC systems are available in many high and mid-class models such as BMW 7 series, Jaguar (XKR, XK6), Cadillac (STS, XLR), Audi A8, VW Phaeton, Mercedes E, CL, CLK, SL class, BMW 5 and 6 series, Audi A6, Nissan (Cima, Primera), Toyota (Harrier, Celsior), Lexus (LS, GS), Honda (Accord, Inspire, Odyssey) and their functionality have been gradually extended, including also pre-crash sensing and collision mitigation.

Besides radar-based equipments, also competing and complementing technologies in vehicular surround sensing and surveillance like Lidar, ultrasonics, and video-cameras have been tested since the first 1990s to implement parking aid, collision warning, and ACC as well. However, radar seems still to be the key technology for driver assistance and safety applications, due to its inherent advantages like weather independence and direct acquisition of range and velocity especially when compared to alternative sensors like video, laser, and ultrasonics.

Fig. 6. DISTRONIC radar sensor mounted on Mercedes Benz S-class vehicles.

More recently, in 2003 – whereas European car manufacturers offer radar systems only for ACC systems so far – their Japanese competitors Honda and Toyota introduced an active brake assist for collision mitigation (in addition to ACC) based on 77GHz long range radar (LRR) technology. In contrast to the only smooth deceleration capability of an ACC system (because ACC is only marketed as a comfort feature), the active brake assist provides much higher braking forces for deceleration, when a threatening situation is identified and the driver starts braking but maybe not as strong as it would be necessary to avoid a crash. This shows the trend from “comfort only” functions to active safety systems with radar sensing technologies that serve both the comfort and the safety domain.

As a matter of fact, car companies and suppliers are currently working on the development of the next generations of LLR at 77GHz that will show improvements with respect to maximum and minimum range, wider field of view, improved range and angular resolution and accuracy, self alignment and blockage detection capability. In addition to forward-looking LRR, ultra-wideband (UWB) short range radar (SRR) sensors with coverage up to 30m are under development for a variety of further applications that will result in significant improvements of road safety.

Short range radar and ultra-wideband sensors can enable a variety of applications such as:

- Support of ACC with Stop&Go functionality
- Collision warning
- Collision mitigation
- Blind spot monitoring
- Parking aid
- Lane change assistant
- Rear crash collision warning

LRR and SRR devices can be used to combine their functionalities creating a protection wall all around the vehicle that is referred to in the literature as a "safety belt" for cars.

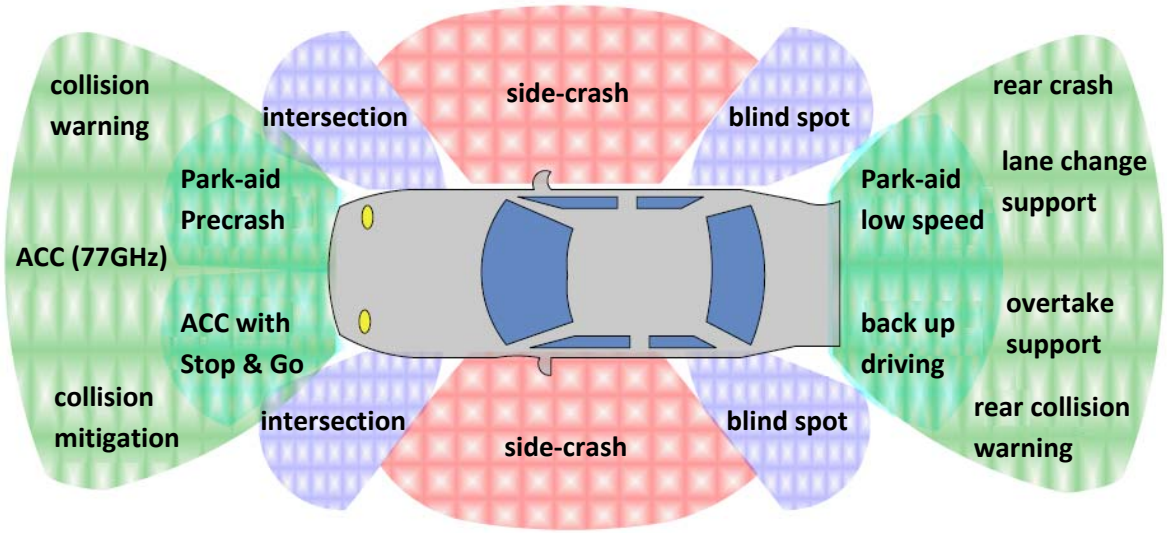


Fig. 7. Possible applications for automotive radars.

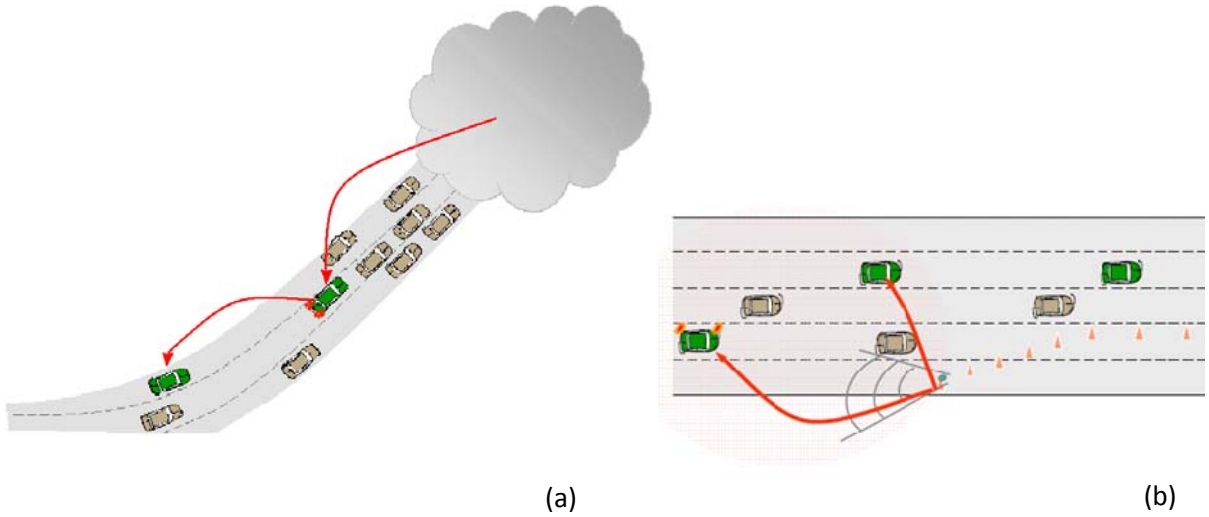
Beside LRR and SRR systems, a complementary branch of potential applications of millimeter wave technology to the automotive industry is given by the next-generation road vehicle communication. Many companies are currently working on the implementation of millimeter wave sensors capable to support inter-vehicle (IVC) and/or vehicle to roadside (and roadside to vehicle) communication (RVC). Next-generation cars will be able to exchange data – concerning, for example, their relative position or information and warning about weather and traffic status – with the surrounding vehicles and with beacons placed on the roadside at regular intervals along all inter-urban trunk routes and at strategic location (e.g. junctions) on more minor roads and in urban areas.

A huge number of applications of inter-vehicle and roadside to vehicle communication can be realized and are investigated by various project and groups; some examples are:

- Traffic information
- Collision avoidance
- Work zone safety warning
- Vehicle and cargo tracking
- Electronic license plate

- Repair-service record
- Parking, fuel, or fast food payment
- Disaster and emergency warning and control

More in general, IVC and RVC systems can be considered as part of a wider context, concerning the so-called Intelligent Transportation Systems (ITS) which represent the application of information and communications technology to transport infrastructure and vehicles, in an effort to manage factors that typically are at odds with each other, such as vehicles, loads, and routes to improve safety and reduce vehicle wear, transportation times, and fuel consumption.



(a) (b)

Fig. 8. An example of IVC: the alarm for an accident is transmitted from a vehicle to another (a). An example of RVC: a warning revealing the presence of a work zone is transmitted to transiting vehicles (b).

An exhaustive description of all the potential applications of millimeter wave to the automotive industry is beyond the aim of this paragraph. Nevertheless, in conclusion of this brief overview, it is important to remark that the applications currently available, as the examples discussed here above, are substantially based on compound-semiconductors technologies. The considerably elevated cost of these technologies has therefore prevented the large diffusion of automotive radar equipments so far, relegating them to the high and mid-class market segments. Moreover, in the particular case of inter-vehicle communication, implementation cost is a concern, as it prevents an adequate diffusion of inter-communicating devices. Nowadays, thanks to the recent improvements in silicon-based technologies it is possible to realize low-cost and high-volume production of systems and circuits for automotive radar operating at millimeter wave frequencies.

1.2.2 Automotive Radar Regulation

The regulation activity for automotive radar started already in the 90s with the allocation of the 76-77 GHz band for LRR systems. Then in 2001, the Federal Communications Commission (FCC) started the regulation for 24GHz SRR sensors in the US, setting their operation in the range of 22-29 GHz with a maximum mean power density of -41.3dBm/MHz, using directional antennas on terrestrial

transportation vehicles and requiring that the center frequency of the emission and the frequency at which the highest radiated emission occurs are higher than 24.075GHz.

Almost at the same time, more than 30 mainly European car manufacturers and suppliers founded the Short range Automotive Radar frequency Allocation consortium (SARA) with the objective to support UWB regulation in the 24GHz band for the European market. The use of the 21.65-26.65 GHz range was adopted by the European Commission for UWB SRR for the short-term period from 2005 to 2022. As a long term solution, in order to avoid interference with other 24GHz communications system, the 77 – 81 GHz band was allocated and made available since 2005.

The current European regulation for automotive radar is reported in the European Telecommunications Standards Institute (ETSI) standard documentation [14]-[17]. The functionality and the technical requirements, as well as the relative measurement conditions, of 76-77 GHz radar are set in [14] and can be resumed as:

- the permitted range of operating frequencies for intentional emissions shall be from 76GHz to 77GHz;
- the transmitted power for equipment with fixed beam antennas measured as mean power (EIRP) shall be less than 50dBm for system others than pulsed Doppler radar (Class 1) and less than 23.5dBm for pulsed Doppler radar only (Class 2);
- the transmitted power for equipment with fixed beam antennas measured as peak power (EIRP) shall be less than 55dBm;
- the transmitted power for equipment with steerable antennas shall be less than the limits shown in Table 8;
- the mean power density radiated outside the 76GHz to 77GHz band shall not exceed 0dVm/MHz;
- for the transmitter, the effective radiated power density of any radiated spurious emission shall not exceed the values given in Table 9 and shall be in accordance to CEPT/ERC/Recommendation 74-01;
- for the receiver, the maximum equivalent isotropically radiated power (max. EIRP) of any spurious emission outside the permitted range of frequencies shall not exceed 2nW (≈ -57dBm) in the frequency range 25MHz ≤ f ≤ 1GHz and shall not exceed 20nW (≈ -47dBm) on frequencies in the range 1GHz < f ≤ 73,5GHz and 79,5GHz < f ≤ 100GHz in accordance to CEPT/ERC/Recommendation 74-01.

Table 8. Limits for transmitted power (steerable antenna). t is the largest dwell time[1] at any angle. D is the ratio of the area of the beam to the total area scanned by the antenna; the power is averaged across one antenna cycle; as D is smaller than 1, the log(D) is negative and leads to a reduction of the 55dBm value.

	Class 1		Class 2	
Maximum antenna signal dwell time[1]	t < 100ms	t > 100ms	t < 100ms	t > 100ms
Mean Power (EIRP)	[55dBm + 10 log(D)] or 50dBm (whichever is the smaller)	50dBm	[55dBm + 10 log(D)] or 23.5dBm (whichever is the smaller)	23.5dBm
Peak Power (EIRP)	55dBm	55dBm	55dBm	55dBm

Table 9. Limits of radiated spurious emissions.

Frequency range [MHz]	Limit values for spurious radiation
47 to 74	-54dBm/100kHz
87.5 to 118	-54dBm/100kHz
174 to 230	-54dBm/100kHz
470 to 862	-54dBm/100kHz
otherwise in band 30 to 1000	-36dBm/100kHz
1000 to 25000	-30dBm/MHz
25000 to 40000	-30dBm/MHz
40000 to 100000 (not applicable when the permitted range of frequencies from 76GHz to 77GHz and in the out of band region)	-30dBm/MHz

In a similar way [15] fixes the technical requirements of SRR receiver, transmitter and integrated transceiver operating in the 22-26.65 GHz band, whereas the functionalities and the technical characteristics of SRR devices operating in the 77-81 GHz frequency range are discussed in detail in [16]. Even if an exhaustive report of these data is obviously beyond the aim of this text, it can be interesting to underline the main features of the 79GHz SRR devices as set by [16] that can be resumed as follows:

- SRR units operate in the 77-81 GHz band;
- operating range can vary from 0 to 30m approximately;
- a precise speed measurement is allowed, with help of a continuous wave (CW) Doppler emission and this speed measurement mode is combined with wideband signals to provide also precise radial range information of objects with a high range separation in order of approximately 5cm to 10cm;
- to obtain the required resolution, the SRR needs a large bandwidth of 4GHz for an accurate measurement of the range;
- several wideband modulation technologies are suggested (eventually combined in hybrid concepts):

[1] The dwell time is the accumulated amount of transmission time of uninterrupted continuous transmission within a single given frequency channel and within one channel repetition interval.

Pulse Modulation

Frequency Hopping Spread Spectrum (FHSS)

Binary Phase Shift Keying (BPSK)

Pulsed Frequency Modulated Continuous Wave (FMCW)

- the peak power limit for the SRR allocation from 77GHz to 81GHz is required to be from 46,2 up to 55 dBm to achieve road safety goals (power levels required for SRR operation are listed in table 10).

Table 10. Overview of SRR and LRR parameters defined by ETSI [16].

	SRR	LRR	
		FMCW	Pulsed
Frequency	77-81 GHz	76-77 GHz	
Worst Case Mean EIRP Spectral Density @ 79GHz	< -15dBm/MHz to -3dBm/MHz		
Mean Power	18-30 dBm	50dBm	23.5dBm
Worst Case Peak EIRP @ 79GHz	46.2-55 dBm	55dBm	
Operating Distance	30m	150m	

Beside LRR and SRR in the 79GHz band, ETSI provides also a preliminary regulation of ITS equipments for IVC and RVC, setting their operational bandwidth in the 63-64 GHz range [17].

1.2.3 VeLo Project

The European Commission has adopted a policy goal to reduce road fatalities in Europe by 50% by 2010. The objective and focus of "The EC Approach to Road Safety and Intelligent Transport Systems (ITS)" are defined as to "Improve Safety, Security, Comfort and Efficiency in all Transport modes" and "Focusing on Advanced Pilot/Driver Assistance Systems".

European industry responded to these goals and, in order to promote the development of the necessary silicon technology, a first joint research project on "Automotive high frequency electronics - KOKON" was started in September 2004, funded by the German Ministry of Education and Research (BMBF), putting together two semiconductor companies (Atmel and Infineon), two automotive radar sensor manufacturers (Bosch and ContiTemic), and one automotive manufacturer (DaimlerChrysler), supported by institutes and universities [18]. The KOKON project investigated the feasibility of 79GHz UWB SRR and of 76.5GHz advanced LRR on silicon-germanium heterojunction bipolar transistor (SiGe HBT) technology that was identified as the best-suited for low cost solutions in the automotive radar context.

In 2006, the French National Research Agency (ANR) launched a new research project focused on millimeter wave automotive radar known as "Communication Inter Véhicules et Localisation Relative Précise en gamme millimétrique à 79GHz - VeLo".

Several French laboratories (CEA-Leti, IEMN, IMS, Lab-STICC, and LAAS) are involved in the project as well as the semiconductor company ST Microelectronics, the automotive manufacturer Autocruise, and the French National Institute for the Research on Transports and Transport Security (INRETS).

The aim of the project is the development of a radar sensor completely integrated on a silicon-based technology capable to support millimeter wave operation.

According to the definition given in the project documentation, the VeLo sensor should operate as an ultra-large band (ULB) SRR in the range of 77-81 GHz, providing a precise speed measurement based on Doppler emission, combined with a precise radial range detection of objects with a resolution between 5cm and 10cm; as a third function, the SRR should provide a lateral detection of a $1m^2$ target at a distance of 30m, with a spatial resolution of 5cm.

At the same time, it should support also inter-vehicle communication, establishing a link with the surrounding vehicles and detecting and constantly monitoring their relative position. A vehicle equipped with a VeLo sensor will therefore analyze the position of other vehicles in the environment considering both other radar-equipped vehicles and passive ones. Furthermore, to complete its perception of the environment, it will estimate its own absolute position on the basis of the localization data collected by surrounding vehicles and supplied by a satellite navigation system.

Generally speaking, the final target of the project is to give a practical demonstration of the feasibility of a millimeter wave integrated device capable to operate simultaneously as short range radar and as inter-vehicle communication system.

Integration of the complete device in a silicon-based technology is a key element since, as already explained in the previous paragraphs, the main trend in automotive applications of millimeter wave technology is to provide low-cost solutions capable of supporting high-volume production.

The proposed technology in the VeLo context is the BiCMOS9MW from ST Microelectronics that provides SiGe HBTs with maximum transistor f_T and fmax higher than 200GHz. Therefore, from the specific point of view of millimeter wave integrated circuits design, the main goal of the project is to demonstrate the feasibility of a complete receiver front-end (LNA, Mixer, and VCO) and of a power amplifier in the BiCMOS9MW technology. A first solution takes also into account the co-integration of the transmitting antenna with the power amplifier and of the receiving antenna with the receiver front-end leading to a complete system on chip (SoC) implementation. An alternative solution requires the implementation of the antennas as distinct modules on a common substrate, thus avoiding the difficulties related to antenna integration at millimeter wave frequencies.

A summary of the preliminary specifications given for the main building blocks is reported in Table 11 [19]. The VeLo project was concluded at the end of 2010 and our work is a part of it. The design of an 80GHz LNA that can respond to the project requirements is described in chapter 3. The main issues concerning the co-integration of the LNA and the Rx antenna, as well as the design of the receiver front-end are also discussed in chapter 3.

Table 11. Preliminary specifications of the VeLo Project (differential architecture) [19].

	Low Noise Amplifier	Down-converter Mixer	Local oscillator	Power Amplifier	Antenna (Rx/Tx)
Freq	79GHz	79GHz	79GHz	79GHz	79GHz
Gain	14.4dB	11.5dB		15dB	13dBi
Noise Figure	6.5dB	15dB DSB			
S11	-10dB	-10dB		-15dB	
S22	-15dB	-15dB		-10dB	
IIP1dB	-29.6dBm	-27.6dBm		15dBm	
IIP3	-20dBm	-18dBm		25dBm	
Output Power		0dBm LO	3dBm	15dBm Min	
Phase Noise			-79dBc/Hz at 50kHz -105dBc/Hz at 1MHz -120dBc/Hz at 5.6MHz		
Directivity					19.6dBi
Efficiency					22%
Insertion Loss					-25dB
LO/RF Isolation		57dB			

1.3 Millimeter wave imaging in the 94GHz band

1.3.1 Overview

The upper portion of the millimeter wave spectrum has been traditionally exploited for applications in the field of astronomy. Recently, further improvements of the related technology have turned the interest of industry on the development of terrestrial applications based on millimeter wave imaging [20]-[22].

As a matter of fact, the property of millimeter wave imaging systems to see through materials that are opaque to more conventional imaging wavebands (visible, IR and UV) makes them remarkably useful for many scientific and industrial applications.

In particular, millimeter waves can readily penetrate common clothing materials and are reflected from the human body and any concealed items. For this reason they can be efficiently used in the field of security and defense, for the detection of concealed weapons or explosives. Furthermore, since millimeter wave imaging uses low-power and not-ionizing radiations, it results to be safer when compared to concurrent technologies based on X-rays and is better suited for the implementation of body scanner used for example for airport security screening. In such a context, it is possible to reveal not only weapons and explosives, but also any object hidden on the body, like drugs and

contraband stuff. The personal scanner uses harmless coherent radar waves from a millimeter wave antenna array to illuminate the person under surveillance. Then the reflected signal from the body or from any object on the body is collected by the array and processed to form high-resolution three-dimensional images like that of Fig. 9.

Fig. 9. An example of body scanner operation based on millimeter waves.

Following a similar approach it is possible to realize many other applications of millimeter wave as ground penetrating radar (GPR) imaging , and wall probing systems like inner-wall imaging, through-wall imaging, through-concrete imaging.

Ground-penetrating radar (GPR) is a geophysical method that uses radar pulses to image the subsurface and detects the reflected signals from subsurface structures. GPR can be used in a variety of media, including rock, soil, ice, fresh water, pavements and structures. It can detect hidden objects, changes in material, and voids and cracks, without drilling, probing, or digging. The applications of GPR imaging cover a large number of fields as engineering, where it is used for non-destructive testing of structures and pavements, for example to map defects such as voids, moisture and cracking or to determine pavement type and thickness in the context of highway and airport runway inspection without traffic interruption; archeology, where it is used to map archeological structures and sites; military, for landmine detection; geophysical, for ground investigation in order to trace foundations and other obstructions, and to locate geological hazards that may pose a risk to construction activity or human habitation; or environmental protection, where GPR can be used to define landfills, contaminant plumes, and remediation sites.

Also wall probing systems can be useful for non-destructive structural test, to map buried utility cables, pipes and ducts or to measure the depth to buried utilities, so making maintenance interventions simpler, reducing costs and the risk of accidental damage to power and gas lines.

Fig. 10. Applications of GPR. Surface mapping of an underground storage tank (left). Roadway inspection (right).

Many of the applications discussed so far are currently realized using concurrent technologies (like UWB in the range from 30 MHz to 12.4 GHz in the case of GPR). Nevertheless, all of them can be potentially supported by millimeter wave systems operating in the frequency range around 94GHz.

The importance of 94GHz band is given by the fact that the atmospheric absorption reaches a minimum level precisely at 94GHz, as depicted in Fig. 2.

Besides imaging, millimeter wave technology can be applied also to the implementation of a cloud profiling radar (CPR) employed in satellite missions to investigate cloud structure and its variability. By using the 94GHz frequency range it is possible to penetrate ice clouds with negligible attenuation and obtain a profile of cloud characteristics.

Moreover, the capability of 94GHz signals to propagate through fog, clouds, rain, and even sandstorms with irrelevant attenuation can be exploited to design millimeter wave sensor capable to improve the aviation safety and facilitate airport ground control in extremely poor visibility.

1.3.2 Direct detectors for Imaging systems

In general, imaging systems can be divided into two categories: active and passive.

Passive (known also as radiometric) imaging is based on the principle that any object not at the absolute zero temperature emits electromagnetic energy at all wavelengths. This energy can be detected by an appropriate receiver and can be used to produce an image. It is the preferred method for outdoor stand-off detection and for systems working at distances of a few meters to several kilometers. Such imagers rely on temperature differences of the target, where the cold sky helps to improve the image contrast. The main benefit is that no high-power source is needed to illuminate the scene. Some problems associated with active illumination, such as poor visibility of metal surfaces due to specular reflections, are also eliminated. As a drawback, due to the lack of phase information in the received signal, the ability to perform radar operation is limited.

Active imaging systems use dedicated millimeter wave sources that project a narrow beam of energy against the target in order to achieve a higher contrast or better signal to noise ratio. It distinguishes

itself from passive imaging, because it often detects the radiation as electric field amplitude and phase in a way similar to radar applications. This enables imaging with three-dimensional and tomographical content and enhances the object detection capabilities.

In the case of active imaging systems, receivers can be implemented as direct detectors or, alternatively, as indirect (coherent) receivers which are commonly based on the heterodyne architecture. Passive imaging, on the contrary, commonly uses direct detectors.

A typical direct detection receiver is made of an antenna, a LNA, a square-law diode detector and dc amplifiers. A block diagram is shown in Fig. 11. The average output voltage Vout (V) is the product of the RF bandwidth Δf, the receiver responsivity $\mathcal{R}$ (V/W) and the equivalent input noise power spectral density (PSD) Pin (W/Hz) which is directly related to the temperature of the observed scene.

$$Vout = \mathcal{R}\,\Delta f\; Pin \tag{1}$$

The usual detector figure of merit is the noise equivalent power (NEP), defined as the measured output noise voltage spectral density divided by the responsivity [22].

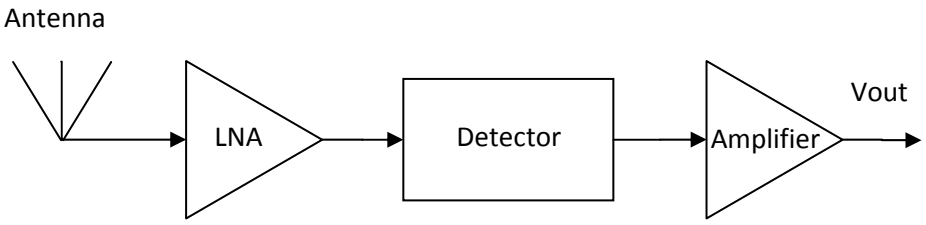


Fig. 11. Block diagram of the direct detection receiver

Present LNA-based passive imaging detectors are mainly implemented using GaAs or InP devices. Nevertheless, due to the recent advances in silicon and SiGe technologies, low-cost systems with high temperature resolution can be designed. Recently, a first 94GHz passive imaging direct detector with state-of-the-art NEP implemented on a SiGe BiCMOS process has been reported [23]. It consists of a five-stage common-emitter LNA and a square-law detector implemented using a common-emitter BJT. The chip occupies less then 0.26mm^2, consumes 34.8mW and has ~ 4000kV/W peak responsivity, a 3-dB bandwidth of 84-103 GHz, and NEP of 0.014 pW/Hz$^{1/2}$ at 94 GHz.

A first example of 90GHz-band passive imaging receiver implemented in a standard 65nm CMOS process is given in [24]. It consists of a 5-stage cascode LNA and a square-law power detector implemented in differential form. A transformer converts the LNA single-ended output into a differential signal that drives the detector. The peak gain of the LNA, measured as stand-alone, is over 27dB at 88GHz and coincides with its minimum NF of 6.8dB. The responsivity of the receiver peaks at over 220kV/W and the best NEP, measured at an IF frequency of 400 kHz, is below 0.1pW/Hz$^{1/2}$, both at 86GHz. Chip dimensions are 0.665 × 0.47 mm^2.

1.4 Building blocks and design solutions for the implementation of millimeter wave systems

1.4.1 Possible architectures for integrated transceivers: fundamental concepts

The common trend in the implementation of millimeter wave circuits, as it results from literature survey, does not consist in the development of ad-hoc circuital architectures, but relies on the adoption of common solutions traditionally used in the radio-frequency domain, extending their validity at higher frequencies.

Generally speaking, a standard choice for the implementation of radio-frequency front-ends is the super-heterodyne architecture, shown in the block diagram of Fig. 12. This architecture can be applied, with minor modifications, to several RF receivers and transmitters, as well as to integrated transceivers, regardless of their specific operation, their operative frequencies or targeted applications.

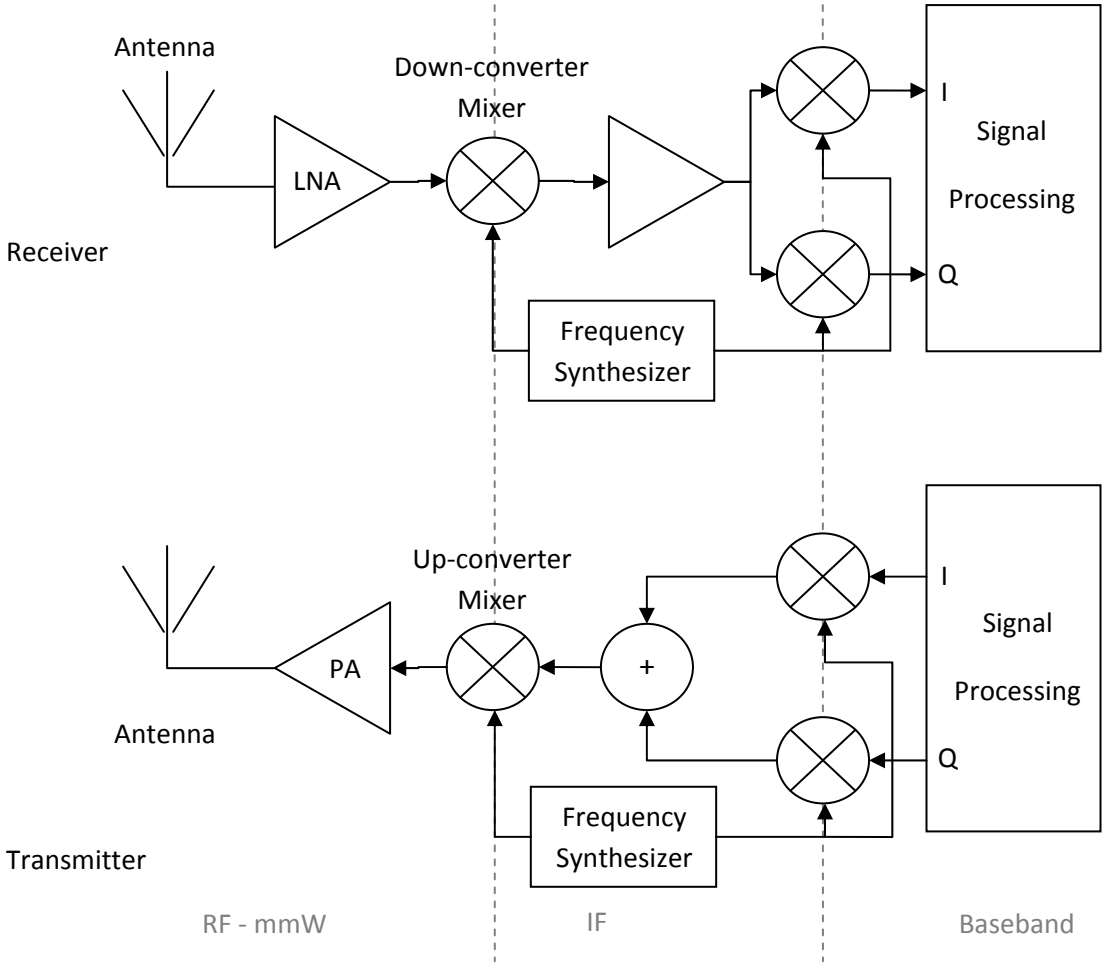


Fig. 12. Simplified block diagram of a super-heterodyne architecture.

As shown in the Figure, the signal at the output of the receive antenna is amplified by a low noise amplifier (LNA) with sufficient gain to establish the system noise figure. The LNA drives a mixer that

translates the RF signal to an intermediate frequency (IF), chosen to provide easy image rejection at the RF and to support very high data-rates. A variable-gain IF amplifier can be used to increase the dynamic range of the receiver.

The IF signal is translated in the baseband frequency and eventually split into its in-phase and in quadrature (I/Q) components, in the case of a quadrature reception scheme. Usually a single frequency synthesizer furnishes the LO signal for the conversion mixer as well as for the quadrature mixers.

The transmit path is designed using a dual strategy. The IF signal – eventually formed by its I/Q components – is translated in the RF carrier frequency by a mixer. Then the RF signal is translated to the desired transmit power by the power amplifier (PA) which, in turn, drives the transmit antenna.

The super-heterodyne performs frequency conversion from RF to baseband and vice versa in two steps. In a variant of the super-heterodyne known as homodyne (or direct conversion, or Zero-IF) architecture, the LO is synchronous to the RF signal and this one is converted to baseband with a single mixing step. It means that the IF frequency is chosen to correspond to the baseband and that the IF signal can be processed as a baseband signal.

In some cases the single-element antenna of Fig. 12 is replaced by a group of antennas forming a phased array. The relative phases of the signals feeding the antennas are varied in such a way that the effective radiation pattern of the array is reinforced in a desired direction and suppressed in undesired directions. The directionality of the array can be varied in a controlled manner applying a signal processing technique known as beamforming. When transmitting, a beamformer controls the phase and relative amplitude of the signal at each transmitter element, in order to create a pattern of constructive and destructive interference in the wave front. When receiving, information from different sensors is combined in such a way that the expected pattern of radiation is preferentially observed. Achieving electronic beam steering, beamforming enables directional communication and reduces interference for receivers that are not targeted. Furthermore, it results in a sensible augmentation of antenna gain both in transmission and reception mode and also of the emitted power if many power amplifiers are inserted in parallel into the array.

Several beamforming techniques have been developed so far and some of them can be applied to the millimeter wave context. Many of the applications discussed in the previous section, from 60GHz links to automotive radars, including imaging systems, can take advantage of it.

Examples of millimeter wave transceivers based on super-heterodyne or direct conversion architecture, with or without on-chip phased-arrays, are illustrated in the next section. Hereafter, we will give a brief overview of the most common building blocks of radio frequency systems and a summary of their main features.

1.4.2 Low-Noise Amplifier (LNA)

The first stage of a receiver is typically a Low-Noise Amplifier (LNA), whose main function is to provide enough gain to overcome the noise contribution of subsequent stages. Usually it is located just after the antenna (or after the antenna switch or filter, if they are included in the receiver

architecture). Regardless of the specific application, a wireless receiver must handle the signal coming from the antenna that is characterized by a very low power level. Any further degradation of the input signal must be avoided and an LNA should provide sufficient gain while adding as little noise as possible. LNAs are therefore designed following special techniques in order to reduce their noise contributions.

The role of the LNA inside a receiver system can be more rigorously expressed by means of the Friis equation that links the noise factor of a *N*-stage reception chain *(F)* with the noise *(F_i)* and gain *(G_i)* contributions of the *i*-th stage, as follows:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \cdots \quad (2)$$

A detailed strategy for LNA design will be described in Chapter 3. Now, let us focus on the parameters commonly used to evaluate the performances of a LNA.

First of all a LNA must provide low noise *and* high gain, as results from (2). Besides, it should accommodate large signals without distortion and frequently must present a specific input impedance in order to be matched to the source. Additionally, power consumption is an important consideration in many applications. Furthermore, in the specific context of millimeter wave, also reverse isolation and stability are critical considerations.

A simple way to estimate the linearity of an amplifier is based on gain compression. For an ideal linear amplifier, power gain should be a linear function of input power. Actually, the gain will start to saturate when input signal becomes too large. The input-referred 1dB compression point (IP1dB) is defined as the value of input power that corresponds to a mismatch of 1dB between the real gain and the ideal gain predicted on the basis of the linear assumption. Sometimes, the output-referred compression point (OP1dB) that is the output power corresponding to the 1dB gain compression point is used.

Linearity can also be measured in terms of third-order intercept point (IP3). The output of a linear, memoryless system can be approximated as follows:

$$f[v] \cong c_0 + c_1 v + c_2 v^2 + c_3 v^3 \quad (3)$$

$$v = V\cos(\omega_1 t) + V\cos(\omega_2 t) \quad (4)$$

where v is the input signal, chosen as sufficiently small to make reasonable the approximation with a power series truncated after the cubic term. If v is the sum of two sinusoidal tones of equal amplitude but slightly different frequencies as in equation (4), then the cubic term of the output gives rise to third-order intermodulation product:

$$\frac{3}{4} c_3 V^3 [\cos(\omega_1 + 2\omega_2)t + \cos(\omega_1 - 2\omega_2)t + \cos(2\omega_1 + \omega_2)t + \cos(2\omega_1 - \omega_2)t] \quad (5)$$

Whereas the sum frequency terms can be neglected in a tuned amplifier, the difference frequency terms lie within the bandwidth of RF signal and their amplitude increases as a function of input amplitude *(V)* three times faster than the fundamental term on a logarithmic scale. The intercept point occurs when the amplitude of the intermodulation term equals the amplitude of the fundamental term:

$$|c_1 V| = \left|\frac{3}{4} c_3 V^3\right|$$

$$V^2 = \frac{4}{3}\left|\frac{c_1}{c_3}\right|$$

Then, moving from the input voltage amplitude to the power domain, the input-referred third-order intermodulation point (IIP3) is:

$$IIP3 = \frac{V^2}{2R} = \frac{2}{3R}\left|\frac{c_1}{c_3}\right| \quad (6)$$

Alternatively, the IIP3 can be calculated feeding the amplifier with a single tone input and evaluating the influence on the output of its third-order harmonic. In both cases, the resulting IIP3 value is a purely mathematical concept that relies on the assumption that the higher-order, non-linear terms not considered in (3) are negligible, a condition that is not verified in many practical cases. In both simulations and experiments, the IIP3 is evaluated by extrapolating trends observed with relatively small amplitude inputs, at the smallest possible power level and its value depends on the measurement conditions that need to be accurately documented. Nevertheless, the IIP3 is often reported in the literature without any detail of the measurement conditions, making impossible the comparison between data of different sources.

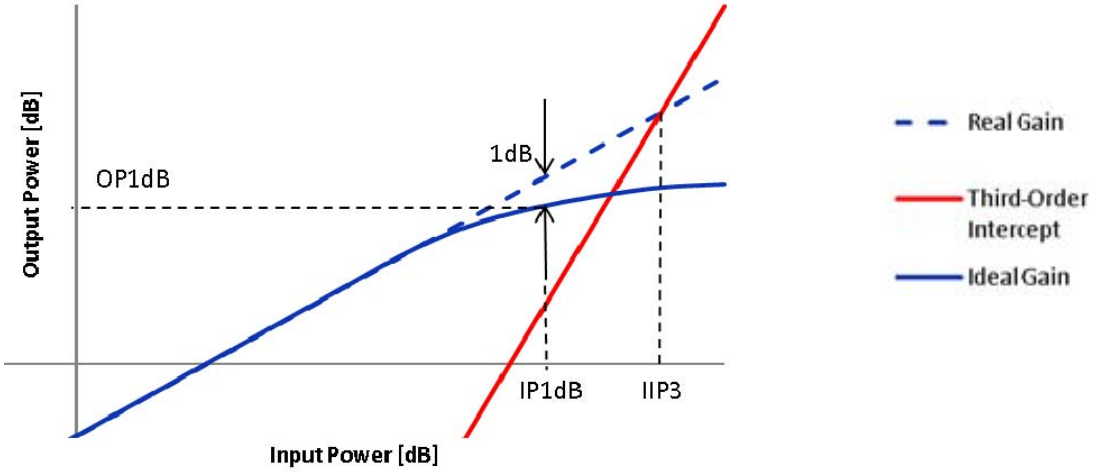


Fig. 13. Linearity.

1.4.3 Mixer

Mixers are frequency-translation devices. In all receivers based on the heterodyne architecture, they are commonly used to convert the radio frequency input signal to the intermediate frequency. Since a linear, time-invariant system cannot produce output signals with spectral components that are not present at the input, mixers must be either nonlinear or time-varying elements in order to perform frequency translation.

Practically speaking, a mixer consists of two inputs, the RF signal and the output of a local oscillator, that are multiplied, giving an output signal at a frequency that is the sum or, as more usual, the

difference between the input frequencies, together with some undesired spurious products. Actually, two input frequencies, frequently referred as *sidebands*, will correspond to a given intermediate frequency since signals both above and below ω_{LO} by an amount equal to ω_{IF} will produce IF outputs at the same frequency. In many cases only one is the desired RF signal whereas the other is a disturb known as *image* signal.

The definition of noise figure in the context of mixers must take this phenomenon into account. Usually a single sideband (SSB) NF is defined considering the signal-to-noise ratio at the RF port, divided by the SNR ratio at the IF port that includes also the noise contribution due to the image signal. On the contrary, when both the main RF and image signals contain useful information, a double sideband (DSB) NF should be used.

Beside NF, an important mixer characteristic is conversion gain which is defined as the power ratio of the IF output to the RF input. It can be greater than unity in active mixers, whereas a passive mixer is usually affected by conversion loss.

As in the case of amplifiers, linearity in the context of mixers can be evaluated using the 1dB compression point of the conversion gain (IP1dB) or the third-order intermodulation intercept point (IIP3) defined considering two input tones with slightly different frequencies, both within the RF range, in a similar manner as explained in the previous paragraph.

Another parameter of great importance in mixer design is isolation. Interaction among the RF, IF, and LO ports must be minimized, to assure that each one of the three concerned signal operates correctly. In particular, to prevent the presence of RF and LO signals at the output port, together with the desired IF signal, a double-balanced mixer can be used. In a double-balanced mixer both the RF and LO signals are applied differentially at the corresponding input ports, adding themselves destructively at the IF port. The most common implementation of a double-balanced mixer as an integrated circuit is the four-quadrant multiplier based on the Gilbert cell, sketched in Fig. 14. This solution, common in traditional radio frequency context, can be efficiently applied also to millimeter wave design.

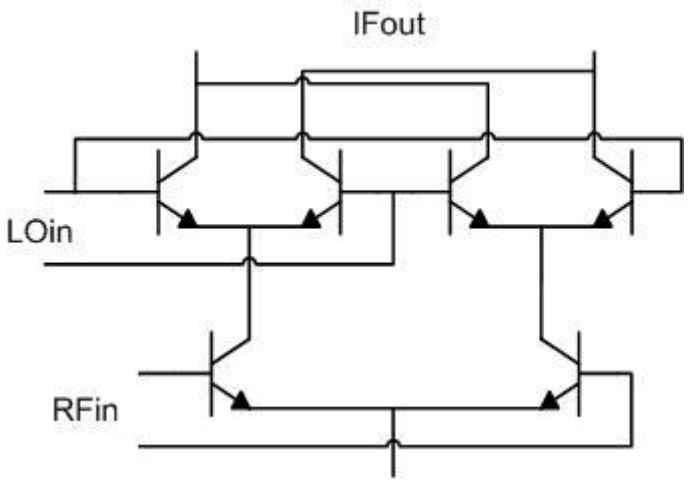


Fig. 14. Double-balanced mixer based on the Gilbert cell.

1.4.4 Frequency Synthesizer

In a RF system, both transmitters and receivers require the implementations of one or more local oscillators capable to provide an output signal whose frequency should be as stable as possible. To provide a sufficient degree of frequency stability, oscillators are usually included into a more complex building block performing as a frequency synthesizer. Practically speaking, the generation of the output signal at the required frequency is based on a reference oscillation at a different frequency (usually a lower one) whose accuracy can be controlled independently.

A frequency synthesizer can operate directly or indirectly. In a direct frequency synthesizer, starting from a unique reference signal, the output signal is obtained by means of subsequent frequency divisions and/or multiplications. An indirect synthesizer, on the contrary, is based on a feedback loop that controls the frequency stability of the output signal by adapting it to the reference. Indirect frequency synthesizers are commonly implemented as phase-looked loops (PLLs).

The performances of a frequency synthesizer are commonly estimated in terms of output-frequency range and resolution, spectral purity and phase-noise. For some application also the establishment time can be a critical factor.

1.4.5 Power Amplifier (PA)

Power amplifiers (PAs) are used in RF transmitters to convert a low-power RF signal into a larger one of significant power that drives the antenna.

The main goal of power amplifier design is to deliver a specified amount of power into a load with the highest possible efficiency consistent with acceptable power gain and linearity. Linearity is given by gain compression as in the case of LNAs, whereas efficiency, in the context of power amplifiers, is evaluated in terms of power-added efficiency (PAE) whose definition takes into account not simply the output power but also the gain of the amplifier:

$$PAE = \frac{Pout - Pin}{P_{DC}}$$

Design of PAs is accomplished using special dedicated techniques, largely different from the design of small-signal amplifiers.

1.5 Examples of millimeter wave integrated transceivers in CMOS and SiGe BiCMOS technologies.

This section contains a list of 9 examples of integrated transceivers, implemented using state-of-the-art, general-purpose, CMOS and BiCMOS technologies, and capable of supporting some of the millimeter wave applications described at the beginning of this chapter.

The demonstration of BiCMOS and CMOS devices capable of millimeter wave operation is a disruptive event in the development of millimeter wave systems. As a matter of fact, millimeter wave circuits are traditionally implemented using dedicated technologies based on compound

semiconductor like III-V elements which, due to their prohibitive costs, have prevented so far the diffusion of millimeter wave systems in an adequate amount.

The possibility to design and realize CMOS and BiCMOS millimeter wave integrated circuits is a consequence of the recent development of such technologies and is a key element in the realization of low-cost products which can support large-volume production as required by the market of consuming electronics.

A large number of the circuits listed hereafter is intended for frequency-modulated continuous-wave (FMCW) radar applications at 77GHz. FMCW is a radar system based on Doppler shift that can be efficiently applied to the millimeter wave context since it does not require high power and large bandwidth operation on the contrary of pulse-radar systems.

Fig. 15 illustrates the principle of FMCW radar. A FMCW signal is a continuous wave whose frequency is modulated in a triangular shape with time. The radar transmitter amplifies the FMCW signal and radiates it via a transmit antenna. Then, the radiated signal reaches the target (located at a distance R from the antenna) that reflects a portion of the signal back to the radar. The reflected signal enters into the receiver via a receive antenna. The incoming signal is then amplified and mixed with a local oscillation generated from the same FMCW generator used in the transmission mode. Since the round-trip time of flight (TOF) of the signal is 2R/c, where c is the speed of light, the frequency difference between the received signal and the LO signal, which results in the frequency of the mixer output (f_{beat}), is related to the TOF. Then, the distance R can be obtained by frequency detection such as a fast Fourier transform (FFT).

Besides, regardless of the system architecture, it is important to observe that two different strategies are applied to passive devices design like inductors or baluns. Some of the design examples proposed hereafter use a more conventional approach and implement passive elements as distributed structures like transmission lines. More recent examples, on the contrary, adopt an innovative approach implementing inductive devices as lumped structures. As it will be explained in the next chapter, inductors design at millimeter wave frequencies is a challenging point and requires special care.

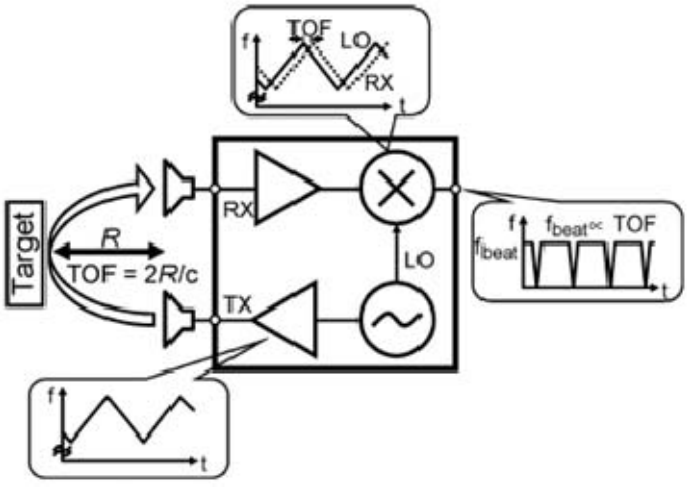


Fig. 15. FMCW radar system.

The first example that we propose concerns a 65GHz Doppler radar transceiver with on-chip patch antenna integrated in an 180nm SiGe BiCMOS process, reported in [25]. The block diagram of the radar transceiver is shown in Fig. 16 and includes a varactor-tuned VCO, a two-stage cascode LNA, a double-balanced down-convert mixer, an IF amplifier, and a microstrip patch antenna. The transmit path of the transceiver consists of a VCO (shared with the receive section), a two-stage emitter follower (not shown in the Figure) and a 65GHz output buffer implemented as a differential cascode with inductive degeneration for improved linearity.

The single-ended receiver measurements without patch antenna show an input-referred 1dB compression point of -30dBm, IIP3 of –20dBm, and single-ended conversion gain of 16.5dB at IF frequencies between 10MHz and 800MHz. The double-sideband noise figure is 12.5-13dB, measured over an IF range of 0.4-2GHz. The transmitter gives a single-ended output power of 1.3dBm (4.3dBm differential) at 64.8GHz. The full transceiver consumes 640mW. Its most relevant features are resumed in Table 12 and compared with those of the following examples.

It is important to remark that the single-chip transceiver makes extensive use of small footprint (less than 30µm per side) inductors and transformers as matching elements to achieve significant area savings, in contrast to more common solution based on distributed elements.

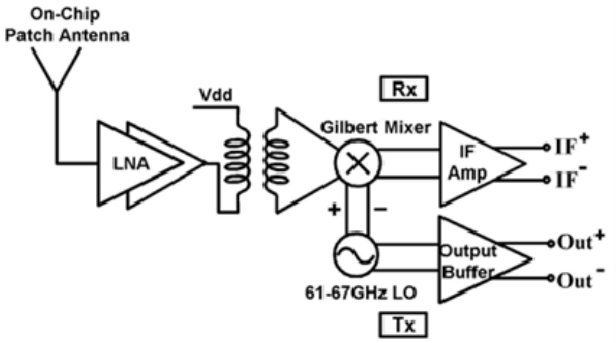


Fig. 16. Block diagram of a 65GHz Doppler radar transceiver with on-chip patch antenna [25].

Ref. [26] proposes a bidirectional architecture based on a 4-element phased array in 90nm CMOS. The use of a common array in the transmit and receive modes can result in a considerable area reduction. As shown in the block diagram of Fig. 17, each element of the array is connected to a bidirectional path formed by a LNA and a PA using separate switch at the input and output node to reduce mutual load. All LNAs and PAs are implemented as 4-stage CS amplifier with similar performances, in order to assure a reciprocal behavior in Rx and Tx operations. The contribution of the four RF unities is combined directly in the RF domain, by means of a Wilkinson device. A passive star mixer driven by a bidirectional amplifier provides frequency conversion.

At 58GHz the measured gain for both Tx and Rx mode is 7dB, with a power consumption of 78mW. Receiver NF is 8.7dB. The IP1dB is of -19dBm and -0.5dBm for Rx and Tx operation, respectively. The maximum output power in Tx mode is 3.5dBm. Chip dimensions are 1.6×1 mm^2.

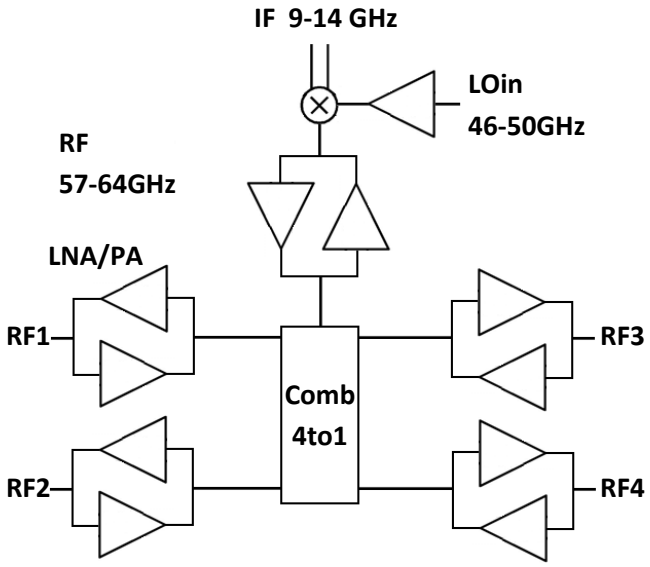


Fig. 17. Block diagram of a bidirectional TX/RX 4-element phased array at 60 GHz in 90-nm CMOS [26].

A directly modulated, 60GHz zero-IF transceiver architecture suitable for single-carrier, low-power, multi-gigabit wireless links in 65nm CMOS technology is reported in [27]. A block diagram of the proposed transceiver is presented in Fig. 18.

This system utilizes direct BPSK modulation at 60GHz, precluding the need for a power amplifier, and operates without requiring image rejection or ADCs in the receiver. The transmitter input accepts baseband digital data at rates beyond 6Gb/s, and the receiver outputs the same digital data stream in a true, single chip bits-in/bits-out radio transceiver. Input data, produced off-chip, is applied to a large power BPSK modulator which directly modulates the 60GHz LO signal and drives 50Ω loads differentially at the transmitter output. The receiver consists of a high-gain, low-noise amplifier that drives a double-balanced Gilbert cell down-convert mixer. No IF amplifier was included, so all the receive-path gain is at 60 GHz. The baseband data is recovered at the IF output of the receiver, without any digital signal processing or analog-to-digital conversion.

The LNA consists of three CS-CG stages. The last stage is loaded with a transformer which acts as single-ended to differential converter between the LNA and the double-balanced mixer. The transmitter consists of a BPSK modulator implemented as a double-balanced Gilbert cell similar to the mixer. A VCO or PLL was not implemented in this transceiver.

Measurements of the LNA breakout display a peak gain of 19.2 dB and a IP1dB of -14dBm at 60GHz. Due to insufficient power at the mixer LO port, the receiver gain measured in the transceiver was degraded. A peak receiver conversion gain of 14.7dB and a noise figure of 5.6dB are measured at 60GHz. The transceiver occupies only 1.28×0.81 mm^2 and consumes 151mW and 223mW in receive and transmit mode, respectively.

