Power amplifiers for 5G phased arrays in ST CMOS065SOIMMW technology

Thomas Capelli

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

SPÉCIALITÉ ÉLECTRONIQUE

Par **Thomas CAPELLI-MOUVAND**

Amplificateur de Puissance pour Réseaux Phasés d'Antenne 5G Multi-bande en Technologie ST CMOS065SOIMMW

Sous la direction de : **Nathalie DELTIMPLE**
Co-directeur : **Anthony GHIOTTO**, Co-encadrant : **Philippe CATHELIN**

Soutenue le 24 Mai 2022

Membres du jury :

M. Denis BARATAUD	Professeur	Université de Limoges	Rapporteur
M. Tan-Phu VUONG	Professeur	Université Grenoble Alpes	Rapporteur
Mme. Florence PODEVIN	Professeure	Université Grenoble Alpes	Examinatrice
Mme. Nathalie DELTIMPLE	Professeure	Bordeaux INP	Directrice
M. Anthony GHIOTTO	MCF, HDR	Bordeaux INP	Co-directeur
M. Philippe CATHELIN	Ingénieur	STMicroelectronics	Co-encadrant
M. Eric KERHERVE	Professeur	Université de Bordeaux	Président
M. Antonio Lisboa DE SOUZA	Professeur	Université Fédérale de Paraiba	Examinateur

"Mieux vaut une page blanche que commencer sur la page de gauche"
P. Cathelin.

"Je n'ai jamais pu résister à la tentation d'aller voir les choses qui n'existent pas."
Geralt de Riv, Le Sorceleur, Tome 1 -Le Dernier Voeu, A. Sapkowski.

Remerciements

Pour commencer je tiens à remercier Monsieur Denis Barataud, professeur de l'Université de Limoges et Monsieur Tan-Phu Vuong, professeur de l'Université de Grenoble Alpes d'avoir accepté d'être les rapporteurs de mes travaux de thèse. Je remercie également, Mme Florence Podevin, professeure de l'Université Grenoble Alpes ainsi que Monsieur Antonio Lisboa De Souza, professeur de l'Université Fédérale de Paraiba d'avoir participé au jury de ma thèse et d'avoir examiné mon travail. Enfin, Monsieur Eric Kerherve, professeur à l'université de Bordeaux, d'avoir présidé de ce jury.

Cette aventure aurait été bien sûr impossible dans mon équipe d'encadrement de choc. Je remercie donc chaleureusement Nathalie Deltimple, ma directrice de thèse, pour m'avoir fait confiance ces trois années qui furent ponctuées par de nombreux défis. Je remercie également Anthony Ghiotto, mon co-directeur de thèse pour son aide et son précieux soutien. Et enfin un grand merci à Philippe Cathelin mon encadrant industriel, pour sa disponibilité et son indéfectible soutien.

Je souhaite remercier toute les personnes qui ont permis, de près ou de loin, venant de multiples entités industrielles ou académiques, la réalisation de mes travaux. Je remercie déjà l'ensemble de l'équipe d'encadrement, administrative et la direction du laboratoire IMS de m'avoir accueilli. Je remercie chaleureusement les doctorants du laboratoire, mon équipe CSH et l'équipe de la BEE branch, trop nombreux pour les citer sans oublier personne, pour leur aide et les bons moments que l'on a pu partager. Mention spéciale à Magali De Matos avec qui on a pu réaliser le défi qu'a représenté la réalisation de mes mesures et à mes colocataires Docteur Alexandre Berthier, Docteur Romane Dumont et Fréderic Munoz avec qui nous avons affrontés le tape-out en confinement.

Un énorme merci à mes collègues de STMicroelectronics, beaucoup ont gravité autour de la réalisation de mon circuits et ont su se rendre disponibles pour partager leur connaissances et leur savoir faire et je leur en remercie. Je tiens à remercier mon équipe RFSS pour son aide précieuse, dont Raphaël Paulin, Nicolas Peloux, Jasmina Antonijevic, Christophe Arricastres, Denis Pache, Patrice Garcia, Antoine Le Ravalec, David Gaidioz et enfin Jacky Uginet pour m'avoir accueilli dans son équipe. Également un grand merci Andreia Cathelin, Frédéric Paillardet et à l'équipe de caractérisation et device qui ont été d'une aide précieuse, notamment Daniel Gloria, Simon Bouvot, Joao Azevedo Goncalves, Cédric Durand et Sébastien Jan.

Un grand merci à tous les doctorants ST que j'ai pu croiser pour leur aide et les échanges que l'on a pu avoir et une pensée particulière à mes co-box Alexandre Flete et Dayana Andrea

Pino Monroy.
Je remercie mes amis pour leur soutien inestimable, qui a été très important, particulièrement en période de confinement malgré la distance, avec des soirées sur Discord qui furent (et sont toujours) de très bons moments.
Et enfin un immense merci à ma famille, mes parents, mon frère et ma grand-mère qui me soutiennent depuis toujours sous maintes formes, pour me permettre, entre autres, de mener à bien mes études et mes travaux.
Pour terminer cette liste non-exhaustive de remerciements, je dédie ces travaux à mon grand-père, Jacques Mouvand, qui fut un exemple d'altruisme, de curiosité scientifique, et une personne dotée d'une gentillesse et d'une culture sans bornes.

Thomas Capelli-Mouvand

Contents

Acronyms

1D 1-dimension.

1G 1st generation.

2D 2-dimensions.

2G 2nd generation.

3D 3-dimensions.

3G 3rd generation.

3GPP 3rd generation partnership project.

4G 4th generation.

5G 5th generation.

6G 6th generation.

A-VSWR active voltage standing wave ratio.

AA antenna array.

ACLR adjacent channel leakage ratio.

ADS Advanced Design System.

AF array factor.

ALM active load modulation.

AM/PM amplitude modulation to phase modulation.

AP alucap layer.

AS-parameters active scattering-parameters.

BEOL back-end-of-line.

BF beamforming.

BGA ball grid array.

BO back-off.

BOX buried oxide.

BS beam steering.

CA carrier aggregation.

CAN Canada.

CDMA code division multiple access.

CG common-gate.

CMOS complementary metal oxide semiconductor.

CPW coplanar waveguide.

CS common-source.

DC direct current.

DK design kit.

DPA Doherty Power Amplifier.

DRM design rules manual.

DUT device under test.

E-field electric field.

EDGE enhanced data rates for GSM evolution.

EIRP effective isotropic radiated power.

EM electromagnetic.

ESD electrostatic discharge.

EU Europe.

FD-SOI fully-depleted silicon on insulator.

FDD frequency-domain division.

FDTD finite time-domain method.

FE front end.

FEM finite element method.

FEOL front-end-of-line.

FSPL free-space path loss.

GaAs gallium arsenide.

GaN gallium nitride.

GPRS general packet radio service.

GSG ground-signal-ground.

GSM global system for mobile communications.

HB harmonic balance.

HFSS High-frequency Structure Simulator.

IC integrated circuit.

IMD intermodulation distortion.

IMD3 third-order intermodulation distortion.

IoT internet of things.

IP intellectual property.

ITU international telecommunication union.

LNA low noise amplifier.

LOS line-of-sight.

LTE long-term evolution.

M1 metal layer 1.

M2 metal layer 2.

M3 metal layer 3.

M4 metal layer 4.

M5 metal layer 5.

M6 metal layer 6.

M7 metal layer 7.

MC mutual coupling.

MIMO multiple-input multiple-output.

MMS multimedia messaging service.

mmW millimeter-wave.

MOM metal-oxide-metal.

MOS metal-oxide-semiconductor.

MOSFET metal-oxide semiconductor field-effect transistor.

MPW multi-project wafer.

NLOS non-line-of-sight.

NMOS N-type metal-oxide-semiconductor.

NVNA non-linear vector network analyzer.

OCP1 1 dB output compression point.

OFDM orthogonal frequency-division multiplexing.

OFDMA orthogonal frequency-division multiple access.

PA power amplifier.

PAA phased antenna array.

PAE power-added efficiency.

PAPR peak power to average ratio.

PCB printed circuit board.

PCR parallel coupled resonator.

PD-SOI partially-depleted silicon on insulator.

pHEMT pseudomorphic high-electron-mobility transistor.

PLS post-layout simulation.

PP push-push amplifier.

PSR peak-to-sidelobe ratio.

PTFE polytetrafluoroethylene.

PVT process, voltage, temperature.

Q quality factor.

QAM quadrature amplitude modulation.

RF radio frequency.

RMS root mean square.

RPA rectangular patch antenna.

Rx reception.

S-parameters scattering parameters.

SE spectrum efficiency.

SM spatial multiplexing.

SMS short message service.

SNR signal-to-noise ratio.

SOI silicon on insulator.

ST CMOS065SOIMMW ST complementary metal oxide semiconductor 65 nm millimeter-wave partially-depleted silicon on insulator.

TDD time-domain division.

TRx transmission / reception.

Tx transmission.

UE user equipment.

UMTS universal mobile telecommunications system.

US United-States.

VCCS voltage-controlled current source.

VGA variable gain amplifier.

VSWR voltage standing wave ratio.

List of Figures

List of Tables

Introduction

5G to the rescue of telecommunications

The emergence of the 5th generation (5G) for mobile telecommunications naturally comes from the continuous growth of telecommunication streams and wireless devices. It is particularly significant as the network gradually includes new applications. The automotive industry is now producing vehicles with smart detection of obstacles for safety and collision avoidance to converge to the driver-less cars. On top of that, the device should provide 100 % reliability with concise response time to ensure the safety of the vehicle occupant together with the other road users (pedestrians...). Automotive needs continuous connection to the network to benefit from cloud computing vehicle-to-vehicle and vehicle-to-infrastructure collaborations. Medical and emergency services need fast and reliable means of communications, and ultra-high-definition videos coupled with virtual reality devices abound to enable distant diagnostic and remote surgery. The 5G should also overcome the cable internet providers' network. Finally, connected devices (internet of things (IoT)) is also part of the 5G bringing more data to the global traffic with nearly 21 billion connected devices. Globally, Ericsson expects the monthly data traffic to be multiplied by five from 2020 and reach 300 EB by 2026 [Eri21].

Putting into perspective the ambitions of 5G, it represents a much more critical step to ensure the requirements of these applications in terms of speed, with latency around the millisecond and bit rates (up to 20 Gbitps) compared to previous mobile communication standards.

0.1 5G, a matter of heritage

The fifth of its name, the 5G is in 2021 the climax of forty years of evolution [GJ15] [Gaw15]:

Succeeding the mobile radiotelephone, the equivalent of what we could call the "1st generation (1G)" was announced in the early eighties with a data rate up to 2.4 kbitps accompanied by many drawbacks. It was the last generation using analog telecommunication, the voice was simply modulated to higher frequencies (hundreds of MHz). Consequently, the calls were unsecured and easily stored and demodulated in radio towers leading to eavesdropping.

In Europe, the 2nd generation (2G) was commercially launched in Finland in 1991. To reinforce the security of telecommunications, encryption was finally digital. Furthermore, the vital introduction of the global system for mobile communications (GSM) allows for the first time the transfer of digital data additionally to voice with bit-rates around 9.6 kbitps using time-domain division (TDD) (1 frequency per cell). However, soon after the launch of GSM, the circuit-switched bearer service appears to be very limiting and unreliable. Furthermore, the connection time to the network is important, and the connection is charged according to the time of usage.

The general packet radio service (GPRS) (2.5G) was then introduced, performing packet switching. The users accessed the public data network and provided bit-rates up to 64 kbitps. Finally, the GSM norm was revised by the 3rd generation partnership project (3GPP) and baptized enhanced data rates for GSM evolution (EDGE) [3GP], reaching 144 kbitps (48 kbitps in practice). The 2G brought new services such as short message service (SMS), e-mail, and multimedia messaging service (MMS).

The next decade comes with essential improvements in the quality of services: frequency-domain division (FDD) (using a different frequency for transmission (Tx) and reception (Rx)), global roaming, voice quality improvement, and bit-rates up to 2 Mbitps (84 kbitps in practice) make the 3rd generation (3G) [ITU03] a big step in mobile telecommunications. In addition, new technologies appeared along, such as universal mobile telecommunications system (UMTS) and code division multiple access (CDMA)2000. Though, the 3G came with some drawbacks. The 3G network is more expansive and requires user equipment power. Today, the services of generations before 3G are about to be or are already shut down.

It is safe to say that the 4th generation (4G) is an essential brick for the 5G foundations. Mixed up under the long-term evolution (LTE) label, the actual 4G standard is the final standard born from the "Long Term Evolution" since 3G. LTE does not mean 4G, the standard specification defined by the international telecommunication union (ITU) requires peak data rates up to 1 Gbitps. The LTE is sub-divided into several phases: The first is the basic "LTE" standard, presented as "faster than 3G". It comes next to the "LTE standard" that delivers downlink peak rates of 300 Mbitps and up-link of 75 Mbitps. Then followed by the "LTE

Advanced" that delivers the proper 4G data rates and finally comes the "LTE Advanced Pro" that is the closest standard to the 5G in terms of data rates. 4G combines abilities from the 2G and 3G by including FDD and TDD with orthogonal frequency-division multiplexing (OFDM) and an important number of carriers. To summarize, the 4G label is far from being obsolete as it will be continuously deployed until 2030 as the primary mobile telecommunication network.

Regarding the 4G, the 5G integrates new technological innovations:

It first pursue the work of the 4G with additionnal frequency bands in sub and millimeter waves and use new techniques such as carrier aggregation (CA). CA is a technique available on the last generations of Wi-Fi. Originally used in some cabled networks, CA combine several carriers with different frequencies for a single user to increase the data rate proportionally with the number of carriers used. The CA is available on the LTE Advanced with an aggregation of 5 carriers and soon up to 13 (up to 32-carrier aggregation is currently allowed by 3GPP). The feature enables the aggregation of carriers with non-adjacent frequency bands, FDD on one carrier, and TDD on another.

4x4 downlink multiple-input multiple-output (MIMO) systems consists in increasing the spatial streams on one or more component carriers that have been aggregated. A 4×4 MIMO uses a 20 MHz component carrier four times to operate the spectrum more efficiently.

The MIMO configuration implies the multiplication of N receiver chains (low noise amplifier (LNA), power amplifier (PA), filters. . .) in the mobile receiver that are coupled with N antennas. In addition, the quadrature amplitude modulation (QAM) up to 256QAM for downlink provides a multiplier of 1.33 on the data rate of each component carrier (and 64QAM for uplink) without changing the radio frequency (RF) front end of the mobile. A theoretical bit rate of 1 Gbitps can be achieved with three aggregated carriers with a 256QAM modulation.

0.2 5G, where and how?

The 1st phase range $<$ 6 GHz has frequency bands around 600 MHz (Canada (CAN) and United-States (US)) and 700 MHz (Europe (EU)), and the second is scattered between 1.4 and 5 GHz. These bands should be dedicated to long-range communications and the IoT. The second phase range will be within the millimeter-wave frequency bands, 24.25 GHz and 29.5 GHz (EU (24.5 – 27.5) GHz with FR @26 GHz). The following 5G phases should be located around 40 and 70 GHz. This broad coverage implies that for mobile applications and worldwide market targets, the devices should support broadband and band switching to evolve within the multiple 5G scenarios presented (see figure 1) [Mat+16]:

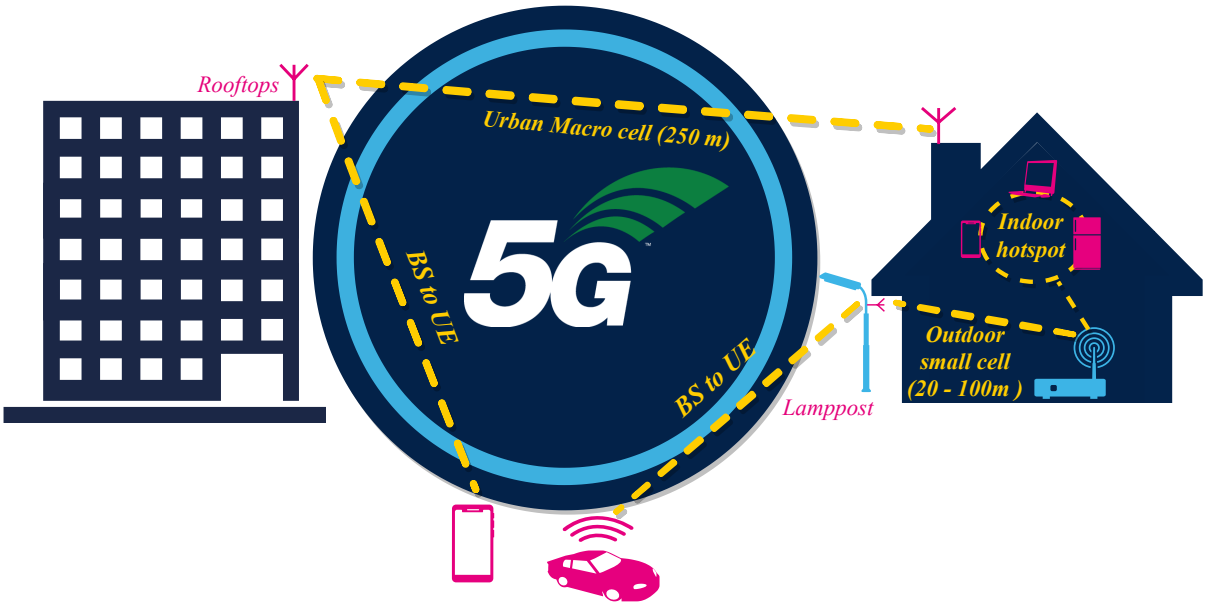


Figure 1. 5G urban scenarios

0.2.1 Urban Macro-cell

A scenario has been studied for macro-cell base stations (Specification on table 1) with antennas placed in a hexagonal or Manhattan (even random) grid at 25 m high above rooftops with a distance up to 250 m between sites. The antenna array elements considered are up to 256 fed by 16 ports with an overall gain of up to 30 dBi. The maximum transmitted power is between 30-40 dBm with 75 dBm effective isotropic radiated power (EIRP).

Parameter	Order of Magnitude
Antenna Element	Up to 256
Antenna Ports	Up to 16
Antenna Gain	(15 – 30) dBi
Antenna Height	25 m above rooftops
Antenna Distance	Up to 250 m
Max Power Transmitted	(30 – 40) dBm
Carrier Bandwidth	Up to 1 GHz

Table 1. Urban macro-cell specifications order of magnitude

0.2.2 Outdoor Small Cell

The primary approach uses a millimeter-waves base station for smaller cells from outdoor cells to house indoor internet boxes (Specifications on table 2). The size of these cells should

be around 100 m. The base station panels are located 10 m above the ground on street lampposts and spaced by 20 to 100 m. The number of antenna elements considered is up to 256 fed by 8 ports. This particular situation can be asymmetrical between BS and the mobile station. The EIRP and the number of elements in reception or emission can differ. The elements gain is around 5 dBi for an overall BS with 30 dBi gain; the transmitted power is estimated to be (14 – 25) dBm. In this situation, line-of-sight (LOS) and non-line-of-sight (NLOS) propagation schemes should be studied. The final purpose of this kind of structure is to provide 100 Mbitps/1 Gbitps without building or street modifications, as optic fiber deployment would require.

Parameter	Order of Magnitude
Antenna Element	Up to 256
Antenna Ports	Up to 8
Antenna Gain	5 dBi
Antenna Height	3 m / mounted on rooftop
Max Power Transmitted	(14 – 25) dBm
Carrier Bandwidth	(1 – 3) GHz

Table 2. Outdoor small cell specifications order of magnitude

0.2.3 Indoor hot-spot

Another case is studied for interconnecting outdoor base stations with indoor equipment such as internet boxes with laptops and smartphones with a 5G link. Considering the new Wi-Fi 802.11.ax [Kho+19] (also called Wi-Fi 6) standard currently under deployment with bit-rates up to 4.8 Gbitps, the 5G may not be competitive in this scenario.

0.2.4 Base Station to User Equipment

Here, the network undergoes significant changes, as it concerns the communication at mmW frequencies between BS and user equipment (UE). While other scenarios imply no or limited movement of the communicating device in space, the user equipment is mobile and with different speeds across the orthogonal network.

The question was studied early by Samsung [RSP14] and Qualcomm [Miy16]. The latter has proposed a 5G 30 GHz transceiver for smartphones integrated with its 2×2 MIMO dual-polarized phased antenna array. Since 2019 these transceivers have been commercially implemented in smartphones (the first one being Samsung Galaxy S10 5G) while mobile services were available

Specifications	Type	Value
Physical	Size	< Notebook (20×30 cm)
	Price	<1 k$ per site
	Weight	<4.5 kg
	Connection	Plug and Play
Electrical	EIRP	(53 – 65) dBm
	Streams	4-16 MIMO
	Nb Panels	1-6 per BS
	Nb Antennas	65-256

Table 3. BS to UE specifications order of magnitude

much later for most countries. The base station (Specifications on table 3) should be compact (less than a notebook), cheap (<1000 $), lightweight (< 4.5 kg), and plug and play.

0.2.5 Mobile communications at millimeter-wave frequencies

The spectrum allocated to 5G mobile telecommunications occupies various frequency ranges, scattered across 39 GHz of spectrum [Mil17].

In mobile telecommunications, the pre-5G and the first sub-6 GHz 5G frequency range (5G FR1), as its name suggests, is located below 6 GHz on the frequency spectrum.

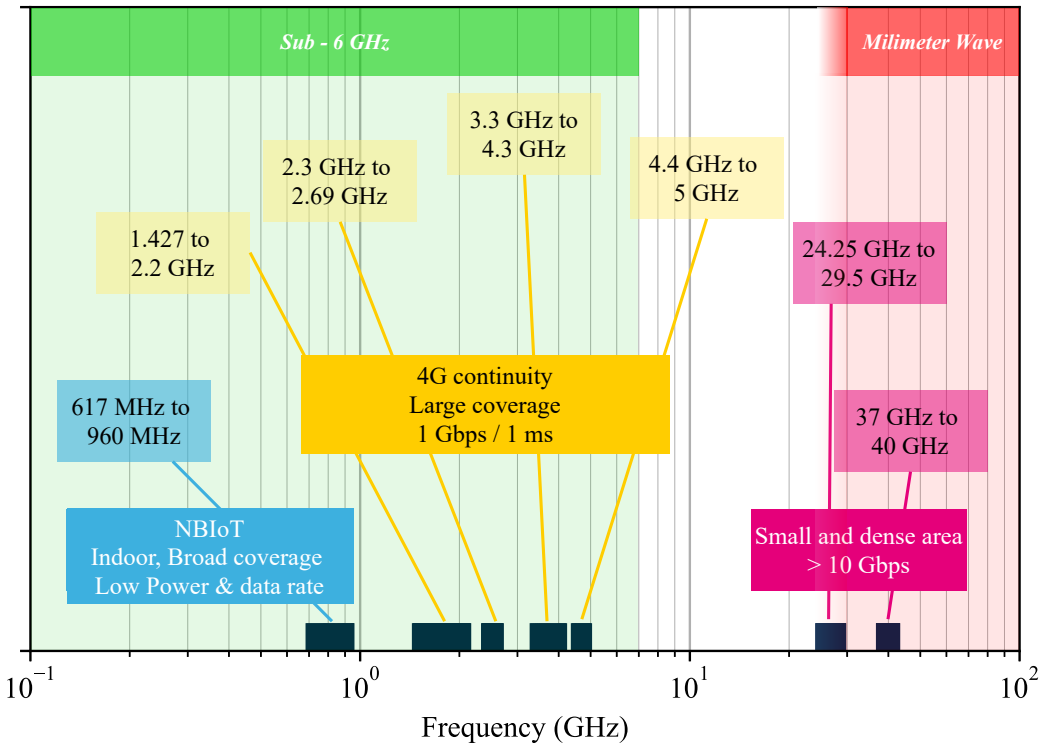


Figure 2. 5G sub-6 GHz and mmW frequency bands

To increase the bitrate of these communications drastically, the 5G standard place a second Frequency Range (5G FR2) around 28 GHz and has to face the demanding constraints of the free-space path loss at this frequency with a free-space path loss (FSPL) above 100 dB on a 100 m link.

$$FSPL = \left(\frac{4\pi d}{\lambda}\right)^2 \tag{1}$$

To mitigate the FSPL, a vital effort on each element composing the Friis equation has to be made to maximize the received power P_r (equation 2) The transmitted power P_t represents the contribution of the PA section essentially, it must provide strong output power, we will focus necessary attention on its efficiency; The deployment of 5G base stations should provide a very dense network and consequently consume a lot of energy (which is the leading critic about 5G and that motivate the research on the 6th generation (6G)), that is why the yield of the RF chain must fully contribute to the output signal power. The term D_r represents the gain of the receiver device antenna (Sub-Base station or user terminal in our case). Finally, D_t represents the gain of the transmitter antenna (Base station).

$$P_r = P_t + D_t + D_r + 20\log_{10}\left(\frac{\lambda}{4\pi d}\right) \quad (2)$$

With such a path loss, mobile communications at 28 GHz become possible with higher gain antenna D_t. To reach such a high gain, the antenna array structure is a good candidate. It is widely used and proposes a substantial gain in a precise direction. Doing so can compensate for a strong path loss. Furthermore, the main criticisms about 5G point out the increase of power consumption of its dense network architecture. This is why a critical focus concerns the efficiency of the communicating systems. All in all, these significant changes will not go without challenges for the front-end designers, especially with PAs designs.

This thesis proposes analysis and solutions to some of the crucial challenges the 5G mmW power amplifier designers have to face. The first chapter will show the main principle and design rules of phased antenna array with different realizations done by the scientific community. Then, two phenomenons that play a crucial role in the PA design are highlighted: The phased antenna array (PAA) parasitic active load variation effect critical for the system efficiency will be characterized and its impact quantified in a second chapter. Then, the third and last chapter will address the particular aspect of the non-linearities radiations in 5G PAA compromising the signal integrity of communications. A solution embedded on ST complementary metal oxide semiconductor 65 nm millimeter-wave partially-depleted silicon on insulator (ST CMOS065SOIMMW) technology with a 65 nm length transistors is proposed to relax the linearity of the PA in a phased array.

References

[3GP] 3GPP. *GPRS and EDGE*. URL: https://www.3gpp.org/technologies/keywords-acronyms/102-gprs-edge (visited on 11/15/2021).

[Eri21] Ericsson. *Ericsson Mobility Report Data And Forecasts - Mobile Data Traffic Outlook*. 2021. URL: https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast (visited on 11/15/2021).

[Gaw15] Ajun Uttam Gawas. "An Overview on Evolution of Mobile Wireless Communication Networks: 1G-6G". In: *International Journal on Recent and Innovation Trends in Computing and Communication* 3 (May 2015), pp. 3130–3133. ISSN: 2321-8169.

[GJ15] A. Gupta and R. K. Jha. "A Survey of 5G Network: Architecture and Emerging Technologies". In: *IEEE Access* 3 (2015), pp. 1206–1232. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2015.2461602.

[ITU03] ITU. *ITU - The Evolution to 3G Mobile — Status Report*. July 29, 2003. URL: https://www.itu.int/itunews/issue/2003/06/thirdgeneration.html (visited on 11/15/2021).

[Kho+19] Evgeny Khorov et al. "A Tutorial on IEEE 802.11ax High Efficiency WLANs". In: *IEEE Communications Surveys Tutorials* 21.1 (Jan. 2019), pp. 197–216. ISSN: 1553-877X. DOI: 10.1109/COMST.2018.2871099.

[Mat+16] Michal Maternia et al. *5G PPP use cases and performance evaluation models*. White Paper. 5G PPP, Apr. 2016, p. 39. URL: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-use-cases-and-performance-evaluation-modeling_v1.0.pdf (visited on 11/15/2021).

[Mil17] Lawrence Miller. *5G for dummies*. John Wiley & Sons. Vol. 1. 1 vols. Qorvo, 2017. ISBN: 978-1-119-42417-8.

[Miy16] Shin Miyazaki. *They said it couldn't be done, so we did it: Introducing the QTM052 mmWave antenna module family [Update]*. Apr. 2016. URL: https://www.qualcomm.com/news/onq/2018/07/31/they-said-it-couldnt-be-done-so-we-did-it-introducing-qtm052-mmwave-antenna (visited on 11/15/2021).

[RSP14] Whonil Roh, Ji-Yun Seol, and JeongHo Park. "Millimeter-Wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results". In: *5G Wireless Communication Systems: Prospects And Challenges* Samsung Research America (2014).

Chapter 1

Phased-Array antennas and power amplifiers at the heart of the 5G mmW

This chapter's purpose is to set the base principles that are needed to derive the reasoning, analysis, and realizations of this thesis. The idea is first to explain the behavior of 5G PAA and the PA and then highlight the issues in these new systems.

1.1 Phased array antennas for 5G

The antenna array proposes high efficiency; it comes with a panel of different configurations, from 1-dimension (1D) to 3-dimensions (3D) structures. The 2-dimensions (2D) planar antenna is suitable for mobile telecommunications because of its compactness, low cost, and beam steering abilities. Available on an extensive range of substrates, the planar antenna has strong flexibility in terms of design; an infinite range of different shapes can be possible as long as it is wisely sized at the right frequency with good impedance matching.

The antenna design must be simple, compact, and low cost for mass production purposes. Generally, rectangular patch antennas (RPAs) (and the circular variant) are used to ease the electromagnetic study and are preferred for their radiation predictability. Hence, the state-of-the-art focused its analysis on commonly used rectangular microstrip antenna patch arrays. They are simple to realize and accompanied with rich literature [Bal05] [Sai96] [Gar+01] describing their behavior and design.

1.1.1 Patch Antennas

1.1.1.1 Characteristics and substrate

Although the have been described quite late (70's), patch antennas are probably one of the simplest radiating structures. They are widely used for various applications with good integration properties using wavelength comprised between centimeters and millimeters. However, their simplicity comes with drawbacks; poor efficiency, spurious feed radiations, and very narrow bandwidth (high-quality factor (Q)). Fortunately, the strong flexibility of micro-strip structures is a helping hand to reduce the loss of performance. For instance, the choice of substrate is crucial, with a thickness usually [Bal05] comprised between:

$$0.003\lambda_0 < h_{sub} < 0.05\lambda_0 \tag{1.1}$$

A thicker substrate can raise the efficiency to 90 $\%$ and the bandwidth up to 35 $\%$. Those substrates are characterized by many parameters, including the relative permittivity often between:

$$2.2 < \varepsilon_{rsub} < 12 \tag{1.2}$$

Based on those two first parameters, we can expect a thick substrate with a low dielectric constant, better efficiency, and large bandwidth at the cost of larger radiating elements. We prefer high dielectric constant and thin substrates for microwave circuitry to minimize undesired radiation, coupling, and elements size. In contrast a thicker substrate is more suitable for antennas to get larger bandwidth (necessary for 5G applications).

1.1.1.2 Fringing effect and patch dimensions

The patch antenna design can easily be made referring to the theory described in [Bal05], introducing different models. Notably, an RPA can be seen as two radiating aperture of width W separated by a length L.

The finite dimensions of the antenna imply some field radiation fringing at the edges of the conductor from either side of L (figure 1.1), where the radiation essentially occurs. As said before, the substrates come with a permittivity larger than the air $\varepsilon_r >> 1$; most of the field lines are located within the substrate. However, small radiations still exist in the air, so this fringing effect must be considered in the design as it represents a small additional electrical length that modifies the resonance frequency of the patch. Basically, the patch looks wider

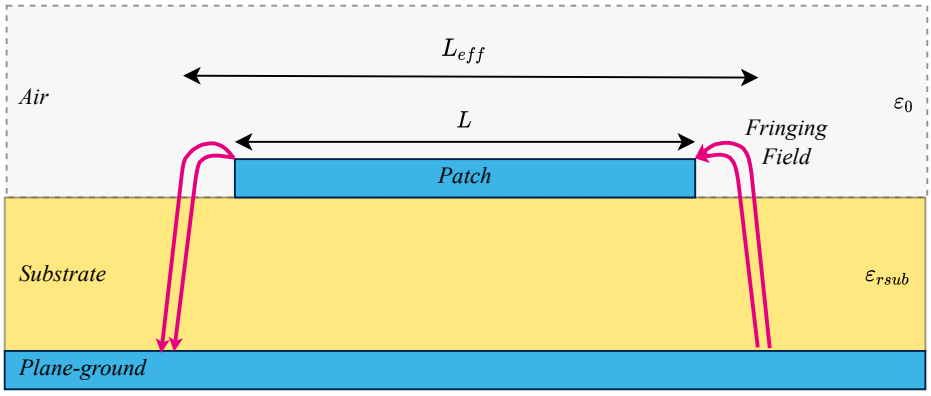


Figure 1.1. Fringing field on a rectangular patch antenna of length L

electrically than physically. When the patch is $\lambda/2$ long, the fringing fields are in opposition of phase at each side of the patch.

The fringing effect mainly represents the energy storage capacitance C given by (1.3)

$$C = \frac{\tan(\beta \Delta L)}{\omega Z_c} \tag{1.3}$$

with ΔL the extra length due to the fringing effect. To help the design, we introduce the relative effective permittivity that unifies the propagation environment. In our case, it is composed of air and substrate, and represents an "average" permittivity between $\varepsilon_0 = 1$ and ε_{rsub}. With $\varepsilon_{rsub} >> 1$, we have:

$$1 < \varepsilon_{reff} < \varepsilon_{rsub} \tag{1.4}$$

with ε_{reff} closer to ε_{rsub}.