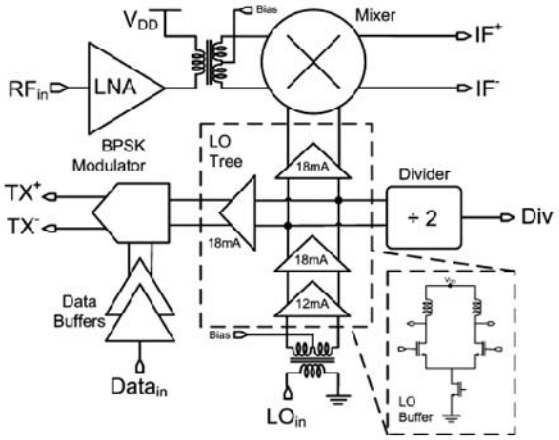


Fig. 18. 60GHz direct modulation BPSK transceiver architecture [27].

A four-element phased-array transceiver with on-chip antennas operating at 77GHz and implemented in a 120nm SiGe BiCMOS process is reported in [28] and [29]. The first paper describes the receiver section of the chip that consists of four down-conversion paths adopting a two-step down-conversion scheme with a RF frequency of 76–81 GHz and an IF in the range of 25–27 GHz.

As shown in Fig. 19, each receiver front-end consists of an on-chip differential dipole antenna, a two-stage differential cascode LNA, a double-balanced RF mixer, and a dual-mode IF amplifier with variable gain. The four-path IF signals are combined using a symmetric active combining amplifier and the combined signal is further down-converted to baseband using quadrature paths.

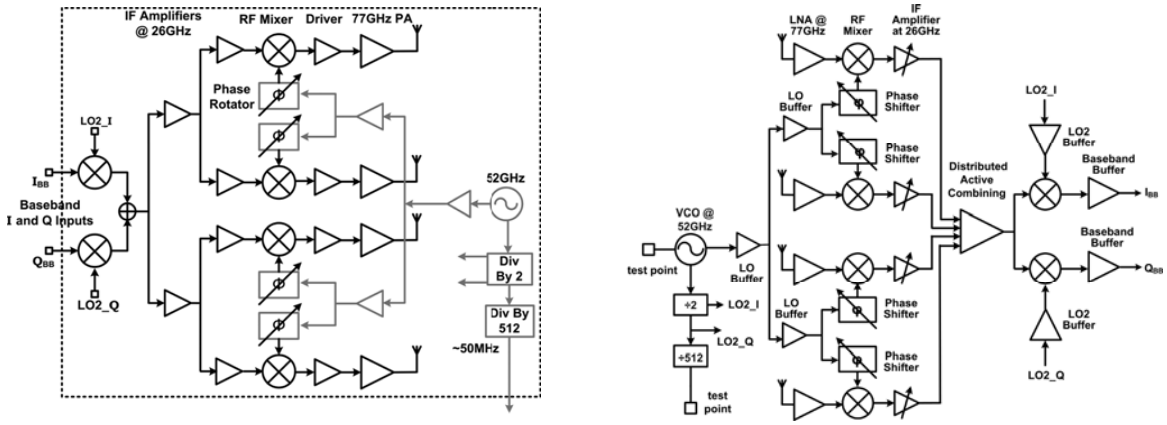


Fig. 19. A 77-GHz phased-array transceiver with on-chip antennas in 120nm SiGe BiCMOS: receiver (right) and transmitter (left) block diagrams [28]-[29].

A 37dB single-path receiver gain with a 2GHz bandwidth is measured at 79.8GHz, after laser trimming of the antenna. The measured antenna gain is 2dBi and the corresponding inferred array gain is 49dB. The minimum receiver noise figure is measured to be 8dB at 78.8GHz.

In the corresponding transmitter section, as described in [29], each element of the array generates up to 12.5dBm with a 1dB compression point of 10.2dBm, leading to a 4-element effective isotropic radiated power (EIRP) of 24.5dBm.

A differential VCO generates the LO signals at 52GHz required for the first RX down-conversion and for the second TX up-conversion. Then, the quadrature LO is obtained dividing by two the 52GHz oscillation and a frequency divider chain is used to further divide the second LO frequency down to 50MHz to be locked to an external reference frequency.

Two versions of a unique direct conversion transceiver intended for Doppler radar and millimeter wave imaging, implemented in two 130nm SiGe BiCMOS technologies, are described in [30]. Due to different f_T/f_{max} in the two processes, the two versions operate at 77GHz and at 82GHz.

A block diagram of the circuit is shown in Fig. 20. The received signal is amplified by a 3-stage LNA and converted to a differential signal. A double-balanced Gilbert cell mixer down-converts the RF signal to baseband, and an IF amplifier provides additional current gain to drive off-chip test equipment. A Colpitts VCO, with differential tuning inputs for the frequency modulation signal, drives a network of buffers and transmission lines which distribute the LO signal to all circuit blocks. A static frequency divider is also included on-chip to monitor the precise VCO frequency.

The 77GHz transceiver has a 26dB peak down-conversion gain and a power consumption of 740mW. The identical transceiver fabricated in the higher ft/fmax process exhibits better performances and has a 40dB peak conversion gain centered at 82GHz with a power consumption of 780mW. Its IP1dB is -35dBm and the DSB NF is of only 3.85dB. The output power is 11.5dBm.

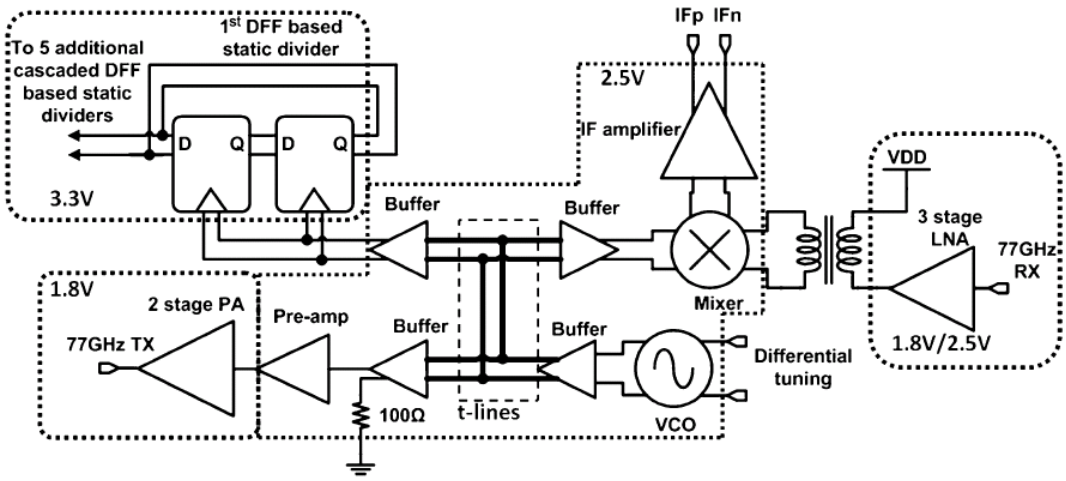


Fig. 20. Block diagram of a Doppler radar transceiver [30].

An example of FMCW radar transceiver implemented in a 90nm CMOS technology is reported in [31]. The proposed transceiver architecture, shown in Fig. 21 consists of a LNA, a down-conversion mixer with an output buffer, a driver amplifier, a power amplifier with power combiner, a LO distributer and a FMCW synthesizer.

The LNA employs five differential stages with a Marchand balun based on transmission lines at the input node to convert the single-ended input coming from the antenna. The down-conversion mixer consists of a double-balanced mixer with a common-gate circuit for the input stage, while the PA consists of 2 differential common-source amplifiers and a power combiner based on a Marchand balun used to achieve differential to single-ended transformation, power combining of two amplifiers output and impedance transformation from the PA output to an external antenna input. A PLL based

on a 77GHz VCO is also implemented on chip to perform FMCW signal generation and to provide LO signal for down-conversion. The chip size is 3.50 × 1.95 mm^2. Measured performances are resumed in Table 12.

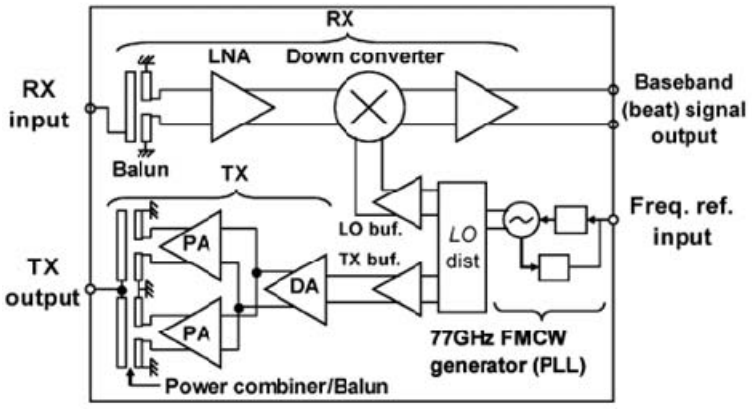


Fig. 21. Block diagram of the fabricated TRX IC [31].

A fully integrated FMCW radar system for automotive applications implemented in a 65nm CMOS process is proposed in [32]. The transceiver architecture is illustrated in Fig. 22 and consists of a RF front-end (PA, LNA, and mixer), two high-gain antennas, a FMCW generator (basically a fractional-synthesizer), and a FPGA-based signal processor. By tuning the divide modulus, the full-rate VCO delivers FMCW carrier signal around 77GHz directly to the PA, the mixer, and the first divider, without requiring any frequency doublers or triplers, simplifying the circuit design.

The LNA is realized as three identical gain cascode stages without source degeneration, whereas the PA is a cascade of five class A amplifiers. The VCO is implemented as a standard tank structure with varactors and the mixer as a single-balanced structure to reduce the LO port loading.

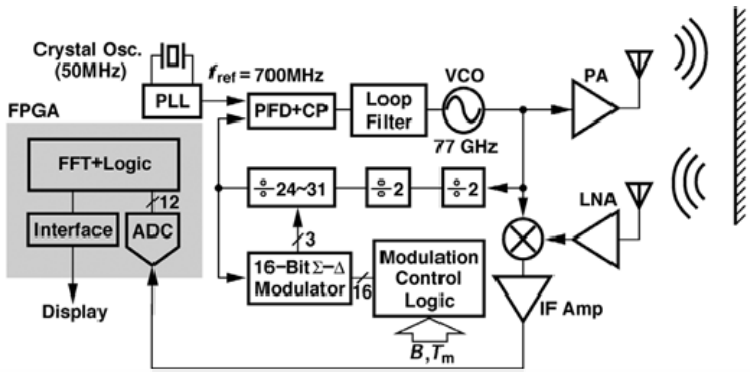


Fig. 22. Block diagram of a FMCW radar transceiver. The antennas, a signal processing FOGA module, and a PLL with crystal oscillator are implemented off-chip [32].

The chip measures 0.95 × 1.1 mm^2 and consumes a total power of 243mW. The LNA achieves 17.5dB gain, 7.4dB NF, -22dBm IP1dB, and -12.5dBm IIP3 at 77 GHz. The PA reveals a peak gain of 13.7 dB, with a IP1dB of 6.7 dBm, a saturation power of 10.5dBm and maximum PAE of 8.4%.

A second example of a 90nm CMOS integrated transceiver for 77GHz FMCW automotive radars is reported in [33]. The transceiver architecture is shown in Fig. 23. A signal generator (SG) based on a 38.5GHz VCO provides a 77GHz oscillation with triangular modulation that, in turn, is divided into two differential signals. One of them drives the transmitter, whereas the other is used as a LO for the down-conversion of the input signal. The transceiver operates in the frequency range between 73.5 and 77.1 GHz, with a total power dissipation of 920mW. The IF amplifier is designed with an output impedance of 1kΩ and therefore the conversion gain of the receiver measured in a 50Ω setup, is only 2±1.5dB.

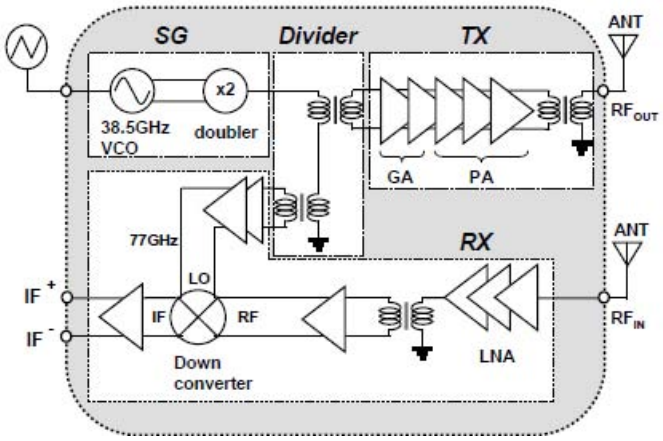


Fig. 23. Block diagram of a 77GHz transceiver for FMCW automotive radars in 90nm CMOS [33].

1.6 Conclusion

The first sections of this chapter have provided a brief overview of the most common applications of millimeter wave technology that can potentially result in large-volume production, leading to a market revolution in the context of consuming electronics. The main features of WPANs and high data-rate wireless communications at 60GHz, short and long range radars at 77-79GHz, and imaging systems at 94GHz have been illustrated.

Traditionally, applications of millimeter waves rely on the use of special technologies based on compound semiconductors which, due to their considerable cost, have prevented a significant diffusion of millimeter wave devices. Nowadays, the recent developments in silicon and SiGe technologies make possible the realization of low-cost systems and circuits operating in the millimeter wave range.

To give a tangible evidence of such a trend, several examples of millimeter wave transceiver fabricated using standard CMOS and BiCMOS processes, taken from the very recent state of art (published since 2006 up to 2010), have been reported in the previous section.

However, even if the latest CMOS and BiCMOS technology nodes have been proved to be capable of supporting millimeter wave operation, nevertheless millimeter wave design still requires special techniques and methodologies to overcome a large number of difficulties and constraints which appear along with the augmentation of the operative frequency.

Trying to answer to this requirement, the next chapter focuses on the development and definition of a design methodology for millimeter wave integrated circuits in the BiCMOS9MW technology.

Table 12. Measured characteristics of millimeter wave integrated transceivers implemented on CMOS and BiCMOS standard processes.

Ref.	[25] 2006	[26] 2009	[27] 2009	[28][29] 2006	[30] 2008	[30] 2008	[31] 2010	[32] 2010	[33] 2009	Unit
Freq.	65	60	60	77	77	82	77	77	77	GHz
Tech.	180nm SiGe BiCMOS	90nm CMOS	65nm CMOS	120nm SiGe BiCMOS	130nm SiGe BiCMOS	130nm SiGe BiCMOS	90nm CMOS	65nm CMOS	90nm CMOS	
Area	6.25	1.6	1.04	25.84	1.17	1.17	3.32	1.04	2.88	mm^2
Arch.	Doppler Radar with on-chip patch antenna	4 ph-array bi-directional Rx/Tx (off-chip LO)	A Zero-IF Transceiver with Direct BPSK Modulation (off-chip LO)	4 ph-array super-heterodyne with on-chip array	Direct-conversion Doppler radar	Direct-conversion Doppler radar	FMCW Doppler Radar	FMCW Doppler Radar	FMCW Doppler Radar	
Induct.	Lumped	Lumped	Distrib.	Distrib.	Mixed	Mixed	Distrib.	Distrib.	Lumped	
Transmitter										
3dB Bw		57-60		77±2.5	76-81	77-85			73.5-77.1	GHz
Output Power	4.3	3.5	2.4	12.5	5.8	11.5	-2.8	5.1	3.3-6.3	dBm
Gain		7		40.6		32	14	13.7		dB
Power Cons.	394	78	223	2945		175	305	115	390	mW
Receiver										
3dB Bw		57-60	54-66	77±2	76-81	77-85				GHz
Conv. Gain	16.5	7	11.3	37	25.6	40	23.1	38.7	2	dB
NF	12.5-13	8.7	5.6	8	9	3.85	15.6	7.4	6.8	dB
Power Cons.	394	78	151	1050		115	111	55	260	mW
Frequency synthesizer										
Tuning range	61-67	46-50		50.35-55.49	77	82	78.1-78.8	75.6-76.4	73.5-77.1	GHz
IF	0.01-2	9-14		25-27						GHz
PhN at 1MHz	-104				-99	-99	-85	-85.33z	-86	dBc/Hz
Power Cons.	394			263		120	101	73	270	mW

Bibliography

On 60GHz applications and standard definitions for 60GHz WPANs:

[1] Ali M. Niknejad, Hossein Hashemi, "mm-Wave Silicon Technology 60GHz and Beyond", Springer, 2008.

[2] Su-Khiong Yong, Pengfei Xia and Alberto Valdes Garcia, "60 GHz Technology for Gbps WLAN and WPAN: From Theory to Practice", John Wiley & Sons, 2011.

[3] Xiaoyi Zhu, A. Doufexi, T. Kocak, "On the performance of IEEE 802.15.3c millimeter-wave WPANs: PHY and MAC," in *2010 6th Conference on Wireless Advanced (WiAD)*, Aug 2010, pp. 1-6.

[4] IEEE std 802.15.3c, "Wireless Medium Acess Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal area Networks (WPANs), Amendment 2: Millimeter-wave-based Alternative Physical Layer Extension", IEEE, October 2009.

[5] WirelessHD, available on http://www.wirelesshd.org/.

[6] Wireless Gigabit Alliance, available on http://wirelessgigabitalliance.org/.

[7] IEEE 802.11 Working Group, Very High Throughput in 60 GHz; available on http://www.ieee802.org/11/Reports/tgad_update.htm.

[8] Standard ECMA-387. High Rate 60 GHz PHY, MAC and HDMI PAL, available on http://www.ecma-international.org/.

[9] ETSI Standards "Electromagnetic compatibility and Radio spectrum Matters (ERM); Technical characteristics of multiple gigabit wireless systems in the 60GHz range System Reference Document (TR 102 555 Ver. 1.1.1)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, Feb. 2007.

[10] ETSI Standards "Broadband Radio Access Networks (BRAN); 60 GHz Multiple-Gigabit WAS/RLAN Systems; Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive (EN 302 567 Ver. 1.1.1)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, Mar. 2009.

On Automotive Radar Applications and Regulation:

[11] M. Schneider, "Automotive Radar – Status and Trends", *German Microwave Conference*, pp 144-147, April 2005.

[12] Dale M. Grimes, Trevor Owen Jones, "Automotive radar: A Brief Reveiw", *Proceedings of the IEEE*, Vol. 62, No. 6, pp. 804-822, June 1974.

[13] J. Wenger and S. Hahn, "Long Range and Ultra-Wideband Short Range Automotive Radar", *ICUWB 2007 IEEE International Conference on Ultra-Wideband*, pp. 518-522, Sept. 2007.

[14] ETSI Automotive Radar Standards "Short Range Devices; Road Transport and Traffic Telematics (RTTT); Radar equipment operating in the 76 GHz to 77 GHz range (EN 301 091 parts 1-2 Ver 1.3.3)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, Nov. 2006.

[15] ETSI Automotive Radar Standards "Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices; Road Transport and Traffic Telematics (RTTT); Short range radar equipment operating in the 24 GHz range; Part 1: Technical requirements and methods of measurement (EN 302 288-1 Ver 1.4.1)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, Jan. 2009

[16] ETSI Automotive Radar Standards "Road Transport and Traffic Telematics (RTTT); Radio equipment to be used in the 77 GHz to 81 GHz band; System Reference Document for automotive collision warning Short Range Radar (TR 102 263 Ver 1.1.2)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, Feb. 2004.

[17] ETSI Automotive Radar Standards "Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices (SRD); Road Traffic and Transport Telematics (RTTT); Technical characteristics for communications equipment in the frequency band from 63 GHz to 64 GHz; System Reference Document (TR 102 400 Ver 1.2.1)", available on http://www.etsi.org/WebSite/Technologies/AutomotiveRadar .aspx, July. 2006.

[18] R. Schneider, H. L. Blocher, K. M. Strohm, "KOKON - automotive high frequency technology at 77/79 GHz", *European Microwave Conference*, pp. 1526-1529, Oct. 2007.

[19] RNTR-VeLo Communication Inter Véhicules et localization Relative Précise – Architecture Système et Specifications électrique (Rapport intermédiaire de sous-projet 2).

On millimeter wave imaging applications:

[20] Duxian Liu, Ulrich Pfeiffer, Janusz Grzyb, Brian Gaucher, "Advanced Millimeter-wave Technologies Antennas, Packaging and Circuits", John Wiley & Sons Ltd, 2009.

[21] Austin Richards, "Alien Vision: Exploring the Electromagnetic Spectrum with Imaging technology", SPIE – The international Society for Optical Engineering, 2001.

[22] Dwight L. Woolard, William R. Loerop, Michael S. Shur, "Terahertz Sensing Technology Vol. 2: Emerging Scientific Applications & Novel Device Concepts", World Scientifc Publishing Co., 2003.

[23] J.W. May, G.M. Rebeiz, "High-performance W-band SiGe RFICs for passive millimeter-wave imaging" *IEEE Radio Frequency Integrated Circuits Symposium*, pp 437-440, June 2009.

[24] A. Tomkins, P. Garcia, S.P. Voinigescu, "A Passive W-Band imaging Receiver in 65-nm Bulk CMOS", *IEEE Journal of Solid-State Circuits*, Vol. 45, No. 10, pp. 1981-1991, Oct. 2010.

Examples of integrated transceivers for millimeter wave applications (CMOS and BiCMOS technologies):

[25] T. Yao, L. Tchoketch-Kebir, O. Yurjevich, M. Gordon, S. P. Voinigescu, "65GHz Doppler Sensor with On-Chip Antenna in 0.18μm SiGe BiCMOS", *IEEE MTT-S International Microwave Symposium Digest*, pp 1493-1496, June 2006.

[26] E. Cohen, C. G. Jakobson, S. Ravid, D. Ritter, "A Bidirectional TX/RX Four-Element Phased Array at 60 GHz With RF-IF Conversion Block in 90-nm CMOS Process", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 58, No. 5, pp. 1438-1446, May 2010.

[27] A. Tomkins, R. Andres Aroca, T. Yamamoto, S. T. Nicolson, Y. Doi, S. P. Voinigescu, "A Zero-IF 60 GHz 65 nm CMOS Transceiver With Direct BPSK Modulation Demonstrating up to 6 Gb/s Data Rates Over a 2 m Wireless Link", *IEEE Journal of Solid-State Circuits*, Vol. 44, No. 8, pp. 2085-2099, Aug. 2009.

[28] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, A. Hajimiri, "A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Receiver and Antennas", *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 12, pp. 2795-2806, Dec. 2006.

[29] A. Natarajan, A. Komijani, X. Guan, A. Babakhani, A. Hajimiri, "A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Transmitter and Local LO-Path Phase Shifting", *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 12, pp. 2807-2819, Dec. 2006.

[30] S. T. Nicolson, P. Chevalier, B. Sautreuil, S. P. Voinigescu, "Single-Chip W-band SiGe HBT Transceivers and Receivers for Doppler Radar and Millimeter-Wave Imaging", *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 10, pp. 2206-2217, Oct. 2008.

[31] T. Mitomo, N. Ono, H. Hoshino, Y. Yoshihara, O. Watanabe, I. Seto, "A 77 GHz 90 nm CMOS Transceiver for FMCW Radar Applications", *IEEE Journal of Solid-State Circuits*, Vol. 45, No. 4, pp. 928-937, April 2010.

[32] J. Lee, Y. A. Li, M. H. Hung, S. J. Huang, "A Fully-Integrated 77-GHz FMCW Radar Transceiver in 65-nm CMOS Technology", *IEEE Journal of Solid-State Circuits*, pp. 1-11, Oct. 2010.

[33] Y. Kawano, T. Suzuki, M. Sato, K. Joshin, "A 77GHz Transceiver in 90nm CMOS", *IEEE International Solid-State Circuits Conference - Digest of Technical Papers*, pp. 310-311, Feb. 2009.

Chapter 2

Design Methodology for Millimeter Wave Integrated Circuits

2.0 Introduction

Until few years ago, integrated circuits operating beyond 30GHz were a prerogative of dedicated technology processes relying on the use of III-V compound semiconductors such as gallium arsenide or indium phosphide. Recently, improvements in silicon-based technologies have led to a sensible extension of the upper frequency limit of silicon integrated devices. Nowadays, as a consequence, it is possible to design low-cost and high-volume systems and circuits operating at millimeter wave frequencies.

As discussed in the previous chapter, new application standards have been recently defined – or their definition is currently in progress – at 60GHz, in the 77-82GHz band, at 94GHz, and even beyond.

However, the augmentation of the operative frequencies and bandwidths entails considerable difficulties, concerning at the same time the development of performing technologies, but also measurement, modeling and design.

Basically, three families of silicon-based technologies can address the target of high volume and low cost production: bulk CMOS, SOI CMOS and BiCMOS. All these technologies have been demonstrated to be potentially capable to furnish active devices operating at millimeter wave frequencies, but with different properties and characteristics. Therefore, the first critical element in millimeter wave design is the choice between different technology processes, and it requires an attentive analysis.

Besides active devices, also the implementation of passive elements is made difficult by the increased frequencies. Indeed, if on one hand higher frequencies result in smaller devices and in a reduced area, then, on the other hand, frequency augmentation unavoidably implies the degradation of device performances.

Accurate modeling of both active and passive devices is another crucial element. In this work, it can be remarked that even a relatively small frequency mismatch due to model inaccuracy, usually negligible at lower frequencies, can have catastrophic effects in the millimeter wave context. Often silicon foundries do not furnish a complete device characterization and the search for efficient models, or modeling optimization at the very least, is left to the designers. This task is usually accomplished by means of an extensive use of electromagnetic analysis on dedicated simulation tools, that contributes considerably to the complexity of the design.

Generally speaking, many of the assumptions currently made in conventional radio-frequency design are not valid at millimeter waves. According to definition of “millimeter wave”, at the concerned frequencies the equivalent electromagnetic wavelength on silicon substrates is of around 1mm and therefore it is comparable with the physical dimension of integrated circuits. As a consequence, the influence of parasitics over intrinsic devices becomes too important to be neglected. Control over parasitics is then the key element for successful millimeter wave design; therefore a large part of the following discussion will be dedicated to investigation of efficient parasitic extraction procedures.

The aim of this chapter is to set a reliable design methodology capable of overcoming all the difficulties depicted so far. Because of its advantageous features, the BiCMOS9MW technology by ST Microelectronics has been chosen to support our investigations. After a general overview of this technology, its active and passive devices are described in the first subsections of this chapter.

Design kit models of each device are closely examined and model improvements, accounting for parasitic or secondary effects, are proposed, if necessary. Then, in section 2.3 the differences between lumped and distributed design are highlighted and in section 2.4 a strategy for lumped device implementation is presented. The proposed design methodology for millimeter wave building blocks is completed by some suggestions on interconnect modeling in section 2.5 and is resumed in a general design flow in section 2.6. Finally, section 2.7 gives a brief insight on the de-embedding of measurements at millimeter wave frequencies.

2.1 Technology Background

The millimeter-waves market segment has been historically held since decades by the III-V semiconductors, thus relegating these applications into some niche market, because of the high manufacturing cost and reduced integration scale of the III-V solutions. Two phenomena have recently transformed full silicon technologies into the preferred integration platform for millimeter-wave design. First of all, the very high frequency performances of silicon active devices (bipolar and CMOS) have dramatically increased over the past years, featuring both f_T and f_{max} close or even higher than 200GHz. And secondly, the market has expressed its need for low-cost consumer products in the millimeter wave range, as explained in the first chapter.

Among the first examples of silicon-based design in the millimeter-wave context reported in the literature prior to 2004, it is possible to find sporadic examples of voltage-controlled oscillators operating between 60 and 100 GHz [1]–[6], low-noise amplifiers, and mixers for 24GHz [7]–[9] and 40GHz [10], and a power amplifier for 77GHz [6] fabricated in SiGe technologies. In the same period, only CMOS oscillators [11]–[13] were demonstrated beyond 30GHz, while CMOS amplifiers [14]–[16] and mixers [16], [17] were respectively limited to 26GHz and 21.8GHz.

An important improvement is made by Reynolds [18] that presents the key building blocks of a 60GHz front-end including an LNA, a direct down-converter, a PA and a VCO implemented in a 0.12µm SiGe technology by IBM featuring NPN devices with f_T and f_{max} both over 200GHz. Almost at the same time Doan [19] reports the design of a 60GHz amplifier using a standard digital 0.13µm CMOS process by ST Microelectronics, together with a detailed study on the optimization of transistors and CPW transmission lines, thus demonstrating that through careful optimization and modeling of the active and passive components, a standard CMOS process is capable of 60GHz operation. A complete 60GHz direct-conversion receiver in 0.13µm CMOS technology is described in [Razavi 20]. In the following years several examples of 60GHz building blocks are realized in the emergent CMOS technology nodes and a design flow of 60GHz LNAs and PAs is defined for the 90nm node [Yao 21].

If CMOS on one hand seems to address the goals of 60GHz applications, on the other hand, the improvements in SiGe BiCMOS technologies extend the design feasibility to higher frequency standards. A complete 77GHz chipset for Doppler radar and imaging applications is reported in 2007 [Nicolson 22], as well as a dual-band imaging transceiver [Laskin 23], transmitting and receiving simultaneously in the 80GHz and 160GHz bands. These examples demonstrate for the first time the feasibility of complete transceivers beyond 60GHz. In both case, an innovative SiGe HBT technology by ST microelectronics is used [26].

The dichotomy between CMOS and BICMOS technology has been present in millimeter-wave design since its very beginning. A third alternative is the use of CMOS processes with Silicon on Insulator (SOI) substrates [24]-[25]. BiCMOS roadmap does not follow Moore's law, which is driven by the shrink of digital functions. The move from a CMOS node to the next one can be motivated by the increase of gate density, while HBT performances may remain unchanged if operation frequency of the targeted applications does not change. The evolution of CMOS technologies will probably end with the dramatic increase of high frequency performances. On the contrary, high-speed BiCMOS roadmap is driven, on one hand by the increase of the optical communications data rate, and on the other hand by the emergence of applications at higher frequencies. The result is a step-by-step evolution aiming at the best combination between CMOS density and HBT performance. This roadmap will go on as long as Si/SiGe:C HBT performances can be pushed forward (with significant advantages over CMOS) to satisfy the requirements of newer applications at higher frequencies.

Focusing on active devices, the last developed BiCMOS technologies are based on the use of bipolar heterojunction transistors fabricated with optimized process steps. As a consequence, they results into a synthesis of the advantages given by high-performing bipolar devices, together with the common benefits of a CMOS background in digital applications.

Table 1. Characteristics of ST Microelectronics technologies.

Technology	BiCMOS9	BiCMOS9MW	CMOS65nm
Node [nm]	130	130	65
Vdc [V]	1.8	1.8	1.2
fT/fmax [GHz]	160/160	230/280	200/220
BEOL	6 + Alucap	6 + Alucap	7 + Alucap
Thin/Thick layers	5/1	3/3	6/1
BVCEO [V]	1.8	1.6	-
BVG Ox [V]	-	-	1.6

Contrary to what usually happens in the analog (and in a given perimeter in RF) design, in the millimeter-wave context, information on active devices by itself is not sufficient to choose the most appropriate technology. For frequencies above 10GHz, the thickness of the Back-End of Line (BEOL) has a significant influence on the electrical behavior of passive devices. Moving from a CMOS node to the next one implies a vertical shrink of the BEOL and a reduction of the metal pitch, required to increase integration density. A reduced thickness of the BEOL is unavoidably reflected into a degradation of performances for passives elements both localized and distributed, such as lumped inductors and transmission lines, because of the increased attenuation and substrate losses. To overcome these limitations, technologies with a BEOL optimized for millimeter-wave designs have been recently developed like the BiCMOS9MW by ST Microelectronics that was available at the beginning of our work. Its BEOL is compared in Fig. 1 with the digital BEOL of a typical CMOS65nm process and to that of the conventional BiCMOS9 technology.

On the basis of the considerations briefly highlighted above, the BiCMOS9MW appears as the best choice for the implementations of the building blocks of a front-end for millimeter wave applications and therefore it will be extensively used in our designs.

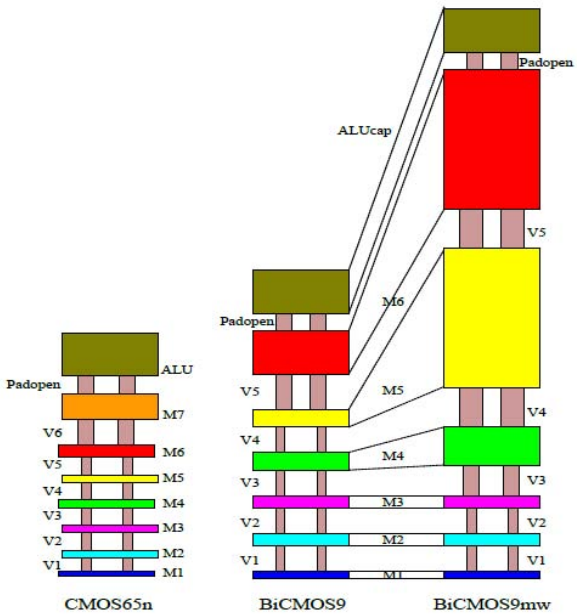


Fig. 1. Comparison between the digital BEOLs of a typical CMOS65nm process and of the BiCMOS9 technology with the millimeter wave dedicated BiCMOS9MW BEOL.

2.2 The BiCMOS9MW technology

2.2.1 General Features of the BiCMOS9MW Technology

The BiCMOS9MW technology by ST Microelectronics is a new platform dedicated to millimeter wave design. It has been developed during the last few years on the basis of the former BiCMOS9 process that was intended for optical networking and wireless applications up to 40GHz – 40Gb/s.

Despite some reported examples, the BiCMOS9 technology can not be efficiently used for circuits design at millimeter wave frequencies because of two important limitations:

- the maximum transition frequency is limited to 160GHz and it strongly affects the transistor behavior at high frequencies;
- the vertical extension of the BEOL is of only 6µm and this prevents the implementation of high-quality passive devices.

To overcome the first constraint the technology process flow has been modified in order to embed a new architecture of the HBT transistor with a fully self-aligned (FSA) implementation using selective epitaxial growth (SEG) of the base. These innovations allow a better control over some parasitic effects and as a consequence the maximum of transistor f_T and f_{max} are increased to 230GHz and 280GHz, respectively [27]-[28].

The BEOL of the BiCMOS9 technology is formed by six copper metal layers eventually overlapped by an aluminum cap layer. As mentioned before, the overall vertical extension of such a profile is of 6µm only. As a consequence, the implementation of lumped inductors and of distributed propagation structures such as transmission lines is affected by strong substrate coupling and

attenuation. To improve the quality of passive devices, the BEOL of the new technology relies still on a six metal structure, but the thickness of the three highest metal levels and of the relative inter metal dielectric layers has been increased. As depicted in Fig. 1, the overall BEOL extension is augmented up to 13µm. As a result, the quality of passive devices becomes close to that of above-IC BCB realizations. For example, on-wafer measurements performed up to 110GHz on a 51Ω characteristic impedance transmission line show a 0.5 dB/mm attenuation at 60GHz (Fig. 2) [27].

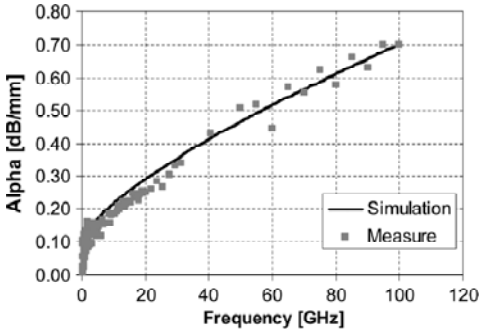


Fig. 2. Attenuation versus frequency of a 11.5µm wide microstrip line of 51Ω characteristic impedance implemented in the BiCMOS9MW technology [27].

Thanks to the innovations briefly highlighted above, the BiCMOS9MW can efficiently support millimeter wave design at least up to 94GHz. Recently, applications at higher frequencies have been also reported [37]. However, the device characterization and the design facilities provided by the design kit are not exhaustive and to get satisfactory results some specific procedures must be put in place as detailed in the next subsections. The main features of the technology are:

- CMOS devices for 1.2V applications.
- 2.5V Capable I/O's.
- High-performance NPN heterojunction bipolar transistors.
- Medium-voltage NPN heterojunction bipolar transistors.
- Shallow trench isolation (STI), triple well (NISO), twin-tub, single poly, for CMOS core process.
- Deep trench isolation (DTI) combined with STI and SiGe-C epitaxial base for NPN transistors
- Optional Dual V_T.
- Back-end with 6 metal levels.
- Damascene Copper for metal 1 to last metal.
- Thin metal layers: Metal 1 to 3.
- Thick metal layers: Metal4, Metal5 and Metal6.
- MIM capacitors

2.2.2 0.13µm BiCMOS Heterojunction Bipolar Transistor

The high-speed NPN HBT transistor is the key element of the technology as it can satisfy the requirements of millimeter wave design, concerning transition frequency, maximum frequency and noise contribution.

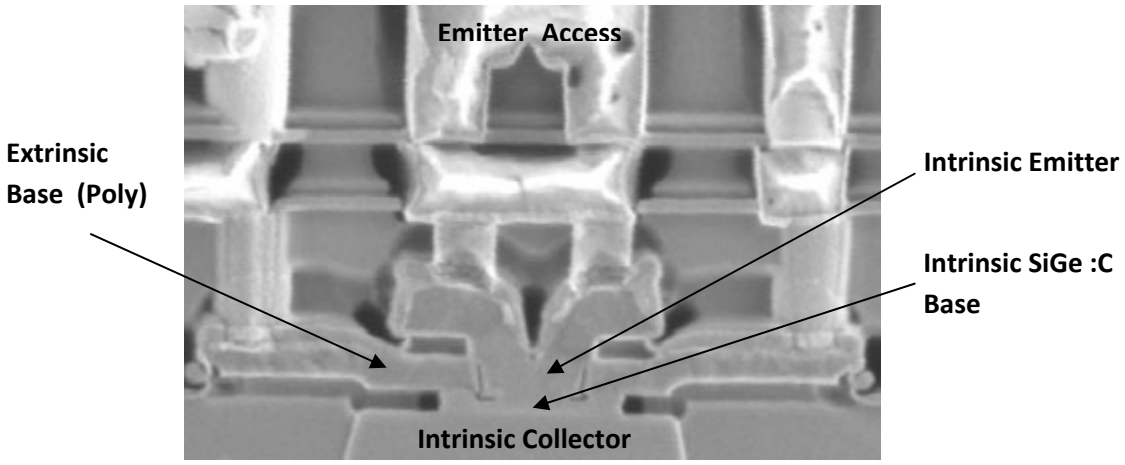


Fig. 3. SEM picture of a Heterojunction Bipolar Transistor.

A microphotograph of a HBT is shown in Fig. 3. The intrinsic device can be easily recognized as it stands vertically from the collector implanted in the substrate up to the polysilicon emitter. The intrinsic base region is made of SiGe:C and forms a double heterojunction with the silicon-based emitter and collector regions. Carbon is added to prevent the diffusion of the boron used for doping.

The main advantage in the use of a heterojunction is that, as a results of the mismatch between the electrical properties of silicon and SiGe, a band offset appears at the interface between the two materials and the energy gap can be set wider in the emitter (and eventually also in the collector) and reduced in correspondence of the base region. Therefore, current gain is increased and it is possible to reach higher level of boron doping into the base, reducing at the same time its width and parasitic resistance. The narrower base width decreases the base transit time, improving the transition frequency and gain. The reduction of the base resistance also contributes to augment the f_{max}.

The concentration of germanium in the base can be gradually modulated from a minimum at the emitter interface to a maximum in proximity of the collector. Thus the band gap in the base region is gradually decreased as shown in Fig. 4. As a consequence, a drift current is induced that drives the electrons across the base towards the collector and the transition frequency is still improved.

The fabrication process of the HBT transistor is depicted in Fig. 5 and is summarized as follows [27]. It starts with the implementation of a n+ buried layer on the p-substrate, followed by thin collector epitaxy and deep trench isolation. After definition of the active area by shallow trench, the collector sinker is implanted. Next, an oxide/poly/oxide/nitride stack is deposited and patterned (a) [26]. A 0.3µm emitter window is patterned and etched in this stack, stopping on the pedestal oxide layer. A selective collector implant (SIC) is performed and nitride sidewalls are formed to protect the polybase from the base epitaxy (b) [29]. The pedestal oxide is then wet etched to open the base cavity and the selective growth of the SiGeC base is done (c). Next are fabricated D-shaped inside spacers to get a final emitter width of approximately 0.1 µm. Finally, arsenic in-situ doped polysilicon is deposited to form the emitter (e).

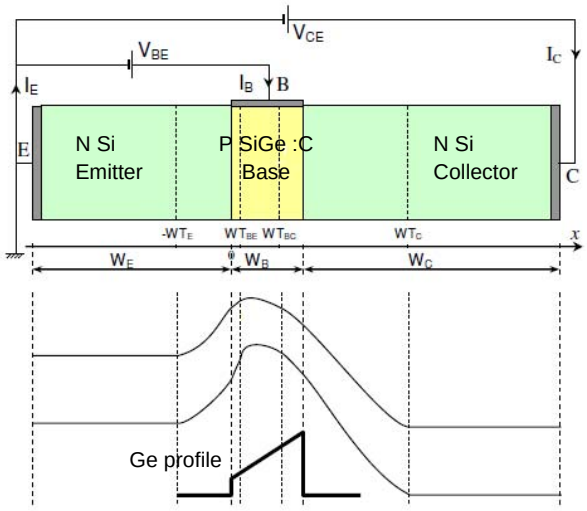


Fig. 4. Band diagram of a HBT device.

The process briefly described so far leads to a so-called fully self-aligned transistor architecture. As a matter of fact, in the definition of the intrinsic device only one mask is used, for the emitter window opening. Then, the collector implant and the base growth are automatically aligned to the emitter window. A crucial step is the wet etching of the pedestal oxide that opens a well defined cavity where the SiGe:C can be selectively deposited and the base growth can be accomplished. The use of only a mask prevents any possibility of alignment mismatching. Therefore, mismatching tolerances can be suppressed and the geometrical dimensions of the structure can be minimized since they are defined with a better resolution.

As a direct consequence of self alignment, parasitic contributions such as extrinsic base resistance, emitter resistance and collector-base capacitance are lowered if compared to the quasi-aligned architecture used in the former BiCMOS9 technology. Moreover, collector-base capacitance is further reduced, thanks to the use of the sacrificial oxide layer that improves the isolation between the extrinsic base and the underlying collector area and also as a consequence of the emitter width reduction due to the presence of the inside spacer modules, together with the emitter resistance.

Collector-base capacitance can be still reduced if the standard collector implantation of phosphorus is replaced by arsenic, as it allows a better control of the implantation vertical profile and prevents lateral diffusion. It should be noted that the use of arsenic corresponds to a measured improvement of 14GHz of the transition frequency [26].

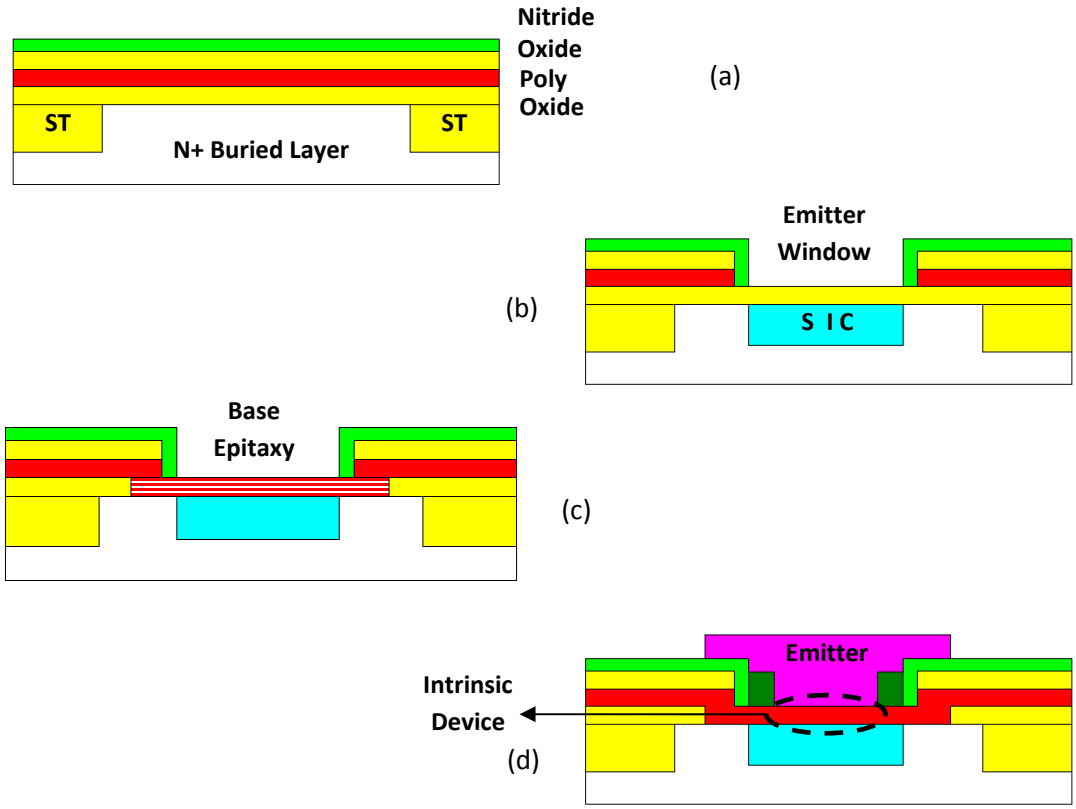


Fig. 5. Process steps for fabricating the SiGe FSA-SEG HBT.

The fabrication of the HBT transistor is part of a more complex BiCMOS process flow. A detailed description of the process and of its related phenomena is beyond the scope of this work.

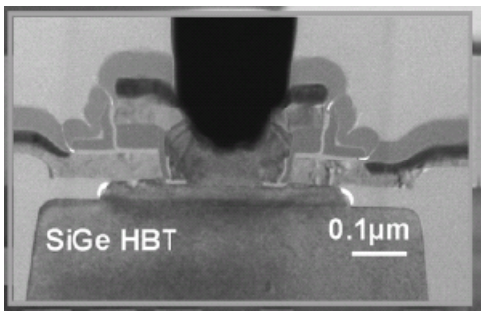


Fig. 6. TEM cross-section illustrating the HBT of the BiCMOS9MW technology [27].

Besides the high-speed HBT, a medium-voltage SiGe HBT can also be fabricated at no additional cost by protecting the device from SIC implantation. The collector-to-emitter breakdown voltage of such devices was raised from 1.56 V to 2.01 V. Maximum f_T dropped to 150 GHz and only little fmax reduction was observed due to concomitant base-collector capacitance reduction. Static and dynamic parameters for both types of SiGe HBTs [27] are summarized in Table 2.

Table 2. Static and Dynamic Characteristics for HS and MV SiGe HBTs with 0.12 × 4.85 µm^2 emitter area.

Parameter	HS	MV	Comments
C_{BC} [fF]	12.2	8.7	V_{CB} 0V
V_{BE} at f_{Tmax} [V]	0.91	0.84	
β at f_{Tmax}	282	442	
BV_{CEO} [V]	1.52	2.01	V_{BE} 0V
f_T [GHz]	240	150	V_{CB} 0.5V
f_{max} [GHz]	270	250	

Designers can use scalable models of both high-speed and medium-voltage transistors. Three possible configurations are supported for each device:

- single-base transistor (CBE)
- symmetric structure (CBEB)
- the previous structure can be multiplied in order to form a multi-emitter transistor (N×CBEB-C).

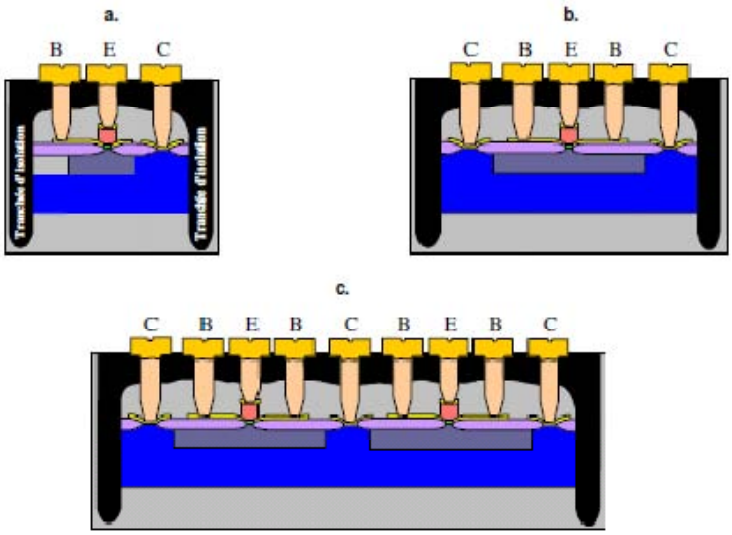


Fig. 7. Three different structures of HBT: single base transistor (a), symmetric structure (b), and symmetric multi-emitter structure (N=2) (c).

A complete model library has been developed for both HBT devices, which offers several SPICE models simulation levels: the classical SGP (SPICE Gummel-Poon) model and the advanced HICUM model (Level2 or Level0 according to the designer's choice). The library is "scalable" over a wide range of emitter lengths and allows the use of bipolar transistors with different configurations (multi-base, multi-collector, multi-finger). A more detailed discussion on transistor models is the subject of the next subsection.

2.2.3 Models of the High Speed Heterojunction Bipolar Transistor

A simplified side view of the transistor physical structure is reported in Fig. 8. It can be used to highlight the main effects on the transistor behavior of its physical structure, as a start point for the electrical characterization of the device.

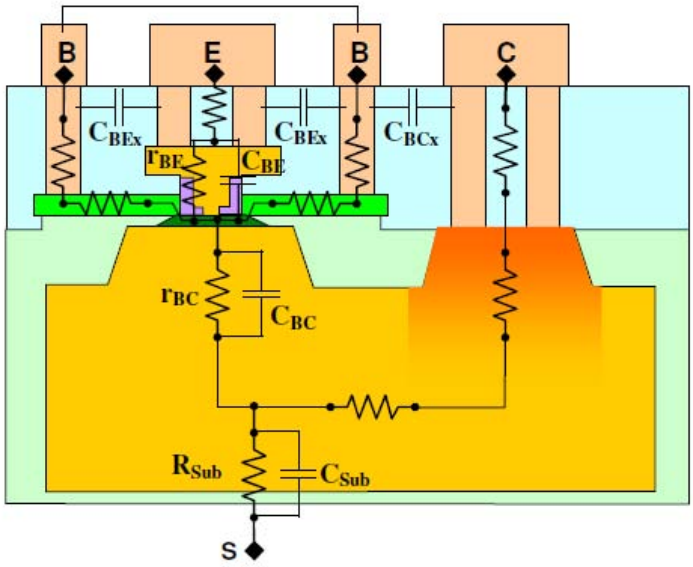


Fig. 8. Physical background of the equivalent-circuit model of HBT.

For the purpose of small signal analysis, the intrinsic effect of base-emitter and base collector junctions in the active region can be represented by means of the traditional equivalent circuit of Fig. 9. The direct biased base-emitter junction is represented by its equivalent dynamic resistance and diffusion capacitance, whereas the transconductance accounts for the overall amplifying effect and the reverse biased base-collector junction is represented by the corresponding transition capacitance.

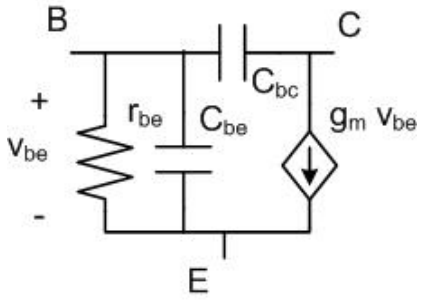


Fig. 9. Equivalent-circuit model for small signal analysis of the intrinsic transistor biased in the active region.

An electrical equivalent model of the transistor in the active region can be obtained by adding the contributions of the elements of Fig. 8 to its intrinsic model. The result is shown in Fig. 10.

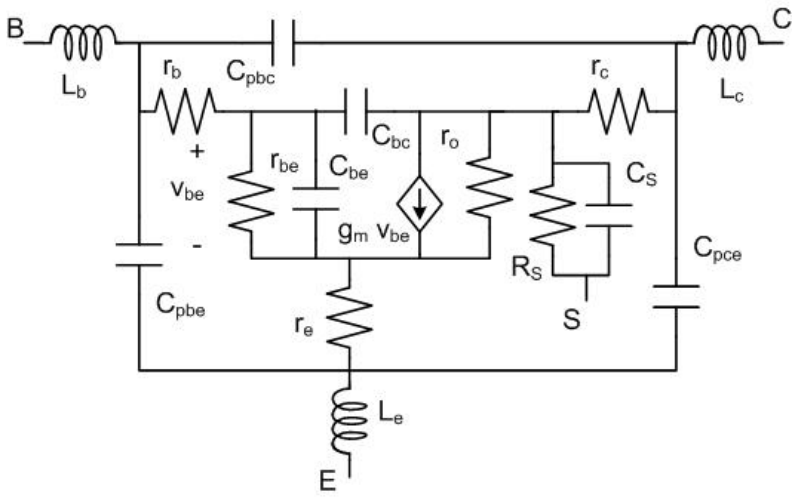


Fig. 10. Equivalent-circuit model for small signal analysis of the HBT biased in the active region.

To obtain an exhaustive characterization of the transistor, noise must be considered. It can be assumed that noise in bipolar transistors has three sources: thermal noise, shot noise and flicker noise also called "*1/f* noise". Each source is modeled with a current noise generator *i*. The mean-square value i^2 of this current is the noise power. The noise power is always associated to a bandwidth Δf. The power spectral density is therefore defined as:

$$S_i = \frac{i^2}{\Delta f} \tag{1}$$

The physical origin of these sources can be summarized as follows:

- Thermal noise is generated by series resistances. It is caused by the random thermal motion of carriers. Thus, it is directly proportional to temperature and frequency independent (white noise). The following equation gives the mean-square value of noise current for a R resistor, where Δf is the bandwidth:

$$\mathrm{i_R^2} = \frac{4\mathrm{kT}\Delta\mathrm{f}}{\mathrm{R}} \tag{2}$$

- Shot noise is generated by direct current flow. It is due to carriers that have to overcome a potential barrier. In a bipolar transistor biased in the normal active mode, carriers (electrons and holes) that go through the direct-biased emitter-base junction have to overcome a certain potential barrier. Consequently, they originate shot noise. This effect is modeled with two noise sources (one for the collector current, the other for the base current). The following equations give the mean-square value of the equivalent noise currents:

$$\mathrm{i_B^2} = 2\mathrm{qI_B}\Delta\mathrm{f} \tag{3}$$

$$\mathrm{i_C^2} = 2\mathrm{qI_C}\Delta\mathrm{f} \tag{4}$$

- Flicker noise is caused mainly by traps (contamination or crystal defects) at the emitter poly/mono silicon interface. These traps capture and release carriers in a random fashion

and give rise to a noise signal with a 1/f dependence (low frequency noise).This signal is always associated with a direct current flow. The following equation gives the mean-square value of the equivalent noise current coming from the base current:

$$i_B^2 = \gamma I_B^{\alpha} \frac{\Delta f}{f} \tag{5}$$

where α and γ are empirical parameters.

As depicted in Fig. 11, noise generators can be added to the transistor equivalent circuit. The mean-square values of the corresponding currents are given by the following equations:

$$i_B^2 = 2qI_B\Delta f + \gamma I_B^{\alpha}\frac{\Delta f}{f}$$

$$i_C^2 = 2qI_C\Delta f \tag{6}$$

$$i_{Rb}^2 = \frac{4kT\Delta f}{Rb} \qquad i_{Re}^2 = \frac{4kT\Delta f}{Re} \qquad i_{Rc}^2 = \frac{4kT\Delta f}{Rc}$$

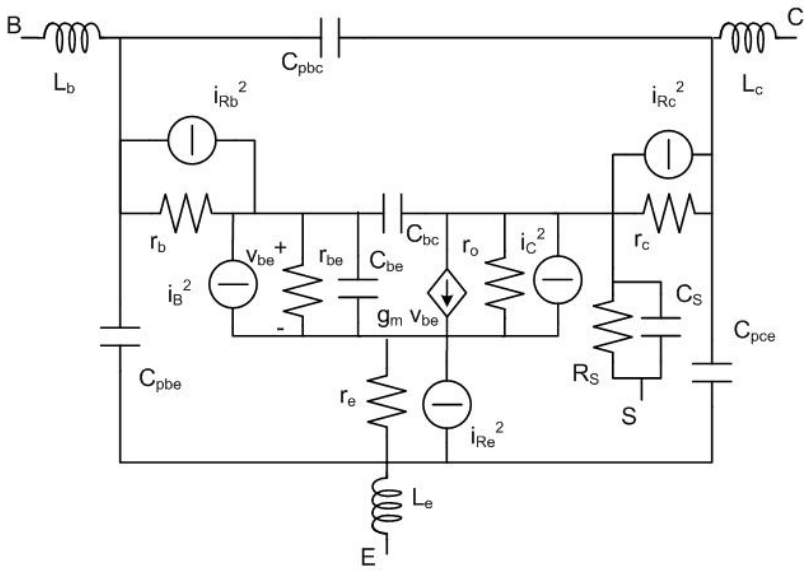


Fig. 11. Equivalent-circuit model of the HBT transistor accounting for noise sources.

The minimum Noise Figure (NFmin) of a HS transistor vs its bias-current density is reported in Fig. 12 for different frequencies: 60, 80, and 94GHz. Also a plot of f_T and fmax is shown for comparison.

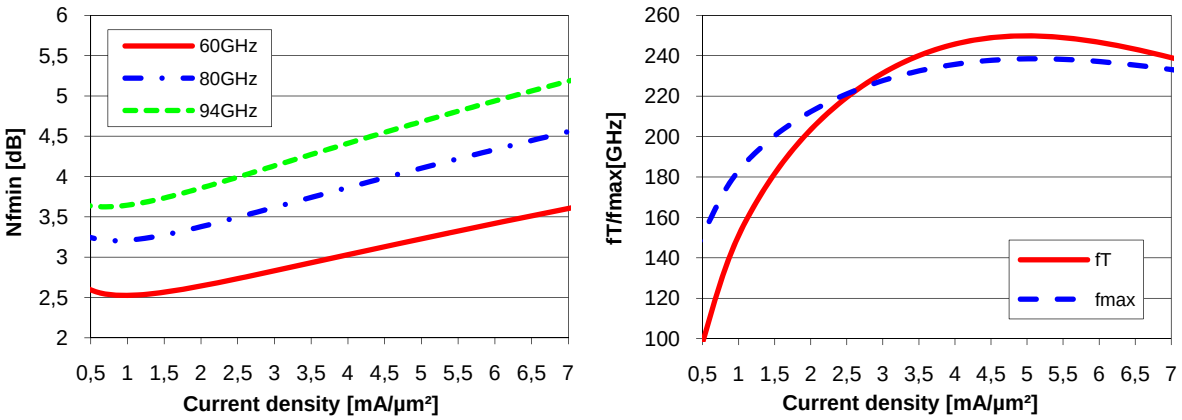


Fig. 12. NFmin, f_T and fmax for a single 1µm emitter HS transistor.

So far, an equivalent circuit for the transistor has been presented. It can be useful to give a first insight on the transistor behavior and to perform a preliminary analysis of circuit schematics. However, the transistor models actually used for simulation can be somewhat different and more complex.

Three different kinds of model levels of both the high-speed and medium-voltage devices are provided among the design facilities of the BiCMOS9MW technology:

- ST-BJT
- HICUM Level 2 (v2.21)
- HICUM Level 0 (v1.2)

All of them are compact models scalable according to the emitter length and can be adapted to each one of the three configurations described in Fig. 7: CBE, CBEB and N×(CBEB)-C. The emitter length can be varied between 0.4 and 15 µm and the emitter number N between 1 and 5. The drawn emitter width is fixed to 0.23µm.

The ST-BJT model level is a modern variant of the classical SPICE Gummel-Poon Model (SGPM) that, in turn, is based on the Ebers-Moll model. SGPM was developed in 1970 and since then has been widely used for bipolar transistors modeling. However, it is affected by some limitations concerning the formulation of charge storage effects, of the transfer current and the resulting transconductance, and of the internal base resistance as well as the missing self-heating and base–collector (BC) avalanche effect.

Advanced compact models such as HICUM Level 2, MEXTRAM, and VBIC have been developed in order to eliminate many of the above issues, but at the price of increased complexity and bigger computational effort.

The HIgh CUrrent Model (HICUM) Level 2 has been developed in 1980s and recently applied to HBTs. It can overcome some of the SGPM deficiencies and in particular can be efficiently applied in presence of elevated bias conditions. It can also take into account self-heating phenomenon, if the correspondent option is selected. According to the design kit documentation, it should be the most accurate model level furnished.

The HICUM Level 0 has been developed on the basis of the full version, HICUM Level 2, and combines the simplicity of the SGPM in terms of equivalent circuit and some of its model equations, with several important features of HICUM Level 2. As a result, it is a more physics-based and accurate model than the SGPM, but also reduces parameter extraction efforts, compared to the aforementioned models.

2.2.4 Corrections for transistor model

The routing of active devices is a critical step in millimeter wave design. At these frequencies the transistor behavior strongly depends on its layout, since the parasitic effects of the metal paths used to route the accesses to the device can not be neglected.

Transistor models provided by the technology always take into account parasitic resistances and capacitances due to the contact layer and, if the corresponding option is set, also the parasitic effects of the structures implemented at Metal1 level included in the standard layout of the transistor cell.

However, all the parasitics due to the Metal2 until the top layer, as well as to the Metal1 added by the designer, are not included in any model, and must be extracted during the post-layout simulation flow. At the beginning of our works no tools for post-layout extraction of the overall design were available among the design facilities and therefore a customized extraction procedure was defined. Capacitance extraction is performed at transistor level by means of a dedicated extraction tool such as Quickcap. Parasitic capacitances between base, emitter, collector terminals and between these terminals and the substrate are deduced on the basis of a simplified layout including only the devices accesses routed on Metal1 up to Metal6, but not the transistor itself.

Resistances can be estimated by hand on the basis of metals sheet resistance and vias resistance data reported in the design kit documentation.

As a result of post layout extraction, the transistor can be represented by the RC model of Fig. 13.

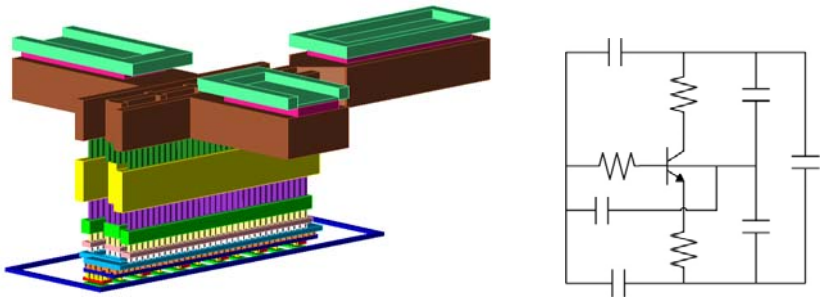


Fig. 13. High Speed HBT layout with 5×1µm emitters and its equivalent model with RC parasitic elements.

2.2.5 Transmission Lines

As discussed erlier, the BiCMOS9MW technology allows the implementation of high-performance transmission lines, thanks to the optimization of the BEOL.

In particular, specific design guidelines are provided for microstrips, together with a dedicated scalable model. Microstrip transmission lines are laid out using the top cupper layer and alucap shunted together. As shown in Fig. 14, a ground plane directly connected to the substrate is routed under the line. Due to the density constraints of the design rules, the ground shield must be implemented as a discontinuous pattern of metal 1 and metal 2 interlaced. Connection to the substrate is assured by taps. The height of such a ground plane is extended up to metal 6 to form a sidewall on both side of the line in order to prevent undesired coupling effects with the adjacent devices. In the case of a line directly connected to RF pads, it also helps the transition from a coplanar excitation mode as provided by the probe to a microstrip propagation mode.

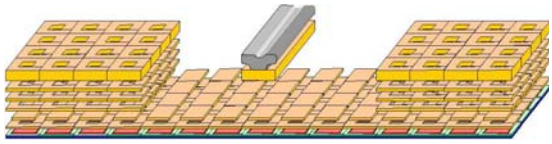


Fig. 14. Microstrip transmission line implemented in metal6 and alucap, with a M1-M2 ground shield connected to the substrate.

From the point of view of a circuit designer, a transmission line can be characterized in terms of its characteristic impedance Zc and its propagation constant γ=α+iβ. Both these parameters can be extracted from an ABCD-matrix representation of the transmission line, according to the following equations:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} cosh(\gamma l) & Zc\ sinh(\gamma l) \\ Zc^{-1} sinh(\gamma l) & cosh(\gamma l) \end{bmatrix} \tag{7}$$

$$Zc = \sqrt{\frac{B}{C}} \tag{8}$$

$$\gamma = \frac{cosh^{-1} A}{l} \tag{9}$$

where l is the physical length of the line. As an alternative, the characteristic impedance and the propagation constant can be also extracted directly by S parameters by means of more involved equations. Moreover, empirical equations have been reported that can approximately express the existing relationships between Zc and γ and the physical dimensions of a microstrip line [30].

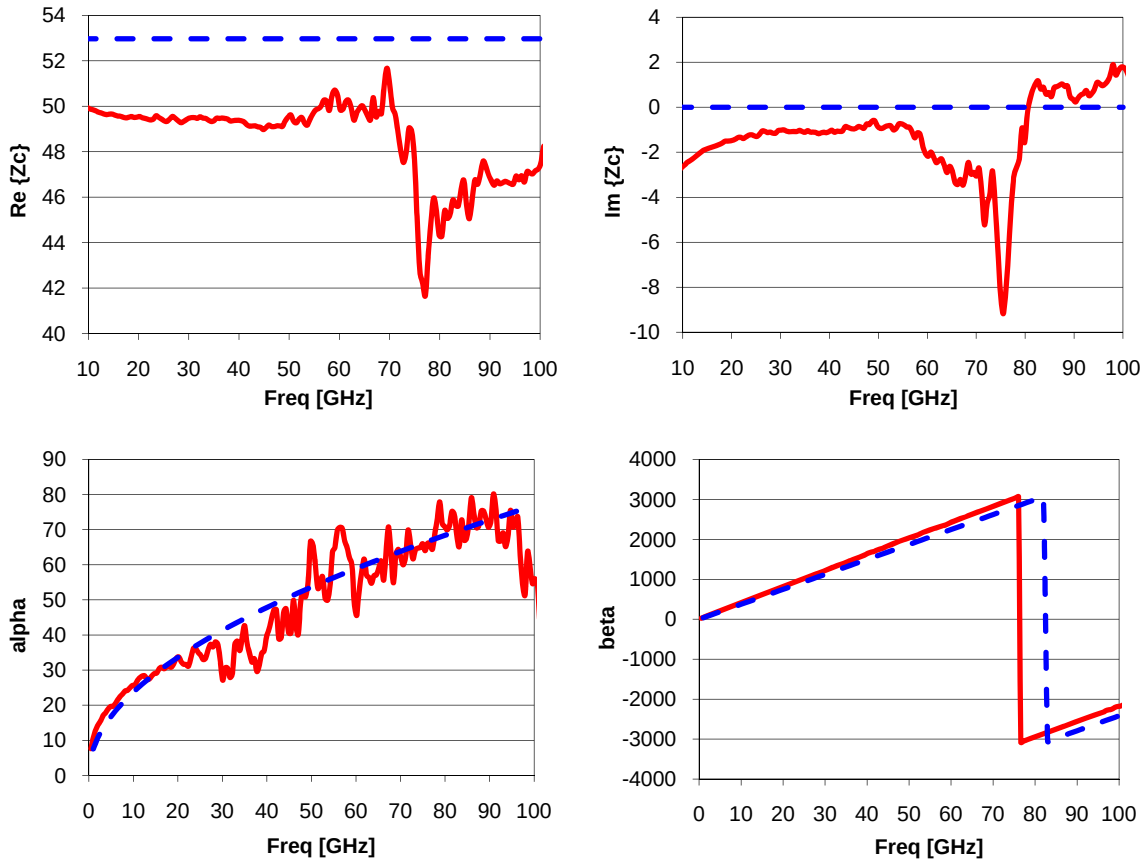


Fig. 15. Characteristic Impedance and Propagation Constant of a 1020µm microstrip transmission line: split-thru de-embedded measures (line) vs models (dash).

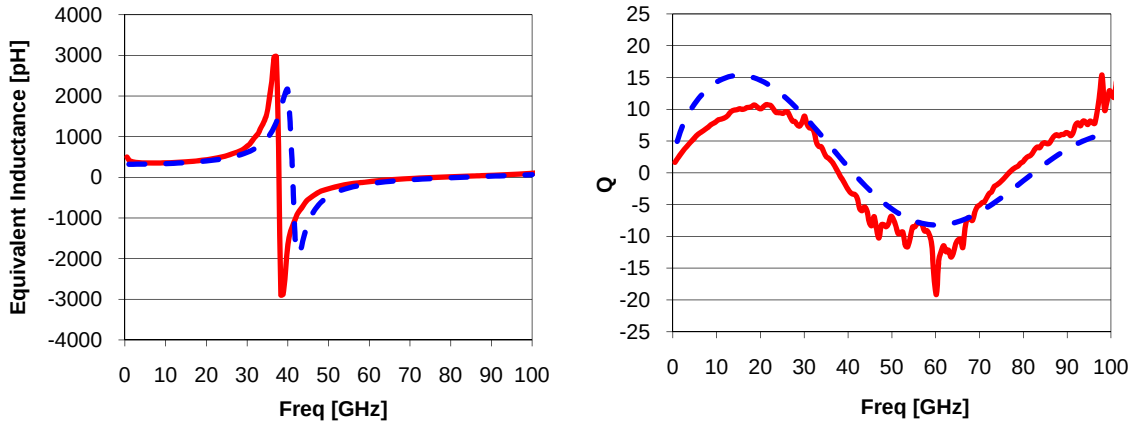


Fig. 16. Equivalent Inductance and Quality Factor of a 1020µm microstrip transmission line: split-thru de-embedded measures (line) vs models (dash).

In order to validate the model provided by the design kit, microstrip transmission lines of different lengths have been fabricated according to the aforementioned design guidelines and tested. Despite of the difference in their lengths, all the lines are implemented in metal 6 and alucap and have the same width (12µm) that approximately corresponds to a characteristic impedance of 50Ω. A first example concerning the characterization of a 1020µm is depicted in Figs. 15-16. To account for the presence of the input and output pads, the line characterization is based on the measured S

parameters of a 1116µm and a 96µm lines, rather than on direct measurement of a 1020µm line, applying the split-thru de-embedding procedure, as described in subsection 2.7.2.

The 1020µm length roughly corresponds to λ/4 at only 37.7GHz and therefore gives the possibility to completely explore the behavior of the line. It should be noted that beyond the first resonance at 37.7GHz up to 76.6GHz, the line exhibits a capacitive behavior. At 76.6GHz (roughly equal to λ/2), a 180° phase shift can be observed that justifies the resonance in the characteristic impedance plot. This last phenomenon is neglected by the model; however a λ/2 line is not used in any circuit implementation. Beside this singularity, the line behavior is correctly reproduced by the corresponding model representation, proving its accuracy all over the bandwidth of interest.

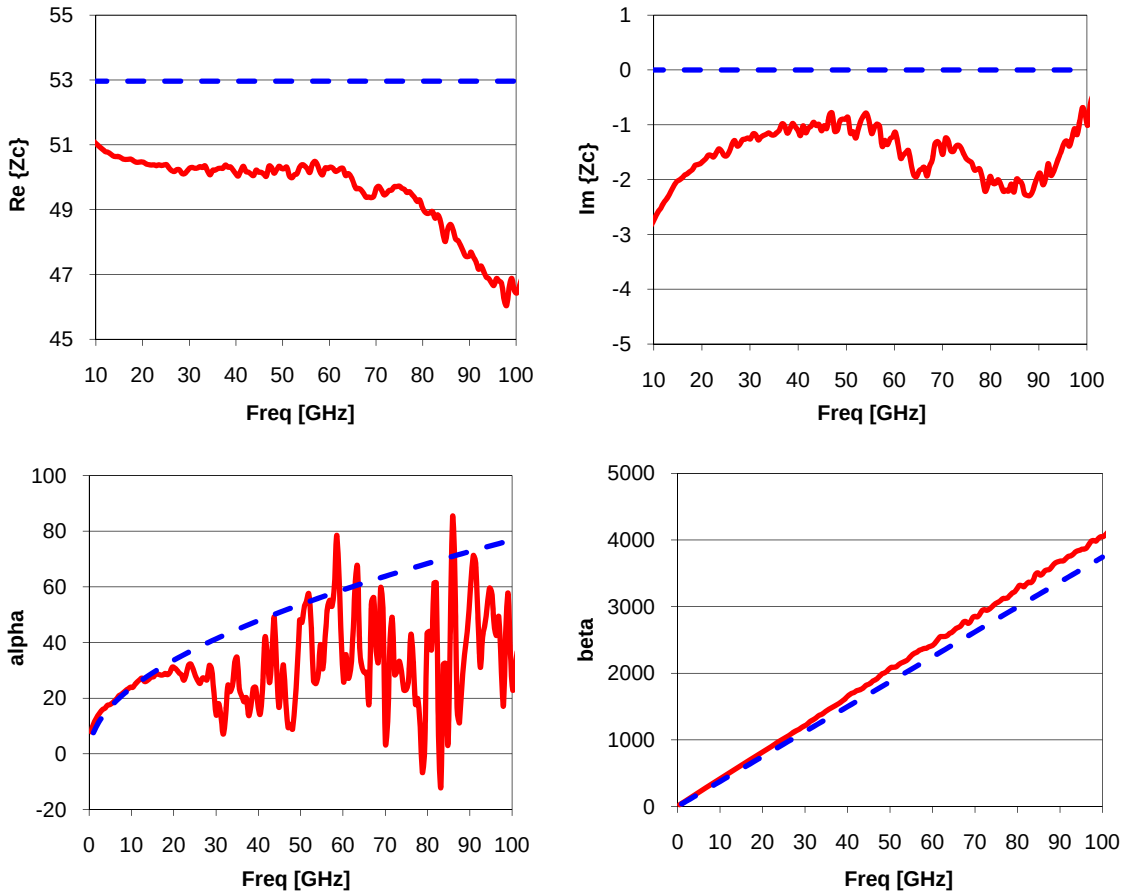


Fig. 17. Characteristic Impedance and Propagation Constant of a 504µm microstrip transmission line: split-thru de-embedded measures (line) vs models (dash).

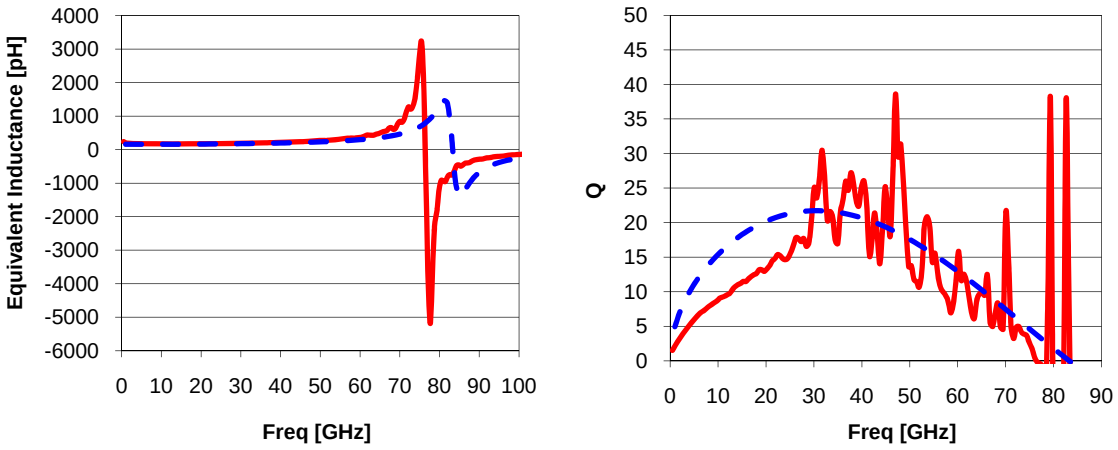


Fig. 18. Equivalent Inductance and Quality Factor of a 504µm microstrip transmission line: split-thru de-embedded measures (line) vs models (dash).

A second example is reported in Fig. 17. It concerns a 504µm microstrip line characterized by means of a similar split-thru de-embedding procedure, on the basis of S parameters measurements of two lines, of 600µm and 96µm, respectively. A 504µm roughly correspond to λ/4 at 78GHz. The equivalent inductance of the line is shown in Fig. 18, together with its quality factor. The measured inductance is 187pH at 30GHz and 365pH at 60GHz; the corresponding values predicted by the model are 178pH and 298pH, respectively. These inductance values are too high and the corresponding self resonating frequency is too low to be efficiently used in a circuit. Nevertheless, this line can be used to implement a λ/4 impedance transformer in 78GHz range applications.

Even if transmission lines realized on metal 6 and alucap provides best performances, a model for microstrips implemented on each metal level is also included in the design kit. Two examples are reported in Fig. 19. These models can be used to account for all the interconnection paths existing in the layout that can affect the circuit behavior. At millimeter wave frequencies the parasitic effects due to interconnections potentially have a strong impact on the devices performances; therefore, the availability of reliable models for interconnections is a crucial element for successful design. As explained in the next paragraphs, transmission line models will be largely used into design.

Fig. 19. Schematic view of transmission lines implemented in metal 1 or metal 2 (left) and in the other metal levels (from metal 3 to alucap) (right).

2.2.6 RF Pads

Pads for the input and output of RF signals are also critical elements because at the concerned frequency, pads can add significant losses that can result, in turn, into a degradation of the overall performances of a circuit.

Two elements must be considered to accomplish pads optimization:

- Losses in the pads must be minimized; the footprint of the pad must be made as small as possible to reduce the capacitance to ground. To prevent losses in the substrate a ground shield can be added under the pad but at the price of an increased equivalent capacitance.
- The behavior of the pads must be as close as possible to that of an ideal capacitor; in this case pads can be accurately modeled as lumped capacitors and their effect can be de-embedded with sufficient accuracy using a lumped elements-based procedure.

To evaluate the benefits of the use of a ground plane, three different structures have been realized and tested. A first pad made of metal 6 and alucap shunted together is implemented as in Fig. 20. The signal pad is placed between two ground pads, according to a ground-signal-ground (GSG) configuration. Both the ground pads are implemented as a stack of all metal layers and are directly connected to the substrate. A ground plane formed by slots of metal 1 and metal 2 interlaced is routed everywhere except under the signal pad that is placed directly over the substrate (a). A direct connection between the ground pads is added on the backside of the structure to provide a more uniform ground contact. In a second implementation of the same structure a shield made of patterned active fingers is added under the pad (b). Finally, in a third variant the ground plane of metal 1 and metal 2 has been extended under the pad (c). For simplicity active fingers and M1-M2 ground plane are represented in the figure as uniform layers.

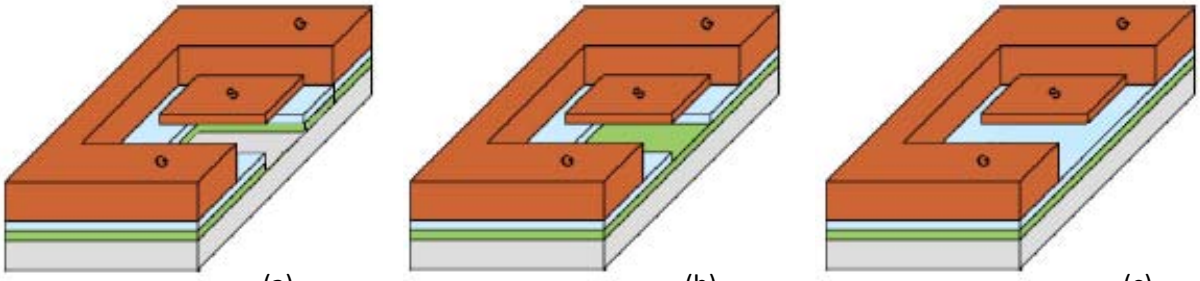


Fig. 20. Simplified view of the RF Pad implemented as M6 and alucap shunted together over the bare substrate (a), with a shield of active finger (b), and on a M1-M2 slotted ground plane (c).

The S parameters corresponding to the three structures have been measured and are reported in Fig. 21. It can be observed that the capacitive behavior of the pad is increased by the addition of a ground plane. In particular, the pad with a M1-M2 ground plane exhibits a better behavior as the corresponding S parameter plot approaches the unity-circle of the Smith chart as an ideal capacitor.

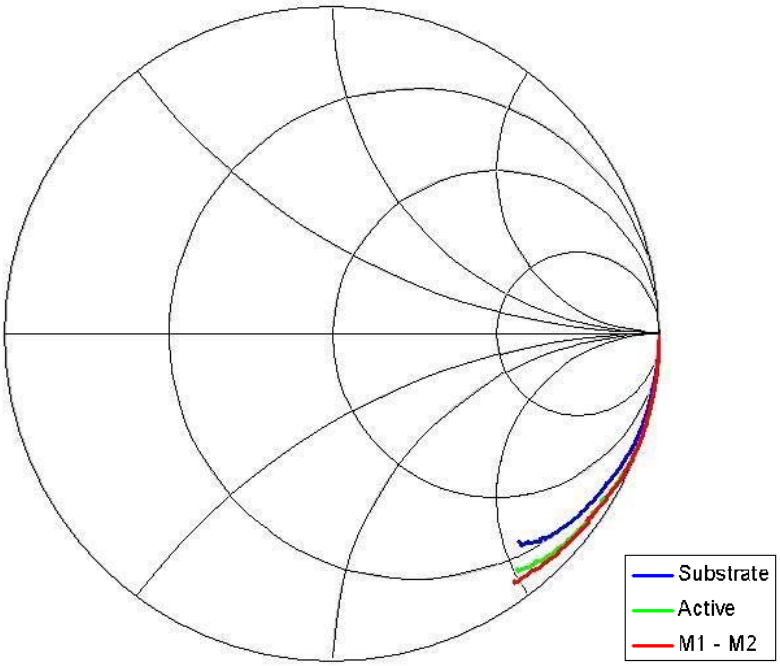


Fig. 21. S parameters of the pads on unity-circle bounded Smith Chart.

The equivalent capacitance of the pad is reported in Fig. 22. It can be extracted from its measured Y parameter according to the equation:

$$C = \frac{Im\{Y\}}{2\pi f} \tag{10}$$

As expected the capacitance is augmented by the presence of a ground plane, however, its value remains reasonable (15.3fF at 80GHz for the pad with M1-M2 ground plane), thanks to the elevated thickness of the BEOL.

In addition to the direct plot of S parameter on the Smith chart, the deviation from an ideal capacitance can be estimated also by means of the quality factor defined as:

$$Q = \frac{Im\{Y\}}{Re\{Y\}} \tag{11}$$

Despite of its slightly bigger capacitance, pad with M1-M2 ground shield exhibits the best performance and therefore must be used when the introduction of uncontrolled non-idealities is critical as in the case of passive devices measurements.

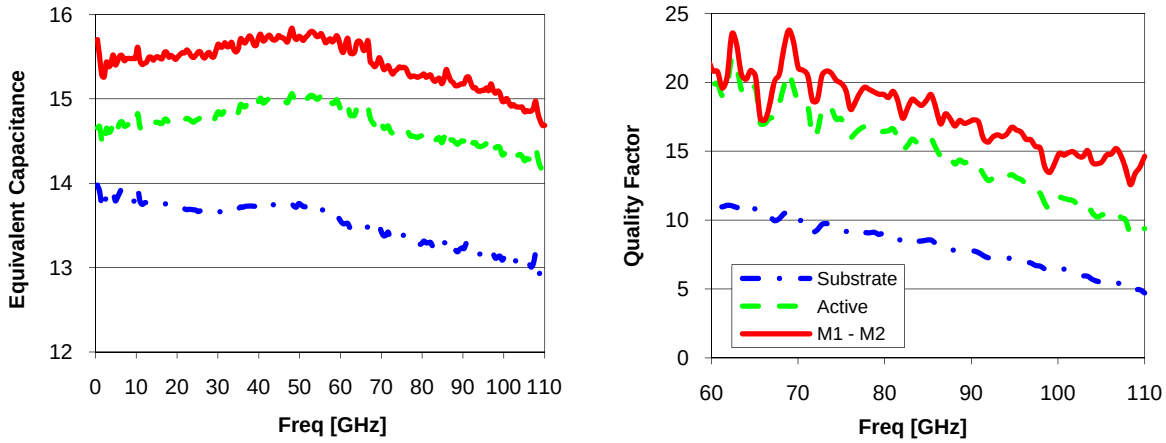


Fig. 22. Equivalent capacitance and quality factor of the pads, extracted from measurements.

Concerning circuit design, RF pads can be simply represented by their equivalent capacitances. Alternatively, they can be modeled by means of two transmission lines. Pads are divided into two parts following an ideal line across the centre, where the test probe is located, as shown in Fig. 23. Then, a transmission line models the connection between the probe and the circuit, whereas an open line represents the outer side of the Pad. DC pads contribution can be neglected if enough DC decoupling is provided.

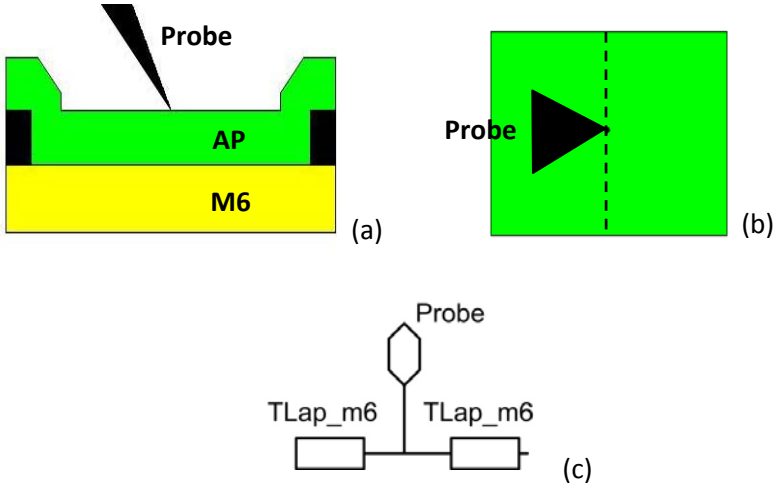


Fig. 23. RF Pad section (a), area (b), and its equivalent model (c). Pads are modeled as two T-Lines, one on each side of the test probe.

2.2.7 MIM Capacitors

The technology includes 2fF/µm^2 Metal-Insulator-Metal (MIM) capacitors using Si_3N_4 as dielectric, situated above the top cupper layer and under the alucap. Accesses to the capacitor plates must be routed in alucap and connected to metal 6.

Two models are available for simulations:

- M1, a simple model which accounts for intrinsic and parasitic capacitances (between the top/bottom plates and bottom plate/substrate or bottom plate/shield);
- M2, a model which accounts for intrinsic and parasitic capacitances and also for series resistances (bottom and top plate access).

Generally speaking these models give relatively similar results. However, it has been observed that the use of the M1 model instead of M2 in simulations can sometimes highlight some effects that are otherwise neglected. Some examples are reported in chapter 3. M2 model is reported in Fig. 24.

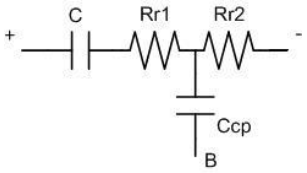


Fig. 24. Schematic view of the M2 model of MIM capacitors.

Despite of their simplicity, modeling of capacitors is one of the most critical aspects of millimeter wave design, since at the concerned frequencies their behavior depends on layout and is influenced by parasitic effects of the metal paths used to implement the capacitor accesses. These effects are not taken into account by any of the aforementioned models, as they represent only the intrinsic device (i. e., the capacitor standard cell). As a matter of fact, each one of the two capacitors terminations must be modeled as a series of two transmission lines. The inner line represents the alucap interconnection, whereas the other one accounts for the outer interconnection path that must be routed on metal 6 and alucap, as required by design rules. An example concerning a 3pF capacitor is reported in Fig. 25.

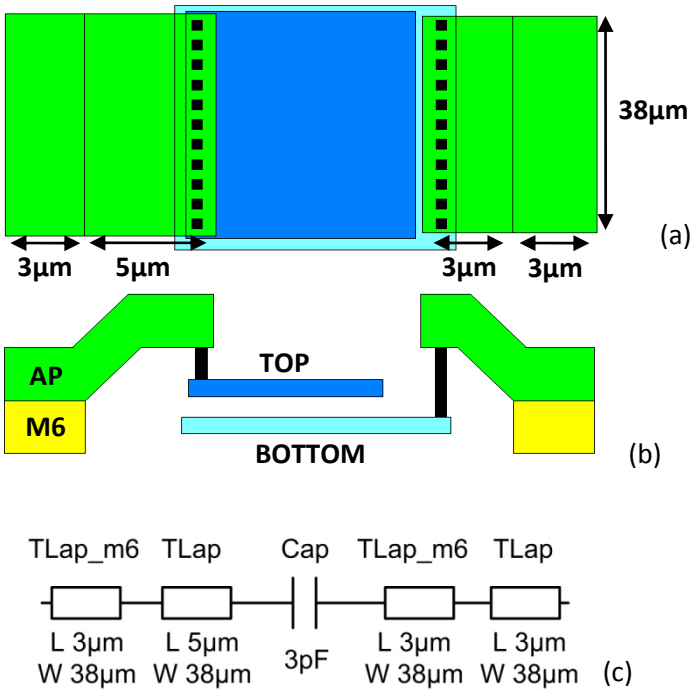


Fig. 25. 3pF MIM capacitor area (a), section (b), and its equivalent model (c).

It should be noted that the vertical profile of alucap layer in proximity of the capacitor plates is different from its standard configuration. As a consequence, the definition of a unique technology description file, valid for both capacitors and other devices such as transmission lines, as required by electromagnetic simulation tools, is difficult. Therefore, accurate electromagnetic simulations of MIM capacitors embedded into a wider circuit are prohibitive.

2.2.8 Resistors

Resistors are implemented in polysilicon layer or in P+ doped active zones. Different kinds of polysilicon resistors are available, with or without silicidation protection. The choice can be made according to the values of resistance allowed for each device. A summary is reported in Table 3.

Table 3. Summary of Resistor available in BiCMOS9MW technology

Device Name	Type	Sheet Resistance [Ω/square]
Rhipob	HIPO RESISTOR	1150
Rpo1b	N+ Unsilicided Poly RESISTOR	125
Rpo1pb	P+ Unsilicided Poly RESISTOR	390
Rpo1sab	N+ Silicided Poly RESISTOR	13
Rpdiffb	P+ Unsilicided Active RESISTOR	150

Two models are available for simulations:

- M1, a simple model which accounts for ideal resistor components;
- M2, an accurate model which accounts for ladder network resistor including bias non-linearity effect when data are available.

A schematic view of these models is reported in the Fig. 26. In the simple model (M1) the total resistance RES accounts for the intrinsic resistance corresponding to the body and for the resistances due to the side contacts. The intrinsic resistance is estimated as:

$$R = Rs \times \left(\frac{L}{W}\right) \tag{12}$$

where L and W are the width and the length of the device, respectively, and Rs is the sheet resistance of material.

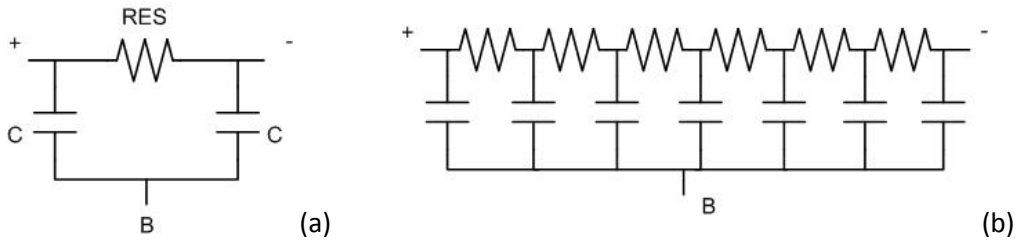


Fig. 26. Schematic view of the M1 (a) and M2 (b) resistor models.

2.2.9 Lumped Inductors

Lumped inductors are included in the design kit. Inductors have octagonal shape and are implemented in a stack of metal 5, metal 6 and alucap. To prevent the induction of eddy currents into the substrate, a pattern ground shield is routed under the inductor, formed by patterned N+ active fingers in Nwell and N+ poly fingers interlaced [35]. For single turn devices, the inner radius and the width of coil can be set by the user, according to the following condition:

- 20µm < Inner radius < 130µm

- 5µm < Width of coil < 29.99µm

The resulting inductance value is then scalable between 74pH and 635pH. Multi-turns inductors are also allowed, with inductance values above 358pH.

The structure of a typical single-turn inductor is shown in Fig. 27. Basically it reproduces the traditional configuration of lumped inductors used in the former BiCMOS9 node, but the frequency validity of models is now extended from DC to Self Resonance Frequency (SRF) of quality factor or to 110GHz, if SRF is higher than 110GHz.

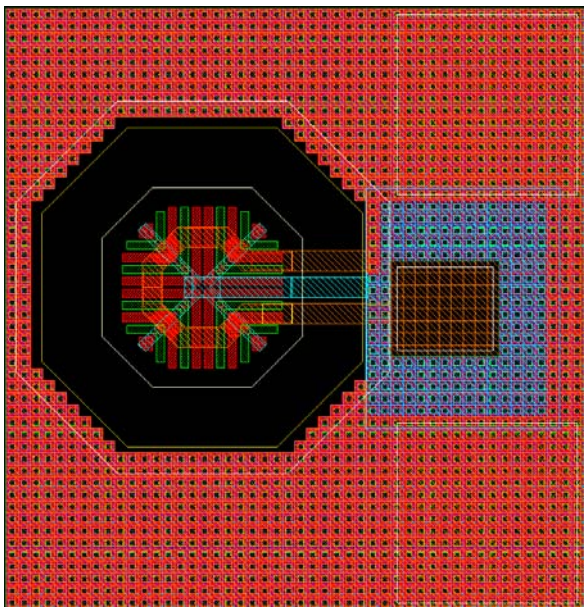


Fig. 27. Layout of a BiCMOS9MW 74pH Inductor with Pad (on the right).

Instead of searching for an optimization procedure dedicated to millimeter wave applications, a common approach has been followed for inductors design. For Instance, the use of a stack of three thick metal layers should reduce the inductor resistance and therefore improve its quality factor. The presence of a pattern ground shield, as well, should assure a better isolation of the device from the underlying substrate and, since losses in the substrate are largely recognized as the main source of quality degradation of inductors at millimeter wave frequencies, it should further improve the devices quality. As a result, quality factors above 20 can be actually achieved in the millimeter wave frequency range, but this seems to be a consequence of the optimization of the BEOL and not an advantage of the applied design strategy.

Unfortunately, inductors provided by the design kit can not be efficiently used, since octagonal inductors with large footprint and line width are not suitable for millimeter wave design. The range of allowed values of intrinsic inductance is not suited to the millimeter wave context, as inductances of less than 74pH are often required in millimeter wave applications.

For lumped inductors in the millimeter-wave regime, loss in the silicon substrate is the dominant mechanism for degradation in the quality factor. Therefore, to minimize these losses the inductor footprint should be made as small as possible. On the contrary the size of inductors included in the design kit is set regardless of such a constraint. Also, the area consumption is itself a concern, as the main advantage of inductor based circuits is that they may allow an important surface reduction, compared to distributed devices solutions.

Moreover, large inductors with both the terminations located on the same side require relatively long metal lines to connect them to the circuit, since the design rules prevent the presence of any other device in their proximity. At the concerned frequencies, the equivalent inductance of the outer interconnections can be comparable to that of the intrinsic device. As a consequence, the overall resulting inductance value is difficult to estimate as it depends on parasitic contributions and can be very different from that of the intrinsic device. Electromagnetic simulations are then required to predict the behavior of the overall structure and therefore the availability of a model of the intrinsic device becomes less attractive. Finally, the efficacy of ground shield at millimeter wave frequencies is not assured.

To overcome to these obstacles, inductors must be designed according to a specific methodology, as described in section 2.4. As an alternative, to take advantage of the technology facilities, inductors can be implemented as transmission lines.

2.3 Lumped or Distributed Approach

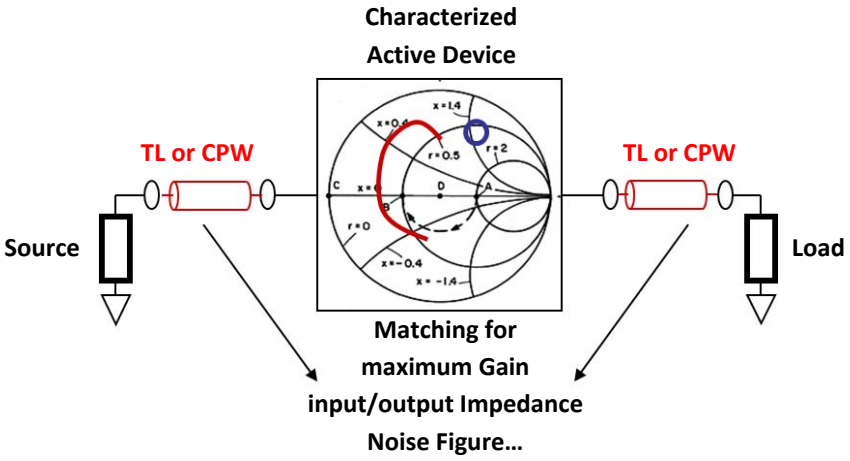


Fig. 28. Distributed design approach.

Traditional designs in the radio-frequency and microwave context can be conventionally divided in two categories according to the approach followed in circuit implementation: distributed approach or lumped approach. The former can be merely described as follows: active devices, characterized by S parameters, are tuned by distributed passive structures such as transmission lines and Co-Planar Waveguides (CPW) as depicted in Fig. 28. It is the original microwave design strategy and its efficiency has been so far proved in many examples, when the signal wavelength is comparable to the physical dimensions of devices. The latter, that is, the lumped approach is typical of radio-frequency/analog design and relies on the use of lumped passive devices combined with transistor models as resumed in Fig. 29.

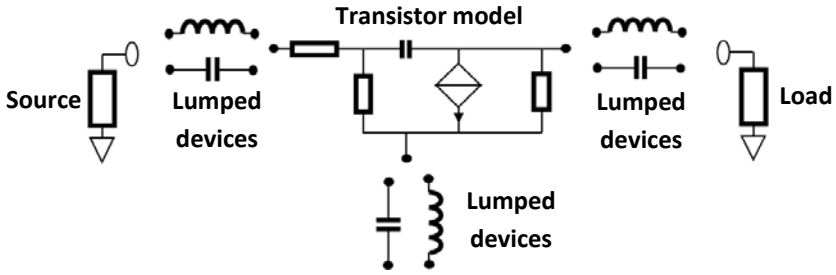


Fig. 29. Lumped design approach.

In the millimeter wave context, the availability of reliable transistor models offers to designers a characterization of active devices, but the dichotomy between lumped and distributed design is still a fundamental issue regarding the implementation of inductive devices, including inductors and transformers.

In the first examples of millimeter waves design implemented in silicon-based technologies, the use of distributed transmission lines was the dominant strategy, whereas the lumped approach suffered from the lack of efficient design and modeling procedures. However the drawback of the use of distributed structure is that they are area expensive, compared to lumped devices.

The search for area optimization led to the first circuit operating at millimeter wave frequencies entirely based on lumped inductors, a 52GHz two-stage cascode LNA implemented in a 0.18µm SiGe BiCMOS process, published in 2004 [Gordon 36]. Since then, many examples of lumped designs have been reported, extending the frequency range of validity of the lumped approach up to 140GHz [37].

Generally speaking, two main trends can be recognized in millimeter wave design: on one hand, the search for an optimization of distributed devices and, on the other hand, the setting of a reliable procedure for lumped inductors design.

In the first case, the optimization of distributed propagation structure usually implies a comparative analysis of the different kinds of available distributed devices such as microstrip transmission lines and CPW, and the evaluation of their aptitude to address the design specifications, on the basis of the different technologies, according to the properties of each BEOL. Besides the conventional microstrip and CPW transmission lines, also dedicated structures such as, for example, slow-wave transmission lines in CMOS back-end have been developed [31]. More details on the comparative design of distributed devices for millimeter wave application can be found in many examples reported in the dedicated literature [32]-[34].

In this work, design of distributed propagation structures is based exclusively on the models of microstrip transmission line provided by the BiCMOS9MW technology. Thanks to the availability of these devices, any further search for alternative distributed element seems to have no practical interest. To overcome the limitations of distributed approach (namely, to reduce area consumption), a design methodology for lumped inductors that has proved its efficiency will be described in the next section. A comparison between a lumped version relying on this methodology and a microstrip-based implementation of a LNA for 80GHz applications is reported in the next chapter.

2.4 Inductor design

2.4.1 Design strategy

Lumped inductors are largely used in traditional radio-frequency and microwave design and are usually included in all the building blocks, for on-chip matching networks and inductive peaking. In the first designs reported in millimeter wave context inductors are replaced by distributed elements such as microstrip or coplanar transmission lines. However, this approach proves to be area intensive, as quarter-wavelengths in silicon dioxide at the concerned frequencies are approximately 500µm. Given the prohibitive mask and fabrication costs in the deep-submicrometer technologies at which silicon becomes commercially feasible for millimeter-wave circuits, area consumption should be minimized if inexpensive components are to be realized. The use of lumped devices can strongly reduce the surface occupation since mutual coupling in spiral coils allows for larger inductance in less area than transmission line structures. The main drawback of using lumped inductors for millimeter

wave applications is the difficulty to realize inductors with satisfactory performances at the concerned frequencies. At 60GHz or beyond, the inductor behavior is affected by non-idealities such as skin effect and proximity effect, by energy losses in the substrate, and by inductive and capacitive coupling with the adjacent devices.

The aptitude of an inductor to properly work at the intended frequency can be estimated by means of its quality factor defined as the ratio between the electromagnetic energy stored by the device in a frequency cycle and the energy losses. The frequency at which the quality factor becomes equal to zero is its self-resonance frequency (SRF). Circuits employing inductors should operate below the peak-Q frequency of the inductor in order to ensure proper behavior, which emphasizes the need for high SRF.

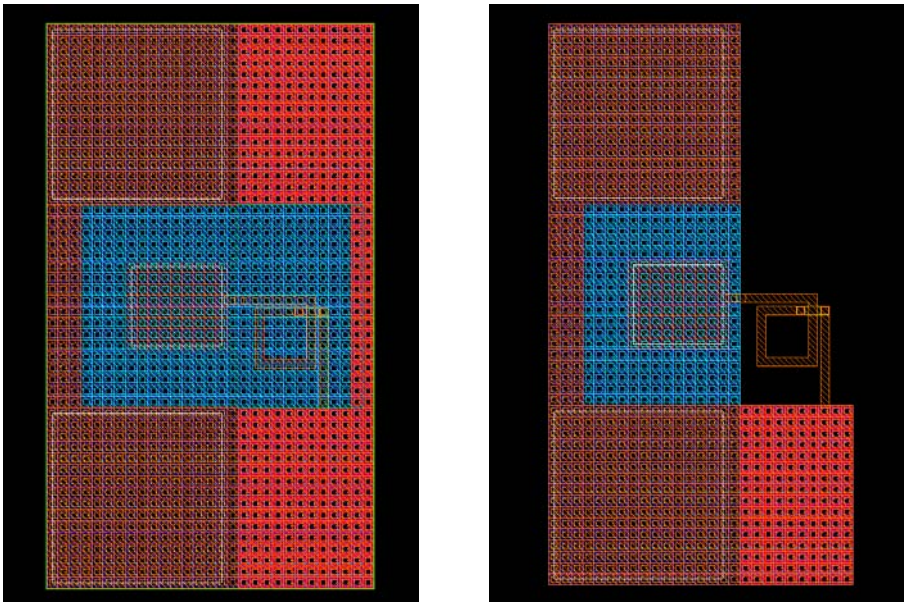
A procedure to design inductors with satisfactory millimeter wave frequency performances has been established since 2005 [38] and has been applied to several examples, both in CMOS and BiCMOS technologies. It relies on custom design of monolithic spiral inductors implemented in a single metal layer (planar inductors) or as three-dimensional structures (stacked inductors).

For silicon-based inductors in the millimeter-wave regime, loss in the silicon substrate is the dominant mechanism for degradation in the quality factor. To minimize these losses, the inductor footprint must be made as small as possible. This involves reducing outer diameter, as well as metal linewidth. While inductor design at radio frequencies emphasizes the use of wide metal lines to reduce series resistances, this approach increases capacitance to substrate and reduces the self-resonance frequency. The lower limit on linewidth can be therefore set according to reliability concerns due to electromigration effects.

Both square and octagonal spiral can be realized. Square inductors with accesses on the opposite sides are preferred since they can easily satisfy the design rules, and can be straightforward reproduced in electromagnetic simulation tools. If accesses are disposed on the opposite sides, the inductor can be easily connected to other devices, with no need of long interconnection paths that can add parasitic contributions to the intrinsic inductance value. Turn-to-turn spacing is made roughly equal to the dielectric thickness to substrate to lessen the impact of interwinding capacitances and further improve SRF. Wide metal spacing also diminishes the frequency dependence of the inductance set forth by the proximity effect. To take advantage of the thickness of the BEOL, planar inductors are implemented on the top copper layer (metal 6) or, as for transmission lines, on metal 6 and alucap shunted together, and with a metal 5 underpass. However, since the quality factor of inductors is determined by substrate losses and not by series resistance, alucap seems to be unnecessary.