ε_{reff} depends on the frequency; as this latter increases, the electric field lines tend to concentrate more and more into the substrate. Behaving more like a homogeneous line of one dielectric. Nevertheless, ε_{reff} is constant at lower frequencies and can be expressed by:

$$\varepsilon_{reff} = \frac{\varepsilon_{rsub} + 1}{2} + \frac{\varepsilon_{rsub} - 1}{2} \frac{1}{\sqrt{1 + 12\frac{h}{W}}} \tag{1.5}$$

Then, at the intermediate frequency, it monotonically increases to reach the value of ε_r. The fringing factor q (equation 1.6) can be used to evaluate the importance of the fringing effect. As the effective resonant frequency of the patch is defined by its effective electrical length, q is the ratio between the resonant frequency computed with and without the fringing effect:

$$q = \frac{(f_{rc})_{010}}{(f_r)_{010}} \tag{1.6}$$

with $(f_r)_{010} = \frac{1}{2L\sqrt{\varepsilon_r}\sqrt{\mu_0\varepsilon_0}}$

and $(f_rc)_{010} = \frac{1}{2L+2\Delta L\sqrt{\varepsilon_{reff}}\sqrt{\mu_0\varepsilon_0}} = q\frac{1}{2L\sqrt{\varepsilon_r}\sqrt{\mu_0\varepsilon_0}} = q\frac{c_0}{2L\sqrt{\varepsilon_r}}$.

With ε_0 and μ_0 the vacuum dielectric permittivity and magnetic permeability respectively, and c_0 the speed of light in vacuum. With that in mind, the patch dimensions are given by the following equations for a given substrate with a permittivity ε_{rsub} and a thickness h at the resonant frequency f_r. The width of the patch W is defined by (1.7); its variations will affect the radiation pattern and also control the input impedance: The wider is W, the lower is the impedance.

$$W = \frac{1}{2f_r\sqrt{\mu_0\varepsilon_0}}\sqrt{\frac{2}{\varepsilon_r+1}} = \frac{c_0}{2f_r}\sqrt{\frac{2}{\varepsilon_r+1}} \tag{1.7}$$

As said earlier, the length of the patch is artificially extended due to the fringe effect, the physical length L, and electrical length L_{eff} linked by the relation (1.8).

$$L_{eff} = L + 2\Delta L \tag{1.8}$$

with the additional electrical length ΔL due to the fringing field:

$$\Delta L = 0.412h\frac{(\varepsilon_{reff}+0.3)(\frac{W}{h}+0.264)}{(\varepsilon_{reff}-0.258)(\frac{W}{h}+0.8)} \tag{1.9}$$

The physical length drawn by the designer is then defined by (1.10), the resonant frequency is controlled by varying L.

$$L = \frac{1}{2f_r\sqrt{\varepsilon_{reff}}\sqrt{\mu_0\varepsilon_r}} - 2\Delta L \tag{1.10}$$

The patch parameters and its electromagnetic behavior are correlated by different relationships:

-The substrate height h affects mainly the patch bandwidth; higher the height, the broader is the bandwidth;

-The relative permittivity ε_r influences the fringing effect. Lower the dielectric constant wider is the fringing effect, and broader is the radiation pattern;

-The input impedance increases with the relative permittivity ε_r;

-The resonant frequency increases with ε_r, hence the size of the patch is smaller on a high ε_r substrate (a factor of 4 on the permittivity result in a twice smaller patch).

The patch antenna designs studied in this thesis are fed with a probe excitation (using a via from an inferior printed circuit board (PCB) layer to the top patch layer) as they are common and easy to simulate. To excite a horizontal polarization mode, the probe is centered on the width of the patch. Its position y on the length of the patch controls the matching of the antenna. When y is null, the probe is in the middle of the patch, and the input resistance is low. The closer the probe gets to the edges around $L/2$, the higher the resistance is.

1.1.2 5G phased array antenna

Planar arrays are widely used in various applications under very different forms: Radio astronomy uses the principle of planar arrays in an original way. For instance, the Karl G. Jamsky Very Large Array in New Mexico uses 28 m high radio telescopes in a Y-shape array to get interferometry measurements between 73 MHz and 50 GHz. Also, the still under construction Atacama Large Millimeter Array is planned to set up 66 radio telescopes in Chile. On a much smaller scale, the military avionic industry has an essential use of planar arrays for radar detection. They usually need to fit in plane's front-end and consist of a circular plate realized on PCB or analogous technology with a phase control technique that may vary.

1.1.2.1 Base Principle

The general principle of a PAA is easier to understand while working in reception and 1-dimension: Let's consider an incident wave coming toward a linear array of N elements separated by a distance d.

If the received wave hits the antenna directly perpendicularly with an angle $\theta = 0$, each element composing the array gets the signal simultaneously: The array effect builds a received signal with more power than a single antenna; hence the array gain is higher than a single element.

If the received signal hits the array with $\theta \neq 0$ (figure 1.2), the signal is not simultaneously received by the elements. One will first receive the signal and the other with a delay that increases with the distance and the element number. To build a coherent signal from all the elements, delays/phase shifts must be introduced. As the same length distances the elements, the relative phase shift $\Delta\varphi$ between them is constant for a value of θ and defined by (considering no phase progression):

$$\Delta\varphi = kd\sin\theta \tag{1.11}$$

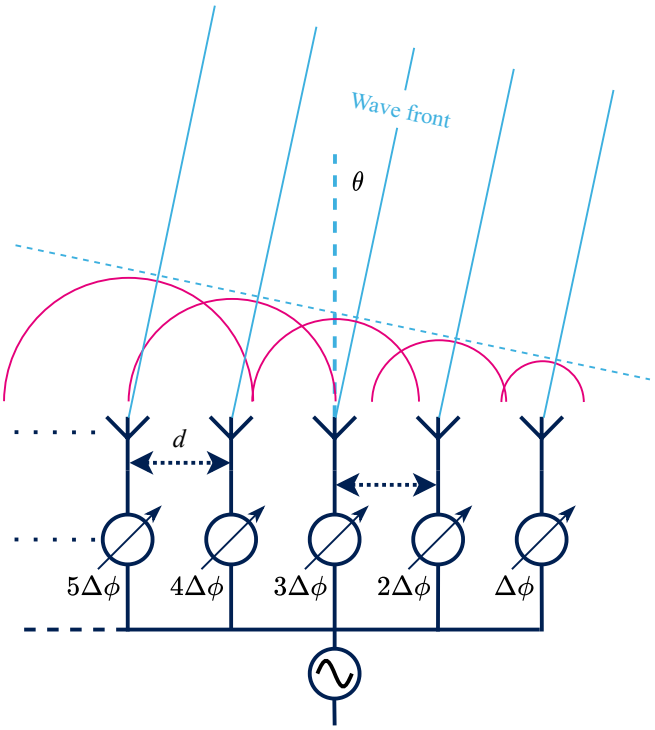


Figure 1.2. Phased array antenna beamforming

with $k = 2\pi/\lambda_0$ the wavenumber and λ_0 the free-space wavelength. It is worth noting that the θ achievable is not infinite; it is first limited by the architecture of the antenna (a plane ground typically forbids any radiations beyond 90°), and also restrained by the spacing d between the antenna:

$$\theta_{max} = \arcsin\left(\frac{\lambda}{d} - 1\right) \tag{1.12}$$

The distance d must be chosen appropriately. When θ is getting closed to θ_{max}, grating lobes appear, degrading the peak-to-sidelobe ratio (PSR) and drastically reducing the antenna's directivity. Beyond θ_{max}, the main beam finally reappears on the opposite side of the radiation pattern.

The array effect is characterized and can be plotted by a quantity such as the array factor (AF). Its general form is:

$$AF = \sum_{n=0}^{N-1} e^{jknd\cos\theta} \tag{1.13}$$

AF represents the sum of the N wave vectors created by the N phased elements phase-shifted by $\Delta\varphi$. Assuming that the sources behind each antenna are identical and independent, its normalized expression for a linear array is:

$$AF_n = \frac{1}{N}\left[\frac{\sin\left(\frac{N}{2}\Delta\varphi\right)}{\frac{1}{2}\Delta\varphi}\right] \tag{1.14}$$

We can find the zeros of the radiation pattern (angles where minimum radiation occurs) by solving $(AF)_n = 0$.

For higher gain and more beam orientation versatility, which is needed in 5G to search for users in a 2D space, a planar array constituted of $N \times M$ elements (distanced by d in both dimensions) is necessary. Then, we use two variables to drive the phase shift between elements along the x and y-axis. The beam steering is then possible along θ and ϕ:

$$\Delta\varphi_x = kd\sin\theta\cos\phi$$
$$\Delta\varphi_y = kd\sin\theta\sin\phi$$

Then the array factor becomes:

$$AF_n(\theta,\phi) = \left\{\frac{1}{N}\frac{\sin\left(\frac{N}{2}\Delta\varphi_x\right)}{\sin\left(\frac{1}{2}\Delta\varphi_x\right)}\right\}\left\{\frac{1}{M}\frac{\sin\left(\frac{M}{2}\Delta\varphi_y\right)}{\sin\left(\frac{1}{2}\Delta\varphi_y\right)}\right\} \tag{1.15}$$

The lobes in the radiation pattern occurs for:

$$\Delta\varphi_x = \pm 2m\pi, m = 0, 1...$$
$$\Delta\varphi_y = \pm 2n\pi, n = 0, 1...$$

While the maximum of the AF that gives the main beam is for $m = 0$ and $n = 0$ hence $\Delta\varphi_x = \Delta\varphi_y = 0$.

The strength of the array theory is that it is applicable no matter the nature of the radiating element used in the array.

1.1.2.2 5G millimeter-wave phased-array antenna

With the recent introduction of PAA to mobile telecommunications comes new technical constraints and additional costs that must be balanced with relatively simple designs and low-cost technologies. The use of frequencies around 28 and 39 GHz implies news strategies of PAA designs to achieve successful high data rate transfers, despite a much higher path loss than previous generations. This is where large antenna arrays (AAs) associated with beam steering (BS) are brought into play. While applications above 60 GHz can allow on-chip

antennas, the 28 GHz 5G with a wavelength around 1 cm imposes large antennas elements that represent around a few millimeter squares each and need the cohabitation of integrated circuits (ICs) and PCBs as it impose a vast network of interconnections between the IC and the PCB.

The main bottleneck for 5G PAA at mmW frequencies relies on the integration density. Even though patch antennas are "large" enough to be implemented on PCB, they are, at the same time, small enough to represent a challenge of interconnections with the IC that feed them.

Another challenge remains the multi-user access of the BS. As PAAs are very directive, multi-beam techniques are necessary to communicate with multiple users at the same time associated with adaptive algorithm techniques [LFY07] to keep track of the movements of the UEs. The management of PAAs for multi-beam is possible in different ways:

-Full array: Each antenna is preceded by transmission / reception (TRx) chains (most of the time switched); the signal of each element on the entire array can contain multiple tones with multiple phase shifts. Doing so, the antenna uses the entire gain available to generate numerous beams with different orientations.

-Segmented array: Again, with switched TRx front end (FE), each beam can be created by different sub-divisions of the array. The sources in each sub-array region deliver a signal destined to a single user location with a single and same phase shift. This setup shows less system complexity as each RF chain must provide a single phase shift. However, each subdivision of the array has a fraction of the total gain of the array, reducing the total EIRP provided by each beam.

-Segmented TRx array: The PAA can be segmented according to the reception and transmission functions; it can be performed dynamically, meaning that subsections are allocated to transmission and other to reception (and inverse, as each FE possess a complete TRx RF chain). On the other hand, it is also possible to physically separate the transmission and reception by either designing sub-arrays only equipped with the wanted function or by alternating an antenna for transmission and an antenna for reception within the array.

For many reasons (integration, measurements, number of connections, thermal issues) it isn't easy to find large PAA realization in laboratory configuration, meaning with monitoring and flexible control to get a maximum insight on the behavior of the devices. Therefore, the industry proposes several solutions [Ano21] [Ltd18], but of course, with minimal details on their product architecture for intellectual property (IP) protection reasons.

Anokiwave [Ano21] proposed in 2021 a 256 element active antenna (figure 1.3) working between 24.25 and 27.5 GHz and able to generate 60 dBm of EIRP with 2D beam steering

Figure 1.3. Anokiwave AWA-0142-IK commercially available 256-element 24.25-27.5 GHz active antenna array

thanks to analog RF beamforming. The product can generate a single beam using the 256 elements of the array and up to 4 beams using 64 elements with a maximum angle of steering of ±60°. Finally, the product presents a weight of 3 kg in a package of $26 \times 15 \times 4$ cm.

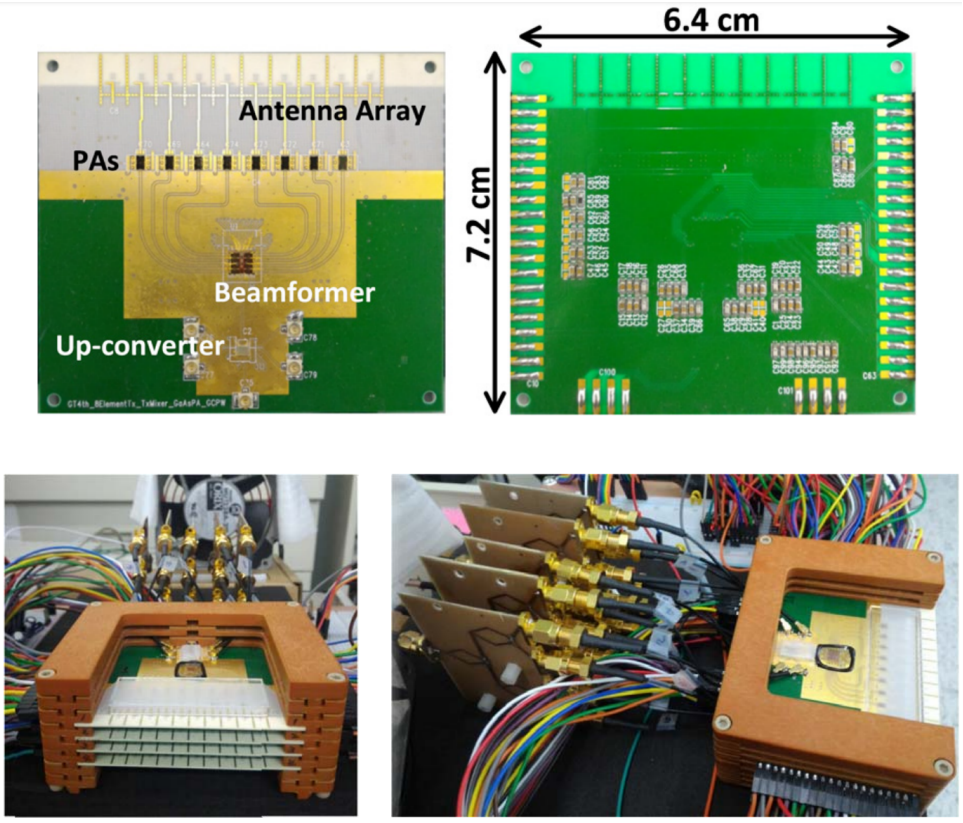


Figure 1.4. Realization of a scalable 38 GHz 4×8 element Tx PAA on PCB

An example that reflects the challenge of realizing a "lab-friendly" PAA is this scalable 5G 32-element array that has been recently proposed by [Che+21] at 38 GHz. A single chip containing 4 TRx modules with switched PAs and LNAs feeds a cluster of 4 antennas. Of course,

the integration of many ICs on a small surface represents thermal issues that are difficult to mitigate while being constrained by the placement of devices to achieve a controlled RF phase path. This analysis has shown that using multiple scalable PCB modules in a "brick" end-fire array configuration (shown in figure 1.4) enabled proper thermal dissipation. With the critical EIRP that phased antenna arrays must deliver, a split (figure 1.5b) configuration is preferred. The shared TRx configuration (figure 1.5a) is then reserved for user equipment setups. A measurement bench has been assembled with the Tx module of 32-element put 30 m away from a Rx 18-element brick array. Concerning the active modules, a 65 nm complementary metal oxide semiconductor (CMOS) technology has been chosen for the Tx and Rx beamformers and up/down converters. In contrast, the PA and LNA have been fabricated in 0.15 µm gallium arsenide (GaAs) pseudomorphic high-electron-mobility transistor (pHEMT).