According to the conventional radio-frequency approach, a patterned ground shield is added under the inductor to prevent the induction of eddy currents and therefore reduce the electromagnetic loss into the substrate [35]. At millimeter wave frequencies, however, the patterned ground shield should not be used, because it increases the capacitance to ground and reduces the SRF. For the same reason, even a continuous metal ground plane should be avoided. Nevertheless, a ground ring or even a ground wall around the inductor is useful to prevent undesired coupling effects with the adjacent components. The distance of the ground ring from the inductor must be optimized in order to reduce capacitive coupling and to provide a sufficiently short return path for signal currents. Indeed, a well defined grounding structure is necessary also to perform correct electromagnetic

simulations of the inductor because simulation tools could operate in arbitrary mode if no grounding structure is provided by the user.



(a) (b)

Fig. 30. Layout of a square spiral inductor for millimeter wave application implemented over a M1-M2 ground shield (a) and directly over the substrate (b) with Pad (on the left).

To evaluate the effects of the presence of a ground plane, the square spiral inductor of Fig. 30 has been designed. Two different versions of the same inductor have been fabricated and tested. Both e inductors are implemented in metal 6 with an underpass in metal 5. An inductor geometry with 4.5µm metal width, 2µm spacing, and an inner diameter of 23µm has been chosen to get an inductance value typical of millimeter wave applications. A ground shield made of metal 1 and metal 2 interlaced slots has been added under one of the inductors (a), whereas the other one is placed directly over the substrate (b). The equivalent inductance of the two devices has been extracted according to (13) from single port measurements of the S parameters after an open-short de-embedding of the pad and is reported in Fig. 31:

$$L = \frac{Im\left\{\frac{1}{Y}\right\}}{\omega} \tag{13}$$

$$Q = \frac{Im\{-Y\}}{Re\{Y\}} \tag{14}$$

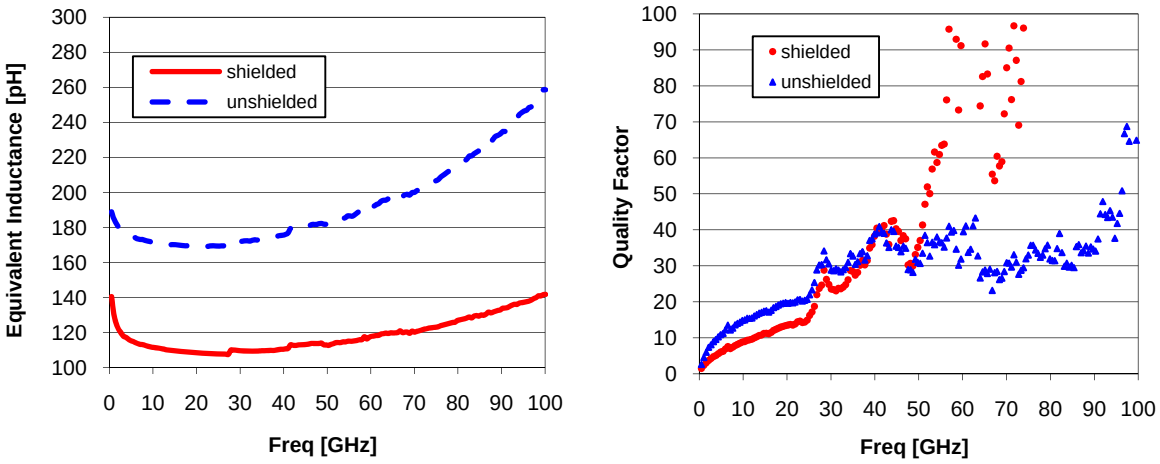


Fig. 31. Equivalent Inductance and quality factor of the square spiral inductor of Fig. 30 implemented in the two version, with (red) or without the ground shield (blue).

As expected, the main effect of the addition of the ground shield is a reduction of the equivalent inductance from 192pH to 118pH at 60GHz, and from 216pH to 127pH at 80GHz. The quality factor extracted from measurements, according to (14), is reported in Fig. 31. Below 40GHz, both the devices exhibit a similar behavior but the presence of the ground plane leads to a reduction of the quality factor, due to the increased capacitance to the ground. Beyond 40GHz no conclusion can be drawn as measurements of the series resistance are strongly affected by the incertitude due to the contact resistances, as explained in subsection 2.7.4. This effect is more pronounced in the case of the inductor with the ground shield because of its lower inductance. On the basis of these results we can conclude that inductors over substrate can provide higher inductance values up to 110GHz and that the introduction of a ground plane, on the contrary, reduces the inductor quality factor, at least below 40GHz.

Concerning the design methodology, it must be observed that the effects of the ground structure are differently modeled by the different electromagnetic simulation tools. For example, if simulations are performed on ASITIC, then a ground structure can be explicitly defined but it is taken into account only to fix a reference for capacitive substrate coupling, whereas resistive and inductive contributions due to the ground path are neglected. HFSS and Momentum simulators, on the contrary, account for the grounding structure, according to its geometry and material properties. In particular, when using HFSS, the presence of a grounding structure corresponding to the real circuit conditions is necessary, otherwise simulation results are affected by the ideal ground path required to connect the lumped ports.

Generally speaking, an accurate choice of the electromagnetic simulation tools is crucial for inductors characterization, as well as a proper setting of the simulation environment and parameters. More details on electromagnetic simulation setting procedures are reported in [39].

Besides planar inductors, also stacked inductors have been reported in millimeter wave applications [38]. Implemented in two or three metal layers, these structures benefit from strong mutual coupling between vertically adjacent metal layers, and can generate the same inductance in less area as compared with planar inductors. The use of stacked devices is advisable when high inductance values are required.

2.4.2 Equivalent-circuit models for lumped inductors

Electromagnetic simulators outputs usually consist of formatted files containing the S (or Y or Z) parameters of the simulated structure. These files can be imported in Cadence and used as input for circuit-level simulations such as DC and SP Analysis. However, simulations performed on the basis of external S parameter files are computationally expensive and can sometime result into convergence problems, namely if complex simulations such as Transient Analysis or PSS are demanded.

To facilitate the use of electromagnetic simulations results in circuit analysis, each device should be represented by a lumped equivalent-circuit model. The model parameters, in turn, are extracted from electromagnetic simulations outputs.

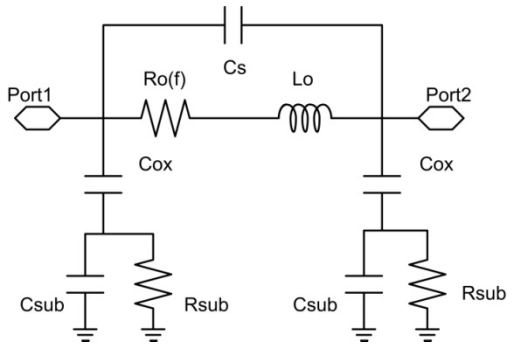


Fig. 32. Equivalent single-π circuit for lumped inductors.

Traditional radio-frequency approaches simply represent the inductor as a lumped single-π equivalent circuit, as shown in Fig. 32. In this scalable model, series metal resistance and inductance, feedthrough capacitance, dielectric isolation, and substrate effects are modeled.

Although physical considerations are included in such a structure, the single-π model lacks the following important features:

- strong frequency dependence of R and L as a result of current crowding in the conductor, i.e., both the skin effect and the proximity effect, which leads to significant degradation of the quality factor at gigahertz frequencies;
- distributed characteristics to match high-frequency behavior, especially for inductors with large dimensions; in addition, half of the capacitance to substrate is artificially neglected in simulations if one terminal of the inductor is connected to a small-signal ground;
- frequency-independent circuit elements for compatibility with transient analysis and broad-band design.

To overcome these inaccuracies, a wide-band equivalent 2π-circuit model for lumped inductors is reported in Fig. 33. For the model to be applicable over wide bands, the frequency dependence of the various loss mechanisms must be accurately captured.

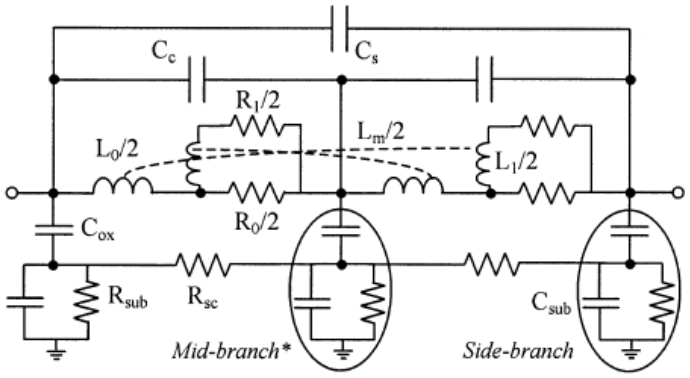


Fig. 33. Equivalent 2-π model for lumped inductors.

The energy loss in the metal is one of the primary sources of energy dissipation and is intrinsic to the spiral structure. As the frequency increases, the depth of current penetrating into the metal becomes comparable to or even smaller than the cross-sectional dimensions of the metal line, due to the skin effect. To accurately account for the resulting frequency-dependent resistance, a ladder circuit is developed to replace the series *R* and *L* in the single-π model.

In addition to the skin effect, the magnetic field generated by neighboring lines further changes the current distribution and results in a higher current density at the edges of the metal lines. This is described as the proximity effect and can be modeled by adding a mutual inductance between L0 and L1. As mentioned in the previous paragraph, the impact of the proximity effect can be neglected if turn spacing is sufficiently large.

Besides the energy loss in metal structures, also substrate loss must be considered:

- Ohmic loss in the conductive substrate, which is due to the displacement current conducted through the metal-to-substrate capacitance. It is modeled by a substrate RC network comprised of Cox, Csub , and Rsub.
- Loss due to eddy current in the underlying substrate, induced by the penetration of the magnetic field into the conductive silicon. This could be modeled by mutual inductance between metal and substrate. However, it can be neglected if high resistivity substrate is used (e.g., >10 Ω-cm).

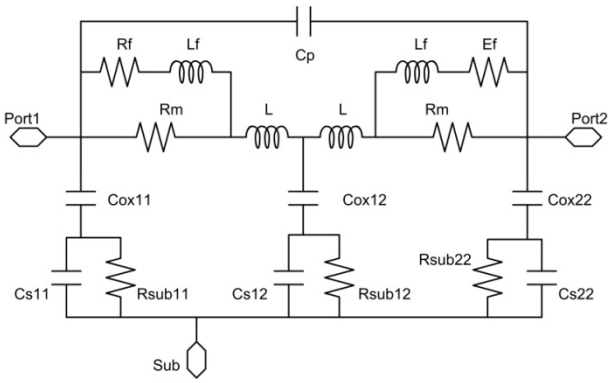


Fig. 34. Compact 2-π model for lumped inductors.

A closed-form solution for each element of the 2π-circuit model based on its physical origin is reported in [40]. However, from a practical point of view, this circuit can be simplified as shown in Fig. 34 without losing in accuracy [38]. The value of each component of the latter circuit model can be evaluated on the basis of 2-ports Y parameters extracted from electromagnetic simulations, according to the following equations:

$$2 \cdot L = \frac{Im\left\{-\frac{1}{Y_{21}}\right\}}{\omega}$$

$$2 \cdot R = Re\left\{-\frac{1}{Y_{21}}\right\}$$

$$C_{OX1} = \frac{\left[Im\left\{\frac{1}{Y_{11}+Y_{12}}\right\}\right]^{-1}}{\omega}$$

$$C_{OX2} = \frac{\left[Im\left\{\frac{1}{Y_{22}+Y_{21}}\right\}\right]^{-1}}{\omega} \tag{15}$$

$$R_{SUB1} = Re\left\{\frac{1}{Y_{11}+Y_{12}}\right\}$$

$$R_{SUB2} = Re\left\{\frac{1}{Y_{22}+Y_{21}}\right\}$$

$$C_{OX11} = \frac{C_{OX1}}{2}$$

$$C_{OX22} = \frac{C_{OX2}}{2}$$

$$C_{OX12} = C_{OX11} + C_{OX22}$$

$$R_{SUB11} = 2 \cdot R_{SUB1}$$

$$R_{SUB22} = 2 \cdot R_{SUB2}$$

$$R_{SUB12} = 2R_{SUB1}//2R_{SUB2}$$

The substrate capacitances C_{S11}, C_{S12}, and C_{S22} are determined from the dielectric relaxation time of the substrate and their associated substrate resistances R_{SUB11}, R_{SUB12}, and R_{SUB22}, respectively, such that:

$$R_{SUB}C_S = \varepsilon_r \varepsilon_o \rho_{Si} \tag{16}$$

where ε_r is the relative permittivity of the substrate (11.7 for silicon), ε_o is the permittivity of free-space, and ρ_{Si} is the substrate resistivity in Ωm.

The Y parameters to be inserted in the former equations are calculated from a single step frequency simulation at low frequency (0.1GHz). Then series inductance L and skin-effect parameters R_f and L_f are optimized, as well as C_P, to obtain broad-band agreement between the equivalent circuit and simulated or measured -parameters. The optimization procedure can be automatically executed on a dedicated tool such as ADS.

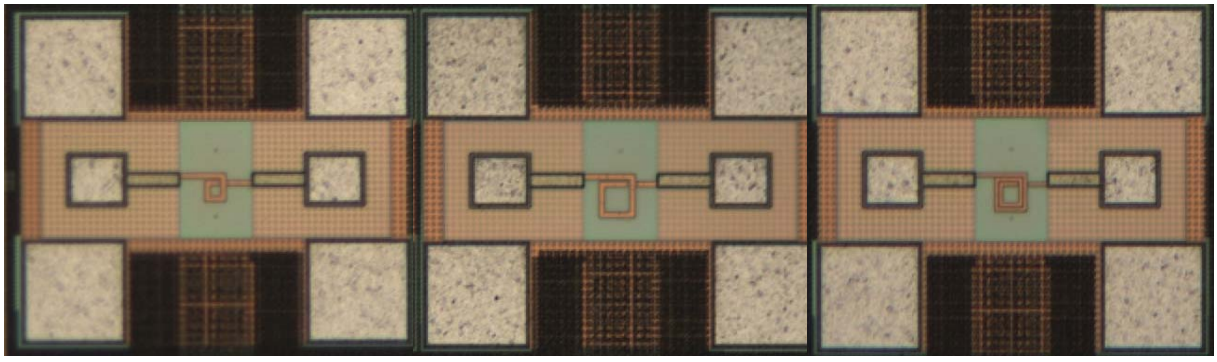
2.4.3 Overview of the design flow

So far the inductor design flow has been described. It can be summarized as follows:

1. Inductance value is chosen according to circuit-level simulations;
2. Inductor layout parameters (shape, diameter, linewidth and spacing) are fixed in order to reach the expected inductance value and sufficient quality factor at the operative frequency; SRF should be as high as possible;
3. S parameters are extracted from electromagnetic simulations performed on the whole structure, including the intrinsic device, the interconnections and the ground ring, if any;
4. S parameters are converted to Y and the elements of the equivalent circuit are evaluated from Y parameters, according to equations (15-16); optimization by means of dedicated tools is still possible;
5. The equivalent circuit is included in the building block and circuit-level simulations are performed to validate the inductor design.

Unfortunately only general criteria are available to predict the dependence of the inductor behavior from its layout. Therefore it is difficult to choose the inductor geometrical structure that corresponds to the desired inductance value. As a consequence, steps from 2 to 5 must be executed iteratively to obtain the desired inductance.

2.4.4 Examples of inductor design



(a) (b) (c)

Fig. 35. Chip microphotographs of spiral inductors fabricated in BiCMOS9MW technology. At 80GHz, inductance values of 81pH (a), 122pH (b), and 202pH (c) are obtained, respectively.

To validate the design methodology described above, three square spiral inductors of different geometries and values have been designed and tested (Fig. 35). All these devices are implemented in

metal 6 over substrate, with a metal 5 underpass. Inductance values have been chosen to cover the range of inductance usually required in millimeter wave applications, roughly between 50pH and 300pH. The smallest fabricated inductor is a single-turn spiral with 4.5µm metal width, 2µm spacing, and an inner diameter of 10µm. Interconnects, always included into device design and modeling, are realized by extending the most external spiral segment of 20µm and adding a 19µm metal 6 segment in series to the underpass (a). The middle-value inductor (b) is also designed as a single-turn spiral with 4.5µm metal width, 2µm spacing, and an inner diameter increased to 23µm. Interconnects of 11µm and 15µm are added respectively in series with the spiral and with the underpass. The biggest inductor (c) is a two-turn spiral with 3µm metal width, 1µm spacing, an inner diameter of 15µm, and 15µm symmetrical interconnects added on each side.

$$\mathrm{L} = \frac{\mathrm{Im}\left\{-\frac{1}{Y_{21}}\right\}}{\omega} \tag{17}$$

$$Q = \frac{Im\{-Y_{11}\}}{Re\{Y_{11}\}} \tag{18}$$

The measured equivalent inductances and quality factors, corresponding to each device have been extracted from two-port measurements of S parameters according to (17-18). They are reported in Figs. 36-38 and compared to electromagnetic simulations performed on ASITIC and Momentum and to the equivalent 2-π model extracted from simulations (Annex 1). It can be remarked that the plots of quality factor for the two smallest inductors as shown in Fig 36 and 37, beyond 40GHz are somewhat irregular and give inaccurate results. This is probably due to the inaccuracy introduced by the contact resistance between the probe tips and the aluminum pads. Indeed, accuracy requirements in the case of quality factor measurements of small devices at millimeter wave frequencies are extremely high, as resistance values around one ohm have to be determined precisely. As a consequence, the impact of the contact parasitic resistance becomes crucial and should be corrected before measurements by a calibration procedure based on aluminum standards that, however, are not available at IMS laboratory. Comparing the plots of Figs. 36-38, it can be noticed that the incidence of this phenomenon is inversely proportional to the physical dimension of the device under test: it is more accentuated for the smallest inductor, it becomes less evident in the case of the middle-value inductor and it disappears almost completely in the measurement of the biggest inductor for which a correct quality factor can be extracted. This behavior can be explained since the series resistance of each device increases with the ratio between the equivalent length and the width of its metal path and an increased series resistance results into a reduced sensitivity to the incertitude due to contact resistance. An approximated method to correct this phenomenon is described in subsection 2.7.4. As demonstrated in Figs. 46-48, satisfactory results can be obtained.

Beside the inaccuracy on the high-frequencies measurement of quality factor for the two smallest devices, a good agreement between measurements and simulations results and models behavior is shown allover the considered bandwidth, proving the efficacy of the proposed design methodology. It can be observed that compared to the other devices, the inductor of Fig. 35 (c) is more sensible to proximity effect and has a more pronounced frequency-dependence due its two-turn implementation with an inter-winding spacing of only 1µm. Equivalent inductance of single-turn devices with 2 µm spacing, on the contrary, is almost constant on a larger bandwidth.

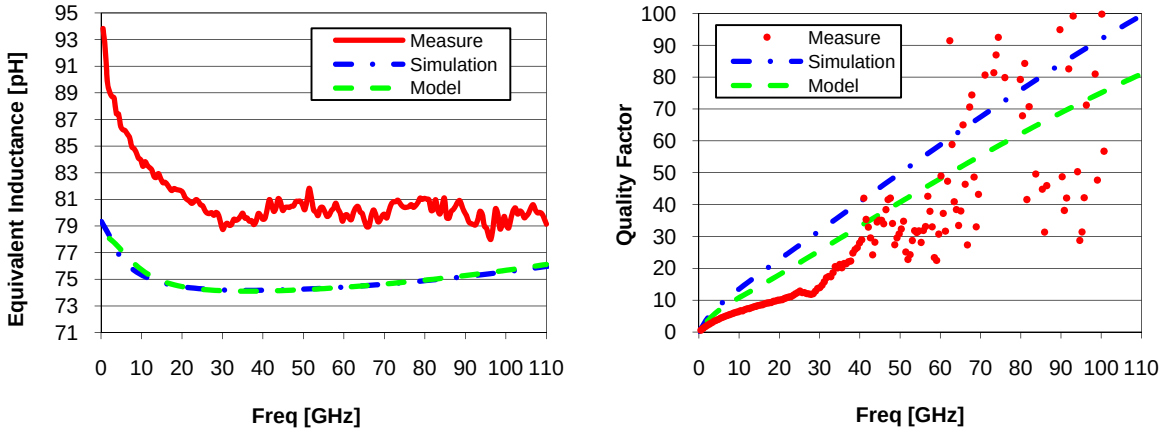


Fig. 36. Equivalent Inductance and Quality Factor for the inductor of Fig. 35 (a). Measurement results (red) vs simulations (blue) and model (green).

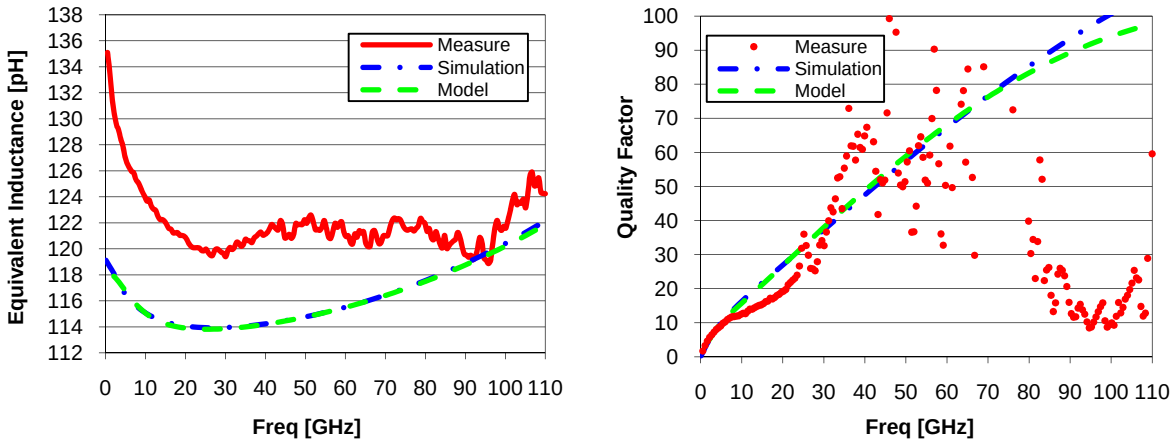


Fig. 37. Equivalent Inductance and Quality Factor for the inductor of Fig. 35 (b). Measurement results (red) vs simulations (blue) and model (green).

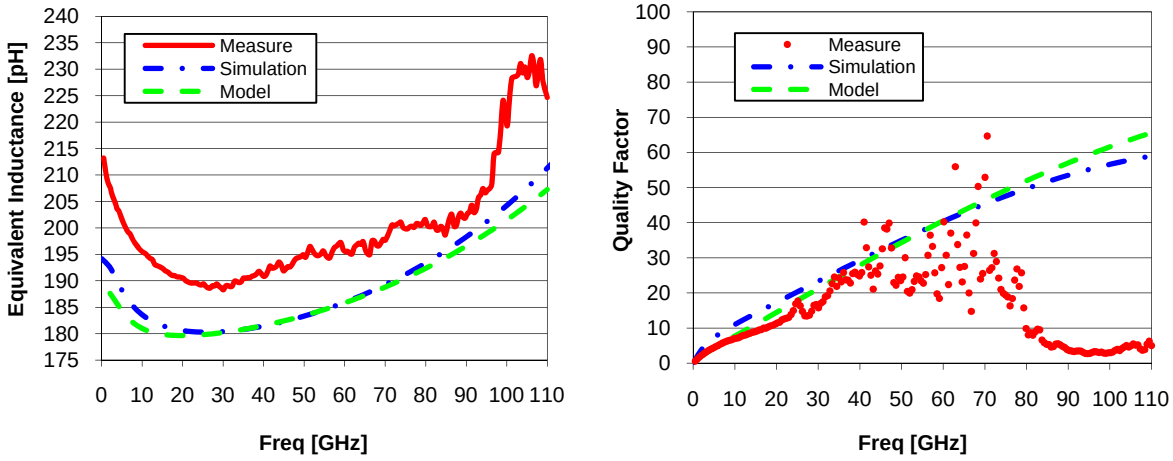


Fig. 38. Equivalent Inductance and Quality Factor for the inductor of Fig. 35 (c). Measurement results (red) vs simulations (blue) and model (green).

2.5 Interconnection modeling

The design facilities provided by the BiCMOS9MW technology have been described earlier in this chapter. So far, it has been proved that this technology can successfully address the main goals of millimeter wave design. At same time it has also been shown that the support provided by the technology is not exhaustive and a special effort is demanded to designers to overcome the technology limitations.

As already explained, if a lumped device approach is followed, then inductors must be designed according to a specific strategy as discussed in section 2.4. In addition, even if inductive devices are implemented as microstrip transmission lines, designers have still another problem to face, that is, the lack of a reliable automatic procedure for parasitic extraction.

At millimeter-wave frequencies, the effects of layout parasitic elements cannot be neglected; otherwise a strong frequency down-shift will occur between simulation and measurement results. In particular, the importance of modeling parasitic effects at transistor level and on the accesses of MIM capacitors has been highlighted in subsections 2.2.4 and 2.2.7, respectively.

The parasitic effects due to interconnections potentially have a strong impact on the global performances of a circuit; as a general rule, none of the interconnections existing in layout, between the different circuit devices, can be considered as a short circuit contact and each of them must be modeled as a distributed element. Layout interconnections can be modeled as transmission lines of corresponding metal levels and geometries. An example, concerning the connection between two lumped resistors, is reported in Fig. 39. The design kit of the BiCMOS9MW technology includes a model for microstrips implemented on each metal level. The availability of reliable models for interconnections is a crucial element for successful design.

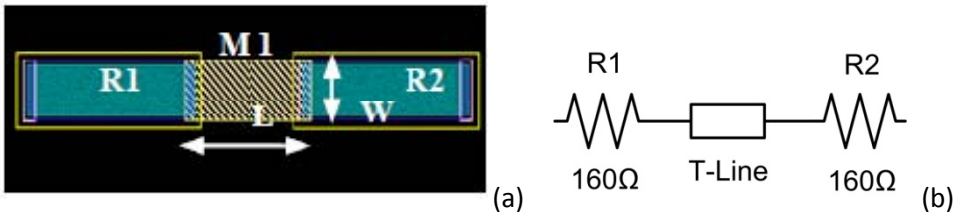


Fig. 39. Interconnection between two resistors (a) modeled as a transmission line (b).

The model of transmission lines included among the design facilities can represent straight connections with sufficient accuracy. Sometimes, in a circuit layout, three metal paths converge into a unique node, creating a tee-junction, as shown in Fig. 40 (a). The simplest way to account for such a structure is to model it with three transmission lines connected together in a common node. However, how to fix the geometry of each line in order to faithfully reproduce the physical structure of a tee-junction is somewhat unclear. As a consequence, the three-line model of Fig. 40 (b) can be inadequate. In particular, the main parasitic contribution of a tee-junction is the capacitance between its middle region and the substrate (or the underlying ground plane). If the junction is modeled as an ideal node between three microstrip lines, like in the circuit of Fig. 40 (b), this capacitance is obviously underestimated. Actually, in many cases this contribution can be neglected

without any critical consequence. Nevertheless, sometimes the effect of a tee-junction can be a crucial element. In such a case, usually occurring in sensible nodes like matching networks, electromagnetic simulations of the complete structure can lead to a more accurate characterization. Simulation results, given in the form of S parameters files, can be then back-annotated at circuit-level.

To avoid any convergence problem that could arise from the direct use of electromagnetic simulation results in circuit-level simulations (as for lumped inductors characterization), the equivalent-circuit model of Fig. 40 (c) can be employed. The proper value of its components can be extracted on the basis of electromagnetic simulations, by hand or with an automatic optimization procedure (ADS). This model is then inserted in the top level circuit to account for the corresponding tee-junction.

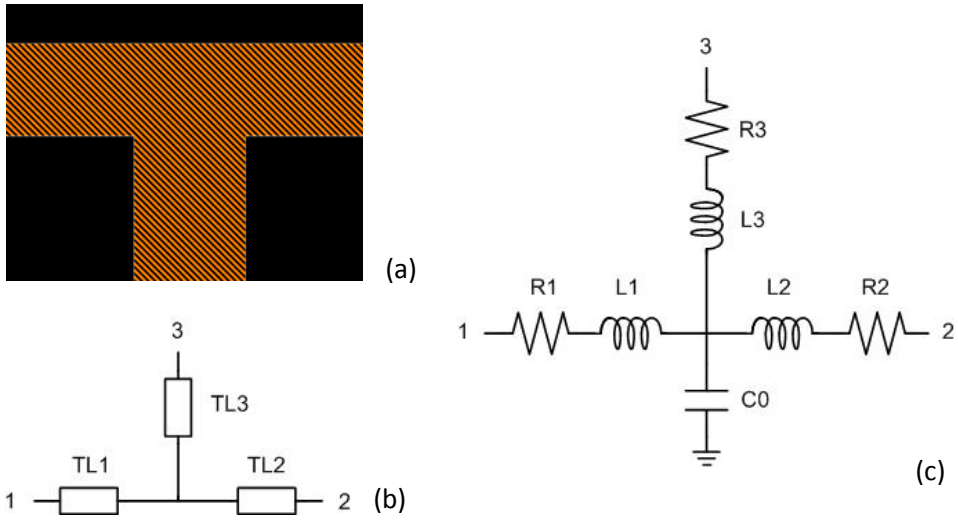


Fig. 40. Layout interconnects forming a tee-junction on metal 6 over substrate(a). Simple equivalent model using three transmission lines (b). Accurate model with lumped elements(c).

2.6 Overview of the Design Flow for mm-Waves building blocks.

An accurate modeling of devices and interconnections is crucial for successful post-layout simulations. However, the need of a complete modeling of all devices and interconnections augments the complexity of the design methodology.

The design of an amplifier for millimeter wave applications usually starts with the definition of its active core. Firstly, transistors sizes and geometries are set. Then parasitics are extracted at transistor level and their effects are annotated on the transistor models by means of the RC equivalent circuit of Fig. 13. Accounting for parasitics can imply a re-design of transistors layout, in order to compensate for the non-idealities introduced by the RC elements.

As a second step, passive devices networks made of capacitors and transmission lines or lumped inductors are designed and added to the amplifier core, to assure gain peaking and input/output matching. Ideal devices are firstly considered to globally address the required specifications and layouts are defined, neglecting parasitic contributions. Then post layout extraction is performed, accounting for capacitor accesses, pads, tee-junction between transmission lines, and

interconnections. Each time, the introduction of a new effect can require a partial re-design of some element in order to preserve a correct behavior of the overall circuit. As a consequence of post layout evaluation, the ideal device-networks set at the beginning must be modified to compensate for layout parasitics. However layout modifications, in turn, require a new post layout extraction and therefore the design flow becomes iterative.

Generally speaking, the design flow of millimeter wave circuits can be represented by the diagram of Fig. 41. Whereas in conventional radio-frequency and microwave contexts, circuit design, layout, and post-layout simulations are considered as three independent and consequent steps of a linear design flow, the design flow of millimeter wave circuits runs iteratively between circuital and layout design. In addition to amplifiers, the design methodology summarized in Fig. 41 can be successfully applied to the other building blocks of a millimeter wave front end like mixers, VCOs, and even PAs.

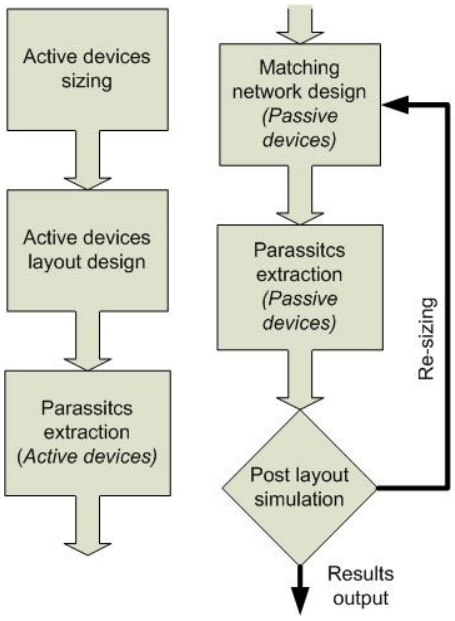


Fig. 41. Design flow for mm-Waves building blocks.

2.7 Tips for measurement de-embedding

2.7.1 Open-source

A de-embedding procedure commonly used is known as “open-short” as it requires two dummy structures, reproducing respectively an open contact between the pads and a short circuit between the accesses to the device under test and the ground. The open dummy reproduces the same configuration of the DUT with identical pads and devices accesses but without the device that is replaced by an open contact. It should account for the influence of pads and terminations that can be represented by an equivalent admittance connected in parallel with the DUT on each one of its terminals. In the short dummy the DUT is replaced by a low-resistive contact between the input and output signal path and the ground. It is used to compensate for the series impedance due to the accesses and to the pads.

The open correction is applied first and it consists of a subtraction of the Y-matrix corresponding to the open dummy from the DUT and the short dummy as well. The resulting Y-matrices of the open de-embedded DUT and short dummy are converted into their equivalent Z-matrices and the short de-embedding is applied, simply subtracting the Z-matrix corresponding to the short structure from the Z-matrix of the DUT.

The use of the open-short correction implies that all the elements to be corrected can be modeled as lumped admittances and impedances with sufficient accuracy. However, this assumption is not always verified in millimeter wave context because at the concerned frequencies the distributed nature of the interconnects linking the pads with the device can not be neglected. To overcome the limits of lumped approximation, more accurate de-embedding techniques based on transmission lines have been proposed. As an example, the "split-thru" de-embedding is described in the next paragraph.

2.7.2 Split-thru

The so-called "split-thru" de-embedding procedure has been recently proposed [41]. It relies on the modeling of device accesses as transmission lines and therefore is not affected by the errors introduced by lumped approximation. Two transmission lines of different lengths are required. The two lines must reproduce the same structure of the DUT terminations and must have the same width, but can have different lengths. If input and output accesses to the DUT are different in their structure or width, then two set of two transmission lines are required.

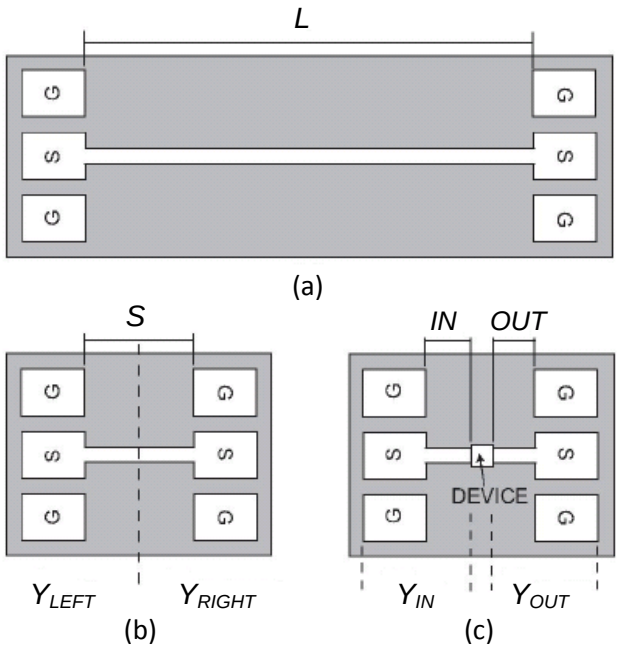


Fig. 42. Test structures required for split-thru de-embedding: long transmission line (a), short transmission line (b), and device under test (c).

The complete procedure can be divided into two parts. The first one leads to a complete characterization of a transmission line corresponding to the difference between the long and short lines used for de-embedding and therefore can be also used as a stand-alone procedure for the experimental characterization of transmission lines [42]. Then, in the second part, the de-embedded electrical parameters of the line are used to account for the DUT interconnections.

The two dummy lines can be represented by their equivalent S, Y, Z or M(=ABCD) matrices. Due to the symmetry of transmission lines the corresponding matrices are also symmetrical and can be swapped. Each line can be considered as the cascade connection of an input pad, an intrinsic (with no pads effects) line and an output pad. In terms of M-matrices the long line and short line dummies can be represented as:

$$M_L^t = M_{P1} M_L M_{P2} \tag{19}$$

$$M_S^t = M_{P1} M_S M_{P2} \tag{20}$$

where M_i represents the intrinsic line segment of structure i, M_{P1} represents the left pad, M_{P2} represents the right pad and *L* and *S* respectively are the length of the short and long lines.

The effect of the right pad can be suppressed by multiplying M_L^t by the inverse of M_S^t:

$$M_{L-S}^h = M_L^t \times [M_S^t]^{-1} = M_{P1} M_L M_S^{-1} M_{P1}^{-1} = M_{P1} M_{L-S} M_{P1}^{-1} \tag{21}$$

where M_{L-S} is the M-matrix corresponding to an intrinsic transmission line of length *L*-*S*.

If the left pad can be modeled by a lumped admittance Y_P, then:

$$M_{P1} = \begin{bmatrix} 1 & 0 \\ Y_P & 1 \end{bmatrix} \tag{22}$$

The hybrid matrix (21) can be converted into its equivalent Y-matrix and expressed as a parallel combination of the intrinsic transmission line and the parasitic lumped pads:

$$Y_{L-S}^h = Y_{L-S} + \begin{bmatrix} Y_P & 0 \\ 0 & -Y_P \end{bmatrix} \tag{23}$$

Then the effects of the pads can be cancelled by connecting the element Y_{L-S}^h with a swapped version of itself:

$$Y_{L-S} = \frac{Y_{L-S}^h + swap(Y_{L-S}^h)}{2} \tag{24}$$

The obtained Y_{L-S} is the Y-matrix representation of the transmission line of length *L*-*S*. It can be converted into its equivalent M form, thus providing an exhaustive characterization of the line of length *L*-*S*:

$$M_{L-S} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} cosh\gamma(L-S) & Zc\ sinh\gamma(L-S) \\ Zc^{-1} sinh\gamma(L-S) & cosh\gamma(L-S) \end{bmatrix} \tag{25}$$

$$Zc = \sqrt{\frac{B}{C}} \tag{26}$$

$$\gamma = \frac{cosh^{-1}A}{l} \tag{27}$$

This characterization technique has been used to extract the electrical characteristics of a 1020µm and a 504µm microstrip lines, from the on chip measurements of two set of dummy structure: 1116µm/96µm and 600µm/96µm, respectively. The results have been reported in paragraph 2.2.5.

For the de-embedding of a lumped device [41] the first step is the splitting into two parts of the short dummy line, obtaining the two matrices Y_{LEFT} and Y_{RIGHT}, as suggested by Fig. 42. If the transmission line is short such that $|\gamma S| \ll 1$, then an approximated expression of Y_{LEFT} can be obtained from the measured Y parameters of the short line, without de-embedding, as:

$$Y_{LEFT} = \begin{bmatrix} Y_{11} - Y_{12} - \frac{\gamma S}{4Zc} & 2Y_{12} \\ 2Y_{12} & \frac{\gamma S}{4Zc} - 2Y_{12} \end{bmatrix} \tag{28}$$

Y_{RIGHT} can be obtained by interchanging the rows and columns of Y_{LEFT} :

$$Y_{RIGHT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \times Y_{LEFT} \times \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{29}$$

Knowing Y_{LEFT} and Y_{RIGHT}, it is possible to deduce Y_{IN} and Y_{OUT} by appending or subtracting to Y_{LEFT} and Y_{RIGHT} a transmission line of length $|IN - S/2|$ and $|OUT - S/2|$, respectively. In terms of M matrices:

$$M_{IN} = M_{LEFT} \times M_{IN-S/2} \tag{30}$$

$$M_{OUT} = M_{RIGHT} \times M_{OUT-S/2} \tag{31}$$

where $M_{IN-S/2}$ and $M_{IN-S/2}$ are the M matrix representations of the transmission lines of length $|IN - S/2|$and $|OUT - S/2|$ obtained applying the equations (21-27) to the short and long dummy lines, rescaled to the proper length. It can be noticed that if the device accesses are symmetrical and their length is equal to *S/2*, then this last step can be skipped.

The M matrix corresponding to the de-embedded DUT can be extracted from the measurements as:

$$M_{de-embedded\ DUT} = M_{IN}^{-1} \times M_{measuredDUT} \times M_{OUT}^{-1} \tag{28}$$

The complete split-thru de-embedding has been applied to the measurement of the lumped inductors reported in paragraph 2.4.4.

2.7.3 Comparison between open-short and split-thru de-embedding

All the measurement results concerning the transmission lines characterization and the lumped inductors design reported in the previous paragraphs were obtained applying the split-thru de-embedding procedure described here above. To gain a further inside into de-embedding issues, it can be interesting to reconsider these results and to compare them to the case of an open-short de-embedding. To make the comparison between the two techniques possible, 48µm terminations with the same structure as the microstrip lines are included in the open and short dummies, as well as in

the test structures used for inductors measurements. Each termination length is therefore equal to one half of the 96µm thru and so the open-short and the split-thru de-embedding can be applied in the same conditions. As a first example, the characterization of the 504µm microstrip line already described in Figs. 17-18 is reported in Fig. 43 and compared to its equivalent open-short de-embedded version. Similar results are obtained with both techniques.

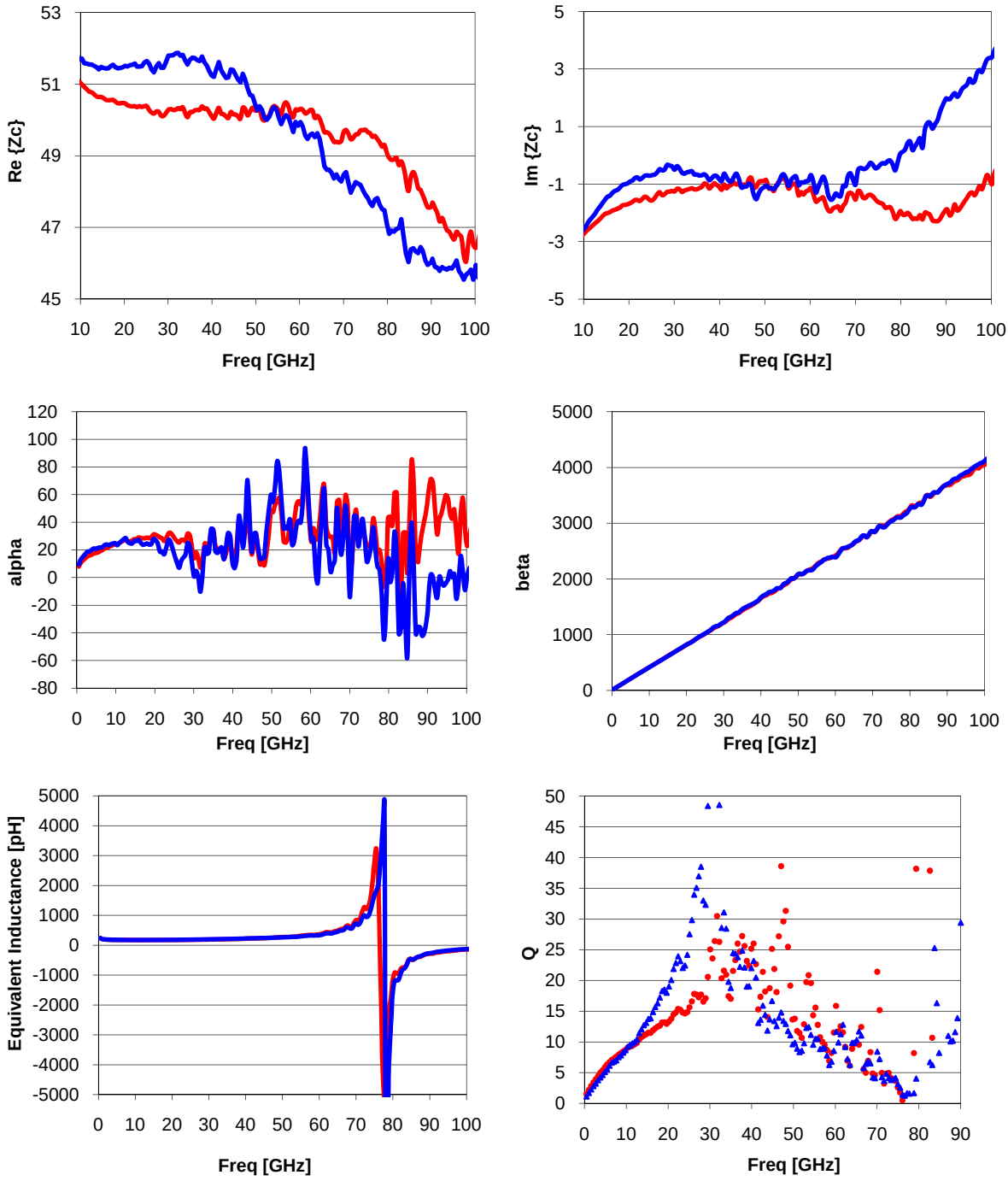


Fig. 43. Split-thru (red) vs Open-short (blue) de-embedding of the measurements of the 504µm microstrip line.

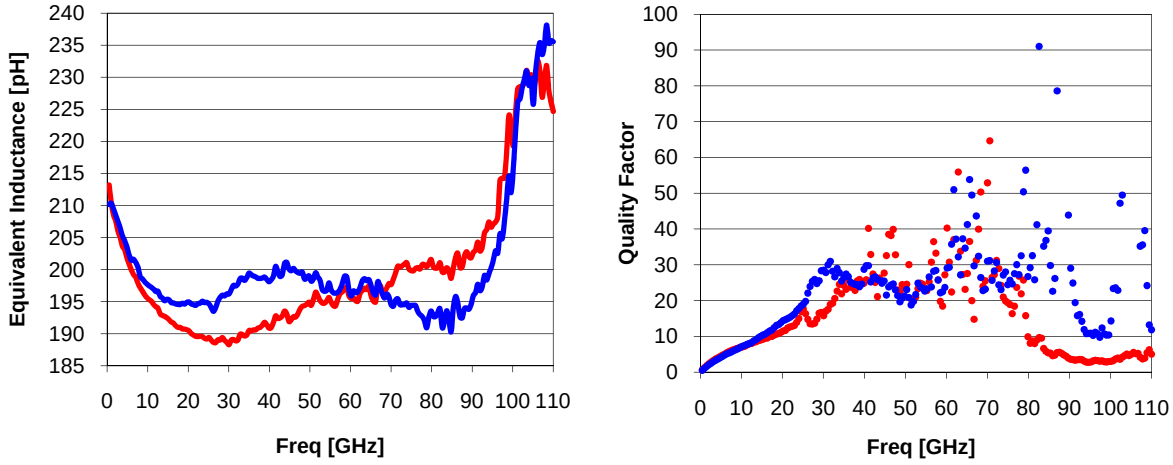


Fig. 44. Open-short (blue) vs split-thru (red) de-embedding of the measurements of the inductor of Fig. 35 (c).

A second example, concerning the inductor of Fig. 35 (c) is reported in Fig. 44. In this case, a less regular behavior in the range between 25GHz and 90GHz can be observed in the inductance resulting from open-short de-embedding. The split-thru technique is more accurate.

2.7.4 Contact resistance

All the measurement results reported in this chapter have been obtained using the E8361A PNA network analyzer, equipped with the N5260A millimeter head controller by Agilent™. Prior to measurement, a SOLT calibration procedure is applied, using an Impedance Standard Substrate (ISS) with high precision calibration standards, like Short, Open, Line and Thru. By this first calibration, the reference plane is shifted to the probe tips. However, it should be mentioned that the calibration is done on a substrate with gold metallization, while the pads on the chip are made of aluminum. Unfortunately, the use of a standard material for the VNA calibration that is different from the aluminum of the pad introduce a parasitic resistance due to the non-ideal contact between the probe tips and the alucap of the pad that can not be properly corrected by a calibration based on gold-made standards. Therefore, in our case, even after the calibration of the VNA, the contact resistance

$$R_{contact,diff} = R_{contact,Al} - R_{contact,Au} \quad (29)$$

is still present and affect the measurement results. The influence of such a parasitic element can be neglected if it is small compared to the series resistance of the DUT, a condition that is easily verified in the case of complex structures such as active circuits or long transmission lines; but it can be extremely important when small passive devices with a small series resistance are to be measured. In particular, when measuring the quality factor of small inductors with a diameter of less than 50µm, the accuracy requirements are extremely high. Since the typical inductor values at millimeter wave frequencies lie between 50 pH and 300 pH and the resulting quality factors can vary from around 10 to 25, resistance values around one Ohm have to be determined precisely. In this case the

incertitude due to the contact resistance can exceed the series resistance value, yielding unphysical, even negative values as already discussed in subsection 2.4.4.

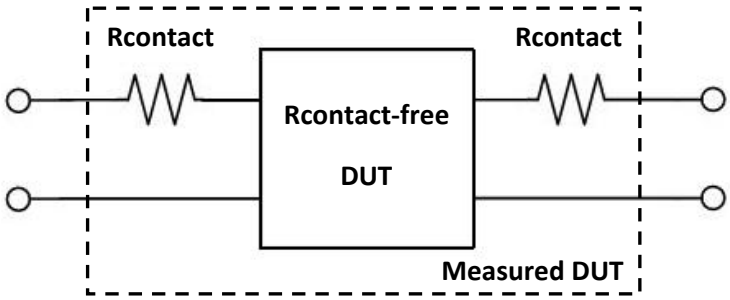


Fig. 45. Equivalent block diagram accounting for contact resistance.

A three-step approximated method that gives an empirical estimation of the contact resistance is reported in [43], using two transmission lines of different length in a similar way as for the split-thru de-embedding technique. Firstly the equivalent series resistance is extracted from the measurement of each line:

$$Rlongline = Re\left\{-\frac{1}{Y_{21longline}}\right\} \tag{30}$$

$$Rshortline = Re\left\{-\frac{1}{Y_{21shortline}}\right\}$$

Then, the difference between the resistance of the two lines is re-scaled to obtain a theoretical value for the long line (or for the short one), according to the following proportion:

$$(Rlongline - Rshortline\,) \div (L - S) = R_{longline}^{theor} \div L \tag{31}$$

The theoretical value of the long line resistance so calculated is finally subtracted from the measured value (30). The result is roughly equal to the contact resistance:

$$Rcontact = Rlongline - R_{longline}^{theor} \tag{32}$$

The test structure can be then represented by the following M-matrix expression, as suggested by Fig. 45:

$$M_{meas} = M_{res} \times M_{res-free} \times M_{res} \tag{33}$$

where M_{res} accounts for the contact resistance.

$$M_{res} = \begin{bmatrix} 1 & Rcontact \\ 0 & 1 \end{bmatrix} \tag{34}$$

The $M_{res-free}$ matrix corresponding to the text structure de-embedded from the effects of the contact resistance can be extracted from measurements, simply inverting the (33):

$$M_{res-free} = M_{res}^{-1} \times M_{meas} \times M_{res}^{-1} \tag{35}$$

From a rigorous point of view, this technique leads only to an approximated result because the assumption made in (31) that the equivalent resistance of a transmission line should depend linearly from its length is not strictly verified. Nevertheless, it can provide reasonable results in several practical cases. For example, if procedure is applied on the measurement of the three lumped inductors presented in subsection 2.4.4, followed by the split-thru de-embedding, then it gives coherent results. The quality factors of the three inductors reported in Figs. 36-38 can be replaced by the more reasonable plots of Figs. 46-48. In the last case, the inductor has a more elevated series resistance and it is less influenced by contact resistance. As a result, the correction seems to be unnecessary. It can be also remarked that the proposed correction for the contact resistance only alters the measurement results that depends on the series resistance, with no significant changes in resistance-independent parameters, as for example, in the case the inductance value reported in Fig. 49.

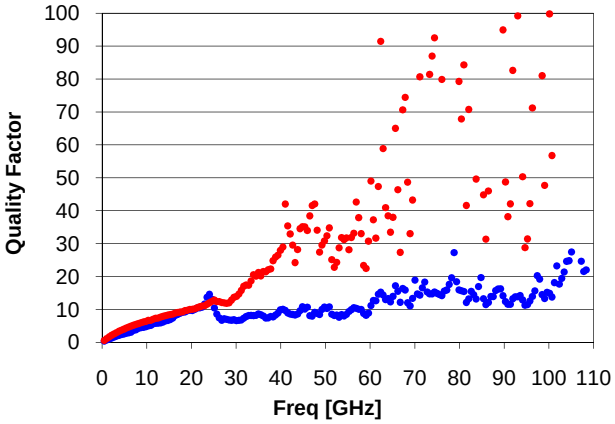


Fig. 46. Quality Factor for the inductor of Fig. 35 (a), before (red) and after (blue) contact resistance correction.