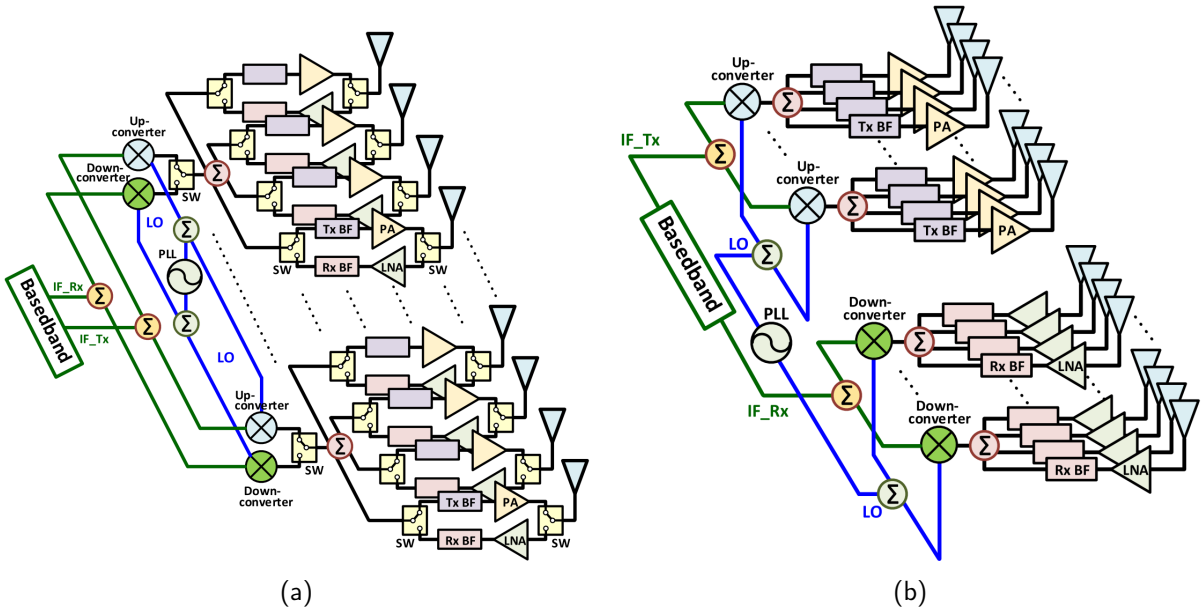


Figure 1.5. Phased array with integrated TRx (a) and separated Tx / Rx (b)

The results show that transmission with a bit rate of 2.4 Gbitps has been achieved with a 64 QAM over 90 dBm of FSPL at 38 GHz. The setup offers capabilities of ±60° of beam steering in azimuth and ±30° in elevation with a step of 22.5° and a peak EIRP of 47.5 dBm.

Each scenario of the 5G deployment finds an adapted solution either for UEs or for BSs; long-distance communications are covered with large arrays with high EIRP in large systems to handle the power dissipation needed. Shorter distances are covered with small modules easier to fit in the landscape and indoors. The size of the transceivers package is not scaled by the array size but essentially by the dissipation of power that they need, as the active or passive cooling systems usually represent a significant part of a module's volume. As it is not our focus

in this thesis, the UE antenna architecture (typically a mobile phone) has not been extensively discussed. However, one can note that for obvious integration constraints, user equipment architectures differ from base stations: Usually, UEs high wireless and thermal performances are achieved with beamforming (BF) modules containing up to 8 elements in mobiles phones [Dun+18] and up to 16 in tablets [HDX17]. They are usually physically dispersed on the device PCB and driven with spatial multiplexing (SM) for higher spectrum efficiency (SE).

1.2 5G power amplifiers and active load modulation

Notwithstanding the higher level of output power, efficiency, and peak power to average ratio (PAPR) requirements with the use of orthogonal frequency-division multiple access (OFDMA), 5G can still benefit from low-cost silicon IC, especially from silicon on insulator (SOI) technologies: One of the advantages of using PAA is that it is sized with M elements driven by M PAs, instead of using a single and significant power source that may come with heavy design constraints, an equivalent EIRP can be achieved with M smaller PAs, giving the room for mmW SOI PAs [Des+18] as it avoids the feeding network loss, and relax the thermal issues. The scientific community has been very active in proposing all kinds of devices to meet the requirements of the incoming 5G; the PA designers are really focused on mainly two types of architectures: Doherty and balanced amplifiers.

1.2.1 Preamble: Active load modulation notion

A simple PA is designed to deliver a specific output power P_{out} (equation 1.16) under a particular supply voltage V_{DD} in a given load Z_L. If this load changes, P_{out} is no longer guaranteed. This mismatch can be caused by the environment of the PA or even controlled in particular cases. It is difficult to talk about 5G PAs without mentioning the active load modulation effect. The phenomenon of controlled impedance variation, commonly called active load modulation, results from different sources delivering power in the same node, modifying the amplitude and phase of the combined currents/voltages. It can be purposely introduced to control the impedance seen by power amplifiers and maximize their performances, as in Doherty amplifiers or again in outphasing Chireix architectures [Chi35].

$$P_{out} = \frac{V_{DD}^2}{2 \cdot real(Z_L)} \tag{1.16}$$

To illustrate the principle, let's take a simple case of two current sources S_{main} and S_{aux} of impedance Z_{Gmain} and Z_{Gaux} that deliver respectively two currents I_{main} and I_{aux}, through two ideal transformers with a coupling coefficient $k = 1$, into Z_{load}.

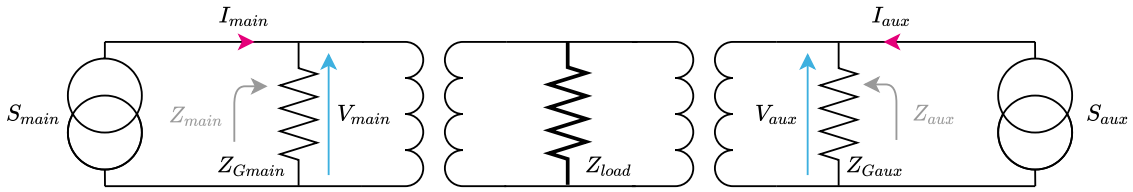


Figure 1.6. Load modulation with main and auxiliary current sources

We can express the impedance Z_{main} seen by S_{main} according to the other elements in the circuit, it gives:

$$Z_{main} = Z_{Gmain}\left(1 - \frac{I_{main}\cdot Z_{Gmain} + I_{aux}\cdot Z_{Gaux}}{I_{main}(Z_{Gmain} + Z_{Gaux} + Z_{load})}\right) \tag{1.17}$$

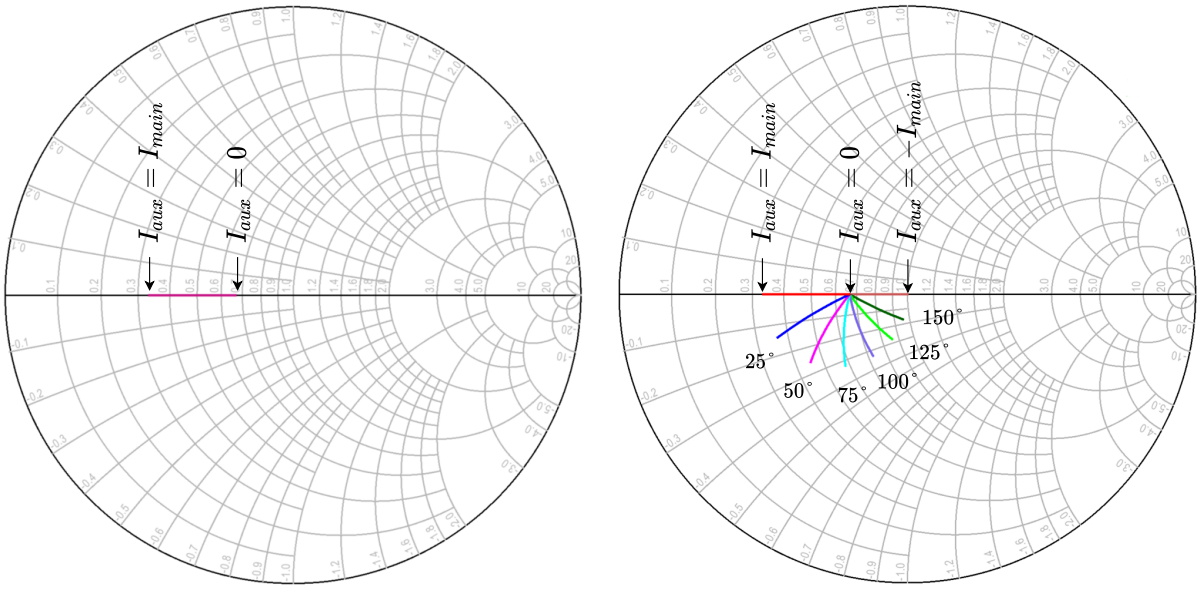
The source S_{main} delivers a constant current while S_{aux} is considered variable between 0 and I_{main}. Although, as shown in equation 1.17, the Z_{main} is dependent on I_{aux}, the range of variation of Z_{main} is:

$$\underbrace{\frac{Z_{Gmain}\cdot Z_{load}}{Z_{Gmain} + Z_{Gaux} + Z_{load}}}_{I_{aux}=0} \leq Z_{main} \leq \underbrace{\frac{Z_{Gmain}(Z_{load} + Z_{Gaux})}{Z_{Gmain} + Z_{Gaux} + Z_{load}}}_{I_{aux}=I_{main}} \tag{1.18}$$

For instance, if the three impedances of the circuit are $Z_{Gmain} = Z_{Gaux} = Z_{load} = 50\Omega$, the range of variation is then between:

$$16.68 \leq Z_{main} \leq 33.34 \tag{1.19}$$

The main impedance can indeed move along the real axis of the smith chart of the figure (1.7a). Assuming we keep the same main source impedance, equation 1.18 shows that a higher Z_{Gaux} or a lower Z_{load} increases the range of load variation. Furthermore, the auxiliary current can be phase-shifted to achieve a specific range of complex impedance. The same circuit is analyzed with a I_{aux} varying from 0 to 180°. Z_{main} is then modulated in amplitude and in phase. The load modulation allows the covering of a particular impedance area in the smith chart (figure 1.7b) directly controlled by a current in amplitude and phase. With a 360°of phase variation, the load variation range would cover a maximum VSWR of 2:1.



(a) I_{aux} magnitude variation (b) I_{aux} magnitude and phase variation

Figure 1.7. Z_{main} active load variation controlled with I_{aux}

The tracking of an optimal load is the direct application of the active load modulation (ALM) that can be easily seen with the example given. A PA designed to deliver its maximum output power and efficiency in a given load will have trouble maintaining its operating state if, for any reason, this load turns out to be variable. The auxiliary source can compensate for the load variations and actively track the best impedance to ensure that the amplifier always sees the optimal load value.

The principle shown in figure 1.6 has other assets that inspire the Doherty PA architecture.

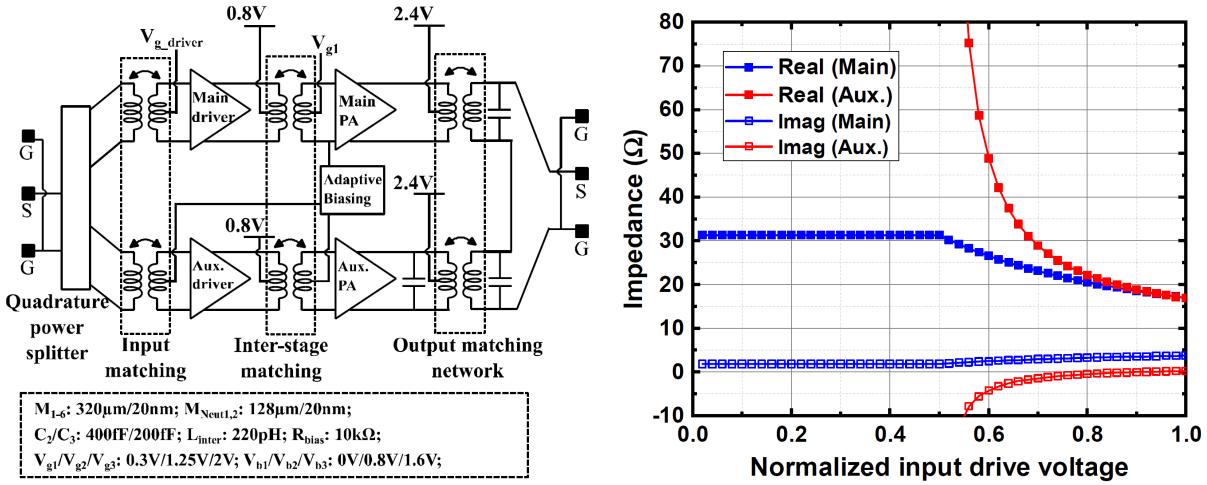
1.2.2 Doherty power amplifier

The Doherty Power Amplifier (DPA) is an architecture developed in 1936 by William H. Doherty [Doh36], an engineer from Bell Telephone Laboratories, and consisted of two tube amplifiers put in parallel biased in class B to deliver tens of kilowatts. Today the Doherty amplifier finds a new place in 5G base station transmitters with high PAPR signal and where significant power and efficiency are required. The literature proposes many solutions with gallium nitride (GaN) technologies for very high output power with P_{sat} between 35 and 43 dBm at mmW [Nak+18] but CMOS and SOI technologies also find their place for moderate power with P_{sat} between 18 to 23 dBm [HWW19].

The main playground for this type of architecture is efficiency improvement, especially in the back-off region. The principle resides in two sources, often with two different profiles of

conduction: a principal and linear source, often called carrier power amplifier of class A or AB, associated with a peaking source of class B or C. The carrier amplifier is always active and deals with the low output power, while the peaking amplifier deals with the high power part when the carrier starts to compress.

Here is an example [Zon+20] of a DPA in 22 nm fully-depleted silicon on insulator (FD-SOI) CMOS technology with a 22.5 dBm output power and 28.5 % PAE at the 1 dB output compression point (OCP1). The design proposed is a fully differential architecture with input drivers and auxiliary adaptive biasing. Both PAs are composed of differential stacked transistors with capacitance neutralization (figure 1.8a). Although one of the main constraints of the DPA is the quarter wavelength impedance inverter required to achieve load modulation, they usually demand some work from the designers to make them the most compact as possible.



(a) (b)

Figure 1.8. Fully differential DPA schematic (a) and ALM with input drive level (b) [Zon+20]

In fully integrated DPAs, as in our example, the inverter is implemented using lumped components: an LC circuit is sized to achieve both impedance inversion and transformation that works with the associated balun. The two L and C components used for inversion are chosen to resonate to interfere in the matching circuit. The output impedance transformer is designed to match $2R_{opt}$ to 50 Ω. By varying the drive voltage of the auxiliary circuit, the real impedance seen by the main circuit is modulated (figure 1.8b).

The power sweep (figure 1.9) shows a pretty typical response of DPAs: at low power, the main amplifier follows a PAE behavior of class B, with the drive level increasing, the auxiliary amplifier turns on to slightly modulate the load of the main amplifier, achieving a PAE enhancement from 5 dBm of P_{out}. While the main amplifier tends to reach its compression

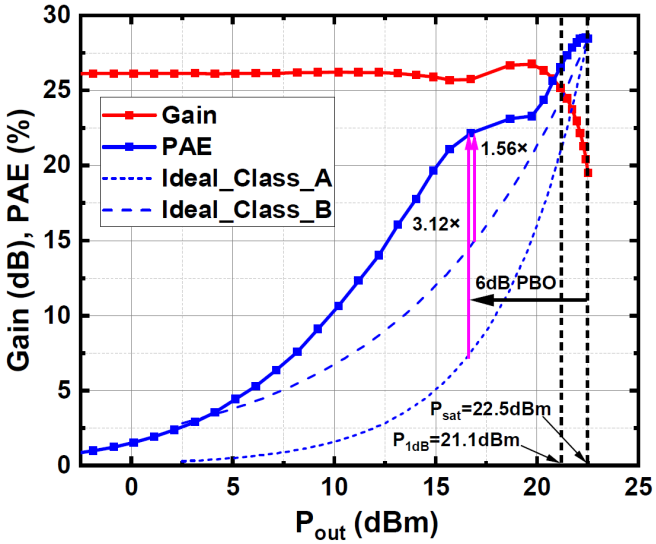


Figure 1.9. Gain and PAE with P_{out} compared to class A and B profiles [Zon+20]

point, the load modulation created by the auxiliary makes a gain expansion, pushing away the circuit compression of 6 dB and creating a typical PAE flat response that characterizes DPAs. Compared to a basic class A the PAE improvement is shown to be a multiplication by three at 6 dB back-off and by 1.5 compared to a class B amplifier.

Although DPAs show substantial assets such as back-off efficiency enhancement, they have certain design complexity. Particular attention is required, especially on the phase recombination of the two amplifiers' paths, which is critical to optimal performance. Furthermore, the integration of splitters and delay lines limits the device's bandwidth.