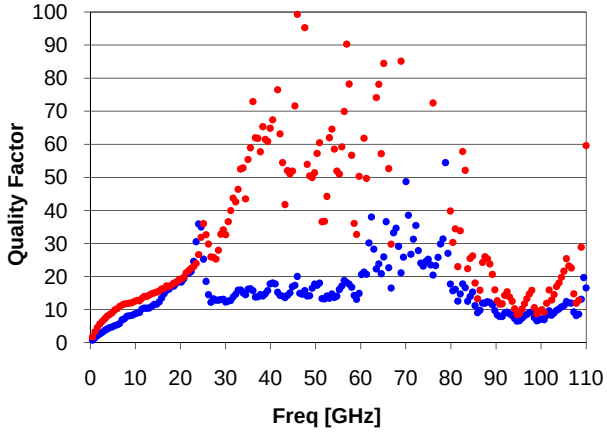


Fig. 47. Quality Factor for the inductor of Fig. 35 (b), before (red) and after (blue) contact resistance correction.

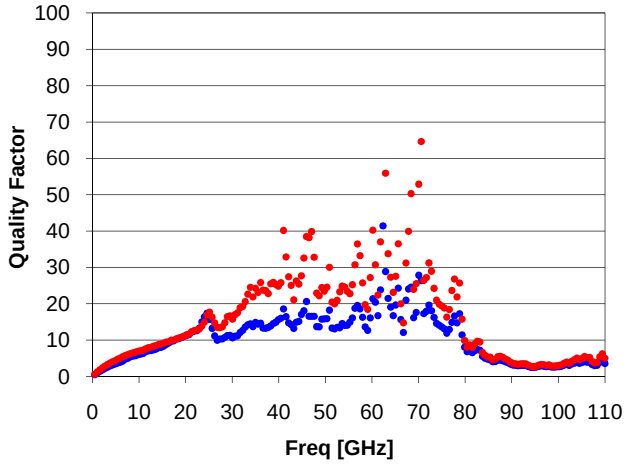


Fig. 48. Quality Factor for the inductor of Fig. 35 (c), before (red) and after (blue) contact resistance correction.

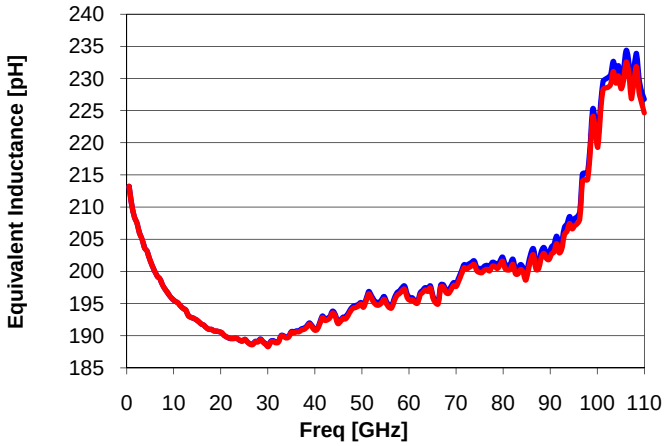


Fig. 49. Equivalent Inductance for the inductor of Fig. 35 (c), before (red) and after (blue) contact resistance correction.

2.8 Conclusion

The key elements needed for successful design of millimeter wave integrated circuits have been presented so far, considering both the lumped and distributed approaches. Starting from a detailed insight on the design facilities provided by the BiCMOS9MW technology, all the aspects of integrated circuit design for millimeter wave applications have been discussed and a complete design methodology has been defined.

Both active and passive devices, including transistors, transmission lines, capacitors, pads, resistors and interconnects, have been closely examined and the convenient model improvements, accounting for parasitic or secondary effects, have been suggested. Concerning lumped design, in particular, three examples of millimeter wave inductors have been reported and an efficient strategy for the design and modeling of lumped devices has been demonstrated. Also, different techniques of measurement and de-embedding of millimeter wave devices have been reviewed and compared.

In the next chapter, the proposed methodology will be applied to the practical case of LNAs design. Examples concerning all the application standards will be presented, at 60GHz, 80GHz and 94GHz.

Bibliography

On the roadmap of millimeter wave design:

[1] W. Winkler *et al.*, "60 GHz and 76 GHz Oscillators in 0.25µm SiGe:C BiCMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech.Papers*, Feb. 2003, pp. 454–455.

[2] Y. Baeyens and Y. Chen, "A monolithic integrated 150 GHz SiGe HBT push-push VCO with simultaneous differential V-band output," in *IEEE Int. Microwave Symp. Dig.*, Jun. 2003, pp. 877–880.

[3] H. Li and H.-M. Rein, "Wide-band VCO's in SiGe production technology operating up to about 70GHz," *IEEE Microw.Wireless Compon.Lett.*, vol. 13, no. 10, pp. 425–427, Oct. 2003.

[4] H. Li, H.-M. Rein, and M. Schwerd, "SiGe VCO's operating up to 88GHz, suitable for automotive radar sensors," *Electron. Lett.*, vol. 39, no.18, pp. 1326–1327, Sep. 2003.

[5] W. Perndl, H. Knapp, K. Aufinger, T. Meister, W. Simburger, and A.Scholtz, "A 98 GHz voltage controlled oscillator in SiGe bipolar technology," in *Proc. Bipolar/BiCMOS Circuits and Technology Mtg.*, Sep.2003.

[6] H. Li, H.-M. Rein, and T. Suttorp, "Design of W-band VCO's with high output power for potential application in 77 GHz automotive radar systems," in GaAs IC Symp. Tech. Dig., Oct. 2003, pp. 263–266.

[7] Y. Li, M. Bao, M. Ferndahl, and A. Cathelin, "23 GHz front-end circuits in SiGe BiCMOS technology," in *IEEE RFIC Symp. Dig. Papers*, Jun.2003, pp. 99–102.

[8] Gresham *et al.*, "Ultra wide and 24 GHz automotive radar front-end,"in *IEEE Int. Microwave Symp. Dig.*, Jun. 2003, pp. 369–372.

[9] H. Hashemi, X. Guan, and A. Hajimiri, "A fully integrated 24 GHz 8-path phased-array receiver in silicon," in *IEEE Int. Solid-State CircuitsConf. (ISSCC) Dig. Tech. Papers*, Feb. 2004, pp. 390–391.

[10] S. Hackl, J. Bock, M. Wurzer, and A. L. Scholtz, "40 GHz monolithic integrated mixer in SiGe bipolar technology," in *IEEE Int. Microwave Symp. Dig.*, Jun. 2002, pp. 1241–1244.

[11] L. M. Franca-Neto, R. E. Bishop, and B. A. Bloechel, "64 GHz and 100 GHz VCO's in 90 nm CMOS using optimum pumping method," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2004, pp. 444–445.

[12] R.-C. Liu, H.-Y. Chang, C.-H. Wang, and H. Wang, "A 63 GHz VCO using a standard 0.25 µm CMOS process," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2004, pp. 446–447.

[13] M. Tiebout, H.-D. Wohlmuth, andW. Simbürger, "A 1 V 51 GHz fullyintegrated VCO in 0.12 µm CMOS," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2002, pp. 238–239.

[14] K.-W. Yu, Y.-L. Lu, D.-C. Chang, V. Liang, and M. F. Chang, "K-band low-noise amplifiers using 0.18 µm CMOS technology," *IEEE Microw. Wireless Compon. Lett.*, vol. 14, no. 3, pp. 106–108, Mar. 2004.

[15] L. M. Franca-Neto, B. A. Bloechel, and K. Soumyanath, "17 GHz and 24 GHz LNA designs based on extended-S-parameter with microstrip-on-die in 0.18 µm logic CMOS technology," in *Proc. Eur. Solid-State Circuits Conf.*, Sep. 2003, pp. 149–152.

[16] X. Guan and A. Hajimiri, "A 24 GHz CMOS front-end," in *Proc. Eur. Solid-State Circuits Conf.*, Sep. 2002, pp. 155–158.

[17] M. Madihian, H. Fujii, H.Yoshida, H. Suzuki, and T.Yamazaki, "A 1–10 GHz, 0.18 µm-CMOS chipset for multi-mode wireless applications," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 2001, pp. 1865–1868.

[18] S. Reynolds, B. Floyd, U. Pfeiffer, T. Zwick, "60GHz Transceiver Circuits in SiSe Bipolar Technology ," *ISSCC Dig. Tech. Papers.*, Feb 2004, pp.442-538, Vol. 1.

[19] Doan, C.H.; Emami, S.; Niknejad, A.M.; Brodersen, R.W.; "Design of CMOS for 60GHz applications," *ISSCC Dig. Tech. Papers.*, Feb 2004, pp.440-538, Vol. 1.

[20] Razavi, B.; "A 60GHz direct-conversion CMOS receiver," *ISSCC Dig. Tech. Papers*., Feb 2005, pp.400-606, Vol. 1.

[21] Terry Yao; Gordon, M.Q.; Tang, K.K.W.; Yau, K.H.K.; Ming-Ta Yang; Schvan, P.; Voinigescu, S.P.; "Algorithmic Design of CMOS LNAs and PAs for 60-GHz Radio," *IEEE Journal of Solid-State Circuits*, May 2007, Vol. 42, No. 5, pp. 1044-1057.

[22] Nicolson, S.T.; Tang, K.A.; Yau, K.H.K.; Chevalier, P.; Sautreuil, B.; Voinigescu, S.P.; "A Low-Voltage 77-GHz Automotive Radar Chipset," *2007 IEEE/MTT-S International Microwave Symposium*, June 2007, pp. 487-490.

[23] Laskin, E.; Chevalier, P.; Chantre, A.; Sautreuil, B.; Voinigescu, S.P.; "80/160-GHz Transceiver and 140-GHz Amplifier in SiGe Technology," *2007 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, June 2007, pp. 153-156.

[24] Cathelin, A. Martineau, B. Seller, N. Douyere, S. Gorisse, J. Pruvost, S. Raynaud, C. Gianesello, F. Montusclat, S. Voinigescu, S.P. Niknejad, A.M. Belot, D. Schoellkopf, J.P., "Design for millimeter-wave applications in silicon technologies," *ESSCIRC 2007. 33rd European Solid State Circuits Conference*, Sept. 2007, pp. 464-471.

[25] Martineau, B. Cathelin, A. Danneville, F. Kaiser, A. Dambrine, G. Lepilliet, S. Gianesello, F. Belot, D., "80 GHz low noise amplifiers in 65nm CMOS SOI," *ESSCIRC 2007. 33rd European Solid State Circuits Conference*, Sept. 2007, pp. 348-351.

On the BiCMOS9MW technology:

[26] P. Chevalier, B. Barbalat, L. Rubaldo, B. Vandelle, D. Dutartre, P. Bouillon, T. Jagueneau, C. Richard, F. Saguin, A. Margain, A Chantre, "300 GHz *fmax* self-aligned SiGeC HBT optimized towards CMOS compatibility" *IEEE BCTM Proceedings*, pp. 120-123, Oct 2005.

[27] Avenier, G. Diop, M. Chevalier, P. Troillard, G. Loubet, N. Bouvier, J. Depoyan, L. Derrier, N. Buczko, M. Leyris, C. Boret, S. Montusclat, S. Margain, A. Pruvost, S. Nicolson, S.T. Yau, K.H.K. Revil, N. Gloria, D. Dutartre, D. Voinigescu, S.P. Chantre, A., "0.13µm SiGe BiCMOS Technology Fully Dedicated to mm-Wave Applications," *IEEE Journal of Solid-State Circuits*, Sept. 2009, Vol. 44, No. 9, pp. 2312-2321.

[28] Chevalier, P. Barbalat, B. Laurens, M. Vandelle, B. Rubaldo, L. Geynet, B. Voinigescu, S.P. Dickson, T.O. Zerounian, N. Chouteau, S. Dutartre, D. Monroy, A. Aniel, F. Dambrine, G. Chantre, A., "High-Speed SiGe BiCMOS Technologies: 120-nm Status and End-of-Roadmap Challenges," *SiRF 2007. Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, Jan. 2007, pp. 18-23.

[29] Chevalier, P. Fellous, C. Rubaldo, L. Dutartre, D. Laurens, M. Jagueneau, T. Leverd, F. Bord, S. Richard, C. Lenoble, D. Bonnouvrier, J. Marty, M. Perrotin, A. Gloria, D. Saguin, F. Barbalat, B. Beerkens, R. Zerounian, N. Aniel, F. Chantre, A. "230 GHz self-aligned SiGeC HBT for 90 nm BiCMOS technology," *Proceedings of the 2004 Meeting on Bipolar/BiCMOS Circuits and Technology*, Sept. 2004, pp. 225-228.

On the design and characterization of transmission lines for millimeter wave applications:

[30] Hammerstad, Erik O., "Equations for Microstrip Circuit Design," 5th European Microwave Conference, Sept. 1975, pp. 268-272.

[31] I. C. H. Lai, Y. Kambayashi, and M. Fujishima, "60-GHz CMOS down-conversion mixer with slow-wave matching transmission lines," *2006 IEEE Asian Solid State Circuits Conf. Dig. Tech. Papers*, Nov. 2006, pp. 195–198.

[32] Repossi, M. Eyssa, W. Vecchi, F. Arcioni, P. Svelto, F., "Design of Low-Loss Transmission Lines in Scaled CMOS by Accurate Electromagnetic Simulations," *IEEE Journal of Solid-State Circuits*, Sept. 2009, Vol. 44, No. 9, pp. 2605-2615.

[33] W. Heinrich, J. Gerdes, F. J. Schmuckle, C. Rheinfelder, and K.Strohm, "Coplanar passive elements on SI substrate for frequencies up to 110 GHz," *IEEE Trans. Microw. Theory Tech.*, vol. 46, no. 5, pp. 2264–2268, May 1998.

[34] A. C. Reyes, S. M. El-Ghazaly, S. J. Dorn, M. Dydyk, D. K. Schroder, and H. Patterson, "Coplanar waveguides and microwave inductors on silicon substrates," *IEEE Trans. Microw. Theory Tech.*, vol. 43, no. 9, pp. 2016–2022, Sep. 1995.

On the design and characterization of lumped inductors for millimeter wave applications:

[35] Cheung, T.S.D. Long, J.R., "Shielded passive devices for silicon-based monolithic microwave and millimeter-wave integrated circuits," *IEEE Journal of Solid-State Circuits*, May 2006, Vol. 41, No. 5, pp. 1183-1200.

[36] Gordon, M. Voinigescu, S.P., "An inductor-based 52-GHz 0.18 μm SiGe HBT cascode LNA with 22 dB gain," *ESSCIRC 2004. 30rd European Solid State Circuits Conference*, Sept. 2004, pp. 287-290.

[37] Laskin, E. Chevalier, P. Chantre, A. Sautreuil, B. Voinigescu, S.P., "165-GHz Transceiver in SiGe Technology," *IEEE Journal of Solid-State Circuits*, May 2008, Vol. 43, No. 5, pp. 1087-1100.

[38] Dickson, T.O. LaCroix, M.-A. Boret, S. Gloria, D. Beerkens, R. Voinigescu, S.P., "30-100-GHz inductors and transformers for millimeter-wave (Bi)CMOS integrated circuits," *IEEE Transactions on Microwave Theory and Techniques*, Jan. 2005, Vol. 53, No. 1, pp. 123-133.

[39] Kraemer, M. Dragomirescu, D. Plana, R., "Accurate electromagnetic simulation and measurement of millimeter-wave inductors in bulk CMOS technology," *SiRF 2010.Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, Jan. 2010, pp. 61-64.

[40] Yu Cao Groves, R.A. Xuejue Huang Zamdmer, N.D. Plouchart, J.-O. Wachnik, R.A. Tsu-Jae King Chenming Hu, "Frequency-independent equivalent-circuit model for on-chip spiral inductors," *IEEE Journal of Solid-State Circuits*, March 2003, Vol. 38, No. 3, pp. 419-426.

On the techniques of de-embedding of measurements at millimeter wave frequencies:

[41] Yau, K.H.K. Mangan, A.M. Chevalier, P. Schvan, P. Voinigescu, S.P., "A Transmission-Line Based Technique for De-Embedding Noise Parameters," *ICMTS '07. IEEE International Conference on Microelectronic Test Structures*, March 2007, pp. 237-242.

[42] Mangan, A.M. Voinigescu, S.P. Ming-Ta Yang Tazlauanu, M., "De-embedding transmission line measurements for accurate modeling of IC designs," *IEEE Transactions on Electron Devices*, Feb. 2006, Vol. 53, No. 2, pp. 235-241.

[43] Kraemer, Michael Dragomirescu, Daniela Rumeau, Alexandre Plana, Robert, "On the de-embedding of small value millimeter-wave CMOS inductor measurements," *2010 German Microwave Conference*, March 2010, pp. 194-197.

Chapter 3

Millimeter Wave Low Noise Amplifiers: Design and Experimental Characterization

3.0 Introduction

A design methodology for millimeter wave integrated circuits was established and described in chapter 2, considering both lumped and distributed approaches. The aim of this chapter is to prove its reliability in the practical case of low-noise amplifiers design. Therefore, three LNAs – operating in the most common millimeter wave standards, namely 60GHz WPAN, automotive RADAR in the 77-80GHz band and imaging applications at 94GHz – have been fabricated in the BiCMOS9MW technology by ST Microelectronics and tested, in order to provide an exhaustive empirical demonstration of the correctness of the design methodology within the millimeter wave spectrum.

After an overview of the general guidelines for LNA design in section 3.1 and of the measurement procedures available for testing in section 3.2, section 3.3 presents the first one of the test circuits, that is a two-stage cascode 60GHz LNA dedicated to WPAN applications. Then, section 3.4 describes a two-stage LNA operating at 80GHz and section 3.5 a single stage 94GHz amplifier for imaging applications.

Section 3.6 is dedicated to a comparison between the lumped and the distributed approach, on the basis of a comparative analysis of two different implementations of a single-stage 80GHz LNA using, respectively, distributed transmission lines and lumped inductors. In section 3.7, the performances of our circuits are compared to the state-of-the-art of millimeter wave LNAs reported in the literature.

Finally – to give a suggestion of the possible applications of millimeter wave LNAs and to introduce the perspectives of our work – the design of an 80GHz receiver front-end and a preliminary investigation on the co-integration of a LNA with a millimeter wave antenna are presented respectively, in section 3.9 and 3.10.

3.1 General guidelines for LNA design

3.1.1 Topology

The first step in the design of an amplifier is to choose the topology. As a result of the comparative analysis of the state-of-the-art of millimeter wave LNAs reported in section 3.7, three basic architectures are used: common emitter (CE), common base (CB), and cascode. These configurations are largely employed in radio-frequency applications and have proved useful also in millimeter-wave context.

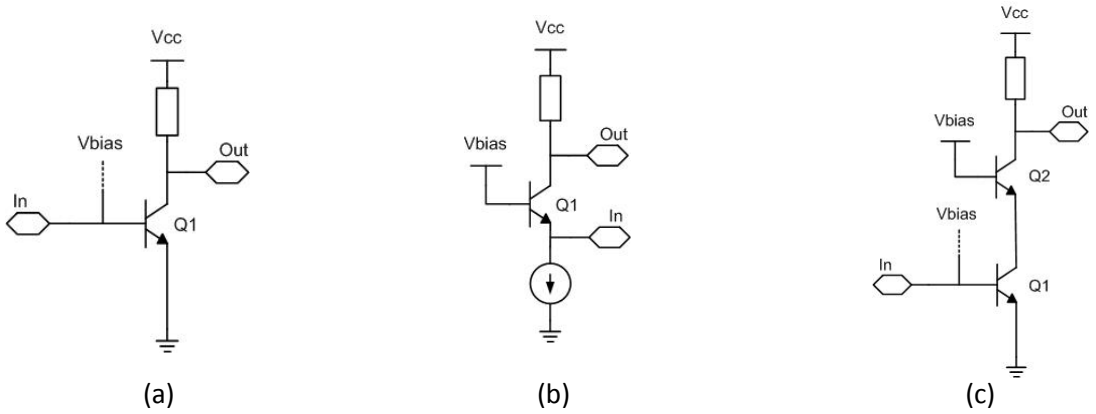


Fig. 1. Possible architectures for millimeter wave amplifiers: Common Emitter (a), Common Base (b), and Cascode (c).

The main disadvantage of designing a CB amplifier is to perform a simultaneous noise and power matching at its input node, whereas a well-established procedure for simultaneous matching can be systematically applied to CE and cascode amplifiers operating at radio-frequency. This same technique can be also extended to millimeter wave applications, with minor modifications, as it will be explained in the next subsection.

Beside the availability of a systematic procedure to perform noise and power matching, the cascode stage can provide higher gain, compared to CE and CB with same bias current, and better isolation.

The small-signal behavior of a two-port device can be analyzed using its S parameters defined in Annex 2. A quantitative estimation of the reverse-isolation, in particular, is given by the S12 parameter. If isolation is not perfect, then S12 is different from zero and therefore the input reflection coefficient Γ_{in} depends on the load at the output port and the output reflection coefficient Γ_{out} depends on the source, as suggested by the following definitions:

$$\Gamma_{in} = S_{11} + \frac{s_{12}s_{21}\Gamma_{load}}{1 - s_{22}\Gamma_{load}}$$

$$\Gamma_{out} = S_{22} + \frac{s_{12}s_{21}\Gamma_{source}}{1 - s_{11}\Gamma_{source}}$$

where Γ_{source} and Γ_{load} are, respectively, the source and load reflection coefficients.

Even at millimeter waves, cascode stages can assure lower S12, compared to CE amplifiers. This consideration is important since it implies less sensitivity to model inaccuracies and higher stability. For these reasons, cascode architecture has been used for the design of all the test circuits reported hereafter.

3.1.2 Cascode Amplifier with Intersgtage Matching

In the traditional design of cascode LNAs, impedance matching between the CE and the CB stages is usually neglected. This is not desirable for the maximum power transfer. The performances of a

cascode amplifier can be enhanced by inserting a matching network between the CE and the CB transistors. The input and the cascoded transistors are then considered as two separated stages and a matching network between them can improve the interstage power transfer, contributing to ameliorate the gain and the noise figure of the overall amplifier. Since both the input impedance of the CE stage and the output impedance of the CB are capacitive, the matching network can be efficiently implemented simply by inserting an inductor in series at the cascode node.

To understand the effect of such an element, let us first consider the equivalent small-signal circuit of a cascode stage shown in Fig. 2 (b). It can be observed that parasitic capacitors between the transistor terminals and between the transistor terminals and the substrate can be included in the schematic simply adding them in the expressions of capacitors C_π and C_μ that are defined as follows:

$C_{\pi 1} = C_{d1}+C_{be1}+C_{bs1}$

$C_{\mu 1}=C_{t1}+C_{bc1}$

$C_{\pi 2}=C_{d2}+C_{be2}+C_{ce1}+C_{esub2}+C_{csub1}$

$C_{\mu 2}=C_{t2}+C_{bc2}+C_{csub2}$

where C_{di} and C_{ti} are respectively the diffusion and the transition capacitances of the junctions included in the model of the i-th transistor. Only C_{ce2} has been neglected to simplify the analysis.

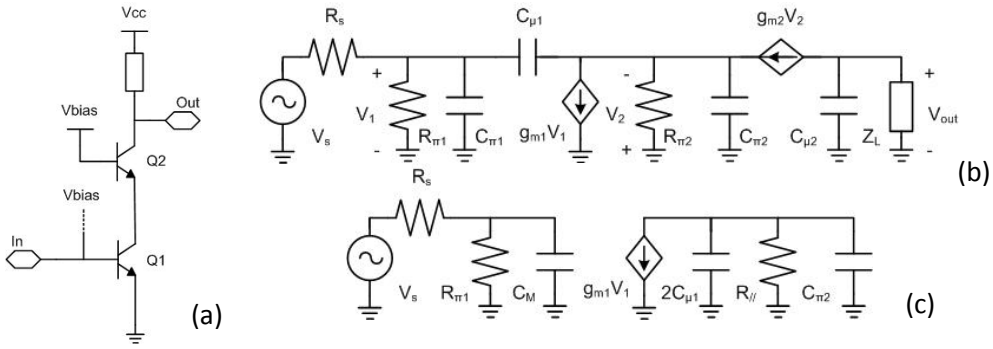


Fig. 2. Small-signal equivalent circuit of a cascode amplifier.

A useful simplification that can be made is to replace the controlled current generator of the CB stage by its equivalent impedance $1/g_m$. Such an impedance appears in parallel to $R_{\pi 2}$ resulting into a total impedance $R_{//}$:

$$R_{\parallel} = \frac{R_{\pi 2}}{1 + g_{m2}R_{\pi 2}}$$

It possible to apply the Miller's Theorem to $C_{\mu 1}$. Neglecting the current contribution of $C_{\mu 1}$ that is much lower than $g_{m1}V_1$, we have:

$$V_{2M} \cong -g_{m1}\left[R_{\parallel} // \frac{1}{j\omega C_{\pi 2}}\right] V_{1M}$$

$$A_M = \frac{V_{2M}}{V_{1M}} = -\frac{g_{m1}R_{\pi 2}}{1 + g_{m2}R_{\pi 2} + j\omega C_{\pi 2}}$$

Then, neglecting the contribution of $C_{\pi 2}$ and observing that $g_{m1} = g_{m2}$ since both transistors are biased with the same collector current, we obtain:

$$A_M \cong -1$$

and $C_M = C_{\pi 1} + 2C_{\mu 1}$

The circuit of Fig. 2 (b) is therefore equivalent to that of Fig. 2 (c) except for the output node that is not represented in the latter. The left part of this circuit contributes to the global frequency response of the amplifier with a pole that is located approximately at:

$$\omega_{P1} \cong \frac{1}{[R_S // R_{\pi 1}]C_M}$$

Since the Miller effect is minimum in a cascode amplifier, this pole is usually a dominant one and accounts the mostly for the bandwidth of the amplifier.

The capacitive elements in the right part of the circuit generate a second pole whose approximated frequency is:

$$\omega_{P2} \cong \frac{1+g_{m2}R_{\pi 2}}{R_{\pi 2}[2C_{\mu 1}+C_{\pi 2}]}$$

This second pole is located at higher frequency than ω_{P1} but it is lower than the other singularities generated by the reactive elements at the output node like $C_{\mu 2}$, therefore ω_{P2} is sometime referred to as the middle pole of the cascode configuration.

Intuitively, it does mean that the pole introduced by the CB transistor in the frequency response of the cascode depends on the transconductance of the CB transistor itself and on the parasitic capacitance at the cascode node. As a consequence, at high frequencies, the power gain decreases and the noise contribution of the CB stage is enhanced. Thus, the input referred noise of a cascode amplifier will rise considerably at high frequencies.

This effect is more accentuated for a cascode stage with inductive degeneration, since degeneration implies a further reduction of gain.

The use of an inductor to tune out the parasitic capacitance at the interstage node increases the value of ω_{P2}. As illustrated in Fig. 3, this inductor creates an artificial transmission line with C_{ce1}, C_{csub1}, C_{be2}, and C_{esub2}. The characteristic impedance of such a line is:

$$Z_O = \sqrt{\frac{L}{C_{be2} + C_{esub2}}}$$

And, in order to assure interstage matching it must satisfy the following conditions:

$$Z_O = Z_{out1}$$

$$Z_O = Z_{in2}$$

Unfortunately these expressions are scarcely exploitable and, from a practical point of view, the optimum value of the interstage inductance must be set by simulations. As suggested in [1], a simple way to operate is to plot the f_T of the cascode versus the inductor value. Its optimum corresponds to the maximum of f_T.

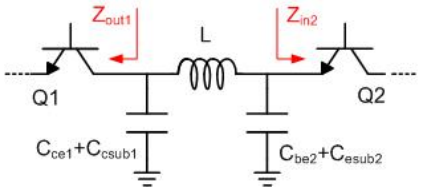


Fig. 3. Transmission line generated by the inductor added at the cascode node.

In the case of a MOSFET cascode the expression of ω_{P2} can be obtained assuming $R_{\pi 2}$ infinite:

$$\omega_{P2} \cong \frac{g_{m2}}{2C_{\mu 1} + C_{\pi 2}}$$

The middle pole is typically at lower frequency than in the case of HBT devices and the efficiency of the interstage matching becomes more evident. For these reason, this matching technique has been proposed for MOSFET cascode. In the case of bipolar devices such as SiGe HBTs, even if its efficacy is less evident, it has been used to partially compensate the parasitic capacitance existing at the cascode node in the higher frequency circuits reported in this thesis.

3.1.3 Design strategy for Noise optimization

The main goal of low-noise amplifiers design is the minimization of the overall noise figure. Noise optimization at circuit level firstly implies noise matching at the input node of the circuit. In the early examples of LNAs and mixers reported for RF and microwave applications, noise matching was usually obtained as a trade off with input impedance (power) matching, by means of a passive matching network. A more efficient methodology, consisting in simultaneous noise and power matching, was proposed in [2]. This technique has been largely applied both in bipolar and CMOS RF context and can provide the general guidelines also for millimeter wave design. It relies on the availability of noise models of active devices scalable with the transistor geometry. Instead of adding a passive network around the active device to achieve noise and/or impedance matching, the transistor is sized to achieve noise matching. Compared to passive matching, this leads to lower noise figure, according to Friis analysis. Since the task of noise matching is removed from the passive network, the latter becomes simpler and input loss can be reduced. Then, to complete the matching procedure, a minimal passive network with two low-loss inductors is designed to provide the desired input impedance with the lowest possible degradation of the overall noise figure.

The noise factor of a two-port device can be expressed as:

$$F = Fmin + \frac{Rn}{Rs}\left|\frac{Z_S}{Z_S opt} - 1\right|^2 \quad (1)$$

where Fmin is the minimum noise factor of the transistor, Rn is its noise equivalent resistance and $Zs = Rs + iXs$ is its source impedance. $Z_{Sopt} = R_{Sopt} + iX_{Sopt}$ is the value of the source impedance that corresponds to minimum noise factor.

To minimize the overall noise figure it is necessary to minimize both the terms in (1).

As the first term depends only on the bias conditions and not on the source configuration, the first step in LNA design flow is the minimization of NFmin. Since NFmin depends on the bias current density as depicted in Fig. 4, its minimum value can be guaranteed by using the corresponding optimum value for the bias current density.

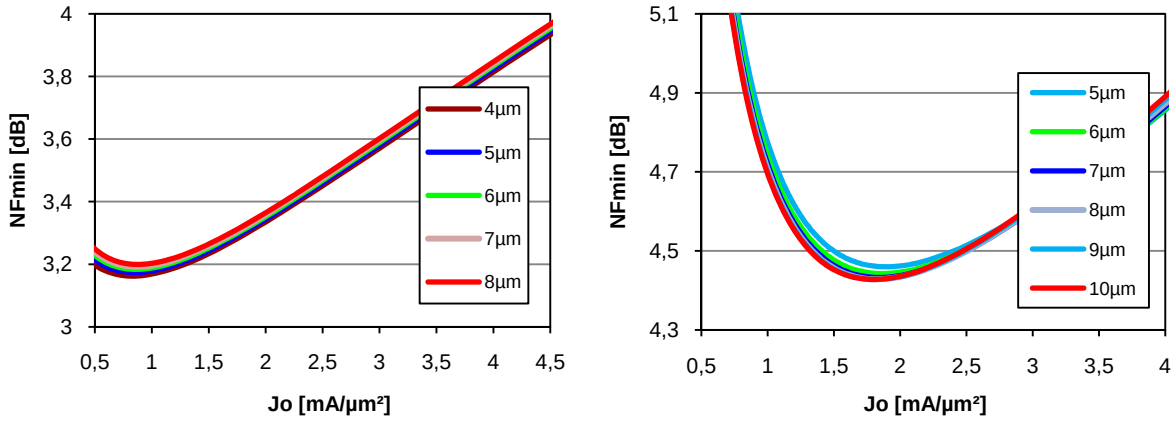


Fig. 4. NFmin of a CE single transistor (left) and cascode (right) vs bias current density, for different values of emitter length.

The plot of NFmin in Fig. 4 is calculated at 80GHz on a HS transistor for various emitter lengths, with minimum emitter width (0.27µm). The minimum value of NFmin is approximately independent from emitter length, at least in the range between 5 and 9 µm, for the two configurations. As a consequence, the optimum current density can be fixed independently from transistor size. Comparing the two plots of Fig. 4, we can observe that the CE performs a lower NFmin compared to the cascode: 3.2dB and 4.4dB, respectively. Furthermore, the optimum current densities for the lowest NFmin are different: Jopt = 0.73mA/µm^2 and Joptcasc = 1.85mA/µm^2.

The study of the NFmin as a function of the transistor current density is a key element in the efficiency of the proposed matching technique. Besides CE amplifiers, the design strategy can be applied to cascode stages with CE and CB transistors equally sized and biased at Joptcasc.

In Fig. 5 NFmin plot is compared to f_T and available gain (Ga) for the same cascode stage, at 80GHz. It must be observed that the current density value of 1.85mA/µm^2 for the minimum value of NFmin unfortunately does not correspond to the maximum of f_T and Ga. Since minimum NFmin and maximum f_T or available gain can not be achieved at the same time, a trade off is usually observed. For now, we try to complete the design of a cascode stage with minimum noise figure.

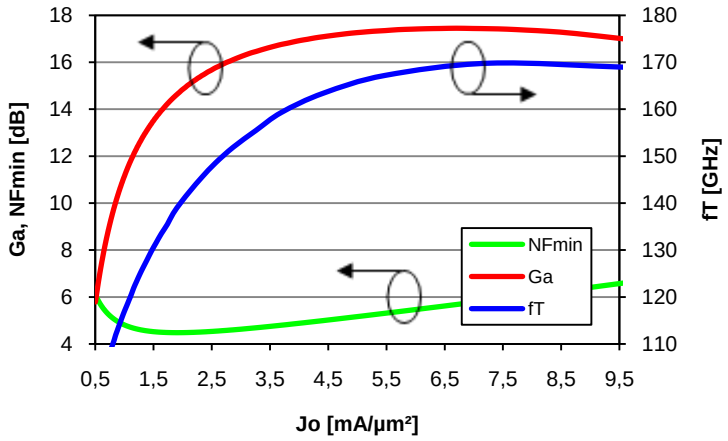


Fig. 5. Available Gain (red), NFmin (green), and f_T (blue) of a cascode stage vs bias current density.

3.1.4 Matching of the real part of the Noise impedance

Once the optimum current density is fixed for each active device as described above, it is possible to match the real part of the noise impedance of the overall amplifier to the source, by minimizing the second addend in (1). Transistor size must be set in order to make Rsopt equal to the real part of the source impedance, usually 50Ω. Using transistors with a fixed emitter width (i.e. 0.27µm) the emitter length can be modified to achieve noise matching. As resulting from Fig. 6, in the case of a cascode biased for minimum NFmin at 1.85mA/µm^2, the most appropriate choice for emitter length is 6µm that corresponds to an optimum noise resistance of 44Ω. We can remark that an emitter length of 5µm will result in an optimum noise resistance closer to 50Ω, however, it must be considered that, after the introduction of the emitter degeneration, the optimum noise resistance will be oversized.

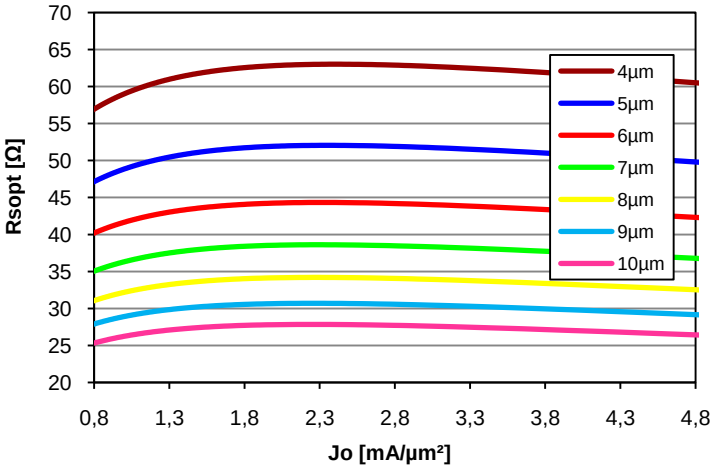


Fig. 6. Rsopt vs emitter length.

Obviously, the choice of a fixed value for the current density and for the emitter area does mean a choice for the bias current. In the given example the resulting bias current is Io = 3mA.

An estimation of the noise matching can be given in terms of noise reflection coefficient Gmin (or Γmin), that is defined as:

$$Gmin = \Gamma min = \frac{Z_{Sopt} - Z_0}{Z_{Sopt} + Z_0}$$

where Z_0 is a normalization impedance usually assumed equal to 50Ω, as well as Zs.

In the ideal case of a perfectly matched device, Gmin is zero (as well as the source reflection coefficient $\Gamma source$) and the resulting noise factor is identically equal to minimum noise factor, since:

$$F = Fmin + \frac{4Rn}{Z_0} \frac{|\Gamma_{source} - \Gamma_{min}|^2}{|1 - \Gamma_{min}|^2 (1 - |\Gamma_{source}|^2)}$$

A real amplifier is usually characterized by a non-zero value of Gmin, due to unavoidable imperfections in noise matching, exactly as a non-zero S11 parameter accounts for power insertion loss.

From a practical point of view, Gmin is automatically calculated by Spectre assuming $\Gamma source = 0$. The real part of the noise impedance Rsopt can be extracted as a function of Gmin according to the following equation:

$$Rsopt = \frac{Z_0 \Re e\left\{\frac{1-Gmin}{1+Gmin}\right\}}{\left[\Re e\left\{\frac{1-Gmin}{1+Gmin}\right\}\right]^2 + \left[\Im m\left\{\frac{1-Gmin}{1+Gmin}\right\}\right]^2} \quad (2)$$

Alternatively, when using ADS, the plot of Rsopt is directly generated by the simulator. In both cases, it is possible to plot the real part of the noise impedance versus the transistor emitter length, always keeping the optimum bias current density previously determined. This leads to the plots of Fig.6.

At this point, it can be observed that in the case of a MOSFET LNA with multi-finger gate transistors, the finger width, beside current density, must be chosen in order to minimize NFmin, and then the number of fingers can be varied to adjust Rsopt to 50ohm. A similar procedure can be eventually applied also to BiCMOS9MW transistors with multi-emitter configuration but it is not strictly necessary, since the value NFmin at optimum current density is approximately independent of emitter length, as already highlighted.

After the transistor sizing, an inter-stage inductor can be added at the common node between the input and the CB transistor, to improve the f_T of the overall cascode stage. Its value can be calculated performing a parametric SP analysis vs f_T, as suggested by Fig. 7. The optimum inductor value, corresponding to maximum f_T, is 31pH.

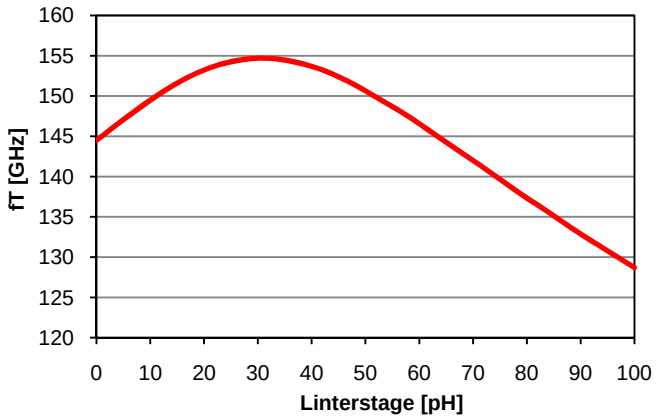


Fig. 7. The introduction of an inductor at interstage node in a cascode stage contributes to augment its f_T.

3.1.5 Real impedance matching with inductive degeneration

So far, the transistor size and its bias current have been determined. The real part of the optimum noise impedance is now matched to 50Ω without having degraded the minimum noise figure.

Now it is necessary to match the real part of the amplifier input impedance Rin to the source, by means of a matching network. Usually a CE or a cascode stage with a minimum noise figure is characterized by a very low value of input resistance (Rin ~ 10Ω). For this reason the input matching can not be easily achieved and emitter degeneration is often required.

An ideal lossless inductor Lemitter is added at the emitter node. As a consequence, the input impedance increases because a real term appears in its expression, calculated from the schematic of Fig. 8 neglecting C_μ and R_π:

$$Z_{in} = \frac{g_m L_{emitter}}{C_\pi} + j\left[\omega L_{emitter} - \frac{1}{\omega C_\pi}\right]$$

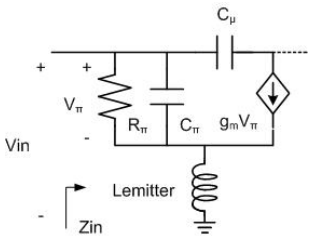


Fig. 8. Cascode with inductive degeneration: equivalent circuit of the input stage.

The introduction of the emitter degeneration slightly modifies also the value of the noise impedance (Annex 4) and a re-sizing of transistor length can be required. To prevent this phenomenon a value of Rsopt slightly lower than 50Ω has been chosen in the previous step. Optimum value for Lemitter is simply calculated according to the simulation results reported in Fig. 9. An inductance of 60pH results into 50 and 59 Ω of Rin and Rs, respectively.

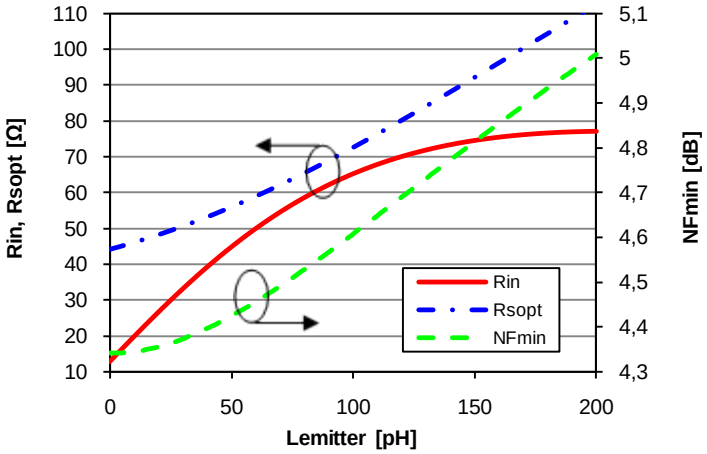


Fig. 9. Cascode input resistance (Rin) vs emitter degeneration.

Beside Rsopt and NFmin, the introduction of the emitter degeneration also affects the available gain of the amplifier that is dramatically reduced as depicted in Fig. 10. At Jo = 1.85mA/µm^2, Ga is decreased from 14.5 to 11.7dB.

An approximated expression of input-stage transconductance can be deduced from the schematic of Fig. 8. Neglecting the contribution of the input current at the emitter node, the collector current is:

$$ic \cong g_m Vin - j\omega g_m L_{emitter} ic$$

that is approximately equal to the cascode output current. Consequently, the equivalent transconductance decreases with the frequency as a single-pole function and this justifies, at least as a first-order approximation, the reduction of the gain observed in the figure:

$$\frac{iout}{Vin} \cong \frac{g_m}{1 + j\omega g_m L_{emitter}}$$

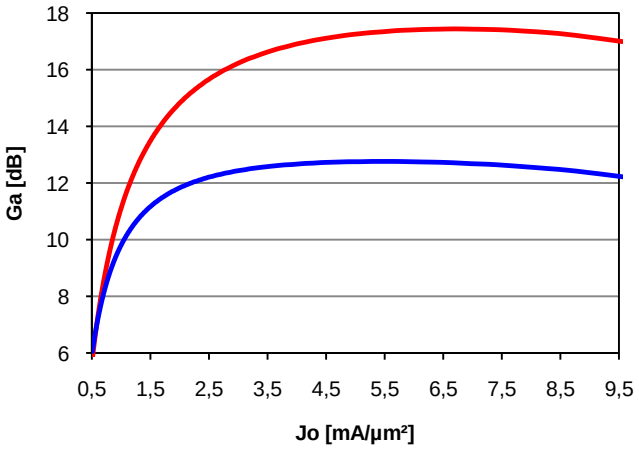


Fig. 10. Available gain of the cascode amplifier before (red) and after (blue) the introduction of emitter degeneration.

3.1.6 Tuning of the imaginary part of the input and noise impedances

To complete the matching procedure, a series inductor Lbase is added on the transistor base to compensate for the imaginary part of both input (Xin) and noise (Xsopt) impedances.

The resulting input impedance is:

$$Z_{in} = \frac{g_m L_{emitter}}{C_\pi} + j\left[\omega L_{base} + \omega L_{emitter} - \frac{1}{\omega C_\pi}\right]$$

and its imaginary part can be cancelled by setting:

$$L_{base} = \frac{1}{\omega^2 C_\pi} - L_{emitter}$$

where the value of Lemitter has been fixed in the previous step to match the real part to 50Ω.

Practically speaking, the value of Xsopt can be extracted from Gmim using expression (3), whereas the imaginary part of the input impedance Xin is automatically calculated by the simulator.

$$Xsopt = \frac{Z_0 \Im m\left\{\frac{1-Gmin}{1+Gmin}\right\}}{\left[\Re e\left\{\frac{1-Gmin}{1+Gmin}\right\}\right]^2 + \left[\Im m\left\{\frac{1-Gmin}{1+Gmin}\right\}\right]^2} \quad (3)$$

In many cases the two reactances exhibit a capacitive behavior and can be cancelled by a proper value of the base inductor. Best value for Lbase can be easily detected by means of a parametric sweep analysis. In the example of Fig. 11, Xsopt and Xin are respectively equal to -22 and -43 Ω at 80GHz and they can be approximately compensated by a base inductor of 50pH.

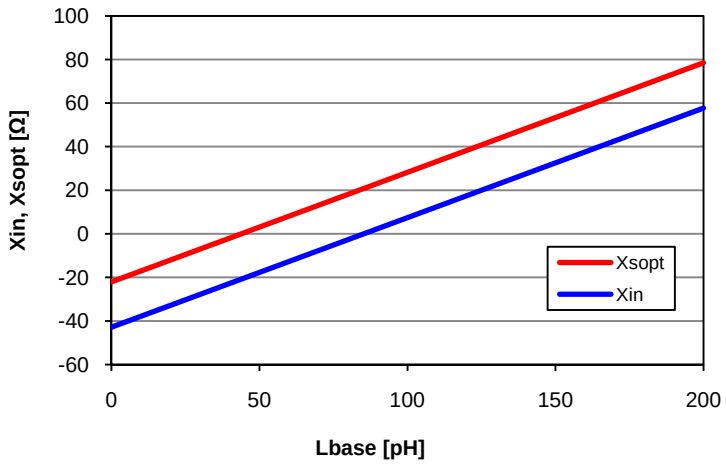


Fig. 11. Imaginary part of input impedance (Xin) and noise impedance (Xsopt) vs base inductor.

Simultaneous noise and input matching is finally obtained and the following conditions are satisfied:

$$\Re e\{Zin\} = Rsopt = 50\Omega \quad (4)$$

$$\Im m\{Zin\} = Xsopt = 0 \quad (4bis)$$

Input impedance matching is commonly evaluated by means of the S11 parameter (reflected power insertion loss), whereas the quality of noise matching can be estimated comparing NF and NFmin plots at the desired frequency or alternatively using the Gmin parameter, as suggested in Fig. 12. Following the methodology described so far, excellent values can be obtained for both noise and power matching in simulation. When moving from the ideal case to a real cascode stage, three effects contributing to the degradation of matching level should be taken into account:

- Parasitics capacitors due to transistor interconnects are not included in the intrinsic models of transistors and must be extracted using a dedicated procedure as explained in paragraph 2.2.4. Since these elements are layout dependent, they can be extracted after transistor sizing. Therefore, a complete transistor model accounting for parasitics should be used in the design of emitter and base inductors.
- Ideal inductors are not available in practice, so Lemitter and Lbase must be implemented as distributed transmission lines or alternatively as lumped elements if a lumped approach is applied to the design. In both cases, the introduction of non-idealities due to the emitter inductor can alter the value of Rsopt set previously to emitter degeneration, requiring a re-sizing of transistor emitter length. Furthermore a lossy inductor in series to the base contributes to degrade the noise factor.
- A third source of losses is the input pad. Usually modeled as an ideal capacitive element, it can be compensated by the inductors used for input matching. As suggested in [3] the influence of the input pad can be taken into account including it in the source impedance and adapting the amplifier to the resulting impedance value instead of 50Ω. However losses due to the unavoidable mismatching between the pad and an ideal capacitance are critical as they concern the very input of the circuit.

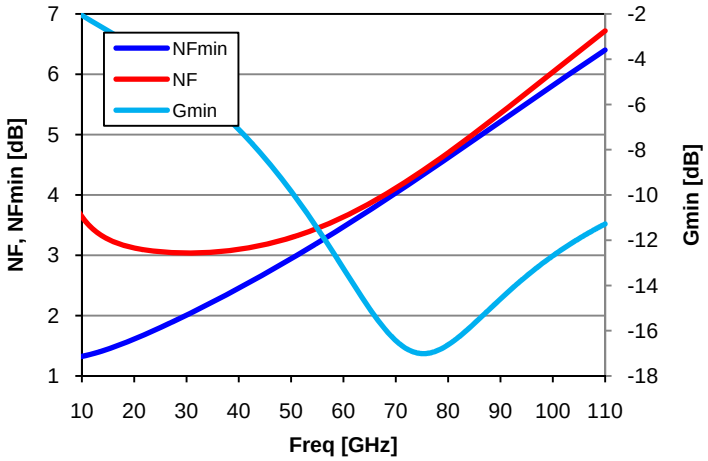


Fig. 12. Noise matching of the cascode amplifier.

3.1.7 Design of the output network

After matching the noise and input impedances, the next step in the LNA design flow is the optimization of gain within the bandwidth of interest. In the case of a cascode amplifier, the shape of the available gain plot depends on the peaking element.

If a single-stage approach is followed, then the design flow can be completed with output impedance matching, realized by means of a passive network that can be designed according to the conventional methodology based on the Smith chart. Practically speaking, the output matching network of our cascode can be realized by a peaking transmission line connected through a series capacitor C1 to the output node; a second capacitor C2 shunted to ground is added to force C1 within a reasonable range of values. For example, with C1 fixed to 48fF, optimum values for C2 and the length of the peaking line can be set with a parametric simulation. As shown in Fig. 14, with 140fF and 149µm excellent output matching is obtained. The final schematic of the amplifier is reported in Fig. 13.

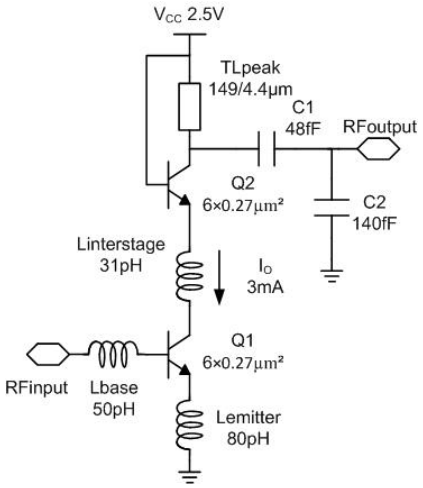


Fig. 13. Schematic view of a single-stage cascode LNA designed for minimum Noise Figure at 80GHz.

The introduction of the peaking line and of the output network can alter the input impedance matching set previously. If necessary, input mismatching can be corrected by a re-sizing of emitter and base inductors. In our example, an emitter inductor of 80pH has been used to maintain a reasonable value of S11. As shown in Fig. 15, the resulting gain (S21) of the input/output matched cascode stage is 11.6dB, approximately equal to its available gain, as expected.

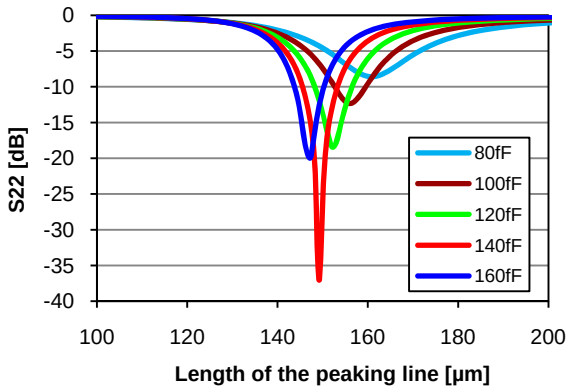


Fig. 14. Design of the output matching network. A 4.4µm-wide microstrip is used.

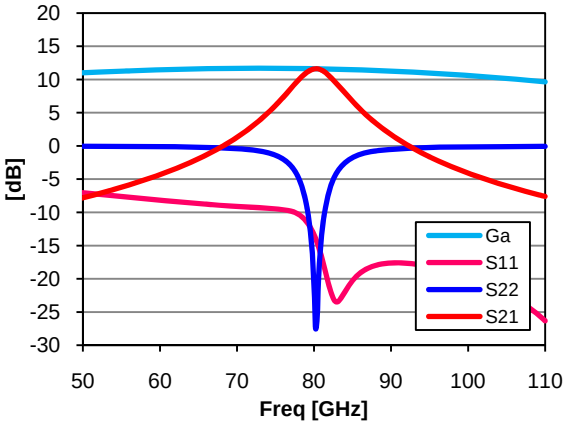


Fig. 15. Power gain (S21), input insertion loss (S11), and output return loss (S22) of the cascode stage.

In the case of the amplifier of Fig. 13, the circuit has been biased at the current density corresponding to NFmin minimization. As a consequence, whereas a minimum noise figure of 4.6dB has been actually achieved, the resulting gain of only 11.6dB is too low to satisfy design requirements in some practical cases. To improve gain, as already suggested at the beginning of this paragraph, current density should be set in the middle region of the Fig. 5, resulting in a trade-off between noise and gain minimization.

If the desired power gain can not be obtained with a single-stage amplifier, then a second cascode stage can be added to the first one in order to achieve the expected performances. The overall noise figure of a two-stage amplifier depends on the noise figure of both first and second stages, but also on the gain of the input stage, according to the Friis equation:

$$F = F_1 + \frac{F_2 - 1}{G_1} \tag{5}$$

where F_n is the noise factor of the *n-th* stage and F is the noise factor of the two-stage amplifier. If the overall noise figure must be minimized, then the first stage must provide, at the same time, minimum noise factor *and* maximum gain. In practice, as explained before, this requirement can not

be satisfied because of the mismatch between optimum bias current density for minimum NFmin and maximum f_T and a tradeoff between noise and gain is necessary, once more.

When using two-stage architectures, an inter-stage matching network must be designed. In many cases a good matching level can be obtained with small capacitors connected in series between the two stages.

3.2 Experimental characterization for millimeter wave LNAs.

The Low Noise Amplifiers presented in the next sections have been experimentally characterized by on-chip measurements of S-parameters and linearity. Measurements of noise figure have also been executed in the case of the 60GHz and 94GHz LNAs.

3.2.1 Small and large signal measurement

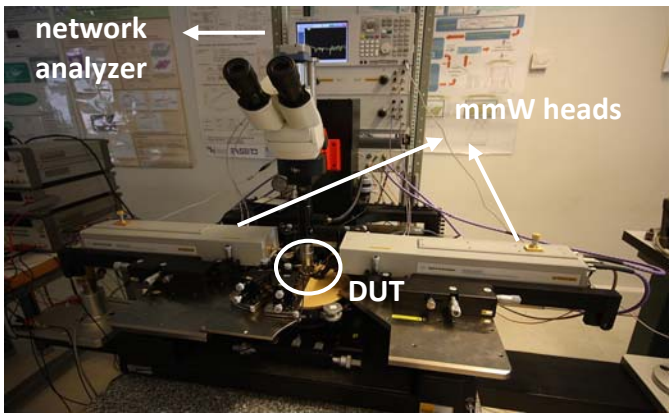


Fig. 16. Test bench for S parameters measurement up to 110GHz.

The measurement of S parameters is a very useful source of information since it gives a complete characterization of the small-signal behavior of the circuit under test.

At IMS Laboratory, we are able to perform measures of one-port or two-port S parameters in the millimeter wave range, up to 110GHz, using a dedicated test bench. As shown in Fig. 16, it consists of an E8361A PNA network analyzer, equipped with the N5260A millimeter head controller by Agilent™.

To investigate the behavior of a circuit in the large-signal domain, we use a load-pull test bench. The main characteristic of a load-pull system is its capability to modify the source *(source-pull)* and load *(load-pull)* impedances seen by the device under test, by means of two dedicated tuners connected respectively at the input and output of the device. The input-side tuner is generally used to match the source to a value of impedance that assures appreciable gain and is maintained at this fixed value, while the output load is varied across the Smith chart, providing a complete characterization of large signal parameters such as compression characteristics, saturated power, efficiency and linearity.

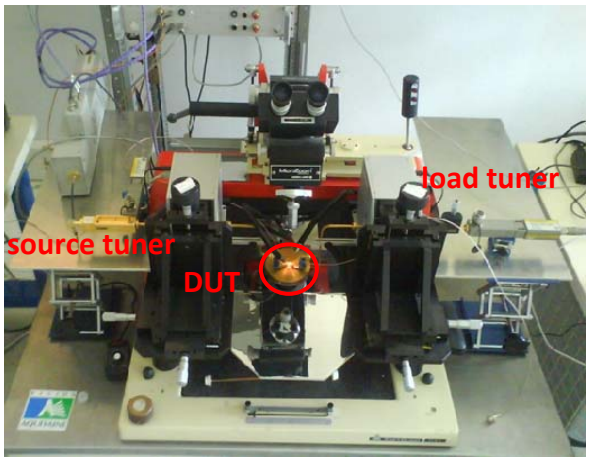


Fig. 17. Load-pull Test bench.

In case of a LNA, a load-pull tool can be used to evaluate the linearity trough measurements of the 1dB compression point and of the IIP3 and, in the context of noise measurements, also to investigate how the noise figure varies with source impedance.

However, the measurement of the IIP3 of a millimeter wave amplifier is prohibitive due to the high frequency of the spectral components of the intermodulation product. As a matter of fact, the load-pull bench has been only used to determine the 1dB compression point.

If the concerned power rates are reasonably low, the 1dB compression point of a LNA can be also measured by means of a power meter and of an E8361A PNA network analyzer, following the three-step procedure suggested in Fig. 18. The power meter and the network analyzer are firstly connected to a thru and the output power is measured (Pout1). Then the power (Pin2) at the output of the network analyzer is measured connecting it directly to the power meter and, as the last step, the thru is replaced by the LNA and its output power (Pout3) is measured. Assuming that the cable used to connect the input and the output of the devices under test (thru and LNA) are identical and that power losses are symmetrically distributed at the input and at the output, their contribution can be estimated as Pcable = (Pout1 – Pin2)/2. The power emission of the LNA at its very output (that is not affected by cable loss) can be calculated as Pout = Pout3 – Pcable and the input power on the very input of the LNA is given by the following equation:

$$\mathrm{Pin}|_{\mathrm{dB}} = \mathrm{Pout}|_{\mathrm{dB}} - \mathrm{Gain}|_{\mathrm{dB}} \tag{6}$$

where the gain of the LNA (in dB) is calculated as Gain = Pout3 – Pout1.

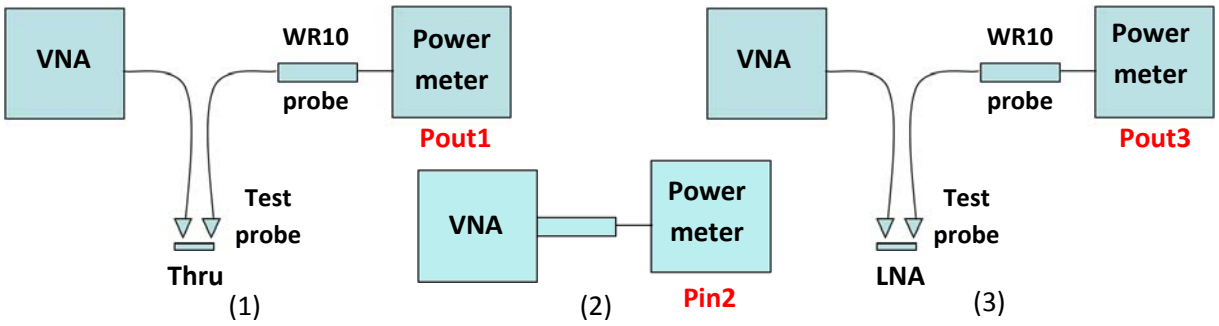


Fig. 18. A three-step procedure for 1dB Compression Point measurement.

The direct measurement of 1dB compression point by means of a power meter is a more involved procedure compared to the use of the load-pull test bench since it is not supported by the automatic setup of the latter. However, it becomes the only choice when the concerned frequencies are higher than the limit of operability of the load-pull that is 90GHz.

3.2.2 Measurement of Noise Figure

Noise measurement tools are not yet available at IMS Laboratory and a preliminary measurement session of NF was performed at IEMN Laboratory.

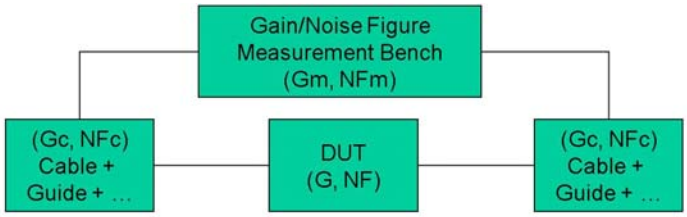


Fig. 19. Block diagram of the test bench for NF measurement at 60GHz

A simplified diagram of the NF test bench available at IEMN is reported in Fig. 19. The device under test (DUT) is analyzed by on-chip probes, connected through a couple of waveguides to the very measurement system that is formed by a noise source and a mixer-based down-converter. Calibration performed prior to measurement accounts for losses in the system, setting the reference for measurements at the outer connectors of each guide (and not directly on the probe tips). As a consequence, losses due to waveguides amounting to 1.7dB are not included in the calibration step and must be corrected apart. The correction procedure relies on the Friis equation applied to the overall system:

$$Fm = Fc + \frac{Fdut-1}{Gc} + \frac{Fc-1}{Gc \cdot G} \qquad (7)$$

where Fm and Gm are respectively the values of the overall noise factor and gain as issued by the measurement, G is the expected LNA gain resulting by independent measurements, and Gc is the loss due to the guideline (-1.7dB).

3.3 Two-stage cascode differential LNA for 60GHz WPAN

The first circuit has been designed and fabricated in order to test the design methodology described in the previous chapter and to prove its efficacy. It is a low noise amplifier operating at 60GHz that can be used for WPAN applications. Since the conventional topologies of millimeter wave mixers are commonly of differential type, a differential architecture was used in this LNA design to facilitate the co-integration within a complete front-end.

3.3.1 Theoretical design and circuit implementation

The schematic of the 60GHz LNA is depicted in Fig. 20. Cascode stages have been here selected to increase reverse isolation and stability. The first stage, (Q1, Q2), performs both 50Ω input impedance and noise matching. The second stage, (Q3, Q4), improves voltage gain and provides 50Ω output matching. Both transistors of the first stage are implemented with five emitters of 1μm length and are biased with 3mA. Second stage transistors are sized as single 10μm emitters biased with 5.5mA. Since the NF of a two stage amplifier depends on the NF of the two stages and on the power gain of the first stage, transistors Q1 and Q2 are sized on a tradeoff between minimum NF and maximum fT. Also the second stage transistors (Q3, Q4) are biased at almost the same value of current density (~2mA/μm²) in order to reduce their contribution to NF and to provide sufficient gain. The number of emitters in Q1 is set to achieve noise matching, whereas inductive degeneration (Lb, Le) achieves 50Ω input matching. Lpck1 is tuned to optimize available gain (Ga) at the operating frequency, when Cm and Q3 are connected. The output network consists of a 300pH series transmission line and a 45fF shunt capacitor. Lpck2 and the output network assure 50Ω output matching. Biasing of the cascode stages is assured by two single-transistor current mirrors (not shown of Fig. 20). With VCC set to 1.2V, the DC power consumption for the LNA core is 10.2mW.

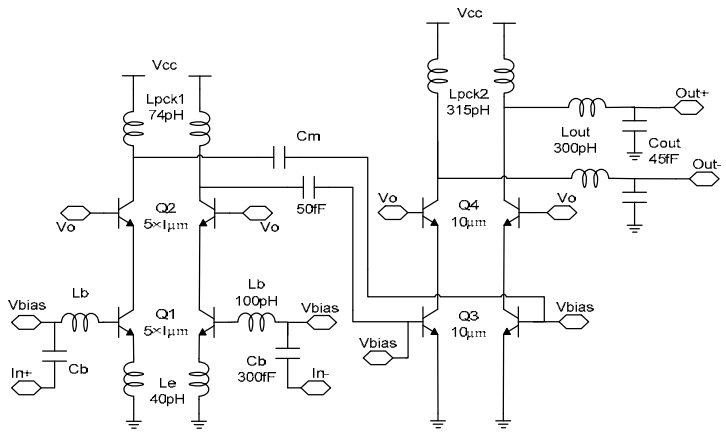


Fig. 20. Schematic view of the differential two-stage cascode LNA for 60GHz applications.

All the ideal inductors in the schematic are realized as transmission lines implemented in metal 6 and alucap shunted together, according to the design guidelines discussed in subsection 2.2.5. The minimum width allowed by the design rules (4.4μm) has been used in order to reduce area

occupation whereas the length of each line has been fixed to obtain the desired inductance value at 60GHz. The resulting dimensions of microstrips are listed in Table 1.

Table 1. Ideal values of the inductors in the schematic of Fig. 20 vs physical lengths of the equivalent microstrip lines. Microstrips width is 4.4µm.

	Inductance [pH]	Microstrip Length [µm]
Le/TLe	40	90
Lb/TLb	100	330
Lpck1/TLc1	74	167
Lpck2/TLc2	315	360
Lout/TLout	330	539

A snapshot of the layout is shown in Fig. 21. The two differential side of the circuit are symmetrically located along the horizontal axis. However, differential test at the concerned frequencies is impossible in practice, due to the lack of millimeter wave differential probes. Therefore only the upper half-side of the circuit is provided with input and output pad for single-ended measurement, whereas the input and output of the remaining half-side are connected to on chip 50Ω resistors. All the active elements, that are cascode and bias transistors, are placed in the center of the circuit; input and output pads are connected to the core through the microstrip lines corresponding to Lb and Lout respectively, which can be easily recognized in Fig. 21 as they stand along the horizontal axis in the middle of the picture. Concerning bias feeding, both first and second cascode stages on each one of the differential side are biased through a single Vcc plot and partially share a common DC path formed by transmission lines and decoupling MIM capacitors. To simplify the layout, all the bias cells are connected to a single Vbias plot, asymmetrically located at the bottom side of Fig. 21.

The chip dimension is 1254µm × 1185µm. Active elements have a negligible footprint compared to transmission lines which are implemented as straight lines. Indeed, the size of microstrips can be partially reduced simply by introducing one or more 90 degrees corners in metal path, without any considerable degradation of the circuit performances. However, the equivalence between straight and multi-corner microstrips was not demonstrated prior to the circuit design and therefore the more suitable solution (that is the use of straight microstrips) was adopted at the price of area consumption. In this work, it should be observed that this kind of modifications in path shape of transmission lines can result only in a slight reduction of silicon area; a more efficient area reduction (of 50% and even more, depending on circuit topologies) can be achieved when a lumped design strategy is used, replacing microstrips with lumped inductors. An example is reported in section 3.6.

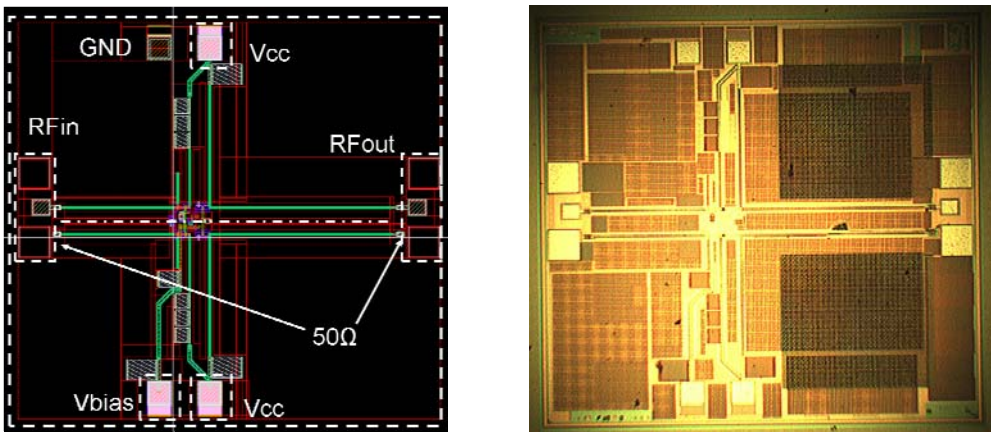


Fig. 21. Layout (left) and microphotograph (right) of the differential two-stage cascode LNA for 60GHz applications. Chip dimension is 1254µm × 1185µm.

As it was detailed in the previous chapter, we focus on three effects in the implementation of millimeter wave building blocks to improve circuit behavior. They are listed below:

- Parasitic capacitances at transistor level, between base, emitter, collector terminals and between these terminals and the substrate. In order to take these elements into account, transistor layouts have been simulated with Quickcap, a capacitor extraction tool.
- All the interconnections existing in layout, between the different circuit devices, have been modeled as distributed elements. Indeed, any portion of metal path should be considered as a distributed passive device that could potentially alter the circuit behavior and must be modeled as a transmission line of corresponding metal level and geometry.
- MIM capacitors terminations as well as RF Pads must be considered as metal6 and alucap transmission lines as explained in subsections 2.2.6 and 2.2.7 and modeled consequently.

Taking into account the parasitic extracted model of each device, the schematic of the LNA presented in Fig. 20 is modified as in Fig. 22. The fine tuning of the LNA design finally runs over this schematic.

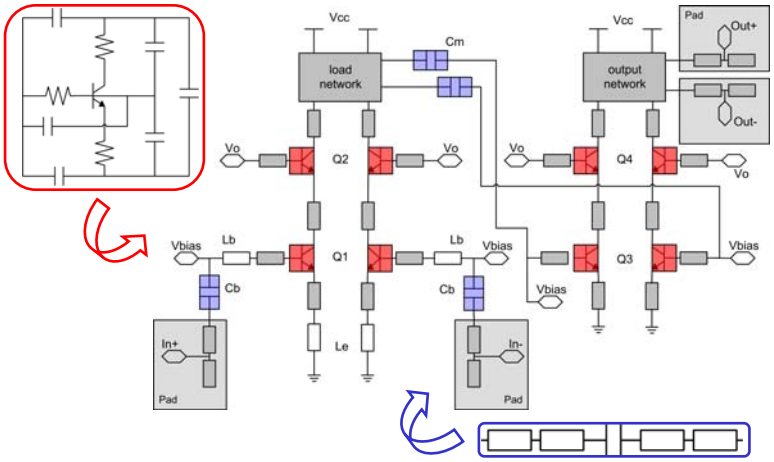


Fig. 22. Schematic view of the differential two-stage cascode LNA for 60GHz applications. Starting from the ideal circuit of Fig. 20, parasitic extracted model of transistors (red), capacitors (blue), and interconnections and RF Pads (gray) are added.

As a last step – in order to increase reverse isolation and to prevent undesired coupling effects between the different layout elements – a ground plane, directly connected to the substrate, is routed under the whole circuit at metal 1 and metal 2 levels. In proximity of microstrips, it reaches the highest metal level, forming a unique structure with the ground plane of microstrips, providing a return path for RF currents.

3.3.2 Experimental Results

The differential two-stage cascode 60GHz LNA has been characterized by on-chip measurements of S-parameters, noise figure and linearity. Since no differential GSG probes are nowadays available in the mm-Waves range, the LNA has been characterized with a single-ended configuration. To do so, one side of the differential structure of Fig. 20 is terminated with 50Ω on chip resistors. Results of post layout simulation and measurement of S-parameters are depicted in Figs. 23-24.

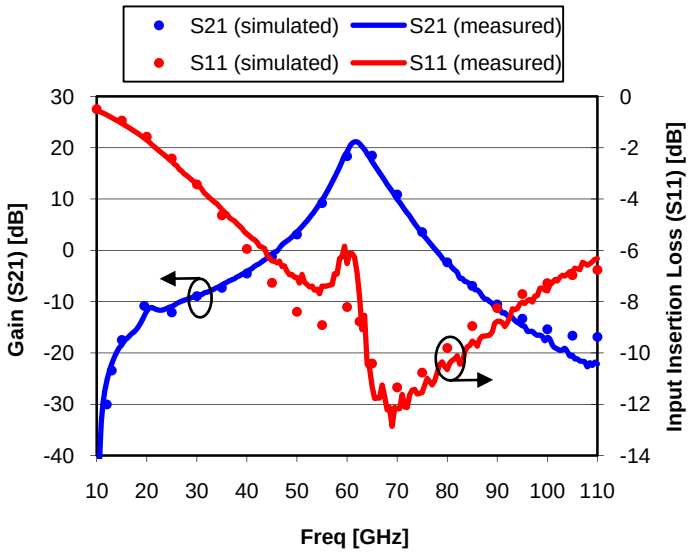


Fig. 23. Measured (line) and simulated (dot) Input Matching (S11) and Power Gain (S21).

In excellent agreement with post-layout simulations, the power gain exhibits a maximum of 21.14dB at 61.5GHz. S11 parameter plot shows a good input impedance matching over a large bandwidth (less than -6dB between 44GHz and 110GHz) and reaches its optimum (-12.9dB) at 69GHz. The peak of S11 rising at 59GHz is due to input transistor parasitic elements (Q1). Even if its extent has been slightly underestimated by simulations, this phenomenon does not affect significantly the performances of the amplifier. It can be definitely removed if a more accurate routing of input devices is accomplished.

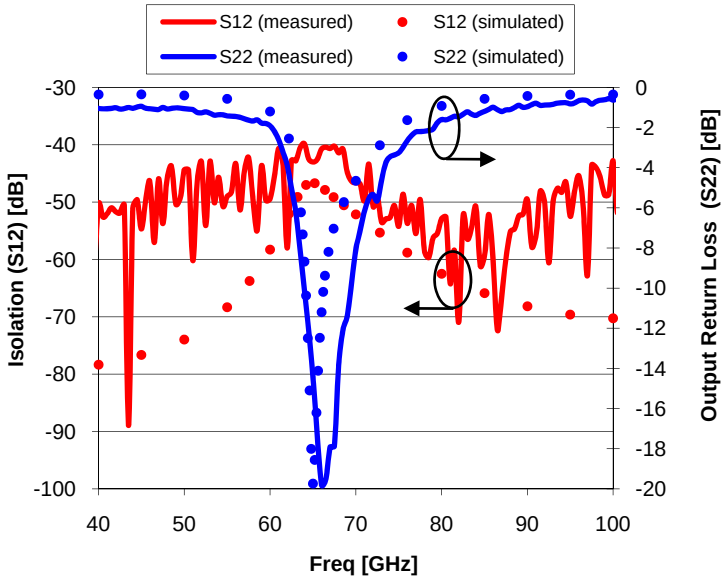


Fig. 24. Measured (line) and simulated (dot) Output Matching (blue) and Reverse Isolation (red).

Optimum measured output return loss is -20dB at 67GHz and measured reverse isolation is higher than 30dB up to 110GHz. The simulated noise figure is shown in Fig. 26. Its minimum value is 5.6dB at 60GHz.

Noise measurement tools are not yet available at IMS Laboratory and a preliminary measurement of NF was performed at IEMN Laboratory following the methodology illustrated in section 3.2.

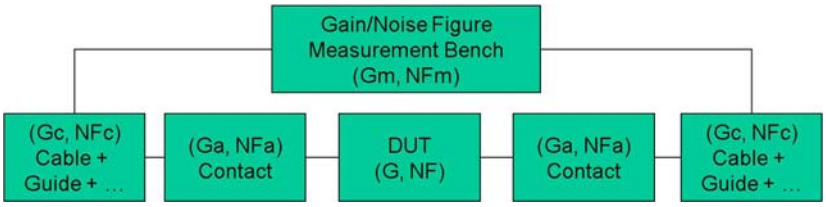


Fig. 25. Block diagram of the test bench for NF measurement at 60GHz.

The LNA was biased in order to provide 19.19dB of power gain at 60GHz (as in the case of S parameters measurement), however when performing noise measurements the resulting gain was underestimated, probably due to a bad contact of test probes. This effect introduced some incertitude in the measurement and has required a modification of the deembedding procedure described in subsection 3.2.2. Assuming that the entire mismatch in gain measurement is localized on probe contacts, the block diagram of Fig. 19 is modified as in Fig. 25.

Then the NF of the LNA can be extracted from measurement according to the Friis equation:

$$Fm = Fc + \frac{Fa-1}{Gc} + \frac{Fdut-1}{Gc \cdot Ga} + \frac{Fa-1}{Gc \cdot Ga \cdot G} + \frac{Fc-1}{Gc \cdot Ga \cdot G \cdot Ga} \qquad (7bis)$$

where Fm and Gm are respectively the values of the overall noise factor and gain as issued by the measurement, G is the expected LNA gain resulting by independent measurements (19.19dB), Gc is

the loss due to the guideline (-1.7dB) and Ga is the loss introduced by imperfection in probe contact and can be estimated as follows:

$$Ga|_{dB} = -\frac{Gdut-Gm}{2}\Big|_{dB} - Gc|_{dB} \quad (8)$$

This measurement procedure was repeated four times, obtaining the values reported in Table 2. The average value of the de-embedded NF is 6.05dB.

Table 2. NF Measurement Results.

LNA (raw data)		Correction Factor	LNA (deembedded data)	
Gain [dB] (Gm)	NF [dB] (NFm)	Total Loss [dB] 2|Gc+Ga|	F (Fdut)	NF [dB] (NFdut)
11.56	9.50	7.63	3.68	5.66
12.60	9.48	6.59	4.14	6.17
12.40	9.50	6.79	4.06	6.09
13.00	9.40	6.19	4.26	6.29

In spite of the uncertainty that affected the testing procedure, this value is relatively close to simulation prediction. The observed mismatch of 0.45dB between measured and simulated NF can be attributed to imperfection in noise models of transistors. In any case, any more reliable estimation of the NF necessary to prove or to discard this hypothesis is still subject to the availability of proper test tools. At this proposal, it should be remarked that noise measurements at millimeter wave is a very critical operation.

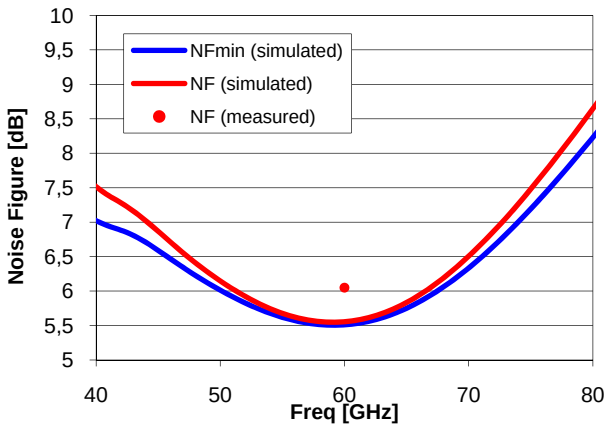


Fig. 26. Measured (dot) and simulated (line) Noise Figures. Measured data is available only at 60GHz.

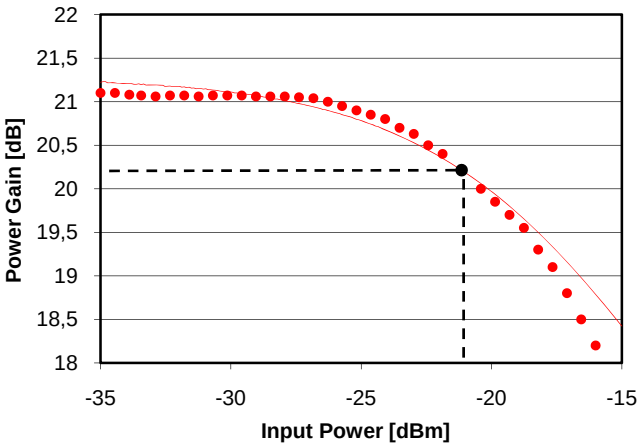


Fig. 27. Measured (line) and simulated (dot) 1dB Compression Point at -21.2dBm of Input Power.

Besides S-parameters and NF measurements, the characterization of the LNA is completed by the measurement of the power gain compression that gives an estimation of its linearity. The 1dB compression point has been measured by means of a load-pull test bench. At 61.5GHz it features -21.2dBm of input power.

As shown in Figures 23-24 and 26-27, simulations results are very close to measured data, both in small and large signal analysis. That is a proof of the correctness of the design method described in the second chapter and its accuracy in the 60GHz band.

3.4 Two-stage LNA for 80GHz applications

The search for optimum transistor geometry and biasing corresponding to noise optimization at 80GHz has been detailed in section 3.1, using, as an example, the single-stage cascode amplifier of Fig. 13, implemented with ideal passive devices. Starting from the results of paragraph 3.1.7, a complete design of a two-stage 80GHz LNA is developed hereafter. Current density has been augmented from the ideal value of 1.85mA/μm^2 to 2.9mA/μm^2 to compensate for the gain reduction due to emitter degeneration. Parasitcs effects due to transistors and passive devices as MIM capacitors, microstrips and RF Pads are included in the design. As a consequence of these modifications a power gain of 18.4dB and a simulated noise figure of 5.7dB have been obtained.

3.4.1 Theoretical design and circuit implementation

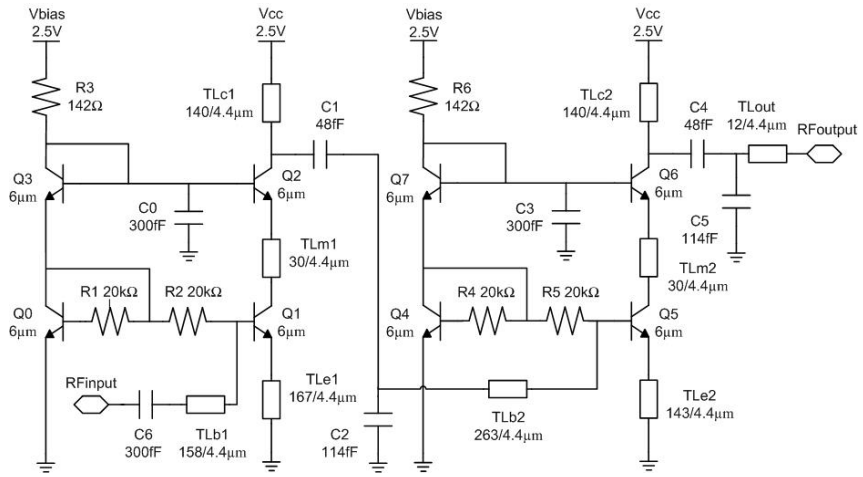


Fig. 28. Schematic view of the two-stage cascode LNA for 80GHz applications.

The schematic of the two-stage LNA is depicted in Fig. 28. All transistors are implemented as single-finger CBEB-C devices. Emitter length is fixed to 6µm to perform 50Ω noise matching. Input power matching is provided by emitter degeneration (TLb1, TLe1) and inter-stage matching is obtained combining the effect of the capacitors C1 and C2 with the emitter degeneration of the second stage (TLb2, TLe2). Each cascode stage is completed adding a 140µm peaking line (TLc1/TLc2) on the CB collector node and a 30µm microstrip (TLm1/TLm2) at the inter-stage node that corresponds to the minimal distance between the CE and the CB transistors allowed by the design rules in the layout. 50Ω output matching is obtained with the network formed by C4, C5 and TLout. It can be observed that, beside the differences in the base and emitter microstrips, the two cascode stages are identical. The LNA biased at 2.5V dissipates 45mW. Due to the symmetry between the core of the circuit and the bias cells, only one half of the power (22.5mW) is dissipated in the core.

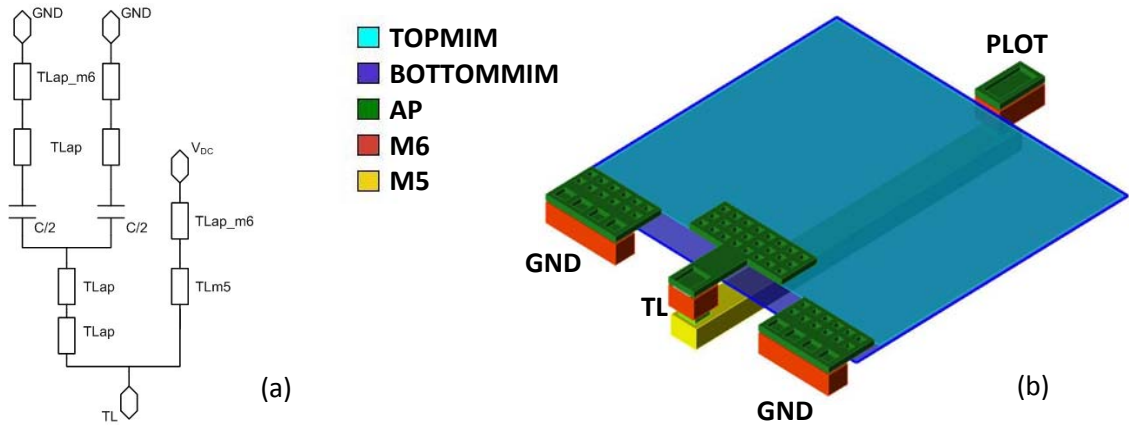


Fig. 29. A DC-decoupling MIM Capacitor. Schematic model (a) and layout (b). All elements are faithfully reproduced. The equivalent capacitance is spitted into two parts as well as the connections to ground (GND).

Once more, all the guidelines suggested in chapter 2, concerning the parasitic extraction at transistors terminals and the modeling of interconnects, RF pads, and MIM capacitors have been strictly applied to layout design. Additionally, in order to avoid any difficulty in post-layout modeling, microstrips are connected directly to transistors or passive devices without any additional interconnect and the surface of tee-junctions is minimized to reduce their impact on the circuit behavior.

The layout of MIM capacitors as defined in the standard cell is not suited for the design of decoupling devices because of the lateral connection between the capacitor itself and a microstrip line. As depicted in Fig. 37, a lateral connection results into an uncertainty on the length of the line. As a matter of fact, the layout of decoupling capacitors has been designed as in Fig. 29. In particular, the connection of the top-plate to the peaking line is located on the very terminal of the line and not on its side like in Fig. 37. As a result, the length of the peaking microstrip can be accurately controlled, avoiding any modeling incertitude.

A snapshot of the layout is shown in Fig. 30. It can be observed that several microstrips are routed as winding elements, with one or two 90° turns, rather than as straight lines. This leads to a reduction of the chip size.

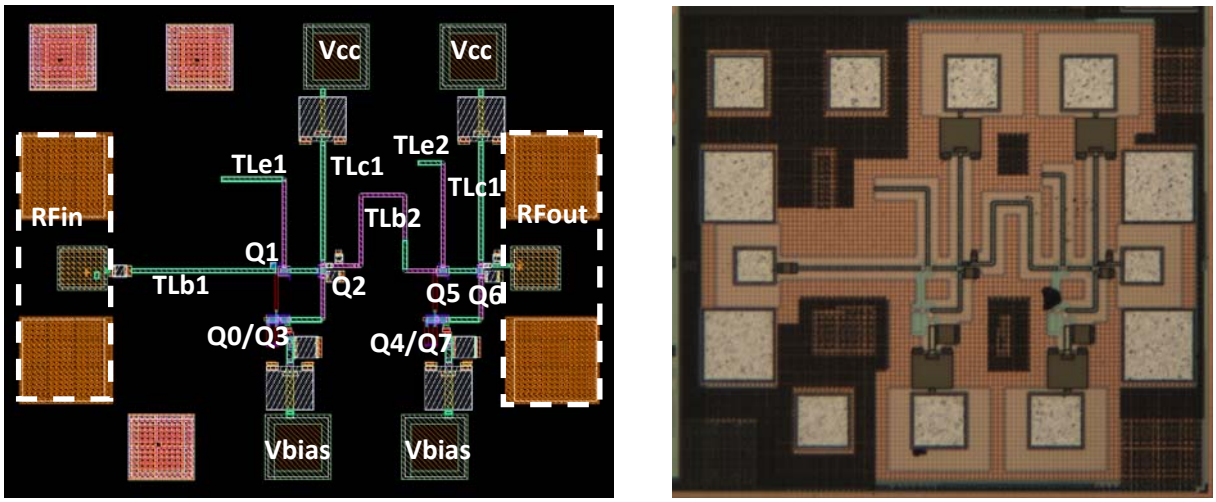


Fig. 30. Layout (left) and microphotograph (right) of the two-stage cascode LNA for 80GHz applications. Chip dimensions are 637µm × 589µm.

3.4.2 Experimental Results

As all the test circuits described so far, the two-stage 80GHz LNA has been characterized performing on chip S-parameters measurements done with an E8361A PNA network analyzer, equipped with the N5260A millimeter head controller by Agilent™, as usual. Post layout simulation and measurement results are compared in Figs. 31-32.

The power gain of the LNA exhibits a maximum of 18.43dB at 80.6GHz as predicted by post-layout simulations. As shown by the S11 parameter plot, the input matching has been overestimated due to an inaccuracy on parasitic extraction at transistor level, however S11 equals -8.1dB at maximum gain frequency and is lower than -10dB over 83.3GHz. Output return loss is -13.7dB at 80.3GHz. Reverse isolation is higher than 30dB up to 110GHz.

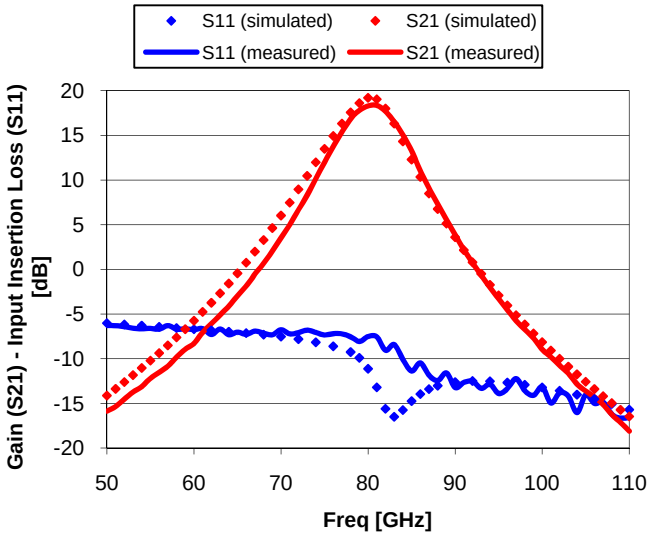


Fig. 31. Measured (line) and simulated (dot) Input Matching (S11) and Power Gain (S21).

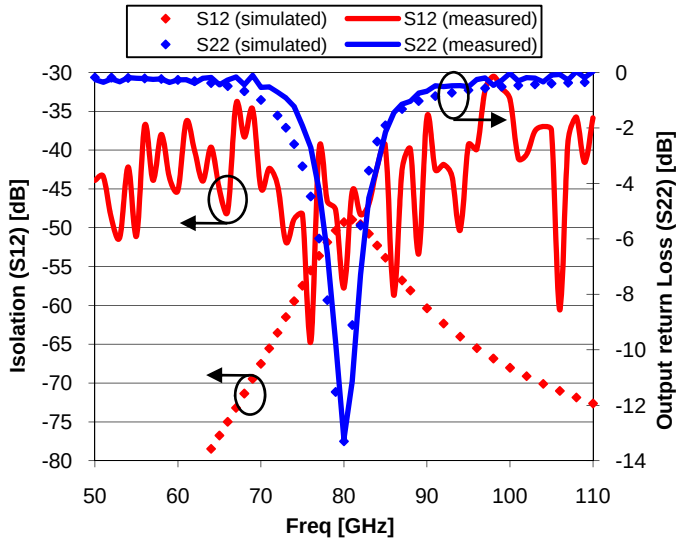


Fig. 32. Measured (line) and simulated (dot) Output Matching (S22) and Reverse Isolation (S12).

Noise measurements have not yet been performed due to the lack of adequate instrumentations. Simulated noise figure features 5.7dB as shown in Fig. 33. It is the best value that can be obtained at 80GHz, consistently with reasonable power gain and accounting for circuit non-idealities.

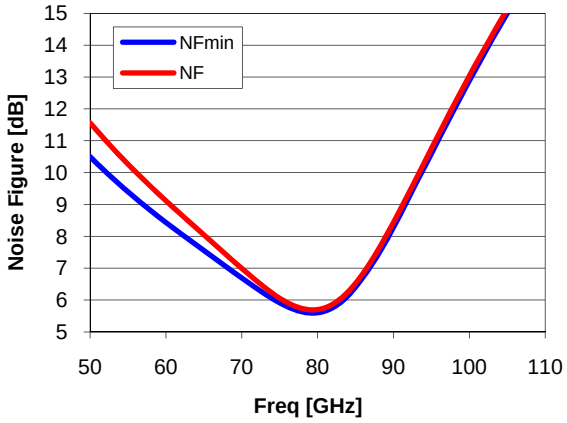


Fig. 33. Simulated noise figure.

Besides small-signal measurements, the LNA characterization is completed by the 1dB compression point measured as described in section 3.2. It corresponds to -21.12dBm of input power at 80.6GHz, as depicted in Fig. 34.

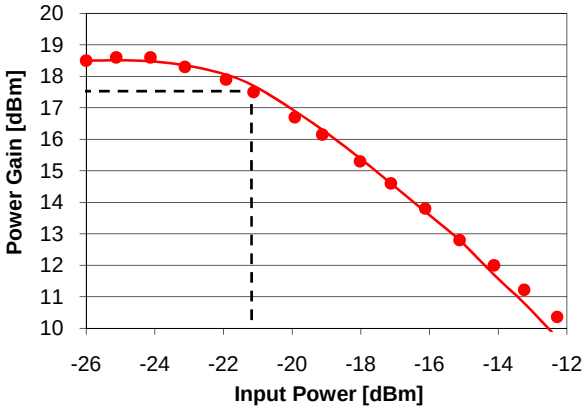


Fig. 34. Measured (dot) and simulated (line) 1dB Compression point at -21.12dBm of Input Power.

The experimental characterization of the two-stage LNA gives a demonstration of the reliability of the applied design methodology in the 80GHz band.

3.5 Single-stage LNA for imaging applications at 94GHz

The design and the experimental results of a two-stage 60GHz LNA on one hand, and of a two-stage amplifier operating at 80GHz on the other hand, have been reported so far. Now, in order to extend the frequency limit of validity of our design methodology up to the 94GHz standard for imaging application, an example of single-stage LNA operating at 94GHz will be presented.

3.5.1 Theoretical design and circuit implementation

The schematic of the circuit is depicted in Fig. 35. Transistor Q1 and Q2 are implemented with five emitters of 1µm length and are biased with 3.6mA. The emitter length as well as the number of emitters is set to achieve 50Ω noise impedance, whereas input impedance matching at 94GHz is obtained by means of an inductive degeneration (TLb, TLe). A 256µm transmission line (TLm) and a 50fF shunt capacitor (C3) feature the inter-stage matching. These values have been chosen in order to maximize the contribution of the input transistor to overall gain and, as a consequence, to get best noise performance on the amplifier, according to Friis equation. Output network devices are tuned to optimize available gain at the operating frequency over a 50Ω load. With Vcc and Vbias set to 3V and 1.3V respectively, the DC power consumption is of 10.74mW for the LNA core and 2.26mW for the bias cell (Q0).

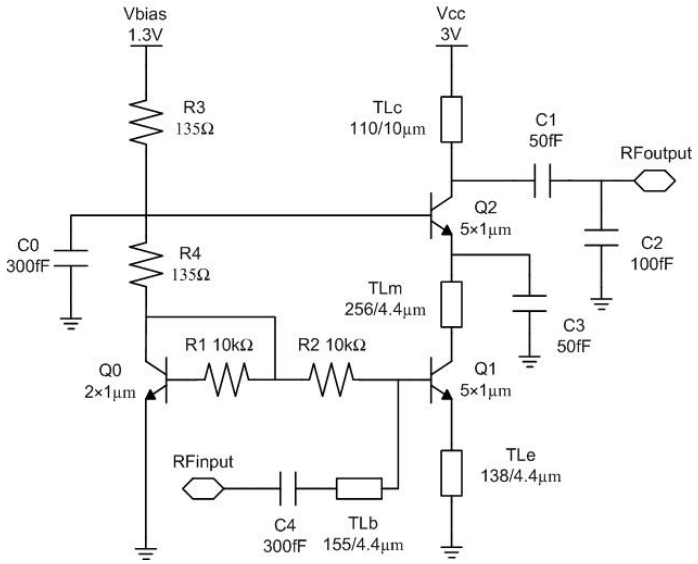


Fig. 35. Schematic view of the single-stage cascode LNA for 94GHz applications.

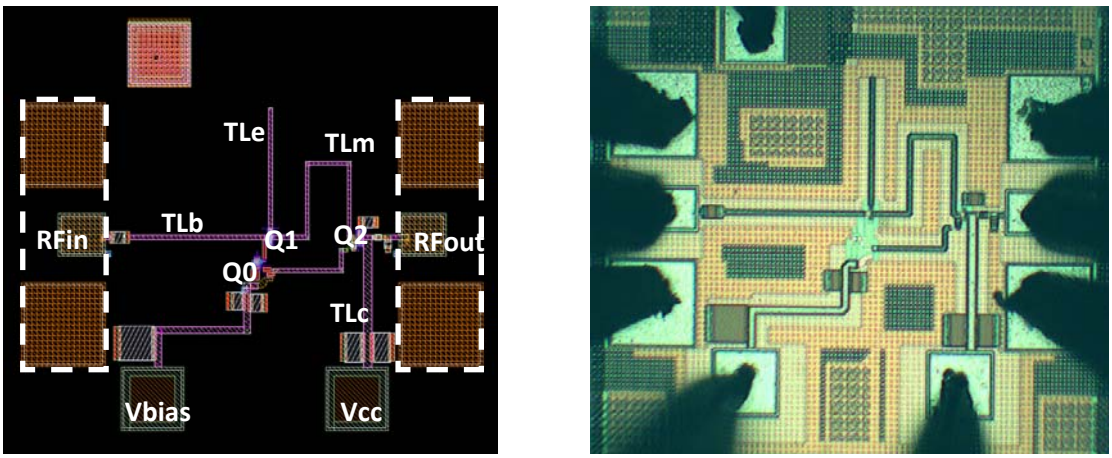


Fig. 36. Layout (left) and microphotograph (right) of the single-stage cascode LNA for 94GHz applications. Chip dimensions are 529 × 535 µm².

As in the case of the other realizations discussed so far, all the design guidelines of chapter 2 concerning the parasitic extraction at transistors terminals and the modeling of interconnects, RF Pads, and MIM capacitors have been here applied. To simplify the layout modeling, some specific configurations have been tested.

Despite of this simplification effort, three elements have been maintained in the layout that can potentially result into undesirable effects at millimeter waves, as shown in the snapshot of Fig. 37:

- two tee-junctions have been used to realize interconnections between transmission lines and capacitors in the output network;
- decoupling capacitors on the biasing plot of the cascode branch (Vcc) are directly connected to the peaking line.

In this particular case, tee-junctions, because of their small sizes, have almost no effect on the LNA performances and can be neglected. The presence of the decoupling capacitors on the peaking line is more critical, because it results into an incertitude of 34µm on the effective length of the line. Since the operating frequency of the amplifier strictly depends on the length of the peaking line, the latter phenomenon must be carefully investigated. As a matter of fact, the circuit is correctly modeled as in Fig. 37 where the microstrip length (110µm, as fixed by design considerations) is calculated from collector node up to the lower edge of capacitor connections. From a practical point of view, whereas it is almost impossible to avoid all tee-junction interconnects in a layout, an alternative way to connect the decoupling capacitors that avoids any modeling incertitude is suggested in section 3.4.

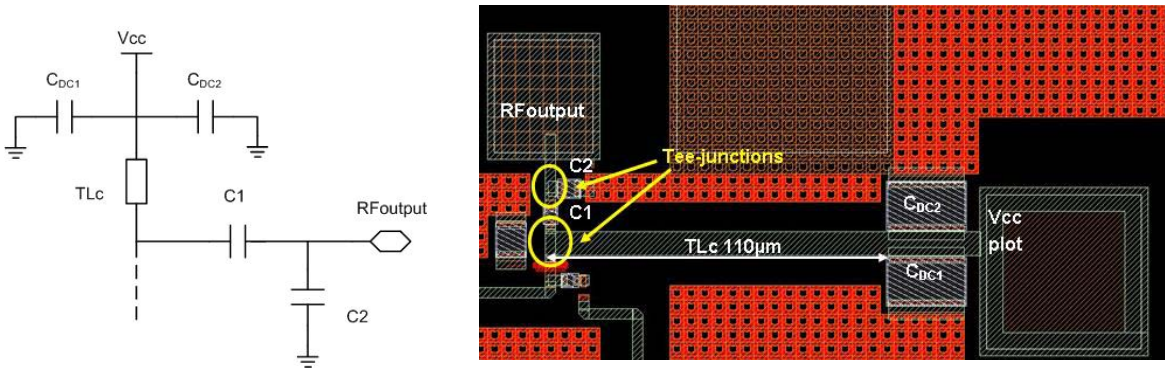


Fig. 37. Output network of the 94GHz LNA: schematic (left) and layout (right). Effective length of the peaking transmission line is 110µm.

3.5.2 Experimental Results

Post-layout simulation and measurement results are depicted in Figs. 40-41. In good agreement with post-layout simulations, the power gain exhibits a maximum of 9.08dB at 94.7GHz. S11 parameter plot shows a good input impedance matching over a large bandwidth (less than -10dB from 50.9GHz to 96.5GHz) and reaches its optimum (-15.5dB) at 75.4GHz. Output return loss is -12dB at 94.7GHz.

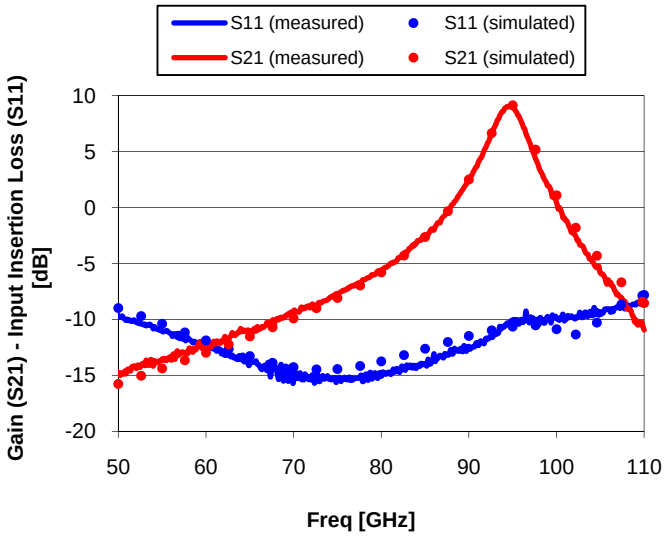


Fig. 38. Measured (line) and simulated (dot) Input Matching (S11) and Power Gain (S21).

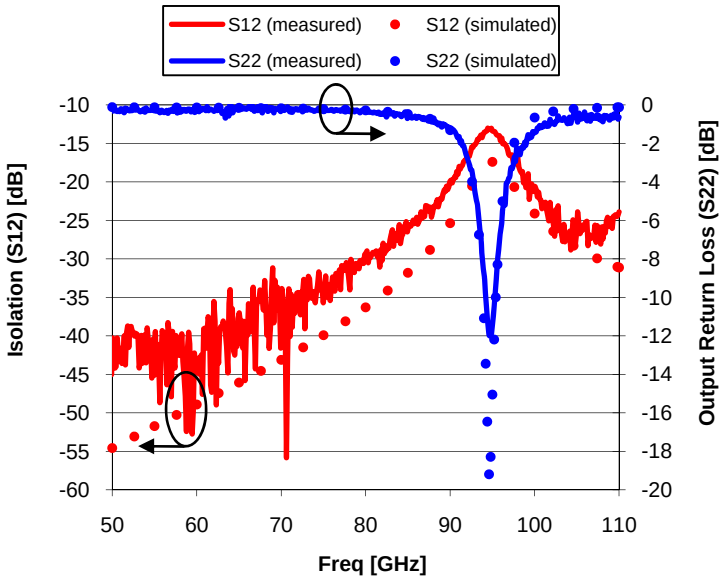


Fig. 39. Measured (line) and simulated (dot) Output Matching (S22) and Reverse Isolation (S12).

The Noise Figure is 8.6dB at 95GHz. The 1dB compression point measured at 94.7GHz features -14.9dBm of input power. Reverse isolation is higher than 13dB up to 110GHz.

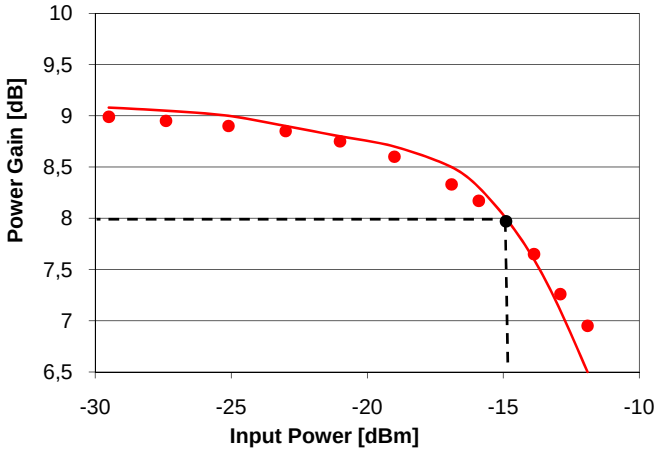


Fig. 40. Measured (dot) and simulated (line) 1dB Compression point at -14.9dBm of Input Power.