1.2.3 Balanced power amplifier

The balanced amplifier was first introduced in 1965 [EEK65]; it is also appreciated and documented by the 5G scientific community. While it does not propose efficiency improvement, the balanced architecture's primary asset is being robust to load variation with a lower design complexity than the Doherty architecture. Again, this architecture relies upon two identical amplifiers in parallel that operate in quadrature. The input and output split and recombination are realized with couplers (with topologies that may be very diverse [Poz11]) that allow the load variation robustness. Compared to a classic single PA, the load Z_L is not directly presented at the output of the amplifier, the properties of phase recombinations of the hybrid coupler allow the termination load Z_T present on the isolated port of the coupler to absorb the variations of Z_L, the reflections due to the mismatch ending up in the termination load. It allows the PA

to work almost constantly around its optimal impedance. Of course, the variation absorption works to some extent; a maximum VSWR is admitted by design to operate correctly.

A 28 GHz balanced power amplifier [MKK17] has been realized on 28 nm FD-SOI CMOS technology. It exhibits an output power of 18.7 dBm with 12.4 % PAE. The circuit (figure 1.10) comprises a differential driver, an input splitter, two class-AB push-pull amplifiers, and an output combiner. The power splitter and output combiner contain impedance transformers, and the hybrid couplers are realized with twisted coupled transmission lines, allowing more accessible connections in the circuit floor plan. The VSWR robustness is shown by plotting

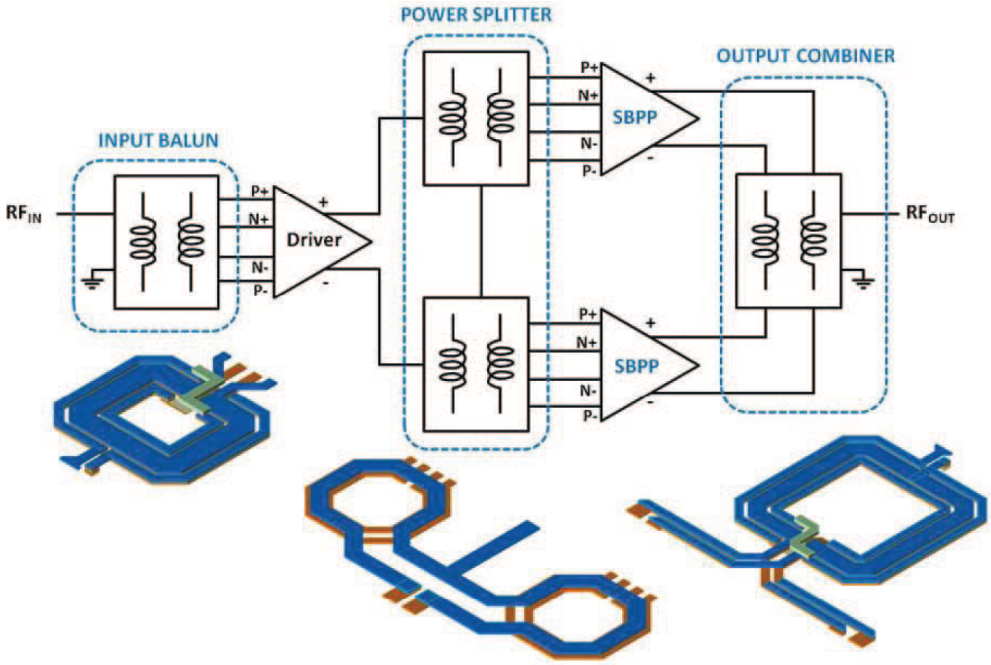


Figure 1.10. Gain and PAE with P_{out} compared to class A and B profiles [MKK17]

the current variations of the PA in balanced configuration versus the individual PA (see figure 1.11). Under a VSWR of 3:1 the balanced PA shows a pretty much constant current. In contrast, the individual PA undergoes much stronger current variations.

The termination load can be challenging to synthesize on a wide bandwidth and must eventually dissipate power in case of substantial mismatch. To summarize, the load variation immunity of balanced amplifiers is a feature that mainly benefits the 5G PAA as they present a peculiar drawback compromising the stability of the output impedance of the PAs that we will discuss in the next section.

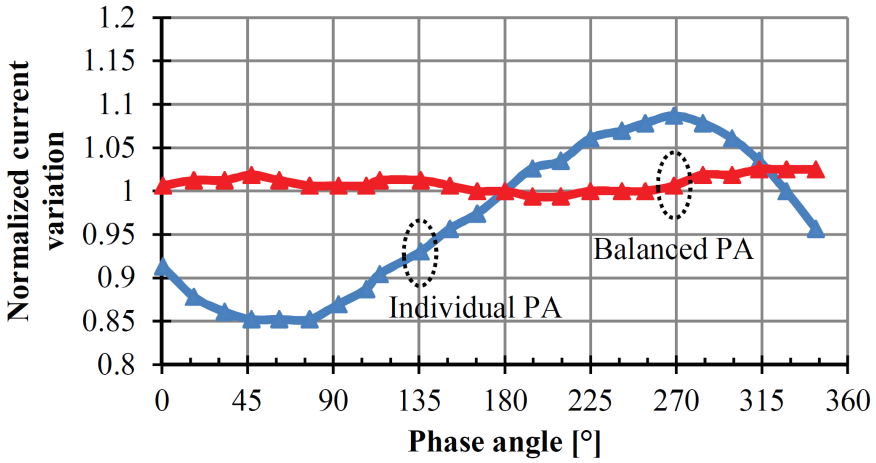


Figure 1.11. Normalized current variations of the balanced PA compared to the individual PA under 3:1 VSWR [MKK17]

1.3 Antenna mutual coupling and parasitic active load modulation

While ALM benefits the designers to meet their specifications, it also exists under an unwanted form in phased antenna arrays.

1.3.1 Mutual coupling, how so?

The analysis of phased antenna array, as seen in section 1.1.2, is frequently simplified by assuming that each source of power exciting the antennas are identical and independent. Unfortunately, in reality, there is no such thing. By construction, we know from section 1.1.1.1 that our patch antennas have many spurious radiations detailed in [Bhu13], degrading their quality factor. Among other issues (distorted radiation pattern, reduced radiation efficiency...) some of these radiations venture to the surroundings of the antenna and cause problems when the elements are disposed of in an array with relatively close spacing. In consequence, the patch antennas are communicating together through an inevitable coupling that occurs through different propagation mechanisms existing within the two mediums of propagation available: the air and the substrate (figure 1.12). We can identify three main mechanisms: free space, surface wave, and near-field coupling.

The free space coupling (as $d > 2D^2/\lambda$ with D the element diameter) is directly due to the natural horizontal radiation of antennas traveling to neighbor's elements.

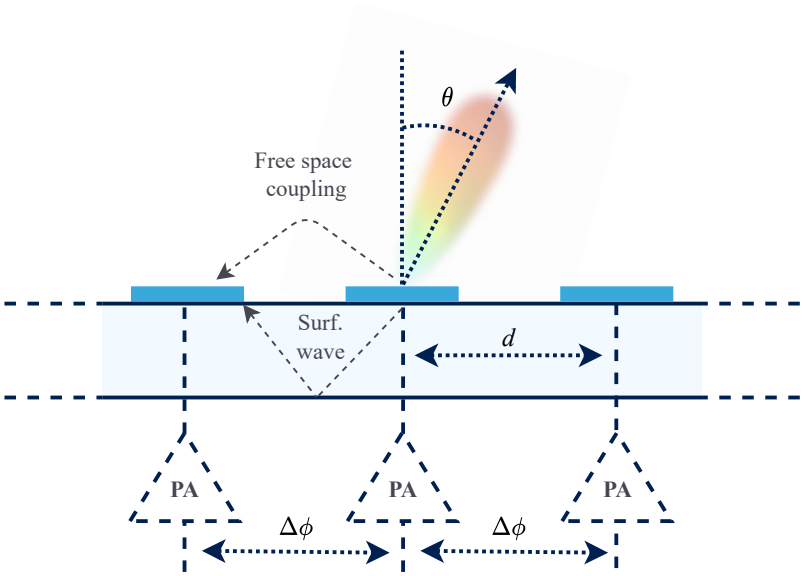


Figure 1.12. Phased-array antenna in Tx mode and main coupling mechanisms

The surface-wave coupling exists when antennas share a common substrate and ground plane. As shown in figure 1.13, surface waves are not radiated. Instead, they are kept trapped within the substrate and are reflected between the ground plane and the substrate/air interface. These waves are created when their reflection angle θ_{sw} is greater than $\theta_{critical}$ that depends on the material relative permittivity ε_r as in (1.20):

$$\theta_{critical} = \frac{1}{\sin \varepsilon_r^{-1/2}} \tag{1.20}$$

The surface waves coupling decays with the distance in $\frac{1}{\sqrt{d}}$. The reactive near-field coupling occurs on a radius of $\lambda/2\pi$ around the radiating element and rapidly decays. It is not significant here as the spacing $\Delta l = \lambda/2$ between the elements is beyond this range.

The resulting parasitic ALM seen in the plan of the power amplifier is then a function of parameters such as the substrate characteristics, the antenna spacing, the array dimensions, and the interconnections between the power amplifiers and the antennas.

The design of an array can not only be based on a single patch antenna analysis. Indeed, the definition of the scattering parameters (S-parameters) used for the antenna matching imposes that all other ports of the system are 'off' and loaded by their characteristic impedance to derive the impedance of an element. However, this metric is insufficient in an array, as all the elements are simultaneously excited. The mutual coupling (MC) influence on the surrounding patches must be considered to match each antenna appropriately. The active scattering-parameters

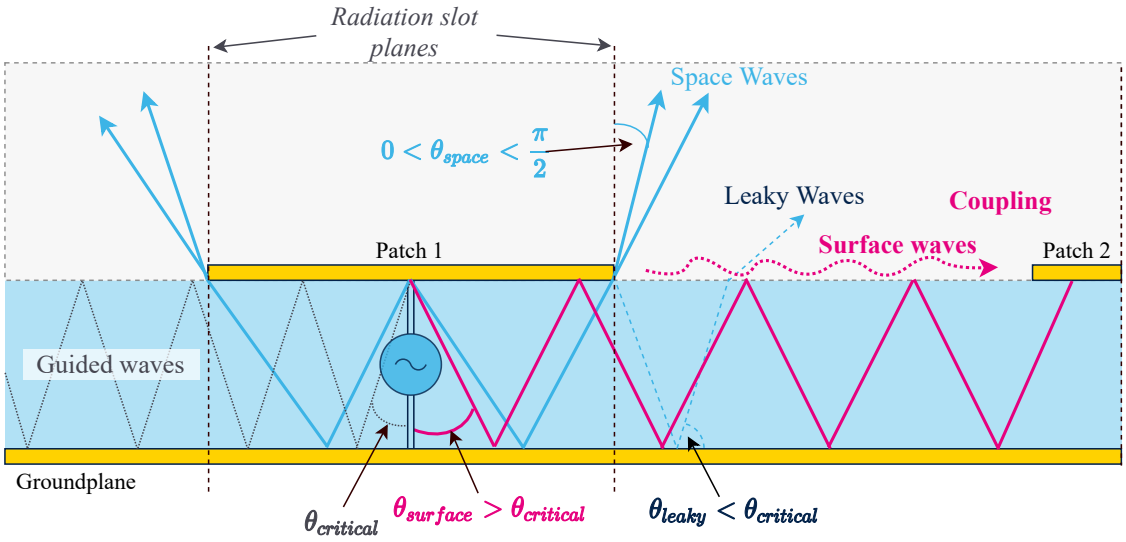


Figure 1.13. Phased-array antenna in Tx mode and main coupling mechanisms

(AS-parameters) notion is then introduced: In a N-network, the impedance seen by the input port of a patch m is composed of the $N-1$ coupled RF path contributions added to the intrinsic and passive impedance S_{mm} such as in (1.21)

$$AS_{mm} = \frac{\sum_{n=1}^{N} A_n e^{j\phi_n} S_{mn}}{A_m e^{j\phi_m}} \tag{1.21}$$

where A_n and φ_n describe the amplitude and phase exciting the n^{th} coupled element, and S_{mn} contains the attenuation and the delay that the coupled signal undergoes before reaching the victim antenna. Usually, the VSWR is a metric widely used to quantify the importance of the mismatch that a device undergoes. In this case, we prefer the A-VSWR notion to discern that the mismatch is dependent on active devices that evolve across time.

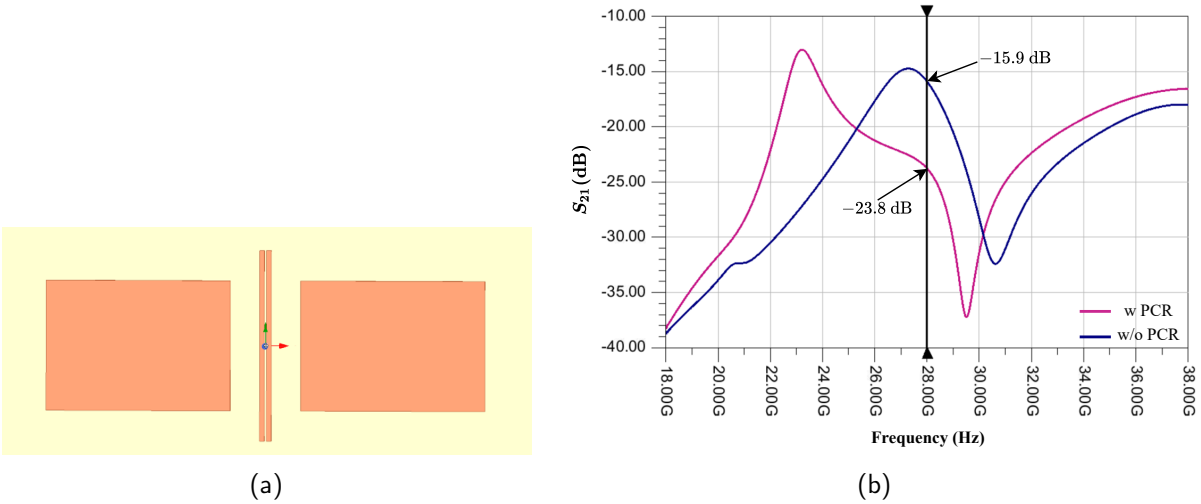
A PA will undergo the most powerfull ALM effect when the active impedance gets far from its original value for which the amplifier has been designed. A substantial mismatch might happen due to different factors: As the terms S_{mn} are complex, the sum of their contributions does not necessarily tend to induce higher and higher mismatch; the phase component is important to consider. On the other hand, to perform beam steering (section 1.1.2), we know that the excitation ϕ_n (and eventually A_n) must be modulated and controlled by $\Delta\varphi$ (that becomes equation (1.22) with half-wavelength spacing $d = \lambda_0/2$). We are inducing parasitics but somewhat "controlled" variations of the PA output impedance.

$$\Delta\varphi = \pi \sin\theta \tag{1.22}$$

Within this matrix that is coupled through dense and complex RF paths, each amplifier's behavior is now dependent on elements that can be physically away from its location due to the domino effect of each source affecting its neighbors. The BF variable θ is directly part of what defines the output impedance of the PA and is susceptible to disturbing its well-being. The level of this disturbance is yet to be determined.

1.3.2 Example of antenna mutual coupling mitigation techniques

The antenna designers have some design tips that can initially help to reduce the MC effect that essentially travels through the air and surface waves. Parallel coupled resonators (PCRs) [Vis+17] are structures that can be implemented to reduce the coupling between patch antennas. The principle is relatively simple; multiple parallel microstrip lines with a determined length, width, and spacing are disposed perpendicular to the direction of propagation (here, horizontal) of the coupling to create a high impedance blocking mainly the surface waves.



(a) (b)

Figure 1.14. Simulated 28 GHz RPA separated with a PCR for isolation enhancement on RO4730G3 substrate

Two RPAs have been simulated at 28 GHz distanced with $\lambda/2$ on RO4730G3 with an $\varepsilon_r = 2.98$ on High-frequency Structure Simulator (HFSS). A single PCR has been sized with the help of Advanced Design System (ADS), the resonator is composed of two parallel microstrip lines sized to create transmission zero (high Z) at 28 GHz. The computed dimensions are a width $W = 0.1$ mm, a length $L = 4$ mm and a spacing $s = 0.045$ mm. A comparison with and without PCR is performed. The S_{21} between the antennas improves from −15.9 dB to −23.8 dB, increasing isolation of +7.9 dB.

One of the main drawbacks of the PCRs is that they may become complex to implement simultaneously on horizontal and vertical coupling as the resonators are usually longer than the patch antennas; the risk is that the space between antennas may not be sufficient for this type of implementation. Furthermore, it is also essential to verify that the resonators do not significantly impact the array's radiation pattern.