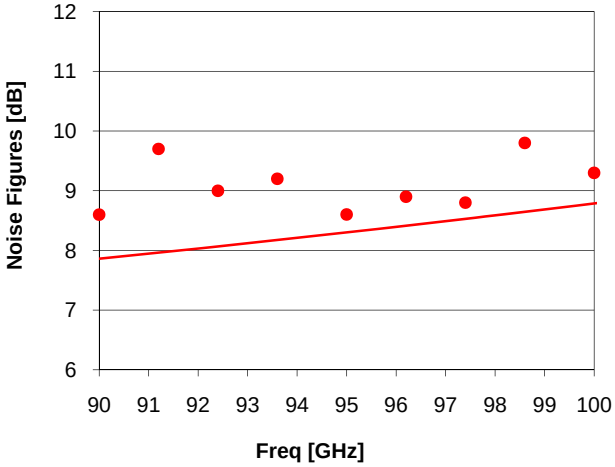


Fig. 41. Measured (dot) and simulated (line) Noise figure.

As shown in Figs. 38-41, simulations results are very close to measured data, both in small and large signal analysis. This proves that the proposed design methodology can be correctly extended up to 94GHz, providing accurate results over the entire millimeter-Waves context.

3.6 Single-stage 80GHz LNA: Comparison between lumped and distributed implementations

Three examples of LNAs, designed according to a distributed approach, have been described in the previous sections, proving the correctness of the design methodology based on distributed passive devices.

The aim of this section is to compare the efficacy of both distributed and lumped approaches. The design of a single-stage cascode LNA operating at 80GHz is therefore detailed hereafter. Two version of the same amplifier are presented:

- a first version is designed following a conventional distributed design approach as explained in section 2.3 and is implemented using distributed passive elements such as the microstrip lines provided by the BiCMOS9MW technology;
- a second version of the same amplifier is then proposed, where distributed microstrips are replaced by custom-designed lumped inductors.

As it will be proved by measurements results, both the distributed and the lumped LNA implementations give satisfactory results. Finally, a comparative analysis will highlight the advantages of lumped design over the conventional methodology based on distributed elements.

3.6.1 Theoretical design and circuit implementations

The schematic used for the implementation of the 80GHz LNA is based on that of the 94GHz LNA reported in Fig. 35 and described in details in section 3.5. The advantage of such a circuit is that its S11 parameter plot shows a good input impedance matching over a very large bandwidth (less than -10dB from 50.9GHz up to 96.5GHz), with a minimum of -15.5dB at 75.4GHz and therefore it can be used as a start point to implement a 80GHz amplifier with minor modifications.

Practically speaking, to move the operating frequency from 94GHz to 80GHz the length of the peaking transmission line TLc has been augmented from 110 to 147µm. A simplified schematic of this single-stage LNA is presented in Fig. 42. With Vcc and Vbias set to 1.5V, the DC power consumption is of only 5.4mW for the LNA core and 3.3mW for the bias cell (Q0). A snapshot of the layout is reported in Fig. 43.

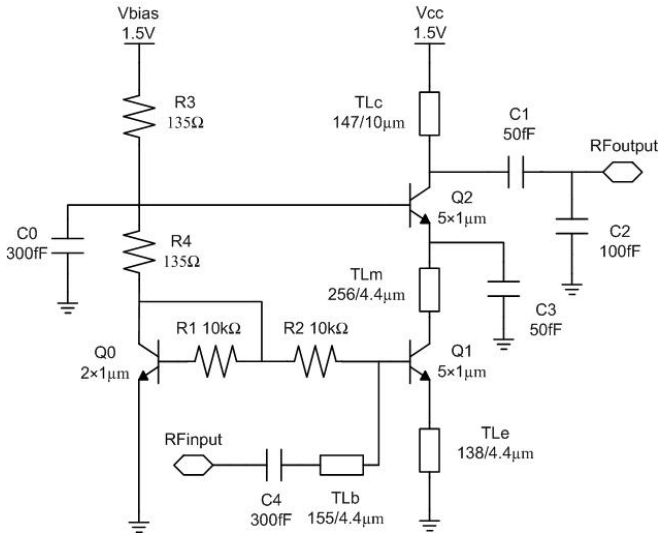


Fig. 42. Schematic view of the single-stage cascode LNA for 80GHz applications.

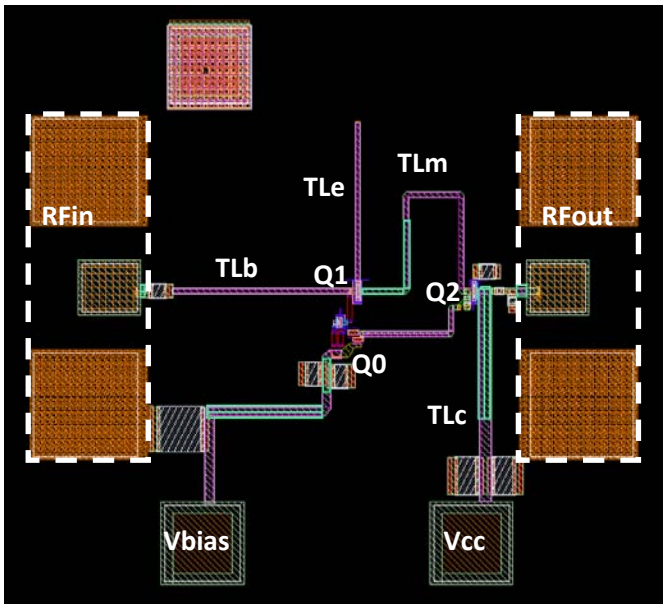


Fig. 43. Layout of the single-stage cascode LNA for 80GHz applications. First version, with distributed transmission lines. Core circuit dimensions are 336 × 344 µm².

A second version of the same circuit has been fabricated by replacing three of the four transmission lines of the first version with lumped inductors, namely base, emitter and inter-stage inductors.

As depicted in Fig. 44, TLb microstrip at the base node corresponds to an equivalent inductance value of 74pH at 80GHz, so it can be replaced by the lumped inductor of Fig. 2.35 (a), featuring 74pH as well, as already explained in paragraph 2.4.4. This inductor was fabricated and characterized prior to the LNA design and its value has been confirmed by measurement.

The 256µm microstrip used for inter-stage matching (TLm) corresponds to a 142pH inductance. It can be replaced again by the inductor of Fig. 2.35 (a) connected in series with a transmission line, tuning the resulting inductance value to 142pH, as required. However, the length of such a line resulting into 143pH is 120µm, somewhat exaggerated compared to inductor size. Simulations demonstrate that the inter-stage inductance can be reduced with no degradation of the LNA performances and then a 33µm microstrip connected to the lumped device of Fig. 2.35 (a) has been retained in the final version of this design, resulting in a 95pH inductance and corresponding to the optimum device footprint in the layout.

It can be remarked that the use of the same inductor both at the base and at the inter-stage nodes overcomes the need for a new device design and avoids any uncertainty on inductor value prediction.

Emitter degeneration, on the contrary, requires a smaller inductance value, namely 67pH, so a new inductor has been designed (Annex 1) using the methodology described in section 2.4.

Finally, in the case of the peaking line, the microstrip implementation is to be preferred because its footprint is controlled by the distance between the cascode transistor and the biasing plot. This distance is fixed by pads location on layout and therefore can not be reduced replacing the microstrip with a lumped inductor.

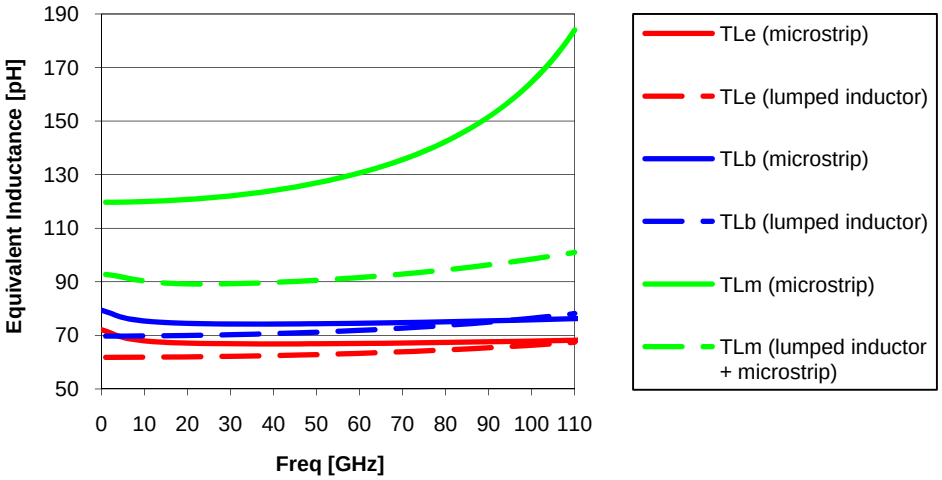


Fig. 44. Equivalent Inductance of the microstrip lines in the schematic of Fig. 42 compared to their corresponding lumped implementations used in the second version of the LNA.

As demonstrated by post-fabrication measurements reported in subsection 3.6.2, despite the small mismatch that can be eventually observed between the nominal values of the three inductors implemented as microstrips or as lumped devices, the two circuit exhibit very similar performances. This means that the circuit topology is not sensible to small variation of inductors values and therefore the incertitude that results from inductors modeling has no consequence on the overall performances of the amplifier. Generally speaking, this is a key element for successful lumped design of millimeter wave building blocks.

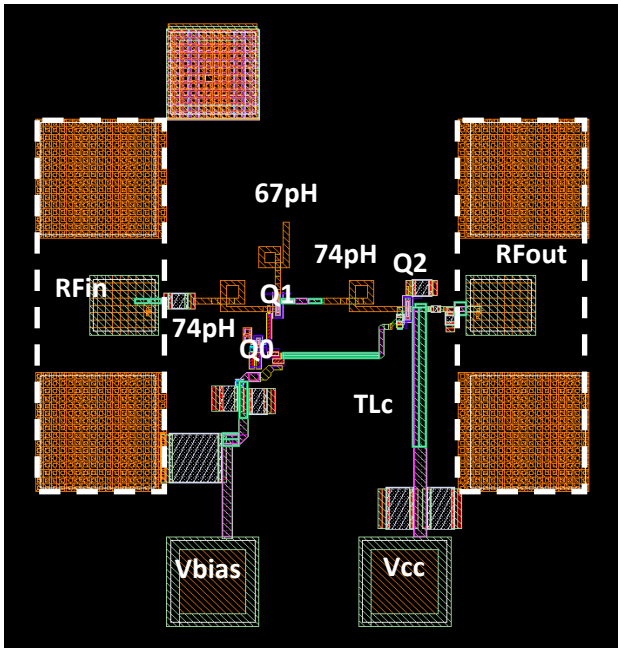


Fig. 45. Layout of the single-stage cascode LNA for 80GHz applications with lumped inductors. Core dimensions are only 240 × 270 µm^2. The use of lumped devices leads to a reduction of 44% compared to the distributed version of Fig. 43.

3.6.2 Experimental Results

The distributed and the lumped versions of the single-stage 80GHz LNA have been characterized by means of on chip S-parameters measured with an E8361A PNA network analyzer, equipped with the N5260A millimeter head controller by Agilent™. Post layout simulation and measurement results are compared in Figs. 46-47.

In good agreement with post-layout simulations, the power gain of the distributed LNA exhibits a maximum of 10.7dB at 82.05GHz. S11 parameter plot shows a good input impedance matching over a large bandwidth (less than -10dB from 62.5GHz to 106GHz) and reaches its optimum (-14.5dB) at 88GHz. Output return loss is -16.9dB at 82GHz.

The lumped LNA have a very similar behavior, with the maximum power gain of 11.3dB at 81.5GHz. Due to the slight difference in the value of the inter-stage inductor, the input insertion loss has a less regular behavior, rising slightly above -10dB below 81.7GHz and reaching its optimum (-19.9dB) at 86.5GHz. Optimum output return loss is -11.4dB at 81.5GHz. Reverse Isolation is higher than 14dB for both the circuits.

Noise Figure has not yet been measured due to the lack of measurement tools. As shown in Fig. 48, simulated NF is, respectively, 7.2dB at 82GHz for the distributed and 7.3 at 81.5GHz for the lumped amplifier. The latter exhibits a higher noise matching level.

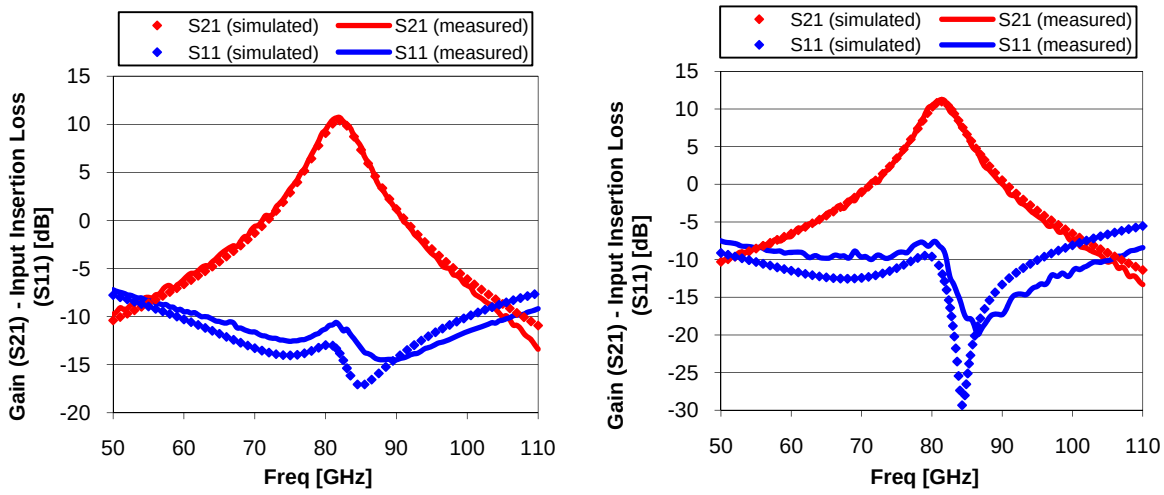


Fig. 46. Measured (line) and simulated (dot) Input Matching (S11) and Power Gain (S21) of the distributed (left) and the lumped (right) version of the single-stage 80GHz LNA.

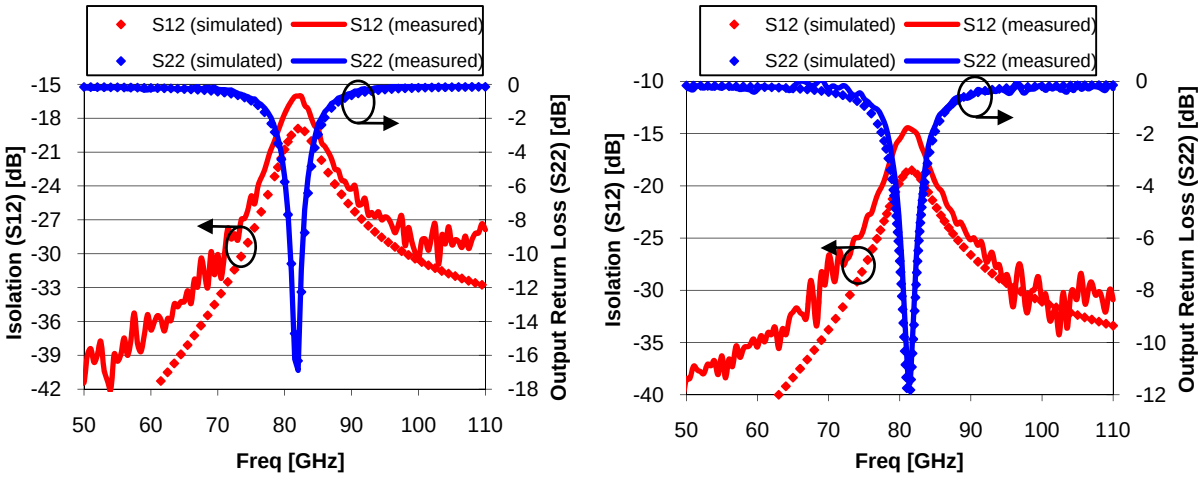


Fig. 47. Measured (line) and simulated (dot) Output Matching (S22) and Reverse Isolation (S12) of the distributed (left) and the lumped (right) version of the single-stage 80GHz LNA.

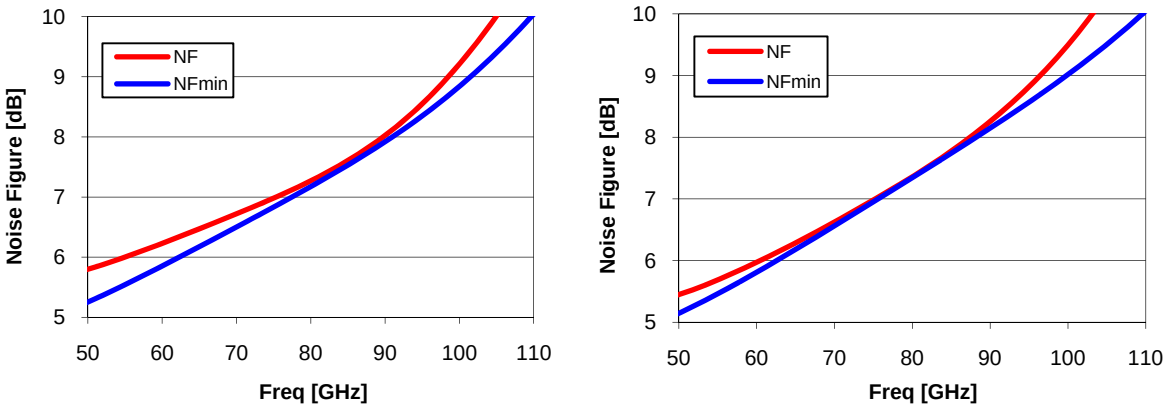


Fig. 48. Noise figure of the distributed (left) and the lumped (right) version of the single-stage 80GHz LNA.

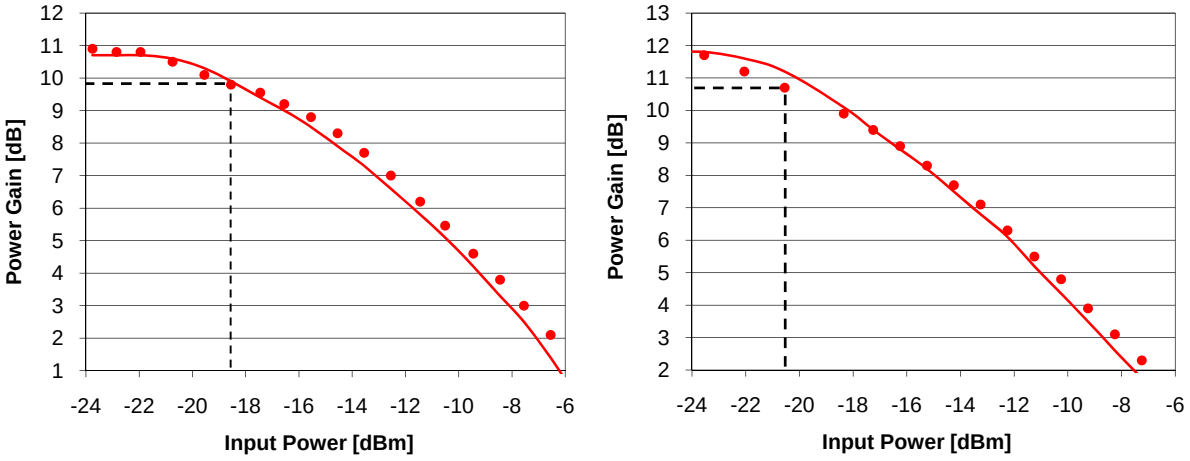


Fig. 49. Measured (dot) and simulated (line) 1dB Compression point for the distributed (left) and the lumped (right) version of the single-stage 80GHz LNA is at -18.55dBm and -20.55 of Input Power, respectively.

The 1dB Compression Points reported in Fig. 49 have been measured by means of a power meter as described in section 3.2.

S parameters of the distributed and lumped LNAs are compared in Fig. 50. As already observed, besides the excellent agreement between post-layout simulation and measurement results for each one of these circuits, they both exhibits very similar performances. The small mismatch that can be observed between the two plots of S11 parameter is justified by the use of a 95pH lumped inductor to replace the 256µm microstrip at the inter-stage node. At the same time, this can also explain the reduced optimum value of output return loss in the lumped version, since output matching depends on the inter-stage node and – due to the low isolation level of single-stage architecture – also on the input node, and not only on the output network that is exactly the same for both circuits.

The microphotographs of the two LNAs are reported in Fig. 51. Focusing on the core of each circuit (that is, neglecting pad contributions), it can be remarked that dimensions are reduced from 336 × 344 µm² to 240 × 270 µm² when moving from the distributed version to the lumped one. In other words, the lumped LNA provides a 44% reduction of area occupation compared to the distributed version, and this with no degradation of the electrical performances.

We can therefore conclude that the design methodology described so far can be successfully applied to distributed and lumped design. Furthermore, design based on lumped inductors has the precious advantage of allowing a significant reduction of silicon area occupation, compared to conventional design implementations based on distributed elements such as transmission lines.

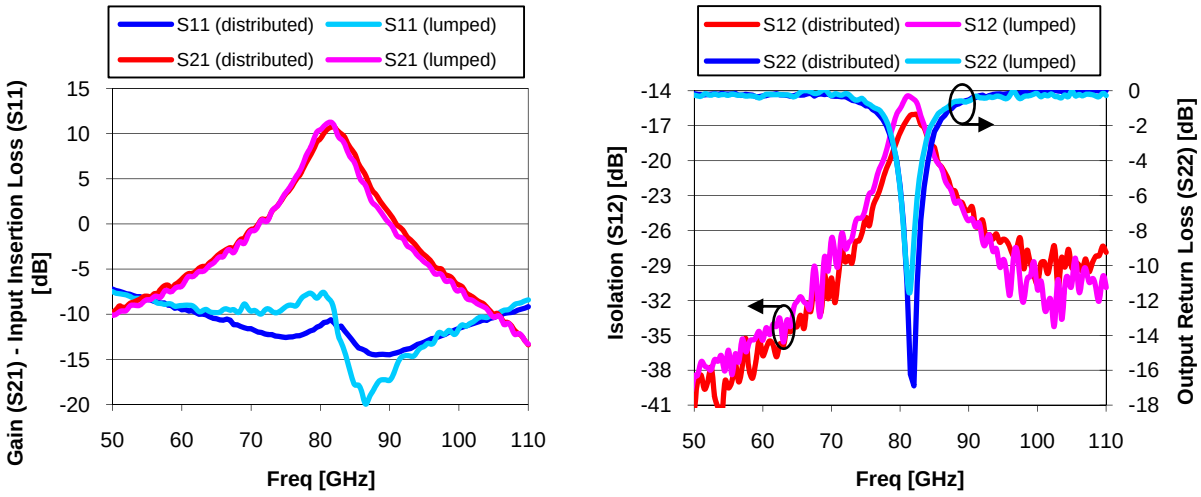


Fig. 50. S-parameters of the single-stage 80GHz LNA: distributed vs lumped implementation.

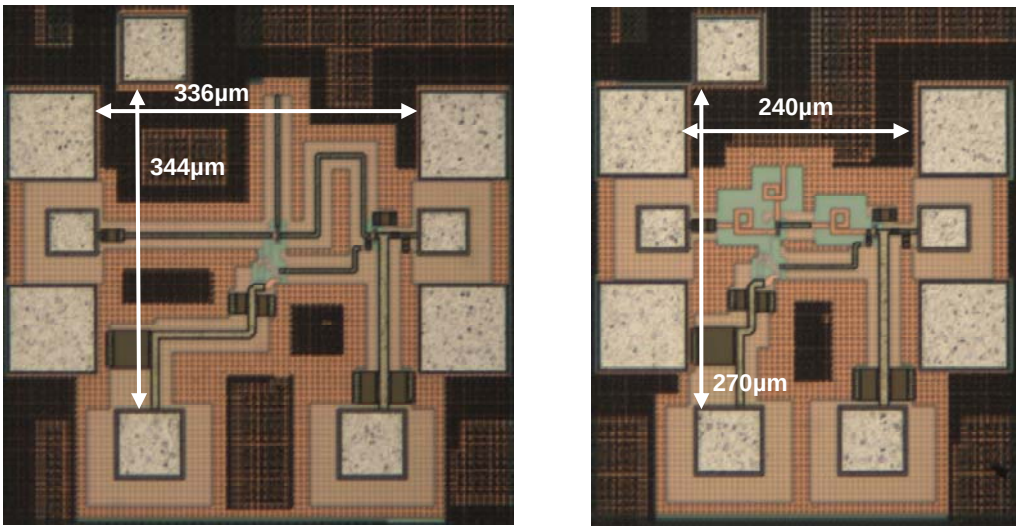


Fig. 51. Chip photos of the single-stage 80GHz LNAs: distributed (left) and lumped (right) version.

3.7 Comparison with the State of the Art

The design and experimental results of a two-stage LNA operating at 60GHz, a two-stage 80GHz LNA and a single-stage LNA for 94GHz applications, designed according to a distributed approach have been reported so far. To evaluate the benefits of the lumped design, two versions of a single-stage 80GHz LNA have been realized using, respectively, distributed transmission lines and lumped inductors.

The aim of this section is to compare the performances of all our circuits to the many examples of millimeter wave LNAs reported in the literature. The simplest way to perform such a comparison is to evaluate the main characteristics of each circuit using a unique Figure of Merit (FoM). A definition of the FoM for LNAs is given by the International Technology Roadmap for Semiconductors (ITRS) Committee. It links the gain, the noise factor (F), the power consumption (Pwr), and the linearity (IIP3), according to (9):

$$FoM = \frac{Gain \times Freq \times IIP3}{(F-1) \times Pwr} \quad (9)$$

However the IIP3 of our circuits has not been measured and is often neglected also in the literature. A simplified expression of the FoM is given in equation (10). It will be retained for our analysis.

$$FoM = \frac{Gain \times Freq}{(F-1) \times Pwr} \quad (10)$$

Examples of LNAs operating at 60GHz, in the 77-80GHz band, and at 94GHz are reported in Table 3-5, respectively. Our circuits exhibit performances in the state of art. In particular, the two-stage 60GHz LNA described in section 3.3 exhibits the best FoM over the 4 reported references.

Amongst our three LNAs operating at 80GHz, the FoM is lower for the single cascode stage (the distributed and the lumped version have almost the same value of FoM). The two-stage circuit has a higher FoM because of its higher gain. At this proposal, it should be remarked that the gain has a predominant role in the definition of the FoM given by (10). Therefore, multi-stage circuits exhibit better FoM, whereas the performances *per stage* are somewhat neglected.

A summary of the performances of our 94GHz LNA is reported and compared with other realizations in Table 5. This circuit exhibits the fifth best FoM over the 7 references. Once more, higher gain and therefore higher FoM are obtained by multi-stage realizations.

Table 3. Comparison with the State of Art of 60GHz LNAs. (* Simulated NF).

Ref	Topology	Tech	FoM	Freq [GHz]	Gain [dB]	NF [dB]	Power Consumption
This work	**2 Cascode**	**130nm**	**255**	**61.5**	**21**	**6**	**1.2V/8.5mA**
[4]	1 CB + 1 Cascode	120nm	99	61.5	15	4.5	1.8V/6mA
[5]	2 Cascode	250nm	55	60	20	6*	3.3V/11mA
[6]	3 Diff CE	250nm	15	60	18	6.8*	2.2V/30mA
[7]	2 Cascode	180nm	49	52	22	7.5*	3.3V/11mA

Table 4. Comparison with the State of Art of 80GHz LNAs. (* Simulated NF).

Ref	Topology	Tech	FoM	Freq [GHz]	Gain [dB]	NF [dB]	Power Consumption
Distributed	**1 Cascode**	**130nm**	**26**	**82**	**10.7**	**7.2***	**1.5V/8.7mW**
Lumped	**1 Cascode**	**130nm**	**29**	**81.5**	**11.3**	**7.3***	**1.5/8.7mW**
Two-stage	**2 Cascode**	**130nm**	**91**	**80.6**	**18.4**	**5.7***	**2.5V/22.5mW**
[8]	1 Cascode + 1 CE	180nm	15	77	14.5	6.9	37mW
[9]	2CE + 1 Cascode	130nm	285	86	25	5.3*	1.8-2.5V/40mW
[10]	2 Cascode	130nm	111	77	23.8	5.7	3.5V/17.5mA
[11]	3 Cascode	250nm	12	79	21.7	10.1*	3.5V/30mA
[12]	1CB + 1 Cascode	120nm	64	77	15	5.6*	1.8V/14.4mW
[13]	2CE	300nm	2	77	8.9	4.8	5.5V/22mA
[14]	2 CE Diff	-	2	76.5	12	9.5	3.3V/79mW

Table 5. Comparison with the State of Art of 94GHz LNAs. (* Simulated NF).

Ref	Topology	Tech	FoM	Freq [GHz]	Gain [dB]	NF [dB]	Power Consumption
This work	**1 Cascode**	**130nm**	**9**	**94.7**	**9**	**8.6**	**3V/13mW**
[15]	2 Cascode	250nm	6	94	16.2	10.6	3.5V/61mW
[16]	5 CE	120nm	106	98	23	8*	1.2V/29mA
[17]	2 Cascode	130nm	63	90	22	7	56mW
[18]	Diff 3 Cascode	250nm	3	94	15	11.3*	3.3V/28mA
[19]	1 Cascode	120nm	100	91	13	5.1	8.1mW
[20]	1 CE	130nm	7	94	5	7*	11.2mW
[20]	2 CE	130nm	10	94	10.5	7.7*	21.3mW

3.8 An application of a millimeter wave LNA: a 80GHz receiver Front-end

The feasibility of millimeter wave LNAs with state of art performances designed using the methodology described in chapter 2 has been demonstrated so far. Furthermore other circuits (namely a mixer, a VCO and a power amplifier operating at 80GHz) have been recently designed at IMS laboratory using the same methodology, extending its validity to the design of all the most common building blocks.

Next step in such a context is the development of a complete receiver front-end in the BiCMOS9MW technology.

The feasibility of millimeter wave building blocks on silicon-based technology, as already explained many time, is a condition for the production of low-cost systems supporting millimeter wave applications. But in order to further reduce production costs, it is necessary to adopt design solutions based on fully-integrated systems, like integrated transceivers, or, at least, integrated receiver front-ends and transmitters.

To investigate the feasibility of a millimeter wave front-end integrated on a single chip, two circuits have been designed and fabricated. The first one is a complete receiver consisting of a two-stage cascode single-ended LNA, a double-balanced differential mixer and a synchronized push-push VCO. Integrated transformers are added at RF and LO inputs of the mixer to operate single to differential conversion.

Since the use of a push-push oscillator can eventually result into a low-power output signal, potentially inadequate to properly drive the mixer, a second circuit has been realized including only the LNA and the mixer, without any local oscillator. With a LO signal generated off-chip, it can be used for a complete characterization of LNA and mixer in optimal conditions, regardless of the oscillator behavior.

The LNA included in the front-end is the two-stage cascode described in section 3.4. Since the design of the front-end has been made prior to LNA measurement, two minor modifications in the circuit schematic of Fig. 28 were made as a precaution:

- capacitors C0 and C3 have been reduced from 300 to 100fF, in order to reduce the risk of instability;
- the length of transmission line TLc1 and TLc2 has been reduced in both circuits from 140 to 134µm to prevent any eventual down-shift of the operative frequency.

However the experimental characterization of the LNA as a standalone building block reported in subsection 3.4.2 have demonstrated that the original circuit works efficiently and therefore all the subsequent modifications were unnecessary.

The mixer used for the co-integration with the LNA is based on a former design described in [21] and [22] as a stand-alone implementation. The schematic, illustrated in Fig. 52, is based on a double-balanced Gilbert cell. Transistors dimensions and current density are chosen to optimize the conversion gain (CG) and to reduce the NF. Two transmission lines TL2 and TL4 are inserted between the transconductor and the switching stages to perform inter-stage matching and compensate for the parasitic capacitances which contribute to lower both the isolation between LO and RF ports and the conversion gain. Single-ended to differential conversion of the input signals at RF and LO ports is made using integrated transformers. The output differential signal at the intermediate frequency is recombined into a single-ended termination for measurement purpose by an off-chip balun.

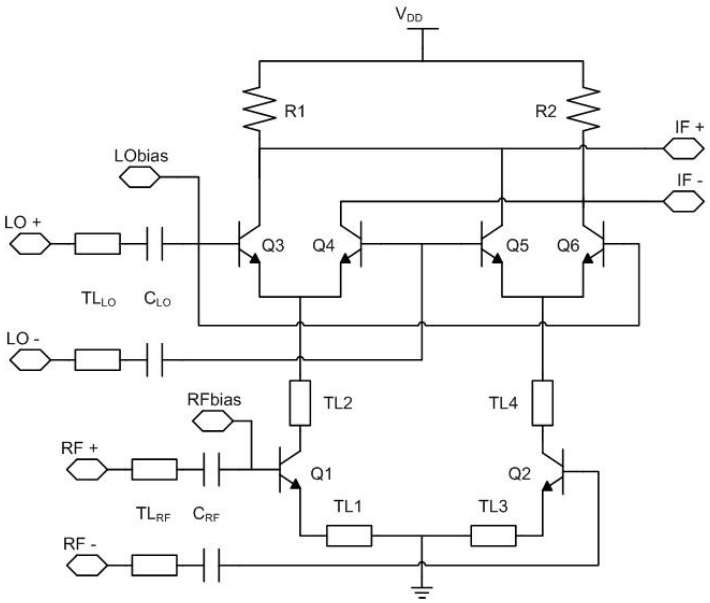


Fig. 52. Schematic view of a double-balanced Mixer (Bias cells not shown).

Concerning the design implementation, all the suggestions of chapter 2 have been followed. Furthermore, electromagnetic simulations have been performed on HFSS for transformers design.

Besides single-ended to differential conversion, the two integrated transformers added at the RF and LO ports ensure input matching over a large bandwidth. As shown in Fig. 53, transformers are routed

with the two top thick metal layers (metal 5 and metal 6), using stacked topology with octagonal shape turn.

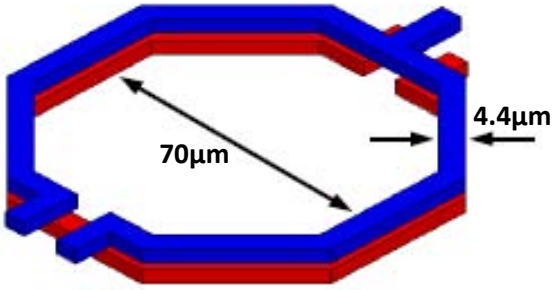


Fig. 53. Layout of the integrated transformer used for single-ended to differential conversion of the input signals at RF and LO ports of the mixer.

The primary and secondary access lines are located on opposite sides of the transformer (flipped topology) and the primary is center-tapped to match the mixer layout. Each winding exhibits a single octagonal turn with an average diameter of 70µm and a trace width of 4.4µm. The phase difference between the differential terminals shows little deviation from the desired 180 degrees over the considered frequency band.

Mixer performances have been measured at IMS. It provides a good trade-off between gain, noise figure and power consumption. Measured conversion gain and SSB noise figure are respectively 18.5dB and 13.8dB, over a 74 to 81 GHz range. The power consumption is only 80mW under 2.5V and the ICP1 is -13dBm for a RF frequency of 77GHz. Measurements results are reported in Table 6.

Table 6. Characteristics of the mixer as a standalone circuit measured with LO signal at 80GHz. (*) NF is reduced to 5.5dB after de-embedding of integrated transformers.

Maximum Conversion Gain	18.5dB at 77GHz
Minimum SSB Noise Figure*	13.8dB at 77GHz
IP1dB	-13dBm at 77GHz
Power Consumption	80mW
Area	0.57mm^2
Required LO power for best operation	~1dBm

The circuit that generates the LO signal in the integrated front-end is a synchronized push-push oscillator. A stand-alone implementation of this block, formerly fabricated and tested, is described in [23].

The core of the push-push oscillator is depicted in the diagram of Fig. 54 and consists of two symmetrical and individual Colpitts sub-oscillators working approximately at 40GHz. The signals generated by the two sub-oscillators have the same spectral components with the same amplitudes and are fed with a 180 degrees phase difference to a common load. As a consequence, the power distributed to the load is delivered only in correspondence of the even-order harmonics. Frequency

contributions of odd harmonics, on the contrary, cancel out at the output node. The second-order oscillation at approximately 80GHz, in particular, is retained as the useful output signal.

In each one of the Colpitts oscillators, two MIM capacitors, Cb and Ce, resonating with a 50Ω transmission line (TLb) implemented in metal 6 and alucap, provide a feedback path for the output signal and generate oscillation. PMOS varactors are connected in parallel to Ce to tune the resonating frequency in the range of 400±250fF according to a control voltage. The VCO has been designed to operate at 82.5GHz with a tuning range of 3GHz.

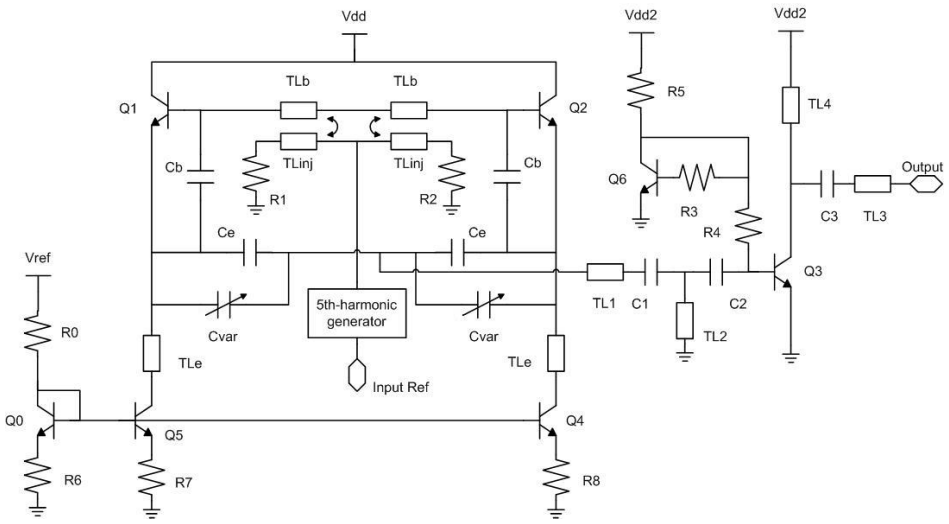


Fig. 54. Schematic view of the Synchronized Push-push Oscillator included in the Integrated Front-end.

This circuit works as a synchronized oscillator according to the principle of operation of injection locked oscillators. It can provide an output signal synchronized to an input reference. The signal reference is injected into each sub-oscillator by means of a transmission line TLinj inductively coupled to TLb, as depicted in Fig 54. The reference is an 8GHz oscillation generated off-chip and subsequently transformed in a square wave. Its 5^{th}-order harmonic is then feed to two transmission lines (TLinj) implemented in metal 5 and respectively routed under the TLb of each sub-oscillator.

The coupling effects generated in the stacked transmission lines, running one under the other over a length of 190µm, was estimated by means of electromagnetic simulations. A lumped-element model of the two coupled lines was then extracted from simulation results and included in circuit-level analysis of the sub-oscillators, following a methodology similar to that proposed in section 2.4 for the design of lumped inductors.

The VCO is biased by a current mirror. Two transmission lines (TLe) designed to operate as λ/4 stubs at 80GHz (corresponding to a length of approximately 560µm for a 6µm-wide transmission line implemented in metal 6) isolate the VCO core from the bias cell.

The VCO has been tested at IMS and its resulting free-running frequency is 84GHz. Measurement results are resumed in Table 7.

The advantage in the use of a synchronized VCO is the easy generation of LO, since it avoids the implementation of more complex circuits like a complete PLL. In particular, the choice of a push-push

oscillator, since it consists of two sub-circuits working at 40GHz rather than at 80GHz, contributes to relax the design constraints. The main drawback is that the low power delivered to the output signal can be insufficient to properly drive the down-conversion mixer. To increase the power of the LO signal in our front-end, a buffer consisting of a single-stage amplifier (Q3) has been connected in series at the output of the VCO.

Table 7. Measured characteristics of the synchronized VCO as a standalone circuit (without output buffer).

Free-running frequency	84GHz
Free-running phase noise	-87dBc/Hz at 10MHz
Synchronous mode phase noise	-108dBc/Hz at 10MHz
Synchronization range	3GHz
Tuning range	3GHz
Output power	-16dBm
Power consumption	108mW
Area	1.1mm^2

The two receivers with and without the VCO have been fabricated and will be tested in a near future. A snapshot of the layout is reported in Fig. 55. A summary of the performances of the complete receiver, as issued by simulations, is reported in Table 8. Fig. 56 shows the conversion gain (CG) as function of the RF input frequency. Since an independent implementation of each building block has been already experimentally validated, satisfactory results can be reasonably expected for the co-integrating circuits.

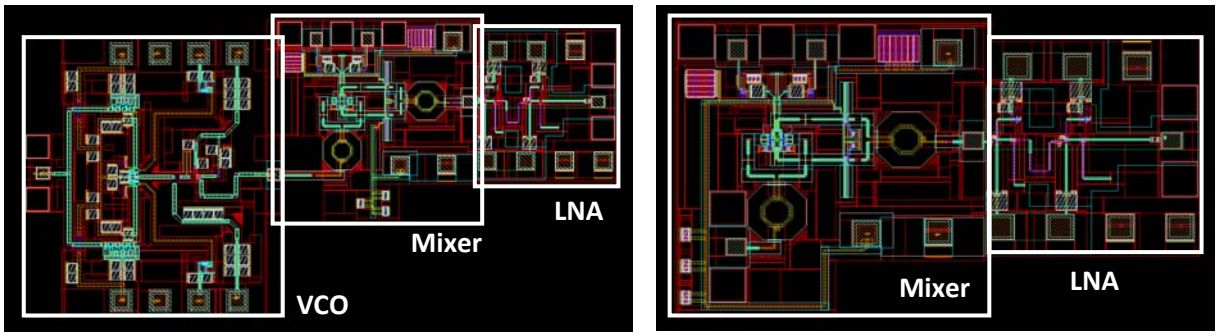


Fig. 55. Layout of the complete 80GHz receiver front-end (left) and of the front-end with off-chip LO generation (right).

Table 8. Simulated performances of the complete receiver front-end.

CG	NF	S11	S22	P_{DC}
39 dB @ (fRF=77GHz & fLO=80 GHz)	6.5 dB @PLO= - 2 dBm	-13.71 dB @ 77 GHz	-20 dB @ 3 GHz	140mW @2.5V

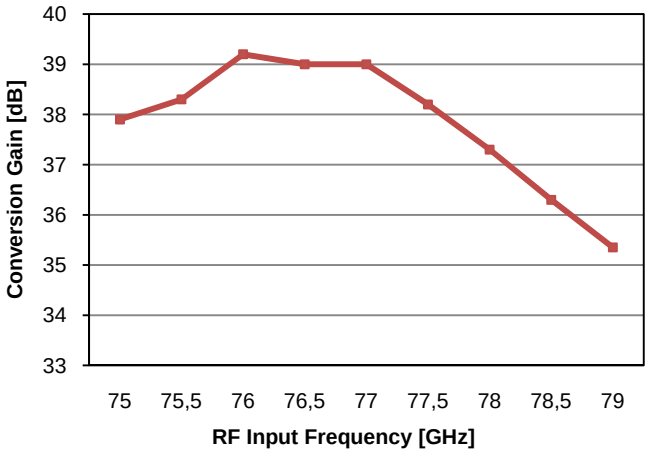


Fig. 56. Simulated conversion gain of the complete receiver front-end vs the RF input frequency with 80GHz LO frequency.

3.9 LNA and Antenna co-integration

After the realization of an integrated front-end, the last step towards the implementation of a complete on-chip receiver is the integration of the antenna.

The possibility to integrate millimeter wave antennas directly on silicon substrates can significantly improve system reliability and repeatability. It also reduces the cost by eliminating the last mm-wave electrical interfaces to the chip.

However, in a standard silicon process the substrate is a lossy element due to its low resistivity, which often is less than 10Ωcm. This is substantially smaller than that of GaAs substrates (10^7-10^9Ωcm). Additionally, the high dielectric constant of silicon (~11.7) results in the substrate absorption of most of the power, preventing an efficient radiation in free-space. For these reasons design and implementation of on-chip antennas at millimeter wave are still a critical bottleneck and very few examples of on silicon integrated antennas have been demonstrated so far. In many cases designers adopt alternative solutions that for the most part are based on above-IC realizations.

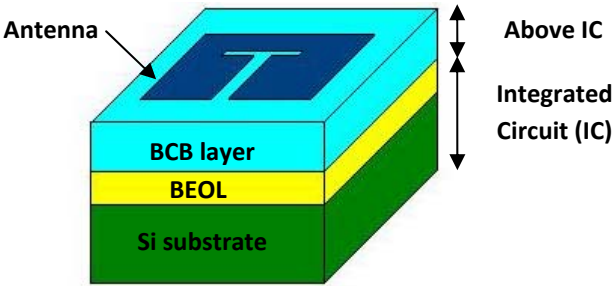


Fig. 57. Example of Above-IC Antenna, implemented on a BCB layer deposited over the chip, after IC fabrication.

The antenna is usually implemented on a BCB layer deposited – after chip fabrication – over the top dielectric layer. The BCB layer (ε_r ~ 2.6) acts as a shield for the substrate and increases the thickness of the BEOL, reducing the effective dielectric constant seen by the antenna.

Despites of these difficulties, we tried to study the co-integration of a LNA with a millimeter wave antenna implemented as a conventional on-chip device in the BiCMOS9MW process flow without any additional fabrication step.

The LNA used for the co-integration is the 80GHz single stage described in section 3.6. The millimeter wave antenna was designed in collaboration with the LabSTICC.

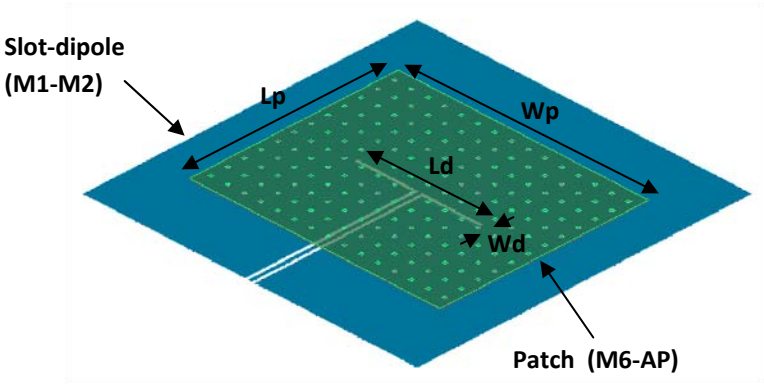


Fig. 58. 80GHz Patch Antenna.

The solution proposed by the LabSTICC is a patch antenna excited by an underlying slot-dipole as shown in Fig. 58. Such a topology is the best-suited for on-chip realization, since it can be easily implemented employing the back-end of line of the BiCMOS9MW technology.

The patch is implemented on metal 6 and alucap and is inductively coupled with the short-circuited slot-dipole implemented on metal 1 and metal 2.

The dimensions of the patch (Lp = 750µm, Wp = 900µm) are chosen to set its resonating frequency at 79GHz. The dimensions of the dipole (Ld = 450µm, Wd = 12µm) control the coupling with the patch and are optimized to match the antenna with the input impedance of the LNA input. The accesses to the dipole are obviously designed to facilitate the connection with the LNA input node.

Electromagnetic simulations of the antenna give a gain of -9dBi. Its efficiency is 15%.

An integrated circuit including the antenna and the LNA has been designed and fabricated. The experimental characterization of such a circuit has not yet been made since the measurement of antennas at millimeter wave frequencies relies on specific procedures and requires a special setup that is not yet available. The test of the fabricated devices will be made at LabSTICC in the months to come.

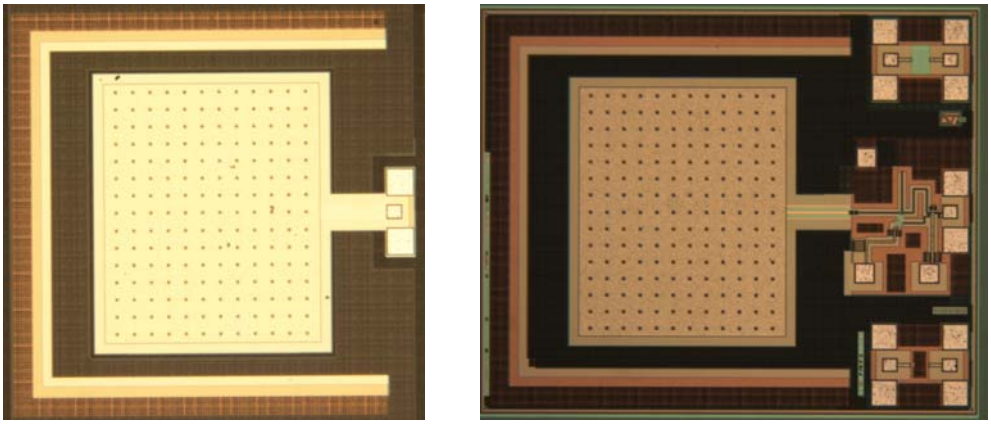


Fig. 59. Microphotographs of the antenna (left) and of the LNA/antenna co-integration chip (right).

3.10 Conclusion

As explained in the introduction, the aim of this chapter was to give, in a sufficient amount, empirical demonstrations of the efficacy of the design methodology for millimeter wave integrated circuits as proposed in chapter 2. It has been applied in the particular context of low-noise amplifiers operating in the most common millimeter wave standards.

The design of a two-stage 60GHz LNA, a two-stage LNA operating at 80GHz, and a single-stage LNA for 94GHz applications, implemented according to a distributed approach have been described in details. For each one of these circuits, an experimental characterization has been provided. On chip measured S-parameters have been presented, as well as large-signal measurements of the 1dB-Compression Point. Preliminary measurements of the noise figure of the 60GHz and 94GHz LNAs are also reported.

All the tested circuits exhibit performances in the state of art. Namely, the 60GHz LNA amplifier has a noise figure of 6dB and a maximum gain of 21.14dB at 61.5GHz; the two-stage 80GHz LNA exhibits a maximum gain of 18.4dB at 80.6GHz and the 94GHz LNA has 9dB of gain and a noise figure of 8.6dB. The core power consumption of these circuits is 10.2mW, 22.5mW and 10.7mW, respectively.

A good agreement between measured data and post layout simulations is observed. Such a close correlation demonstrates the exactitude of the proposed design methodology and its reliability in the design of LNAs operating in the most common standards at millimeter wave.

Furthermore, to evaluate the benefits of the lumped design, two versions of a single-stage 80GHz LNA have been realized using, respectively, distributed transmission lines and lumped inductors. Both the distributed and the lumped implementations of such a LNA have been tested and give very similar results (respectively, a maximum gain of 10.7dB at 82GHz and 11.3dB at 81.5GHz, with only 5.4mW of power consumption).

The direct comparison of a lumped and a distributed circuit proves that the two design approaches have the same potentialities. As a matter of fact, design based on lumped inductors instead of

distributed elements requires some additional effort concerning the inductors design and modeling. However it is to be preferred, since it has the valuable advantage of a significant reduction of the circuit dimensions.

Finally, in the last part of the chapter, the design of a front-end and the co-integration of a LNA with an integrated antenna have been also considered, opening the way to the implementation a fully integrated receiver.

Bibliography

On the design of LNAs:

[1] T. Yao, M. Gordon, K. K. W. Tang, K. H. K. Yau, M. Yang, P. Schvan, and S. P. Voinigescu, "Algorithmic Design of CMOS LNAs and Pas for 60-GHz Radio", *IEEE Journal of Solid-State Circuits*, May 2007, Vol. 42, No. 5, pp. 1044-1057.

[2] S. P. Voinigescu, M. C. Maliepaard, J.L. Showell, G. Babcock, D. Marchesan, M. Schroter, P. Schvan, and D. L. Harame, "A Scalable High-Frequency Noise Model for Bipolar Transistors with Application to Optimal Transistor Sizing for Low-Noise Amplifier Design" *IEEE Journal of Solid-State Circuits*, Sept 1997, Vol. 32, No. 9, pp. 1430-1439.

[3] S. T. Nicolson and S. P. Voinigescu, "Methodology for Simultaneous Noise and Impedance Matching in W-Band LNAs", IEEE Compound Semiconductor Integrated Circuit Symposium, Nov. 2006, pp. 279-282.