1.4 Nonlinearities radiations

1.4.1 Origins of nonlinearities

The nonlinearities are a big deal when working on amplification. They are a source of signal integrity degradation and loss of efficiency. The good operation of 5G OFDMA and its modulation schemes require a reduced adjacent channel leakage ratio (ACLR) and a still strong PAPR that are critically impacted by the PA linearity and efficiency. A small review of the contributor to these nonlinearities in PA is proposed.

1.4.1.1 Small and moderate signal

The physical behavior of metal-oxide semiconductor field-effect transistors (MOSFETs) under small and moderate signals has been widely documented and described [Raz17] [All10]. Consequently, the following section reminds the main contributors of nonlinearities in MOSFETs in moderate-signal.

We can represent any general MOSFET amplifier with a simple common source topology (figure 1.15a) with an NMOS transistor and output load. It can be seen as a basic voltage-controlled current source (VCCS) that draws a current I_D flowing from its drain to its source that is delivered by a static (direct current (DC)) source.

The voltage v_{gs} dynamically modulate i_d between the transistor drain and source. The amount of current i_d driven for a given voltage amplitude v_{gs} is represented by a transconductance g_m (equation 1.23).

$$g_m = \frac{i_d}{v_{gs}} \tag{1.23}$$

Supposing that the VCCS is a perfect source, the shape of the voltage measured on the output load R_L signal has a perfectly matched frequency spectrum with the input voltage, with, in addition, the voltage gain of the amplifier that is defined by:

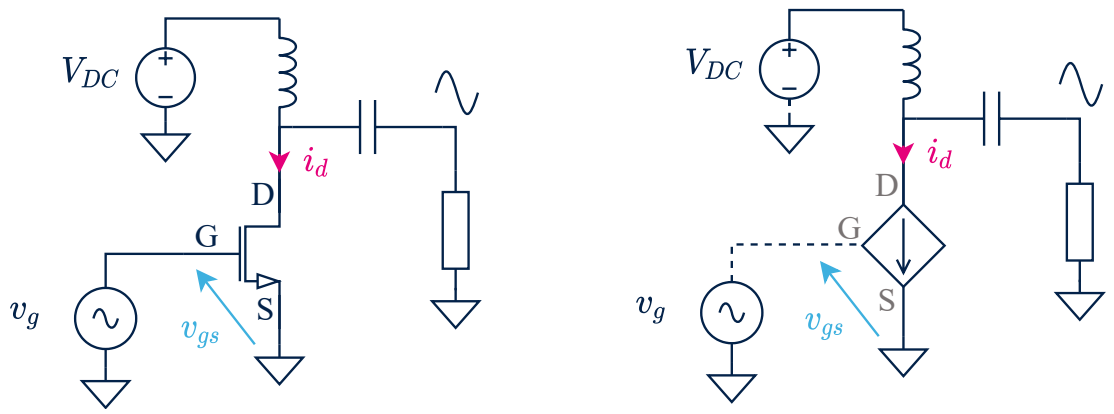


Figure 1.15. Common source amplifier (a) and Amplifier equivalent voltage-controlled current source (b)

$$A_v = -g_m R_L \tag{1.24}$$

Then, if a loaded, the dynamic amplified signal is sent to the output load. However, an actual MOSFET, an essential brick of the amplifier, is far from an ideal VCCS; the current i_d is impacted by diverse physical limitations and imperfections, causing both terms of equations 1.24 to be non-linear. In addition, the transistor geometry is actually the source of many intrinsic parasitic non-linear capacitors (shown in figure 1.16) that cause a non-linear behavior of the drain current.

Those are due to multiple parallel conductive surfaces within the front-end-of-line (FEOL) that are inevitable or due to the process precision. To build a coherent model of the transistor with its imperfections, the simple VCCS must be associated with a (non-exhaustive) list of parasitic capacitances (figure 1.17) that change with the metal-oxide-semiconductor (MOS) polarisation.

Some of those capacitances, such as C_{GS} and C_{GD} are first dependent on the overlap capacitance C_{ov} representing the gate surface overlapping the source and drain. Then, C_{GS} and C_{GD} also rely on channel charges intrinsic to the device operation (cutoff, active, and saturation mode), as shown in figure 1.18. Consequently, they vary in a non-linear way with the semi-conductor channel depletion zone, in which v_{gs} dynamically modulate the conductivity.

In bulk devices, the C_{DS} capacitance increases with the level of doping. As a comparison, an SOI FEOL shows around 20 % less parasitic capacitances [KLW01]. The maximum capacitance that exists between the junctions and the substrate buried oxide (BOX), is the capacitance

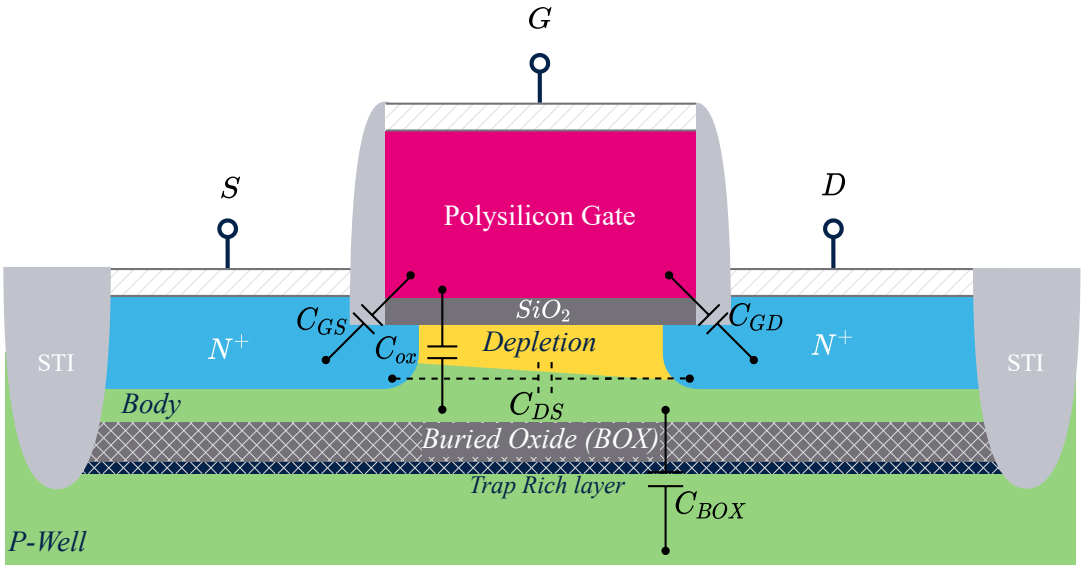


Figure 1.16. A sectional view of a floating body PD-SOI NMOS transistor and its main parasitic capacitances

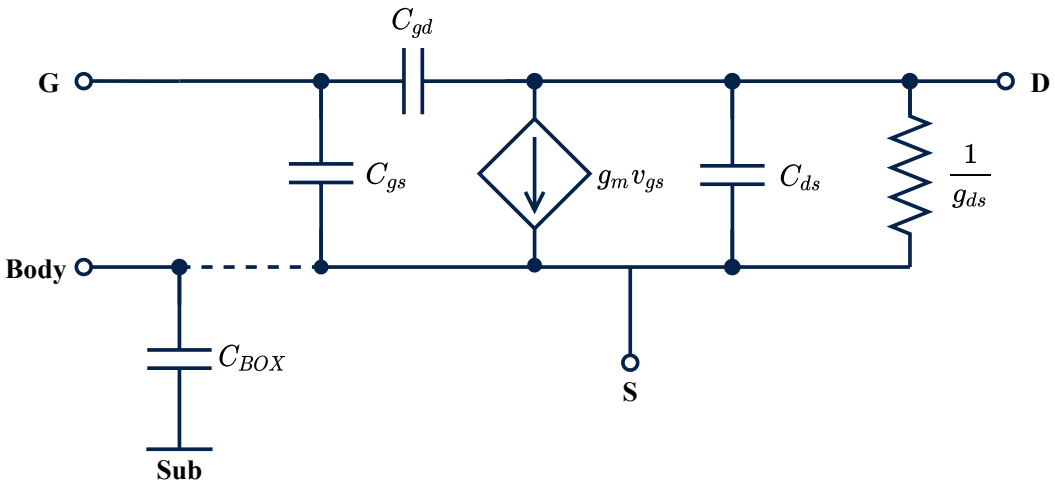


Figure 1.17. Schematic of a dynamic model of a PD-SOI NMOS transistor

C_{BOX} that is defined by the thickness of the BOX (a few hundred nanometers) in PD-SOI.

Furthermore, we note that a short-channel transistor of width W has a transconductance g_m (equation 1.25 in saturation region) dependent on the non-linear drain-source voltage V_{DS} (that undergo non-linear capacitive effects). g_m is also dependent on the oxide capacitance C_{ox} and of the channel modulation coefficient λ.

$$g_m = v_{sat} W C_{ox} \left[1 + \lambda (V_{DS} - V_{D,sat})\right] \tag{1.25}$$

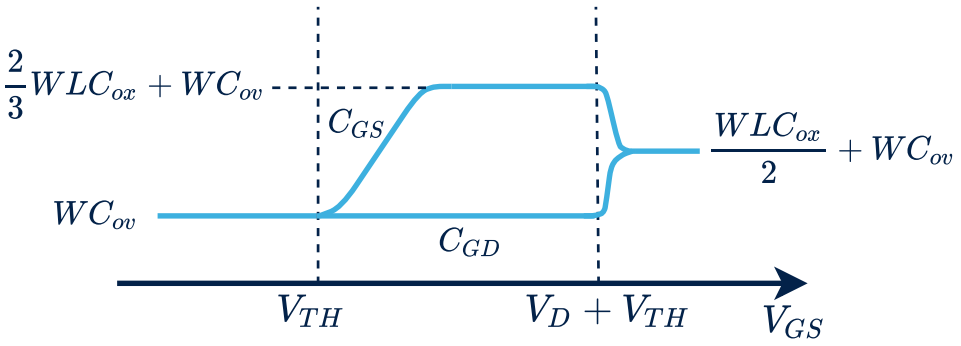


Figure 1.18

To characterize the capacitances and the transconductance nonlinearities, they are decomposed in $n-$order components giving more insight into their contribution to the frequency spectrum.

1.4.1.2 Large signal

The main signal distortion occurring in large signals is clipping. Clipping is a phenomenon that characterizes the waveform of the current i_{ds}.

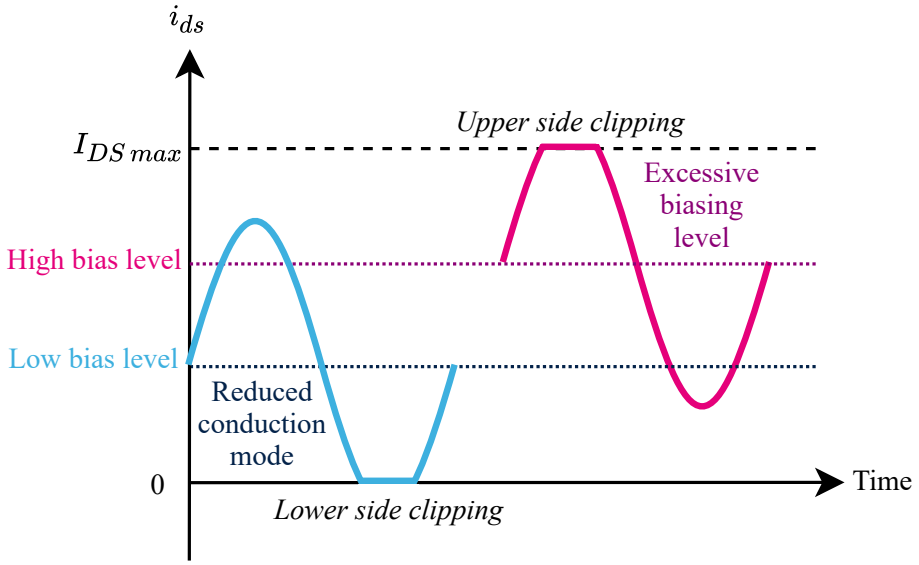


Figure 1.19. Lower and upper side clipping for reduced conduction mode and excessive bias level

A sinusoidal current submitted to clipping has its peaks truncated; it can occur as one-side of two side clipping due to two main reasons:

Lower side clipping is due to the biasing conditions: the amplifier work in reduced conduction mode. The designer purposely chooses this state. The DC biasing current I_{DS} selected by the gate voltage V_{GS} set the quiescent point for the dynamic current to swing around. Ideally, the drain to source current cannot be negative (figure 1.19); different biasing leads to other modes of conduction that cause the current waveform to be truncated. From the most linear class A to non-linear classes such as class C, the current waveform becomes more and more rectified, increasing the apparition of higher-order harmonic content as described in figure 1.20 from [Cri06]. A low biasing level is crucial for reducing DC current consumption and shaping current waveforms that provide higher efficiency despite higher harmonic content.

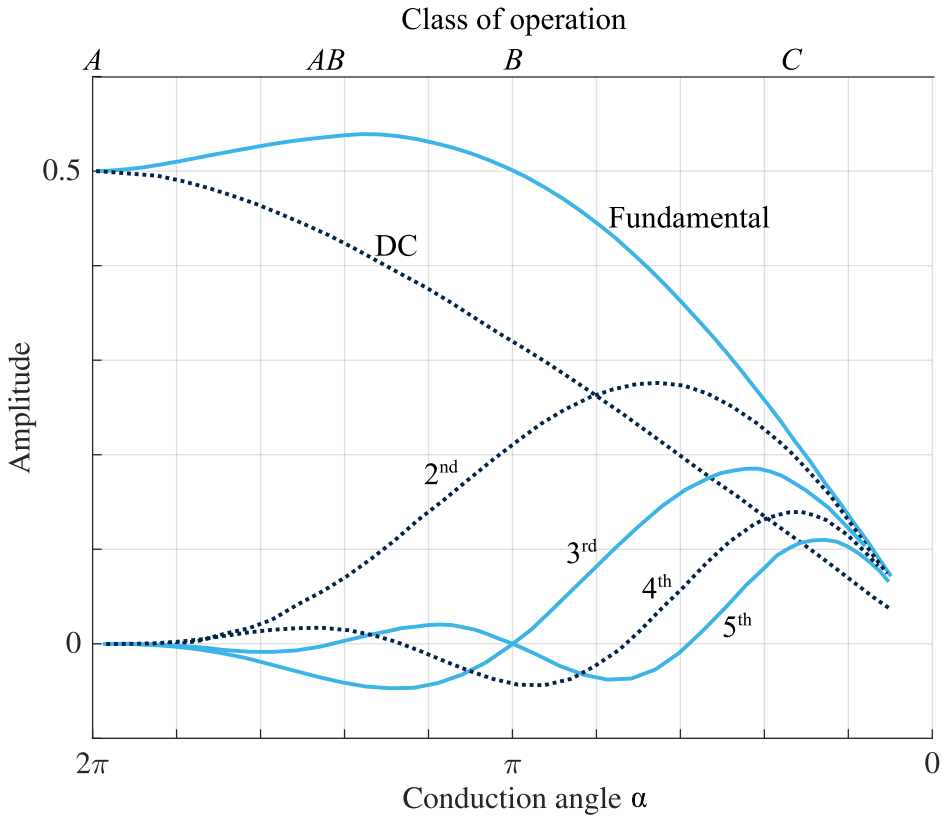


Figure 1.20. Current waveform harmonic content for reduced conduction angles

We can notice different points showing minimum harmonics. Class-B shows no odd harmonic content but high 2^{nd} and low 4^{th} harmonic content. However, a PA might see its polarization change during an input power sweep. Indeed, the shape of a rectified wave at a high power presents a higher mean voltage causing the biasing voltage to increase with the power delivered. Consequently, a PA designed to operate in a specific class will work through

multiple classes across time, especially when dealing with signals with a high PAPR, such as in OFDMA.

Most of the time, the upper side clipping is due to the saturation of the transistor. As the channel is limited, it cannot conduct an infinite current, and saturation occurs. As a result, the top of the waveform becomes truncated, a synonym of high harmonic content. As shown by [Ber19] the one-side clipping is mainly responsible for creating even harmonics. In contrast, two-sided clipping is richer in odd-order harmonics, a class A can eventually show two side clipping for very-large signals.

Globally, the PA designer task will reside mainly in handling the harmonics generated. The designer has many techniques to decrease their impact on efficiency and signal integrity: Differential architecture presents the attractive characteristic to naturally cancel the even harmonics that are an important component, as seen in figure 1.20. However, special care of the output load can short circuit the undesirable harmonics with distributed or lumped (LC) terminations [Par+16].

1.4.2 Harmonics and distortions

As seen in the previous section, many parameters are responsible for the apparition of harmonics. We usually fed its input with a sinusoidal (as possible) signal to quantify how linear a device is. Next, we observe the output signal's frequency spectrum to quantify the contribution of each harmonic. Then an empirical model of the non-linear device is built based on the Volterra kernels. With the PA's bandwidth limitation or different techniques, the generated harmonics can easily be filtered out as they are far from the fundamental frequency. However, with complex, multi-tone signals, distortions are generated, and some of them appear in the operating bandwidth. No direct filtering is possible; then, the designer must work on the linearity of its device or implement additional circuitry such as pre-distorders [Yu+19] to meet the specifications. The introduction of PAA with multi-beam brings up new studies because the power amplifier has to deal with tones with a different phase shift to form an extra beam to fulfill the 5G specifications.

1.4.2.1 Distortions and phase shifting

To understand the behavior and the issues created by these distortions in PAA, let's consider an arbitrary power amplifier feeding a single antenna among a N-array. Its input is a complex signal composed of two tones $f_1 = \omega_1/2\pi$ and $f_2 = \omega_2/2\pi$ separated by a Δf around a central frequency f_0. The two tones are considered with the same amplitude A_n and Δf small

enough so that we can approximate that the tones are synchronous in phase:

$$v_{in} = A_n \cos(\omega_1 t) + A_n \cos(\omega_2 t) \tag{1.26}$$

The power amplifier is seen as a non-linear "black box", the output signal can be decomposed as a sum of n-order components with n kernels k_n of v_{in}.

$$v_{out} = k_1 v_{in} + k_2 v_{in}^2 + k_3 v_{in}^3 + ... \tag{1.27}$$

These Volterra kernels can be derived through different methods based on simulation and measurements [Gha19]. By replacing (1.26) in (1.27), v_{out} can be derived with the cosine exponent simplified [Gao17] (here, the development is limited to the third-order intermodulation distortion difference for convenience) to make the different components appear. In equation 1.28 we find our original signal composed of f_1 and f_2, its 2nd and 3rd harmonics, and 2nd and 3rd order distortions. The output voltage spectrum (figure 1.21) is then composed of the low-frequency component with the 2nd difference. The high frequency components are composed of the 2nd and 3rd harmonics with the sum of 2nd and 3rd order distortions.

$$\begin{aligned}
v_{out} = & \frac{1}{2} k_2 {A_n}^2 && \text{(DC)} \\
& + \left(k_1 A_n + \frac{9}{4} k_3 {A_n}^3 \right) cos(\omega_1 t) && \text{(Tone 1)} \\
& + \left(k_1 A_n + \frac{9}{4} k_3 {A_n}^3 \right) cos(\omega_2 t) && \text{(Tone 2)} \\
& + \frac{1}{2} k_2 {A_n}^2 \cos(2\omega_1) && \text{(2nd order harm.)} \\
& + \frac{1}{2} k_2 {A_n}^2 \cos(2\omega_2) && \text{(2nd order harm.)} \\
& + \frac{1}{2} k_2 {A_n}^2 \cos(\omega_1 + \omega_2) && \text{(2nd order distortion sum)} \\
& + \frac{1}{2} k_2 {A_n}^2 \cos(\omega_1 - \omega_2) && \text{(2nd order distortion diff.)} \\
& + \frac{1}{4} k_3 {A_n}^3 \cos(3\omega_1) && \text{(3rd order harm.)} \\
& + \frac{1}{4} k_3 {A_n}^3 \cos(3\omega_2) && \text{(3rd order harm.)} \\
& + \frac{3}{4} k_3 {A_n}^3 \cos(2\omega_1 - \omega_2) && \text{(3rd order distortion diff. } \alpha) \\
& + \frac{3}{4} k_3 {A_n}^3 \cos(2\omega_2 - \omega_1) && \text{(3rd order distortion diff. } \beta) \\
& + ...
\end{aligned} \tag{1.28}$$

At last, in the bandwidth of the device (and the antenna), we have the desired signal tones, and on both sides, at a distance of Δf, the third-order intermodulation distortions difference that we will call α and β.

It is interesting to note that the 3rd order kernel k_3 affects the fundamental tones. At the same time, while $k_1 A_n$ gives the gain of the amplifier on ω_1 and ω_1; the sign of k_3 can reduce the amplitude of the fundamental and provide gain compression (for $k_3 < 0$). On the other hand, it can increase the fundamentals creating gain expansion (for $k_3 > 0$). In general, most of the devices are compressive, meaning that fundamental power is lost to the harmonics.

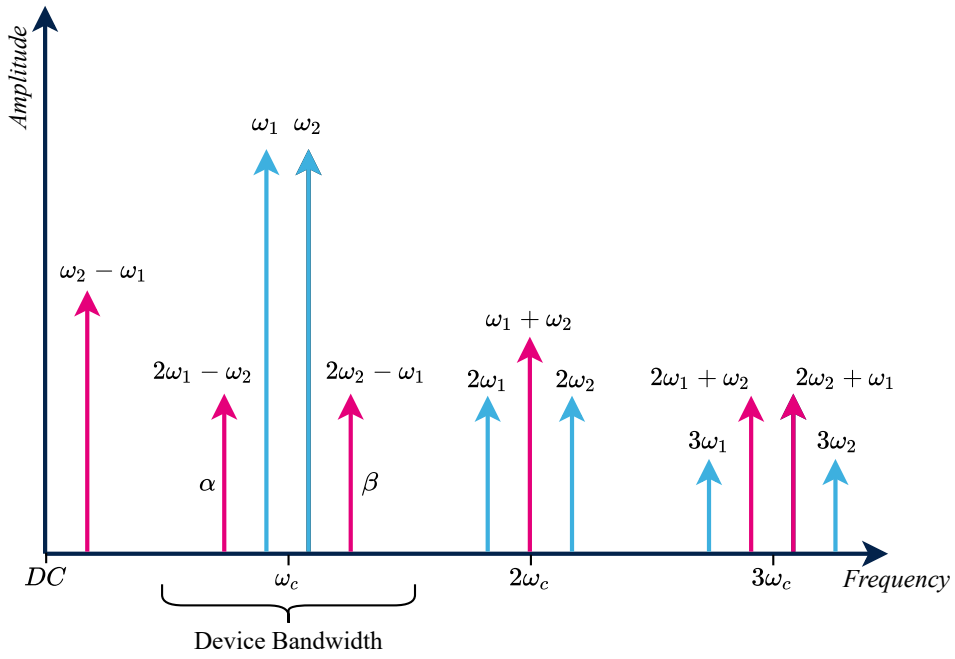


Figure 1.21. Output harmonics and distortions spectrum

For now, the tones ω_1 and ω_2 are considered phase synchronized (neglecting the phase shift due to Δf). When going through a phased antenna array they will be submitted to the same phase shift $\Delta\varphi$ (from equation 1.11) by the beamformer to be finally radiated together in a single beam with all of the harmonics and distortion content. Beamforming excluded, this scenario is very similar to the case of a single antenna transmission of the previous technologies (with much less directivity to communicate with the maximum of users in space), where the base station emits all the terms of equation (1.28) with the same gain through a unique antenna. The consequence is that the users communicating with the station receive all the non-linear content of v_{out}. Hence, the receiver must keep a sufficient signal-to-noise ratio (SNR).

The generation of multiple narrow beams used in 5G PAA needs new variables of phase control to put in equation (1.28).

Let's now consider two UEs UE_1 and UE_2, on a 2-dimensions plane equidistant to r from the base station. UE_1 and UE_2 are located within an aperture angle of 60° at two different positions that form respectively two angles θ_{f1} and θ_{f2} about to the perpendicular axis of the antenna. Each user receives respectively a sub-carrier of data at frequencies f_1 and f_2. The first tone beam f_1 is oriented toward UE_1 with the excitation phase shift $\Delta\varphi$ and the beam of f_2 toward UE_2 with $\Delta\psi$. Our input signal from (1.26) now becomes (1.29)

$$v_{in} = A_n \cos(\omega_1 t - n\Delta\varphi) + A_n \cos(\omega_2 t - n\Delta\psi) \tag{1.29}$$

Each tone is phase-shifted differently so that two distinct beams are generated and controlled via $\Delta\varphi$ and $\Delta\psi$. Each PA being considered identical on the whole array generates substantial non-linear components especially close to their OCP1. As we have introduced two variables of phase shift on the tones, each distortion depending on ω_1 and ω_2 at the same time is itself also dependent on $\Delta\varphi$ and $\Delta\psi$.

The troublesome in-band IMD3 difference component is now defined by:

$$\begin{aligned} IMD3_{diff\alpha} &= \frac{3}{4} k_3 {A_n}^3 \cos\left[(2\omega_1 - \omega_2)t - n(2\Delta\varphi - \Delta\psi)\right] \\ IMD3_{diff\beta} &= \frac{3}{4} k_3 {A_n}^3 \cos\left[(2\omega_2 - \omega_1)t - n(2\Delta\psi - \Delta\varphi)\right] \end{aligned} \tag{1.30}$$

The beamforming relying on the control of a signal in phase, by tuning $\Delta\varphi$ and $\Delta\psi$ to steer θ_1 and θ_2, will result in the creation of two dissociated beams $\theta_{IMD3_{diff\alpha}}$ and $\theta_{IMD3_{diff\beta}}$ of IMD3 with complex patterns entangled with the tones that need to be studied.

1.4.2.2 IMD3 radiation mechanism

Some elements of analysis have been given by [Hem02] and commented by [Loy03] to derive the radiation behavior of the nonlinearities. However, this result is very complex to derive due to the different effects of compression. For instance, the PA amplitude modulation to phase modulation (AM/PM) distortion affect the phase of IMD3s, and the intermodulations manifest themselves, creating additional beams. Despite that, we can approximate the behavior of the main IMD3 with a study of the beams that they generate. Then we will be able to get an idea of their direction of radiation. Let's describe two different cases of BF with our 2 users UE_1 and UE_2, still communicating in line-of-sight with the base station:

1.4.2.2.1 Case 1

In this first scenario, UE_2 is kept on standby, UE_1 communicates with the BS and needs a substantial bit-rate, the two frequencies available f_1 and f_2 are allocated to him (figure 1.22a). The PAA forms a single beam to steer the two tones. However, UE_1 receiver has to deal with an issue of signal integrity: Very similarly to what happened with single antenna propagation, the nonlinearities, including the 2nd and 3rd harmonics generated by the non-linear characteristics of the PAs are sent toward the UE in communication (equation 1.31). The receiver targeted then suffers from a reduced SNR due to the IMD3.

$$\theta_{IMD3_{diff}\alpha,\beta} = \theta_{f1} = \theta_{f2} = \theta_{2f1} = \theta_{2f2} \tag{1.31}$$

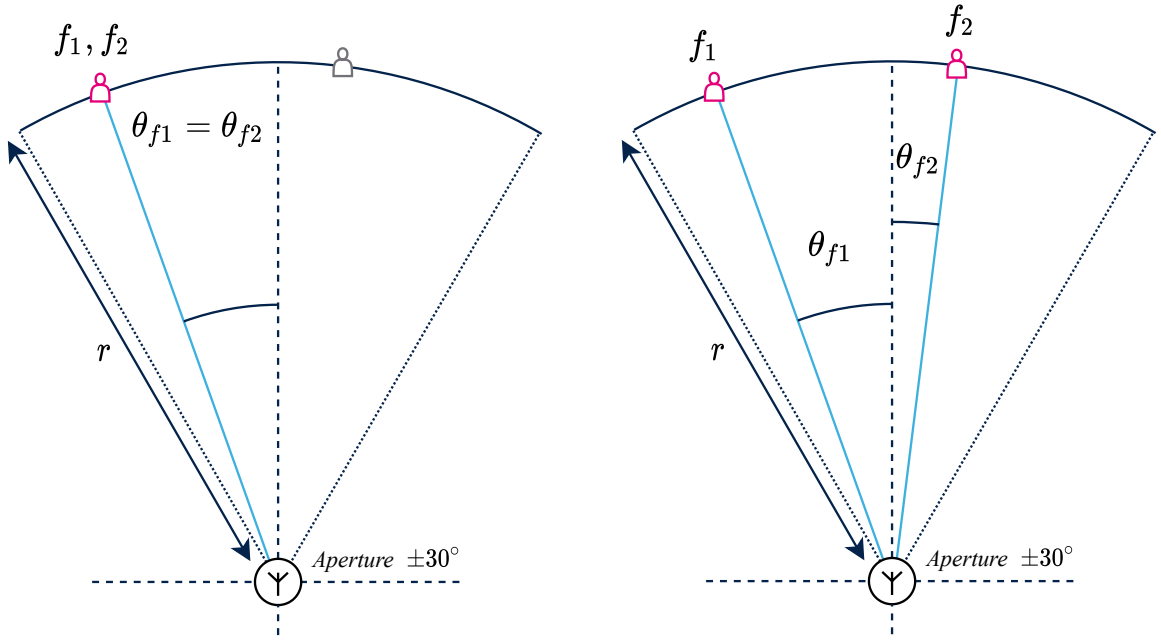


Figure 1.22. Single and two-user communication with single (a) and dual (b) beam

1.4.2.2.2 Case 2

Now UE_2, still located in a different place than UE_1, wakes up, the bandwidth available is split between the two users. To do so, we attribute the frequencies f_1 and f_2 separately to UE_1 and UE_2. Two beams are formed to target their distinct location (figure 1.22b). In this

case, the IMD3 beams α and β follows particular patterns, described respectively in equation (1.32) and (1.33). They depend on the angle that defines the location of both users :

$$\theta_{IMD3,\alpha} = \arcsin\left\{\frac{2f_1 \sin\theta_1 - 2f_2 \sin\theta_2}{2f_1 - f_2}\right\} \tag{1.32}$$

$$\theta_{IMD3,\beta} = \arcsin\left\{\frac{2f_2 \sin\theta_2 - 2f_1 \sin\theta_1}{2f_2 - f_1}\right\} \tag{1.33}$$

In reality, most of the 5G PAA propose 2D beam-steering; hence UE_1 and UE_2 can be located respectively on street level and at the 3rd-floor balcony. The tones beam angles are represented in a (u, v) plane with $u = \sin\theta\cos\phi$ and $v = \sin\theta\sin\varphi$. The beams of f_1 and f_2 have their corresponding vectors (u_1, v_1) and (u_2, v_2). Results the α and β IMD3 beam, formed in this new 2D plane:

$$\theta_{IMD3,\alpha}(u, v) = \left(\frac{2f_1u_1 - f_2u_2}{2f_1 - f_2}, \frac{2f_1v_1 - f_2v_2}{2f_1 - f_2}\right) \tag{1.34}$$

$$\theta_{IMD3,\beta}(u, v) = \left(\frac{2f_2u_2 - f_1u_1}{2f_2 - f_1}, \frac{2f_2v_2 - f_1v_1}{2f_2 - f_1}\right) \tag{1.35}$$

As the intermodulations depend on the two tones, their beams usually point in other directions than the main beams. This aspect most of the time benefits the signal integrity of users 1 and 2. However, the IMD3 creates signal pollution in directions that are not targeted by the array. As a result, we can expect random users to receive UE_1 and UE_2 IMD3 while not even communicating with the base station.

1.5 Conclusion

This chapter has shown some technical solutions that the scientific community has provided for phased antenna arrays and power amplifiers to answer the millimeter-wave 5G needs. Despite the significant challenges that represent the realization of measurable systems, many insightful experiments have been performed.

First, the PA design is currently very active in using active load modulation to enhance efficiency and, at the same time, desperately focused on being robust to the parasitic ALM existing within PAA. However, we can ask ourselves about the relevancy of this precipitation toward complex designs, which often sacrifice efficiency for VSWR immunity, without proposing an extensive analysis of the load variation effect. This work needs to be conducted to fully

understand the essential system parameters that influence the load variation and quantify its impact on PAs to justify using ALM-proof power amplifiers.

On the other hand, concerns on issues have been raised that we're never faced in mobile communications before 5G: The nonlinearities, intrinsic to the use of PAs, are a critical factor degrading the signal integrity. They face a different treatment than in previous communications standards; once radiated by the beam forming techniques, they tend to have a complex propagation scheme with many dependencies with IC behavior and the multi- beam parameters. This behavior may become annoying because it can affect many users in the beam steering range.

References

[All10] P. E. (Phillip E.) Allen. *CMOS analog circuit design / Phillip E. Allen, Douglas R. Holberg*. eng. International 2nd ed. Oxford series in electrical and computer engineering. New York: Oxford University Press, 2010. ISBN: 9780195392463.

[Ano21] Anokiwave. *Anokiwave 5G Portfolio*. 2021. URL: https://www.anokiwave.com/5g/index.html (visited on 08/31/2021).

[Bal05] Constantine A Balanis. *Antenna theory: analysis and design*. Wiley-Interscience, 2005.

[Ber19] Subhash Chandra Bera. *Microwave Active Devices and Circuits for Communication*. 1st ed. Vol. 1. 1 vols. 1876-1119. Springer, Singapore, 2019. 700 pp. ISBN: 978-981-13-3004-9.