On the state of art of 60GHz LNAs:

[4] Brian A. Floyd, Scott K. Reynolds, Ullrich R. Pfeiffer, Thomas Zwick, Troy Beukema, and Brian Gaucher, "SiGe Bipolar Transceiver Circuits Operating at 60 GHz" *IEEE Journal of Solid-State Circuits*, Jan 2005, Vol. 40, No. 1, pp. 156-167.

[5] Y.Sun, F.Herzel, J.Borngräber, and R.Kraemer, "60 GHz Receiver Building Blocks inSiGe BiCMOS" *Proc. of the Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 10-12 Jan. 2007, pp.219-222.

[6] Y. Sun, J. Borngräber, F. Herzel, and W. Winkler, "A Fully Integrated 60 GHz LNA in SiGe:C BiCMOS Technology" *IEEE Bipolar / BiCMOS Circuits and Technology Meeting (BCTM2005)*, pp. 14-17, October 2005.

[7] M.Gordon, S. Voinigescu, "An inductor based 52 GHz 0.18µm SiGe HBT Cascode LNA with 22dB Gain" *ESSCIRC 2004. 30rd European Solid State Circuits Conference*, Sept. 2004, pp. 287-290.

On the state of art of 80GHz LNAs:

[8] Chen, A.Y.-K. Baeyens, Y. Young-Kai Chen Jenshan Lin, "A Low-Power Linear SiGe BiCMOS Low-Noise Amplifier for Millimeter-Wave Active Imaging" *IEEE Microwave and Wireless Components Letters*, vol. 20, no. 2, Feb. 2010, pp. 103-105.

[9] S. Nicolson, K. A. Tang, K. H. K. Yau, P. Chevalier, B. Sautreuil, and S. P. Voinigescu, "A low-voltage 77-GHz automotive radar chipset" *2007 IEEE/MTT-S International Microwave Symposium*, Jun. 2007, pp. 487–490.

[10] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, "A 77-GHz phased-array transceiver with on-chip antennas in silicon: Receiver and antennas" *IEEE Journal of Solid-State Circuits*, Dec. 2006, vol. 41, no. 12, pp. 2795–2806.

[11] Li Wang, Srdjan Glisic, Johannes Borngraeber, Wolfgang Winkler, J. Christoph Scheytt, "A Single-ended Fully integrated SiGe 77/79 GHz Receiver for Automotive Radar" *IEEE Journal of Solid-State Circuits*, Sept. 2008, vol. 43, no.9, pp.1897-1908.

[12] Brian A. Floyd, "V-Band and W-Band SiGe Low-Noise Amplifiers and Voltage-Controlled Oscillators" *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, 2004, pp. 295-298.

[13] Dehlink, B.; Wohlmuth, H.-D.; Aufinger, K.; Meister, T.F.; Bock, J.; Scholtz, A.L.; "A low-noise amplifier at 77 GHz in SiGe:C bipolar technology" *IEEE Compound Semiconductor Integrated Circuit Symposium*, 2005. CSIC '05, pp. 287-290.

[14] Kissinger, D. Forstner, H.P. Jager, H. Maurer, L. Weigel, R., "A differential 77-GHz receiver with current re-use low-noise amplifier in SiGe technology" *SiRF 2010. Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2010, pp. 128-131.

On the state of art of 94GHz LNAs:

[15] Wolfgang Winkler, Johannes Borngräber, Falk Korndörfer, Christoph Scheytt, "94 GHz Amplifier in SiGe Technology" *38th European Microwave Conference*, EuMC 2008, pp. 167-170.

[16] Jason W. May and Gabriel M. Rebeiz, "High-Performance W-Band SiGe RFICs for Passive Millimeter-Wave Imaging" *IEEE Radio Frequency Integrated Circuits Symposium*, RFIC 2009, pp. 437-440.

[17] Johnna Powell, Helen Kim and Charles G. Sodini, "SiGe Receiver Front Ends for Millimeter-Wave Passive Imaging" *IEEE Transactions on Microwave Theory and Techniques*, Vol. 56, pp. 2416-2425, 2008.

[18] Erik Öjefors, Ullrich Pfeiffer, "A 94-GHz Monolithic Front-End for Imaging Arrays in SiGe:C Technology" *38th European Microwave Conference*, EuMC 2008, pp. 422-425.

[19] Javier Alvarado Jr., Kevin T. Kornegay, Brian P. Welch, Yanxin W. Wang, "W-Band SiGe LNA using Unilateral Gain Peaking" *IEEE/MTT-S International Microwave Symposium Digest*, 2008, pp. 289-292.

[20] R. Pilard, D. Gloria, F. Gianesello, F. Le Pennec, C. Person, "94GHz silicon co-integrated LNA and Antenna in a mm-wave dedicated BiCMOS technology" *IEEE Radio Frequency Integrated Circuits Symposium*, RFIC 2010, pp. 83-86.

On the building blocks of the 80GHz Front-end:

[21] A. Mariano, T. Taris, B. Leite, C. Majek, Y. Deval, E. Kerhervé , J.B. Bégueret, D. Belot, "Low power and high gain double-balanced mixer dedicated to 77 GHz automotive radar applications," *Proceedings of the European Solid State Circuits Conference*, pp. 490-493, Sept. 2010.

[22] A. Mariano, B. Leite, C. Majek, T. Taris, Y. Deval, J.-B. Bégueret, D. Belot, "A low power and high gain double-balanced active mixer with integrated transformer-based Baluns dedicated to 77 GHz automotive radar applications," *8th IEEE International NEWCAS Conference*, pp. 81-84, June 2010.

[23] C. Ameziane, T. Taris, Y. Deval, D. Belot, R. Plana, J.-B. Bégueret, "An 80GHz range synchronized push-push oscillator for automotive radar application," *Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 541-544, May 2010.

Conclusion and Perspectives

After a review of the most significant applications of the millimeter wave technology in chapter 1, a complete design methodology for millimeter wave integrated circuits has been defined in chapter 2. The BiCMOS9MW technology has been retained – after a comparative analysis with concurrent bipolar and CMOS nodes – as the platform for our investigations and practical realizations. All the aspects of integrated circuit design have been discussed. Both active and passive devices, including transistors, transmission lines, capacitors, pads, resistors and interconnects, have been closely examined and the convenient model improvements, accounting for parasitic or secondary effects, have been suggested. A particular attention has been addressed to the implementation of inductors as lumped devices.

Subsequently, the proposed methodology has been applied to the design of low noise amplifiers and several examples of LNAs, covering the entire millimeter wave spectrum, have been reported in chapter 3. The design of a two-stage 60GHz LNA, a two-stage LNA operating at 80GHz, and a single-stage LNA for 94GHz applications, implemented according to a distributed approach have been described in details. For each one of these circuits, an experimental characterization has been provided.

A good agreement between measured data and post layout simulations has been repeatedly observed, demonstrating the exactitude of the proposed design methodology and its reliability in the design of LNAs operating in the most common standards at millimeter wave.

All the tested circuits exhibit performances in the state of art. Namely, the 60GHz LNA amplifier has a noise figure of 6dB and a maximum gain of 21.14dB at 61.5GHz; the two-stage 80GHz LNA exhibits a maximum gain of 18.4dB at 80.6GHz and the 94GHz LNA has 9dB of gain and a noise figure of 8.6dB. The core power consumption of these circuits is 10.2mW, 22.5mW and 10.7mW, respectively.

To evaluate the benefits of the lumped design, two versions of a single-stage 80GHz LNA have been realized using, respectively, distributed transmission lines and lumped inductors. Both the distributed and the lumped implementations of such a LNA have given very similar results (respectively, a maximum gain of 10.7dB at 82GHz and 11.3dB at 81.5GHz, with only 5.4mW of power consumption). The direct comparison of a lumped and a distributed circuit has proved that the two design approaches have the same potentialities. As a matter of fact, design based on lumped inductors instead of distributed elements requires some additional effort concerning the inductors design and modeling. However it is to be preferred, since it has the valuable advantage of a significant reduction of the circuit dimensions.

In the future, the experimental characterization of the test circuits realized so far can be integrated and completed by further measurements of noise figure, to be realized as soon as the adequate measurement tools will be available. Further measurements should confirm the noise behavior of each amplifier as suggested by the preliminary results reported in chapter 3.

In a similar manner, the characterization of the 80GHz front-end described in section 3.8 – that is currently in progress – will provide useful information on the co-integration of different building blocks and an insight on the performances of a complete receiver operating at millimeter wave.

Finally, the last realization to be tested is the LNA-antenna co-integration of section 3.9 which requires a special measurement setup currently under development at LabSTICC. Beside the

characterization of the LNA-antenna chip as a receiver, using a power amplifier co-integrated with a transmitting antenna, it will be possible to give a practical demonstration of communication between the two chipsets.

Additionally, an interesting perspective of our work, still concerning on-chip millimeter wave antennas, is the LNA-antenna co-design. As a matter of fact, the most important issue in the co-integration of a LNA with a receiver antenna is the impedance matching between the two blocks that must assure optimum power transfer from the antenna to the amplifier. In many common cases, the LNA is designed in order to be matched to a 50Ω source and the antenna is designed in order to present 50Ω impedance at its terminal. The efficiency of such a strategy has been widely recognized so far; however the design of a 50Ω matched LNA requires a certain effort, as it has been shown in the examples of chapter 3.

Alternatively – starting from the assumption that the antenna impedance can be varied by modifying some of its physical dimensions – it is possible to provide at the antenna's terminals the same impedance required by the amplifier, prior to the application of any matching procedure. As depicted in Fig. 1, in the practical case of the cascode LNAs described in Chapter 3, this means that the impedance existing at the very input of the unmatched cascode stage determines the impedance to be furnished by the antenna. Noise matching can be performed by adjusting both the noise and the input impedance of the amplifier to a common value; then, the antenna is designed in order to provide the same impedance value at the LNA's input node. Both the LNA and the antenna are therefore co-designed, in order to satisfy the overall specifications.

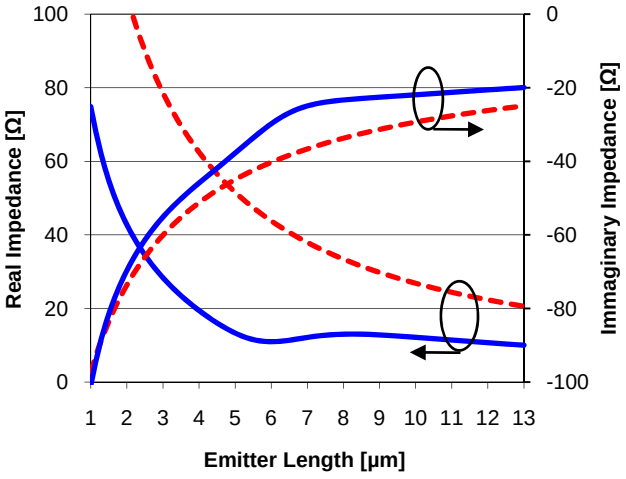


Fig. 1. Input impedance (line) and optimum noise impedance (dot) of an unmatched cascode stage.

The main advantage of co-design over simple co-integration is that design specifications for the amplifier are significantly relaxed. Since no 50Ω matching is needed at its input node, the LNA can assure better noise performances, avoiding any loss that could be introduced by the matching devices. Furthermore, the lack of any matching network results, in many cases, in a consistent area reduction.

The study made at LabSTICC of the patch antenna of section 3.9 demonstrates the feasibility of co-design with a LNA and a co-design can be indeed developed in order to optimize the performance of the overall system.

The advantage of co-design becomes more evident when considering a system-level approach. Thanks to the reduction of silicon occupation, several LNA/antenna co-designed blocks can be easily integrated on a single-chip, in order to implement receivers and transceivers based on a multiple-input multiple-output (MIMO) architecture. In this way, the global performance of the whole communication system can be significantly enhanced with respect to reliability, throughput, and power consumption.

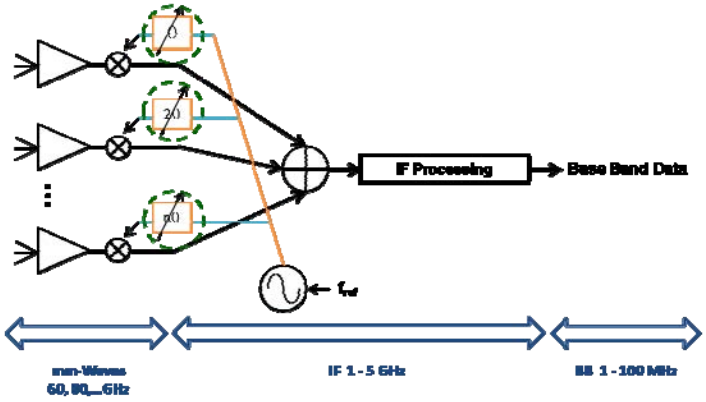


Fig. 2. MIMO Receiver Architecture with LNA/Antenna Co-design.

Publications

[1] R. R. Severino, T. Taris, Y. Deval, D. Belot, J. B. Begueret, "A SiGe:C BiCMOS LNA for 94GHz band applications," *2010 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM 2010)*, pp.188-191, 4-6 Oct. 2010.

[2] C. Majek, R. R. Severino, T. Taris, Y. Deval, A. Mariano, J. B. Begueret, D. Belot, "60 GHz cascode LNA with interstage matching: performance comparison between 130nm BiCMOS and 65nm CMOS-SOI technologies," *3rd International Conference on Signals, Circuits and Systems (SCS 2009)*, pp.1-5, 6-8 Nov. 2009.

[3] T. Taris, Y. Deval, R. R. Severino, C. Ameziane, D. Belot, J. B. Begueret,, "Millimeter-Waves building block design methodology in BiCMOS technology," *16th IEEE International Conference on Electronics, Circuits, and Systems, (ICECS 2009)*, pp.968-971, 13-16 Dec. 2009.

[4] R. R. Severino, T. Taris, Y. Deval, D. Belot, J. B. Begueret, "A SiGe:C BiCMOS LNA for 60GHz band applications," *IEEE Bipolar/BiCMOS Circuits and Technology Meeting, 2009. (BCTM 2009)*, pp.51-54, 12-14 Oct. 2009.

[5] T. Taris, R. Severino, Y. Deval, J. B. Begueret, "mm-Waves design trends in BiCMOS technology," *2008 Joint 6th International IEEE Northeast Workshop on Circuits and Systems and TAISA Conference, (NEWCAS-TAISA 2008)*, pp.375-379, 22-25 June 2008.

Glossary

ACC	Adaptive Cruise Control
ASK	Amplitude Shift Keying
BCB	Benzocyclobutene
BEOL	Back-end of Line
BPSK	Binary Phase Shift Keying
CB	Common Base
CE	Common Emitter
CG	Conversion Gain
CG	Common Gate
CPW	Coplanar Waveguide
CS	Common Source
EHF	Extremely High Frequency
EIRP	Equivalent Isotropically Radiated Power
ETSI	European Telecommunications Standards Institute
FCC	US Federal Communications Commission
FMCW	Frequency Modulated Continuous Wave
FoM	Figure of Merit
FSA	Fully Self-aligned
Ga	Available Gain
GMSK	Gaussian Minimum Shift Keying
GPR	Ground-penetrating Radar
GSG	Ground-Signal-Ground
HBT	Hetero-junction Bipolar Transistor
HICUM	HIgh CUrrent Model
IC	Integrated Circuit
IF	Intermediate Frequency

IIP3	Third-order Intercept Point
IP1dB	Input-referred Compression Point (at 1dB)
ITRS	International Technology Roadmap for Semiconductors
ITS	Intelligent Transportation Systems
ITU	International Telecommunications Union
IVC	Inter-vehicle Communication
LNA	Low Noise Amplifier
LO	Local Oscillator
LRR	Long-range Radar
MAC	Medium Access Control
MIM	Metal-Insulator-Metal (Capacitor)
MIMO	Multiple-input Multiple-output (Architecture)
NEP	Noise Equivalent Power
NF	Noise Figure
NFmin	Minimum Noise Figure
OFDM	Orthogonal Frequency Domain Multiplexing
OOK	On-Off Keying
PA	Power Amplifier
PLL	Phase-locked Loop
PSD	Power Spectral Density
PSS	Periodic Steady State (Analysis)
Q	Quality Factor
RF	Radio Frequency
RVC	Road- vehicle Communication
SC	Single Carrier
SEG	Selective Epitaxial Growth
SFR	Self-resonant Frequency
SGP	Spice Gummel-Poon
SoC	System on-chip
SOLT	Short-Open-Line-Thru (Calibration)
SRR	Short-range Radar

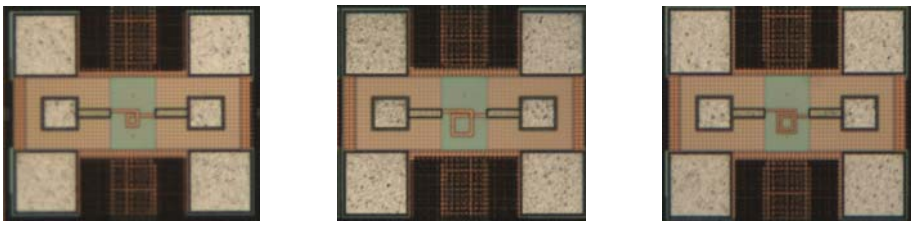
UWB	Ultra-wide Band
VCO	Voltage-controlled Oscillator
VNA	Vectorial Network Analyzer
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network

Annexes

Annex 1

Equivalent Model of Lumped Inductors

Three examples of lumped inductors have been designed and fabricated as explained in section 2.4. A photograph of the three devices is reported in Fig. 1.



(a) (b) (c)

Fig. 1. Chip microphotographs of spiral inductors fabricated in BiCMOS9MW technology. At 80GHz, inductance values of 81pH (a), 122pH (b), and 202pH (c) are obtained, respectively.

As suggested in chapter 2, inductors can be modeled by the 2π-circuit of Fig. 2. The values of the model components extracted from ASITIC simulation of the inductors of Fig. 1 are reported in Table 1. Plots of inductance and quality factor reported in Fig. 36-38 of chapter 2 are obtained using these values.

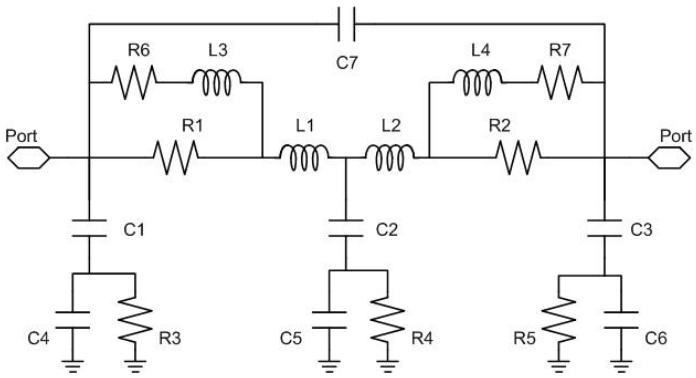


Fig. 2. Equivalent 2π model of the inductor.

Table 1. Elements of the equivalent 2π model of Fig. 2.

	(a)	(b)	(c)
R1	96.0675Ω	96.0675Ω	96.0675Ω
R2	493.483mΩ	493.483mΩ	493.483mΩ
R3	77038.3Ω	77038.3Ω	77038.3Ω
R4	60529Ω	60529Ω	60529Ω
R5	34529.2Ω	34529.2Ω	34529.2Ω
R6	70.6322mΩ	70.6322mΩ	1.1 Ω
R7	404.909mΩ	404.909mΩ	100mΩ
C1	114.125aF	114.125aF	114.125aF
C2	145.241aF	145.241aF	145.241aF
C3	5.69613fF	5.69613fF	5.69613fF
C4	5.99465fF	5.99465fF	5.99465fF
C5	2.9077fF	2.9077fF	2.9077fF
C6	190.684aF	190.684aF	190.684aF
C7	981.452aF	1.4fF	1.7fF
L1	31.4pH	53pH	88.5pH
L2	39.3pH	54pH	88.5pH
L3	2.82479pH	5.82479pH	0.82479pH
L4	15.7354pH	17.7354pH	15.7354pH

The three devices of Fig. 1 cover a large range of inductance, approximately between 80 and 200pH at 80GHz. However, the implementation of a LNA based on lumped inductors described in section 3.6 requires an inductor featuring 67pH to perform emitter degeneration. It can be designed according to the methodology described in section 2.4. The inductor is implemented as a single-turn squared spiral on metal 6, with 4.5µm metal width, 2µm spacing, and an inner diameter of only 8µm. Interconnects, always included into device design and modeling, are realized by extending the most external spiral segment of 20µmand adding a 19µm metal 6 segment in series to the underpass. It has been simulated using ASITIC with the command 2portx, defining a 200 × 200 chip area and with all simulation parameters set as default. The equivalent model of the inductor, extracted from simulation results, is that of Fig. 2. Model components are reported in Table 2. The resulting equivalent inductance and the quality factor are presented in Fig. 3.

Table 2. Elements of the equivalent 2π model of the 67pH inductor used to perform emitter degeneration in the 80GHz LNA of section 3.6.

R1	96.0675Ω	C3	5.69613fF
R2	493.483mΩ	C4	5.99465fF
R3	77038.3Ω	C5	2.9077fF
R4	60529Ω	C6	190.684aF
R5	34529.2Ω	C7	981.452aF
R6	70.6322mΩ	L1	28.15pH
R7	404.909mΩ	L2	35.15pH
C1	114.125aF	L3	2.82479pH
C2	145.241aF	L4	15.7354pH

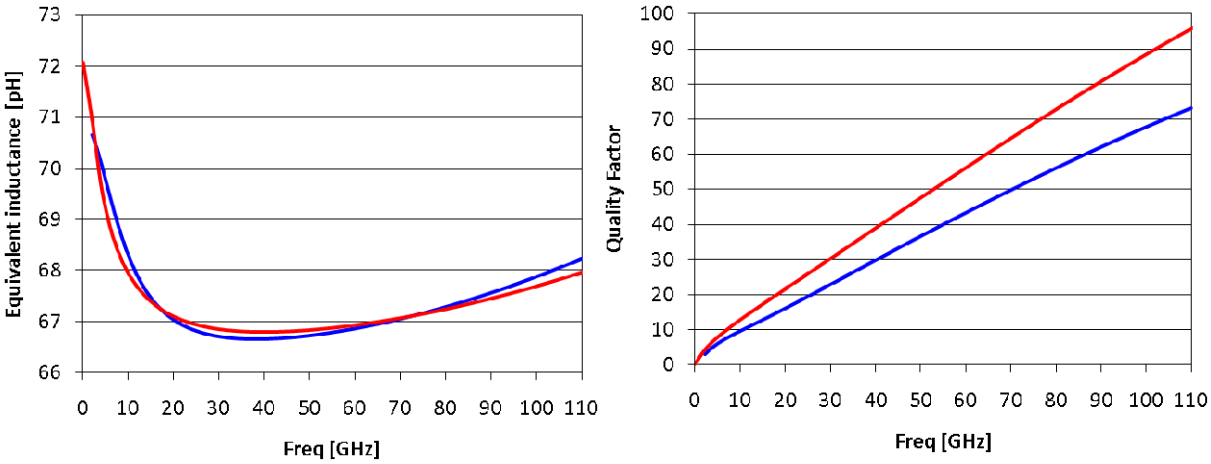


Fig. 3. Equivalent Inductance and Quality Factor of the 67pH inductor: ASITIC simulation (red) and model (blue).

Annex 2

Definition of S parameters and reflection coefficients

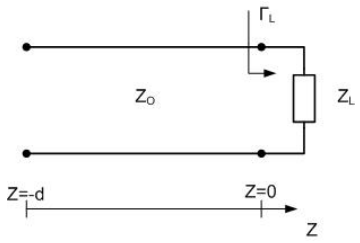


Fig. 1. Transmission line with load impedance.

In a transmission line with characteristic impedance Zo and propagation constant γ as in Fig. 1, the total voltage and current are written as a sum of forward and reflected traveling waves:

$$V(z) = V^+ e^{-\gamma z} + V^- e^{\gamma z} \quad (1)$$

$$I(z) = \frac{V^+}{Z_O} e^{-\gamma z} - \frac{V^-}{Z_O} e^{\gamma z} \quad (2)$$

At the terminals of the load, the total voltage and current are:

$$V = V(0) = V^+ + V^- \quad (3)$$

$$I = I(0) = \frac{V^+ - V^-}{Z_O} = I^+ - I^- \quad (4)$$

Solving from V^+ and V^- we obtain:

$$V^+ = \frac{V + Z_O I}{2} \quad (5)$$

$$V^- = \frac{V - Z_O I}{2} \quad (6)$$

The reflection coefficient associated to the load impedance of the transmission line is:

$$\Gamma_L = \frac{V^-}{V^+} = \frac{Z_L - Z_O}{Z_L + Z_O} \quad (7)$$

In the case of a matched load we have $Z_L = Z_O$ and the amplitude of the reflected wave is zero.

We can introduce the normalization constant $\sqrt{\Re e\{Z_O\}}$, and define the two normalized waves as:

$$a = \frac{V + Z_O I}{2\sqrt{\Re e\{Z_O\}}} \quad (8)$$

$$b = \frac{V - Z_O I}{2\sqrt{\Re e\{Z_O\}}} \quad (9)$$

If we consider the two-port device of Fig. 2, then we can apply (8) and (9) to each port, obtaining two bi-dimensional vectors:

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{11}$$

A matrix of scattering parameters or S parameters can be defined by:

$$b = Sa \tag{12}$$

This definition can be extended to the general case of a *n-port* device, obtaining a *n×n* S matrix. In the particular case of a two-port device of Fig. 2, we have:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{13}$$

Using the S parameters it is possible to establish a relation between the power at the input and output port of a quadripole.

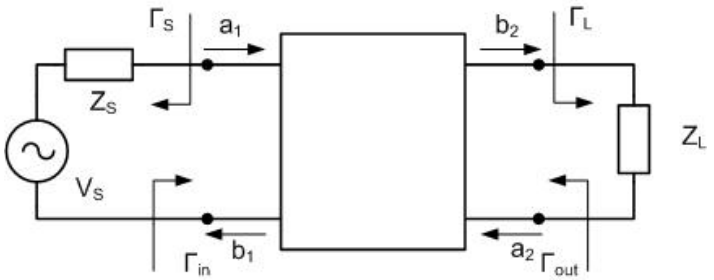


Fig. 2. Definition of S parameters and reflection coefficients for a two-port device.

In relation to the Fig. 2 it is possible to define the reflection coefficients usually used in the analysis of a two-port device. In particular, the source and load reflection coefficients are defined by:

$$\Gamma_S = \frac{a_1}{b_1} = \frac{Z_S - Z_O}{Z_S + Z_O} \tag{14}$$

$$\Gamma_L = \frac{a_2}{b_2} = \frac{Z_L - Z_O}{Z_L + Z_O} \tag{15}$$

The input and output reflection coefficients are defined by:

$$\Gamma_{in} = \frac{b_1}{a_1} = \frac{S_{11}a_1 + S_{12}a_2}{a_1} = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} \tag{16}$$

$$\Gamma_{out} = \frac{b_2}{a_2} = \frac{S_{21}a_1 + S_{22}a_2}{a_2} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S} \tag{17}$$

If both the source and load impedance are equals, then it is convenient to set the normalization impedance as:

$$Z_O = Z_S = Z_L \tag{18}$$

If the (18) is verified, as in many practical cases, then the source and load reflection coefficients are zero. Consequently, the expressions of input and output reflection coefficients are simplified as:

$$\Gamma_{in} = S_{11} \quad (19)$$

$$\Gamma_{out} = S_{22} \quad (20)$$

In this case, parameters S11 and S22 correspond respectively to the reflection coefficients at the input and the output. Therefore, the S11 parameter gives an estimation of the power input insertion loss of the quadripole:

$$Input\ isertion\ loss = |S_{11}|^2 \overset{dB}{\rightarrow} 20\log|S_{11}| \quad (21)$$

In a similar manner, the S22 parameter is a measure of the power return loss at the output port of the quadripole:

$$Output\ return\ loss = |S_{22}|^2 \overset{dB}{\rightarrow} 20\log|S_{22}| \quad (22)$$

The power insertion gain of the quadripole of Fig. 2 (or the insertion loss, if the quadripole is a passive device) can be defined as the ratio between the power delivered at its output port and the power fed at its input port. If the (18) is verified, then $a_2 = 0$ and the insertion gain can be expressed as:

$$Insertion\ gain = |S_{21}|^2 \overset{dB}{\rightarrow} 20\log|S_{21}| \quad (23)$$

In the radio-frequency domain, the power gain of a two-port device is usually defined in terms of available gain, operative gain or transducer gain. It can be demonstrated that, if a quadripole is simultaneously matched at its input and output ports, then all these definitions are equivalent and are equal to (23). In other words, the S21 parameter measures the gain of a perfectly matched quadripole.

In a dual manner, the S12 parameter is an expression of the reverse gain or isolation of the quadripole:

$$Reverse\ isolation = |S_{12}|^2 \overset{dB}{\rightarrow} 20\log|S_{12}| \quad (24)$$

In the ideal case of a unilateral device, the reverse isolation and therefore the S12 parameter is zero. For a unilateral quadripole (19) and (20) are automatically satisfied, even with nonzero source and load reflection coefficient (18).

Generally speaking, the S parameters representation can provide an exhaustive characterization of a n-port devices like passive circuits or amplifiers and for this reason it is widely used in the radio-frequency, microwave, and millimeter wave domain.

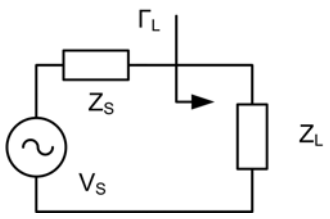


Fig. 3. Maximum power transfer from a source to a load.

To conclude, it can be observed that the definition of the reflection coefficient given by (7) can be extended to the general case of a Z_L impedance as depicted in Fig. 3:

$$\Gamma_L = \frac{Z_L - Z_O}{Z_L + Z_O} \tag{25}$$

In this case, however, the (7) does not respect the condition of maximum power transfer. As a matter of fact, the condition of maximum power transfer is given by the complex conjugate matching:

$$Z_L = Z_O^* \tag{26}$$

But this implies a nonzero reflection coefficient according to (7) or (25). To overcome to this difficulty, a different definition of S parameters and reflection coefficients, based on power waves [1], is given in [2]:

The travelling waves on a transmission line (8) and (9) are replaced by power waves, defined by:

$$a = \frac{V + Z_O I}{2\sqrt{\Re e\{Z_O\}}} \tag{27}$$

$$b = \frac{V - Z_O^* I}{2\sqrt{\Re e\{Z_O\}}} \tag{28}$$

where Z_O is a reference impedance. Therefore the reflection coefficient of Fig. 3 becomes:

$$\Gamma_L = \frac{Z_L - Z_O^*}{Z_L + Z_O} \tag{29}$$

that is zero for a conjugate matched load, as required by (26).

Then an alternative definition of S parameter and of the reflection coefficients considered so far can be given, substituting the (8) and (9) with (27) and (28), respectively.

Practically speaking, the two definitions are substantially equivalent in many practical cases, since the reference impedance Z_O is usually assumed to be real. However, the definition based on power waves can be sometimes more appropriate. A typical example that requires the use of power waves is the measure of the insertion loss occurring at the input of a LNA connected to an antenna characterized by a complex impedance, instead of an ideal 50Ω source.

Bibliography

[1] K. Kurokawa, “Power Waves and Scattering matrix”, *IEEE Transactions on MicrowaveTheory and Techniques*, Mar 1965, Vol. 13, No. 2, pp. 194-202.

[2] J. Rahola, “Power Waves and Conjugate Matching” *IEEE Transactions on Circuits and Systems II: Express Briefs*, Jan. 2008, Vol. 55, No. 1, pp. 92-96.

Annex 3

Definition of Noise parameters

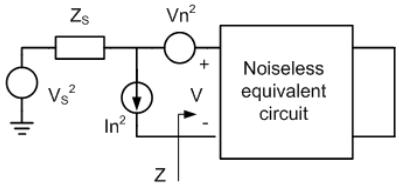


Fig. 1. Equivalent schematic of a noisy circuit.

The noise factor of a circuit is defined by:

$$F = \frac{(S/N)_{in}}{(S/N)_{out}} \tag{1}$$

where $(S/N)_{in}$ and $(S/N)_{out}$ are the signal-to-noise ratio calculated respectively at the input and at the output port. A first consideration is that $(S/N)_{in} > (S/N)_{out}$, meaning that the circuit will contribute with additional noise to the noise power due to the source and therefore the noise factor will be always greater than unity.

By indicating with G the power gain of the circuit and with S_{in} the power of input signal, equation (1) becomes:

$$F = \frac{S_{in}}{N_{in}} \frac{G(N_{in}+N_{add})}{GS_{in}} = 1 + \frac{N_{add}}{N_{in}} \tag{2}$$

where N_{in} is the noise due to the source and N_{add} is the additional noise due to the circuit referred at its input port.

The noise of the source is the noise associated to the source impedance $Z_S = R_S + jX_S$ and can be represented by a noise generator whose mean square value is:

$$\overline{v_S^2} = 4kTR_s\Delta f \tag{3}$$

In a similar manner, it is possible to introduce two noise generators accounting for the noise of the circuit defined as:

$$\overline{v_n^2} = 4kTR_n\Delta f \tag{4}$$

$$\overline{\imath_n^2} = 4kTG_n\Delta f \tag{5}$$

where R_n and G_n are the noise resistance and the noise conductance of the circuit, respectively. Taking into account those elements, the latter can be represented as a noiseless device, as suggested in Fig. 1.

Then, the contribution of the noise generators to the voltage *V* at the input port are:

$$V = \frac{Z}{Z+Z_S} v_S + \frac{Z}{Z+Z_S} v_n + \frac{ZZ_S}{Z+Z_S} i_n \tag{6}$$

where $Z = R + jX$ is the circuit input impedance.

Moving from voltage to power domain, the noise terms defined so far can be written as:

$$N_{in} = \frac{1}{4R} \left| \frac{Z}{Z+Z_S} \right|^2 \overline{v_S^2} \tag{7}$$

$$N_{add} = \frac{1}{4R} \left| \frac{Z}{Z+Z_S} v_n + \frac{ZZ_S}{Z+Z_S} i_n \right|^2 \tag{8}$$

Introducing a correlation impedance $Z_C = R_C + jX_C$ between v_n and i_n the expression of the noise factor becomes:

$$F = 1 + \frac{\overline{v_n^2} + |Z_S|^2 \overline{i_n^2} + 2R_S R_C \overline{i_n^2} + 2X_S X_C \overline{i_n^2}}{\overline{v_S^2}} \tag{9}$$

Substituting the (3)-(5) into (9), we obtain a more eloquent expression:

$$F = 1 + \frac{R_n}{R_S} + \frac{G_n}{R_S}(R_S^2 + X_S^2) + 2R_C G_n + 2X_C G_n \frac{X_S}{R_S} \tag{10}$$

The noise factor depends on the source impedance. In particular, the minimum value allowed for the noise factor corresponds to the optimum value of the input impedance $Z_{opt} = R_{opt} + jX_{opt}$.

The latter can be calculated by minimizing expression (10), obtaining:

$$R_{opt} = \sqrt{\frac{R_n}{G_n} - {X_C}^2} \tag{11}$$

$$X_{opt} = -X_C \tag{12}$$

The minimum value of noise factor can be calculated assuming $Z_S = Z_{opt}$ in equation (10), obtaining:

$$F_{min} = 1 + 2G_n \left[R_C + \sqrt{\frac{R_n}{G_n} - {X_C}^2} \right] \tag{13}$$

The noise factor can be expressed as a function of the source impedance and of the three noise parameters R_n, Z_{opt}, and F_{min}:

$$F = F_{min} + \frac{R_n}{R_S} \left| \frac{Z_S}{Z_{opt}} - 1 \right|^2 \tag{14}$$

Alternatively, it can be expressed in terms of the reflection coefficients Γ_S and Γ_{min} associated to the source impedance and to the optimum source impedance, respectively:

$$F = F_{min} + \frac{4R_n}{Z_0} \frac{|\Gamma_S - \Gamma_{min}|^2}{|1-\Gamma_{min}|^2 (1-|\Gamma_S|^2)} \tag{15}$$

where Z_O is a normalization impedance recurring in the definition of the two reflection coefficients:

$$\Gamma_S = \frac{Z_S - Z_o}{Z_S + Z_o} \quad (16)$$

$$\Gamma_{min} = \frac{Z_{Sopt} - Z_0}{Z_{Sopt} + Z_0} \quad (17)$$

Both equations (14) and (15) mean that – in order to minimize the noise factor – the circuit must be connected to a source presenting the optimum noise impedance. When this condition is satisfied, we have $Z_S = Z_{opt}$ and $\Gamma_S = \Gamma_{min}$ and as a consequence the noise factor reaches its optimum (minimum) value $F = F_{min}$.

In traditional microwave design, noise matching of an active circuit is obtained by inserting a matching network between the source (whose impedance Z_X is different, in general, from Z_{opt}) and the input port of the circuit itself, as illustrated in Fig. 2. The matching network is designed in order to make the impedance Z_S seen looking from the circuit input node equal to Z_{opt}.

In the design approach described in chapter 3, on the contrary, the source impedance is given and is equal to 50Ω and the active circuit itself is designed, by adjusting transistor size and biasing, in order to make the Z_{opt} equal to 50Ω.

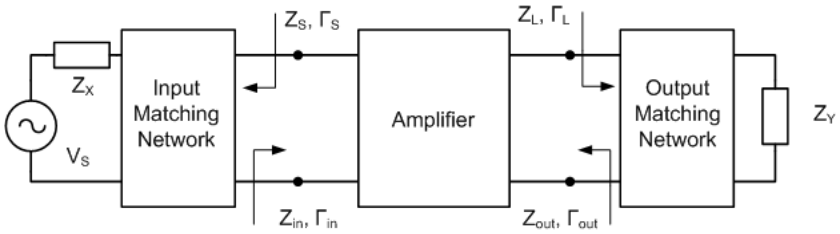


Fig. 2. Input and output matching of an amplifier in traditional microwave design.

It can be observed that noise matching for a low noise amplifier is complementary to input impedance (power) matching. In this case, the source impedance must be equal to the complex conjugate of the input impedance Z_{in}. Following the strategy of section 3.1, it is possible to obtain both noise and power matching, designing the amplifier in order to make both Z_{in} and Z_{opt} equal to 50Ω. As depicted in Fig. 3, since the active circuit itself is dimensioned in order to provide the desired input and noise impedances, no matching network is required.

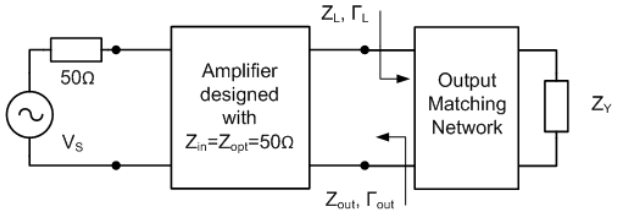


Fig. 3. Alternative input impedance and noise matching.

Annex 4

Considerations on the effects of inductive degeneration on optimum noise resistance

The noise factor of a circuit is defined by:

$$F = \frac{(S/N)_{in}}{(S/N)_{out}} = \frac{Total\ output\ noise}{Output\ noise\ due\ to\ the\ source} \quad (1)$$

The noise of the source can be considered as the noise associated to the source impedance, therefore, indicating with Av the gain of the circuit, equation (1) becomes:

$$F = \frac{Total\ output\ noise}{4kTR_S|Av|^2} \quad (2)$$

where $4kTR_S$ is the power spectral density of the thermal noise due to Rs.

A small signal circuit of a cascode amplifier including noise generators is shown in Fig. 1 (a). R_X and Zs represent the intrinsic base resistance and the source impedance, respectively. In general, the latter can have a non-zero imaginary part Xs.

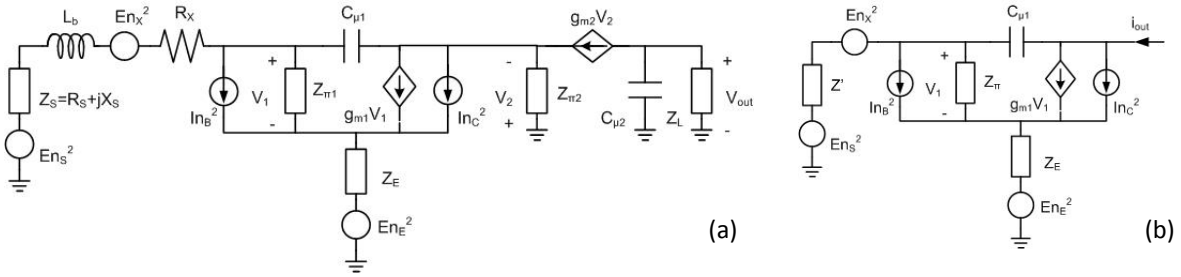


Fig. 1. Equivalent small-signal circuit of a cascode amplifier with noise generators (a) with simplifications (b).

Neglecting the current that flows in $Z_{\pi 2}$ and $C_{\mu 2}$, the CB stage acts as a current buffer. Assuming that the latter has unitary gain as it is usually the case for a CB amplifier, the current in the load is approximately equal to the collector current of the CE stage. The circuit can be consequently simplified as in Fig. 1 (b), with $Z' = R_S + R_X + jX_S + j\omega L_b$.

The contribution of each noise source on the output current is given by a transfer function that can be calculated on the circuit of Fig. 1 (b). Neglecting $C_{\mu 1}$, we find:

$$\left|\frac{\overline{\imath out^2}}{\overline{E_{nS}^2}}\right| = \left|\frac{\overline{\imath out^2}}{\overline{E_{nX}^2}}\right| = \left|\frac{\overline{\imath out^2}}{\overline{E_{nE}^2}}\right| = \left|\frac{g_m Z_\pi}{Z'+Z_\pi+Z_E+g_m Z_\pi Z_E}\right|^2 \quad (3)$$

$$\left|\frac{\overline{\imath out^2}}{\overline{I_{nC}^2}}\right| = \left|\frac{Z'+Z_\pi+Z_E}{Z'+Z_\pi+Z_E+g_m Z_\pi Z_E}\right|^2 \quad (4)$$

$$\left|\frac{\overline{\imath out^2}}{\overline{I_{nB}^2}}\right| = \left|\frac{g_m Z_\pi (Z' + Z_E)}{Z' + Z_\pi + Z_E + g_m Z_\pi Z_E}\right|^2 \tag{5}$$

Equation (2) when applied to the case of Fig. 1 (b) becomes:

$$F = 1 + \frac{1}{4kTR_S}\left[\frac{S_{EnX}\left|\frac{\overline{\imath out^2}}{\overline{E_{nX}^2}}\right| + S_{EnE}\left|\frac{\overline{\imath out^2}}{\overline{E_{nE}^2}}\right| + S_{InC}\left|\frac{\overline{\imath out^2}}{\overline{I_{nC}^2}}\right| + S_{InB}\left|\frac{\overline{\imath out^2}}{\overline{I_{nB}^2}}\right|}{|Av|^2}\right] \tag{6}$$

Where the power spectral densities accounting respectively for thermal noise due to R_X and Z_E, and for shot noise associated to the collector and base currents are defined as follows:

$$S_{EnX} = 4kTR_X$$

$$S_{EnE} = 4kTR_E \tag{7}$$

$$S_{InC} = 2qI_C$$

$$S_{InB} = 2qI_B + \frac{KI_B^2}{f}$$

All the noise sources are considered as uncorrelated. Furthermore, when working at millimeter wave frequencies Flicker noise associated to base current can be neglected.

Since the overall transcoductance Av is calculated as in (3), we have:

$$\left|\frac{\overline{\imath out^2}}{\overline{E_{nX}^2}}\right| = \left|\frac{\overline{\imath out^2}}{\overline{E_{nE}^2}}\right| = |Av|^2 \tag{8}$$

Assuming $g_m = \frac{I_C}{V_T} = \beta_o \frac{I_B}{V_T}$ and substituting (7-8) in (6), the expression of the noise factor becomes:

$$F = 1 + \frac{R_X}{Rs} + \frac{R_E}{Rs} + \frac{g_m}{2R_S}\left|\frac{Z' + Z_\pi + Z_E}{g_m Z_\pi}\right|^2 + \frac{g_m}{2\beta_o R_S}|Z' + Z_E|^2 \tag{9}$$

Equation (9) can be simplified observing that at high frequency:

$$R_\pi \gg \frac{1}{\omega C_\pi} \qquad Z_\pi \cong \frac{1}{j\omega C_\pi} \tag{10}$$

Furthermore, if the quality factor of the degeneration inductor is sufficiently elevated:

$$Z_E = R_E + j\omega L_E \cong j\omega L_E \tag{11}$$

Applying (10-11), the approximated expression of the noise factor is:

$$F \cong 1 + \frac{R_X}{Rs} + \frac{R_E}{Rs} + \frac{1}{2g_m|Z_\pi|^2 R_S}\left[(R_S + R_X)^2 + \left(X_S + \omega L_b + \omega L_E - \frac{1}{\omega C_\pi}\right)^2\right] + \frac{g_m}{2\beta_o R_S}[(R_S + R_X)^2 + (X_S + \omega L_b + \omega L_E)^2] \tag{12}$$

A further simplification can be introduced observing that $R_X \ll R_S$, therefore:

$$F \cong 1 + \frac{R_X}{Rs} + \frac{R_E}{Rs} + \frac{R_S}{2g_m|Z_\pi|^2} + \frac{1}{2g_m|Z_\pi|^2 R_S}\left[X_S + \omega L_b + \omega L_E - \frac{1}{\omega C_\pi}\right]^2 + \frac{g_m R_S}{2\beta_o} + \frac{g_m}{2\beta_o R_S}(X_S + \omega L_b + \omega L_E)^2 \tag{13}$$

The optimum noise impedance is defined as the value of Zs that corresponds to the minimum of F. Optimum value for the real and imaginary part of Zs can be found solving (14) and (15) respectively.

$$\frac{\partial F}{\partial R_S} = 0 \tag{14}$$

$$\frac{\partial F}{\partial X_S} = 0 \tag{15}$$

Equation (15) can be written as:

$$\frac{\partial F}{\partial X_S} = \frac{1}{g_m|Z_\pi|^2 R_S}\left[X_{Sopt} + \omega L_b + \omega L_E - \frac{1}{\omega C_\pi}\right] + \frac{g_m}{\beta_o R_S}\left(X_{Sopt} + \omega L_b + \omega L_E\right) = 0 \tag{16}$$

and the imaginary part of the optimum noise impedance is:

$$X_{Sopt} = \frac{\frac{1}{g_m|Z_\pi|}}{\frac{1}{g_m|Z_\pi|^2} + \frac{g_m}{\beta_o}} - \omega(L_b + L_E) \tag{17}$$

Equation (14) can be written as:

$$\frac{\partial F}{\partial R_S} = -\frac{R_X}{Rs^2} - \frac{R_E}{Rs^2} + \frac{1}{2g_m|Z_\pi|^2} - \frac{1}{2g_m|Z_\pi|^2 Rs^2}\left[X_S + \omega L_b + \omega L_E - \frac{1}{\omega C_\pi}\right]^2 + \frac{g_m}{2\beta_o} - \frac{g_m}{2\beta_o Rs^2}[X_S + \omega L_b + \omega L_E]^2 = 0 \tag{18}$$

Replacing Xs with its optimum value given by (17) and solving for Rs, we obtain the optimum noise resistance:

$$R_{Sopt} = \frac{\sqrt{\left[2(R_X+R_E)\left[\beta_o+g_m^2|Z_\pi|^2\right]+g_m|Z_\pi|^2\right]g_m\beta_o|Z_\pi|^2}}{\beta_o+g_m^2|Z_\pi|^2} \tag{19}$$

The optimum noise resistance given by (19) is independent from emitter degeneration. This is a result of the simplifications made so far and is in accord with classical analysis. However, accurate simulations of a cascode amplifier reveal a dependence of Rsopt from emitter inductor as already observed in paragraph 3.1.5. This dependence becomes stronger at high frequency and cannot be neglected at millimeter wave.

In order to reproduce this phenomenon in our analysis, two elements must be taken into account: capacitance $C_{\mu 1}$ must be included and noise sources of the CB transistor must be added. Concerning $C_{\mu 1}$, to include it in the analysis, it is useful to apply Miller's theorem at its terminals. The expression of Miller's factor, calculated on the circuit of Fig. 1 (a) neglecting the current flowing through $C_{\mu 1}$, is:

$$A_M = -\frac{\frac{g_m Z_{\pi 2}}{1+g_m Z_{\pi 2}}}{1+g_m Z_E + \frac{Z_E}{Z_{\pi 1}}} \tag{20}$$

The Miller capacitance C_M is defined as:

$$C_M = (1 - A_M)C_{\mu 1} = \frac{2g_m + j\omega C_{\pi 2} + j\omega L_E(g_m + j\omega C_{\pi 2})^2}{g_m + j\omega C_{\pi 2} + j\omega L_E(g_m + j\omega C_{\pi 2})^2} C_{\mu 1} \tag{21}$$

The Miller capacitance given by (21) can be inserted in the circuit of Fig. 1 (a) between the base node and ground. Then, it can be included in circuital analysis leading to a new expression of the noise factor:

$$F = 1 + \frac{R_X}{Rs} + \frac{R_E}{Rs}\left|1 + \frac{Z'}{Z_M}\right|^2 + \frac{g_m}{2R_S}\left|\frac{Z' + (Z_\pi + Z_E)\left(1 + \frac{Z'}{Z_M}\right)}{g_m Z_\pi}\right|^2 + \frac{g_m}{2\beta_o R_S}\left|Z' + Z_E\left(1 + \frac{Z'}{Z_M}\right)\right|^2 \quad (22)$$

where Z_M is the Miller impedance corresponding to $C_{\mu 1}$:

$$Z_M = \frac{1}{j\omega C_M} \quad (23)$$

Equation (22) is equivalent to (9) if $C_{\mu 1}$ is neglected, that is considering Z_M infinite. With a non-zero value of $C_{\mu 1}$, on the contrary, equation (22) is somewhat involved. An approximation can be made neglecting the imaginary part of (21). As a consequence, we found a typical result of low-frequency analysis of cascode amplifier:

$$C_M \cong 2C_{\mu 1} \quad (24)$$

Alternatively, a solution based on a single numeric example accounting for the imaginary part of (21) is discussed in [1].

Including (24) in (22), the latter can be used to derive an expression for the optimum noise resistance as in the case of (9), however, due to the calculation complexity, it is more convenient to investigate a practical case, as for example the ideal cascode amplifier described in 3.1 (Fig. 13 of Chapter 3). The parameters of the small-signal model depicted in Fig. 1 of this Annex, deduced from the analysis of the cascode schematic, are reported in Table 1. The optimum noise resistance extracted from (22) is compared to the simulation of the whole cascode amplifier in Fig. 2. A good agreement between the two plots can be observed, proving that, after the introduction of $C_{\mu 1}$, our analysis can correctly reproduce the dependence of optimum noise resistance from emitter degeneration.

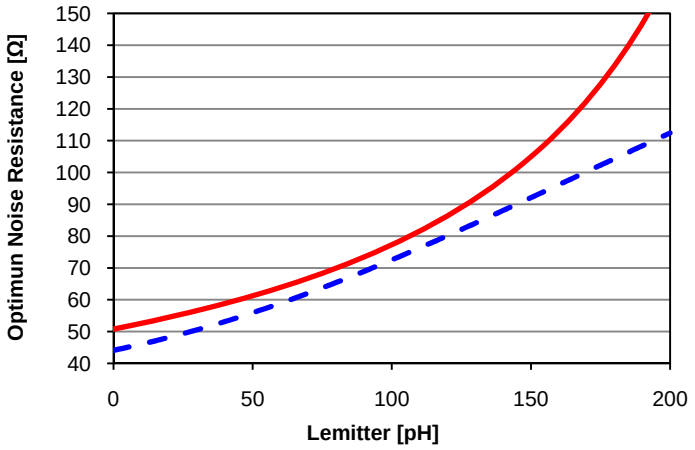


Fig. 2. Optimum noise resistance of an 80GHz cascode amplifier vs emitter degeneration, calculated from (22) (line) and resulting from simulation (dash).

Table 1. Parameters of the small-signal circuit equivalent to the cascode LNA of section 3.1.

Parameter	Value
I_C	3mA
g_m	120mS
βo	698.8
L_b	50pH
R_π	5813Ω
C_π	80fF
C_M	15fF
R_X	23Ω
R_E	3Ω

Bibliography

[1] Byung-Wook Min, and Gabriel M. Rebeiz "Ka-Band SiGe HBT Low Noise Amplifier Design for Simultaneous Noise and Input Power Matching", *IEEE Microwave and Wireless Components Letters*, Dec. 2007, Vol. 17, No. 12, pp. 891-893.