[Bhu13] Sunandan Bhunia. "Microstrip Patch Antenna's Limitation and Some Remedies". In: *International Journal of Electronics & Communication Technology* 4 (Spl - 1 Jan. 2013), p. 3. ISSN: 2230-7109.

[Che+21] Chun-Nien Chen et al. "38-GHz Phased Array Transmitter and Receiver Based on Scalable Phased Array Modules With Endfire Antenna Arrays for 5G MMW Data Links". In: *IEEE Transactions on Microwave Theory and Techniques* 69.1 (Jan. 2021), pp. 980–999. ISSN: 1557-9670. DOI: 10.1109/TMTT.2020.3035091.

[Chi35] H. Chireix. "High Power Outphasing Modulation". In: *Proceedings of the Institute of Radio Engineers* 23.11 (Nov. 1935), pp. 1370–1392. ISSN: 2162-6626. DOI: 10.1109/JRPROC.1935.227299.

[Cri06] Steve Cripps. *RF Power Amplifiers for Wireless Communications, Second Edition*. Artech House, 2006. 469 pp. ISBN: 978-1-596-93018-6.

[Des+18] Thibaut Despoisse et al. "A Comparison of Beamforming Schemes for 5G mm-Wave Small Cell Transmitters". In: *2018 16th IEEE International New Circuits and Systems Conference (NEWCAS)*. 2018, pp. 6–9. DOI: 10.1109/NEWCAS.2018.8585467.

[Doh36] W.H. Doherty. "A New High Efficiency Power Amplifier for Modulated Waves". In: *Proceedings of the Institute of Radio Engineers* 24.9 (1936), pp. 1163–1182. DOI: 10.1109/JRPROC.1936.228468.

[Dun+18] Jeremy Dunworth et al. "28GHz Phased Array Transceiver in 28nm Bulk CMOS for 5G Prototype User Equipment and Base Stations". In: *2018 IEEE/MTT-S International Microwave Symposium - IMS*. June 2018, pp. 1330–1333. DOI: 10.1109/MWSYM.2018.8439517.

[EEK65] K. Eisele, R. Engelbrecht, and K. Kurokawa. "Balanced transistor amplifiers for precise wideband microwave applications". In: *1965 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*. Vol. VIII. Feb. 1965, pp. 18–19. DOI: 10.1109/ISSCC.1965.1157595.

[Gao17] Wei Gao. *Energy and Bandwidth-Efficient Wireless Transmission*. Springer International Publishing. 2017. 499 pp. ISBN: 978-3-319-44222-8.

[Gar+01] R. Garg et al. *Microstrip Antenna Design Handbook*. Antennas and Propagation Library. Artech House, 2001. ISBN: 9780890065136.

[Gha19] Siamack Ghadimi. "Determining Volterra Kernels for Nonlinear RF Power Amplifiers". In: *Microwave&RF* (Sept. 2019). URL: https://www.mwrf.com/technologies/components/article/21849932/determining-volterra-kernels-for-nonlinear-rf-power-amplifiers.

[HDX17] Yiming Huo, Xiaodai Dong, and Wei Xu. "5G Cellular User Equipment: From Theory to Practical Hardware Design". In: *IEEE Access* 5 (2017), pp. 13992–14010. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2727550.

[Hem02] C. Hemmi. "Pattern characteristics of harmonic and intermodulation products in broadband active transmit arrays". In: *IEEE Transactions on Antennas and Propagation* 50.6 (June 2002), pp. 858–865. ISSN: 1558-2221. DOI: 10.1109/TAP.2002.1017668.

[HWW19] Song Hu, Fei Wang, and Hua Wang. "A 28-/37-/39-GHz Linear Doherty Power Amplifier in Silicon for 5G Applications". In: *IEEE Journal of Solid-State Circuits* 54.6 (June 2019), pp. 1586–1599. ISSN: 1558-173X. DOI: 10.1109/JSSC.2019.2902307.

[KLW01] James B Kuo, Shih-Chia Lin, and John Wiley. *Low- Voltage so1 CMOS VLSI Devices and Circuits*. John Wiley & Sons. 2001. ISBN: 0-471-22156-2.

[LFY07] Shenghai Liu, Suili Feng, and Wu Ye. "A Combined Switched Beam-Forming and Adaptive Beam-Forming Algorithm for OFDM Systems with Antenna Array". In: *2007 International Conference on Wireless Communications, Networking and Mobile Computing*. Sept. 2007, pp. 89–92. DOI: 10.1109/WICOM.2007.29.

[Loy03] S. Loyka. "Comments on "Pattern characteristics of harmonic and intermodulation products in broad-band active transmit arrays"". In: *IEEE Transactions on Antennas and Propagation* 51.7 (July 2003), pp. 1683–. ISSN: 1558-2221. DOI: 10.1109/TAP.2003.813644.

[Ltd18] Fujitsu Laboratories Ltd. *Fujitsu Develops World's First Single-Panel Antenna to Simultaneously Support Multiple 5G Communications*. 2018. URL: https://www.fujitsu.com/global/about/resources/news/press-releases/2018/1129-01.html (visited on 08/31/2021).

[MKK17] Boris Moret, Vincent Knopik, and Eric Kerherve. "A 28GHz self-contained power amplifier for 5G applications in 28nm FD-SOI CMOS". In: *2017 IEEE 8th Latin American Symposium on Circuits Systems (LASCAS)*. Feb. 2017, pp. 1–4. DOI: 10.1109/LASCAS.2017.7948059.

[Nak+18] Keigo Nakatani et al. "A Ka-Band High Efficiency Doherty Power Amplifier MMIC using GaN-HEMT for 5G Application". In: *2018 IEEE MTT-S International Microwave Workshop Series on 5G Hardware and System Technologies (IMWS-5G)*. Aug. 2018, pp. 1–3. DOI: 10.1109/IMWS-5G.2018.8484612.

[Par+16] Byungjoon Park et al. "Highly linear CMOS power amplifier for mm-wave applications". In: *2016 IEEE MTT-S International Microwave Symposium (IMS)*. 2016, pp. 1–3. DOI: 10.1109/MWSYM.2016.7540024.

[Poz11] David M. Pozar. *Microwave Engineering, 4th Edition*. Wiley, Nov. 2011. 752 pp. ISBN: 978-0-470-63155-3. URL: https://www.wiley.com/en-ar/Microwave+Engineering%5C%2C+4th+Edition-p-9780470631553.

[Raz17] Behzad. Razavi. *Design of analog CMOS integrated circuits / Behzad Razavi*. eng. Second edition. New York, NY: McGraw-Hill Education, 2017. ISBN: 9780077496128.

[Sai96] Robert A. Sainati. *CAD of Microstrip Antennas for Wireless Applications*. USA: Artech House, Inc., 1996. ISBN: 0890065624.

[Vis+17] Kuttathati Srinivasan Vishvaksenan et al. "Mutual Coupling Reduction in Microstrip Patch Antenna Arrays Using Parallel Coupled-Line Resonators". In: *IEEE Antennas and Wireless Propagation Letters* 16 (2017), pp. 2146–2149. ISSN: 1548-5757. DOI: 10.1109/LAWP.2017.2700521.

[Yu+19] Chao Yu et al. "Full-Angle Digital Predistortion of 5G Millimeter-Wave Massive MIMO Transmitters". In: *IEEE Transactions on Microwave Theory and Techniques* 67.7 (July 2019), pp. 2847–2860. ISSN: 1557-9670. DOI: 10.1109/TMTT.2019.2918450.

[Zon+20] Zhiwei Zong et al. "A 28GHz Voltage-Combined Doherty Power Amplifier with a Compact Transformer-based Output Combiner in 22nm FD-SOI". In: *2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. Aug. 2020, pp. 299–302. DOI: 10.1109/RFIC49505.2020.9218280.

Chapter 2

Phased arrays load variation analysis

This chapter proposes an analysis of the parasitic active load variation in phased arrays, it is focused on its dependence on the system characteristics and on its consequences for power amplifiers

2.1 Array design and electromagnetic simulations

2.1.1 Phased array antenna radiation characterization

A full-wave analysis is commonly used to fully extract the performances of a radiating element without relying on any assumptions. This type of simulation aims solve the Maxwell equations with the currents on the antenna structure as unknowns. Then, looking at the fields produced by these currents to satisfy the boundary and excitation conditions, we get an integral equation given for the currents. The analysis is done with numerical methods, such as the moment method or difference equations methods. The latter converts the Maxwell equations into difference equations that are solved in the time domain. This method is called the finite time-domain method (FDTD). On the other hand, another method called finite element method (FEM) consists in solving the Maxwell equations in the form of vector wave equations.

Nowadays, 3D electromagnetic (EM) software has become indispensable for antenna design. The most popular of them include all of these full-wave solving methods in their solutions and enable a good simulation of thick and complex substrates, and of course, at the cost of non-negligible computation time and, let's face it, poor physical insight.

When dealing with AA, the computation time is a non-negligible issue. Having many elements to analyze requires a large structure with multiple edges that need significant mesh refinement. Secondly, multiple port excitation has a significant impact on simulation time. Actually, to optimize array analysis, particular methods (infinite and finite element array method) are proposed by software such as Ansys HFSS. It uses the symmetry axes

and the repeatable characteristics of the array structure patterns to save considerable time. Unfortunately, this type of simulation doesn't allow co-simulation with active devices and requires additional licenses. The full-wave analysis is then chosen to have a maximum insight of the ALM effect with co-simulation. The size of the arrays simulated has been limited to 25 radiating elements to keep computation time around tens of hours.

The upcoming analyses have been performed with the Keysight ADS suite with modules EMPro and RFPro using the solver FEMs. HFSS has also been partially used for verification and quick finite array simulations.

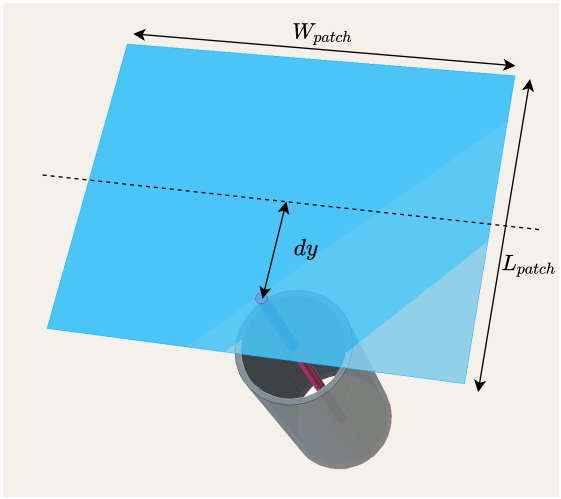
2.1.2 Single antenna

2.1.2.1 Dimensions

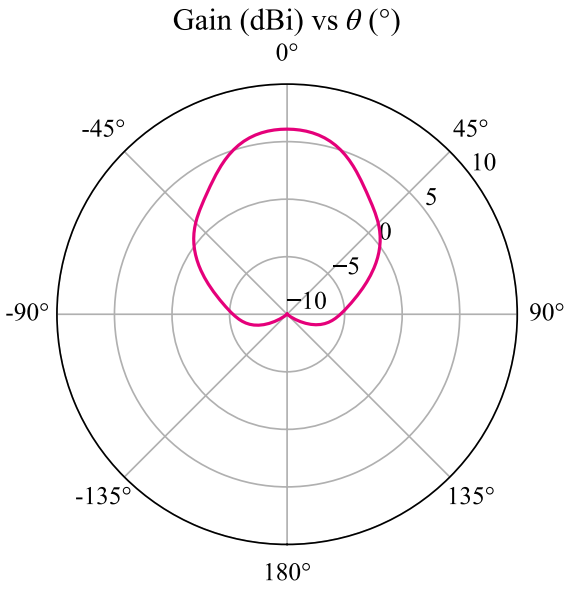
The following designs have been simulated on a substrate that Rogers has recommended for 5G mmW applications. This substrate is the RO4730G3, with a dissipation factor $\tan\delta = 0.0029$, a thickness of $h = 1.034$ mm that is important for high bandwidth, and a permittivity of $\epsilon_r = 2.98$. The antennas designed are copper rectangular patch antennas with a thickness of $t = 35$ µm; the sizing is performed using equations from chapter 1 (1.7) (1.10). The patch is fed by a coaxial probe, with a copper core and shield and a polytetrafluoroethylene (PTFE) insulator of permittivity $\epsilon_{r\,PTFE} = 2.1$. The probe is off-centered on the patch along $-y$ axis by 0.72 mm. The patch width is $W_{patch} = 2.9$ mm and the length is $L_{patch} = 2.25$ mm to resonate at 28 GHz (figure 2.1a).

2.1.2.2 Simulation

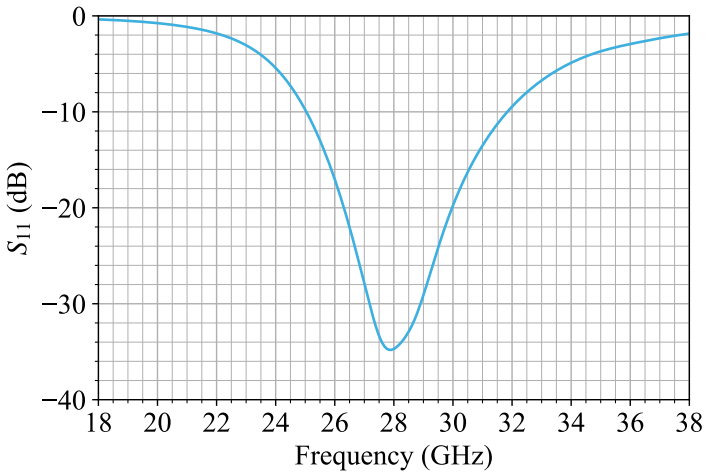
The simulation of the RPA, and the following structures, are performed with a linear sweep between 27.5 and 28.5 GHz. A step of 100 MHz and an adaptive sweep between 100 MHz and 86 GHz are chosen. The choice of the higher frequency defines the frequency for which the mesh is computed, its value is crucial as it must be set where the simulated device does not present any resonance to help the mesh refiner to converge and save some computing time. The solver used can multiply by a factor of 10 to 20 on the simulation time if the mesh frequency is not properly chosen. The delta of S-parameters convergence criteria is selected so that $\Delta S \leq 0.02$ and two consecutive converged steps are required to stop the refinement. The excitation is done with lumped ports with the $(+)$ terminal on the core of the feeding probe, and the $(-)$ terminal on the shield.



(a)



(b)



(c)

Figure 2.1. Designed RPA (a), realized gain in dB pattern (b) according to θ, and matching (c)

The simulation of the designed RPA with the same parameters of simulation shows a realized gain (insertion loss comprised) of 6 dB with a 3 dB aperture of 60° (figure 2.1b).

2.1.3 2D Antenna Arrays

2.1.3.1 Dimensions

From the single patch antenna is derived two antenna arrays of different sizes to study the impact of multiple rows on the antenna active impedance. Two arrays of 9 and $25-$elements are designed; each patch antenna is separated vertically and horizontally by $d = \lambda/2$ relatively to the center of the radiating patch. The maximum beam steering angle is then set from the equation 1.12 to:

$$\theta_{max}\left(d = \frac{\lambda}{2}\right)\arcsin\left(\frac{2\lambda}{\lambda} - 1\right) = \frac{\pi}{2} \tag{2.1}$$

The dimensions of the RPAs have to be adjusted due to the influence of the other elements in the array. L_{patch}, W_{patch} and dy are now respectively 2.33 mm, 2.8 mm and 0.975 mm

The final array design show these dimensions: the $9-$element array is composed of a 3×3 patch structure on a substrate surface of 20.75×20.97 mm with an effective surface occupied by the RPAs of 13.5×13.03 mm. On the other hand, the 25 element array is composed of a grid of 5×5 patches representing an effective surface of 24.42×23.73 mm on a substrate of 31.45×31.67 mm.

2.1.3.2 Simulation results

Each array is simulated with the same setup that in section 2.1.2.2. First, the S-parameters results show that the return loss S_{mm} of each patch takes different values according to the various environment that each patch sees. It is between -21 and -38 dB.

By looking at the radiation diagram, we see that the smaller array achieves a gain of 14.7 dB with a 34° 3 dB aperture angle. The larger array proposes a higher gain of 19.1 dB with a higher directivity hence 20° of 3 dB aperture angle. They show respectively a PSR of 17.5 dB and 14.3 dB. Besides these antenna properties that many techniques could improve, we now have the base elements to converge to this study's main interest, which is the analysis of the impedance that each RPA represents to the PAs feeding the elements.

2.1.4 Phase recombination analysis

To begin the study of the impedance effect, an analysis of the signal recombinations within the arrays is proposed. Using two arrays enable us to derive a comparison to understand how the number of elements and their location influence the load variation. To better find our way

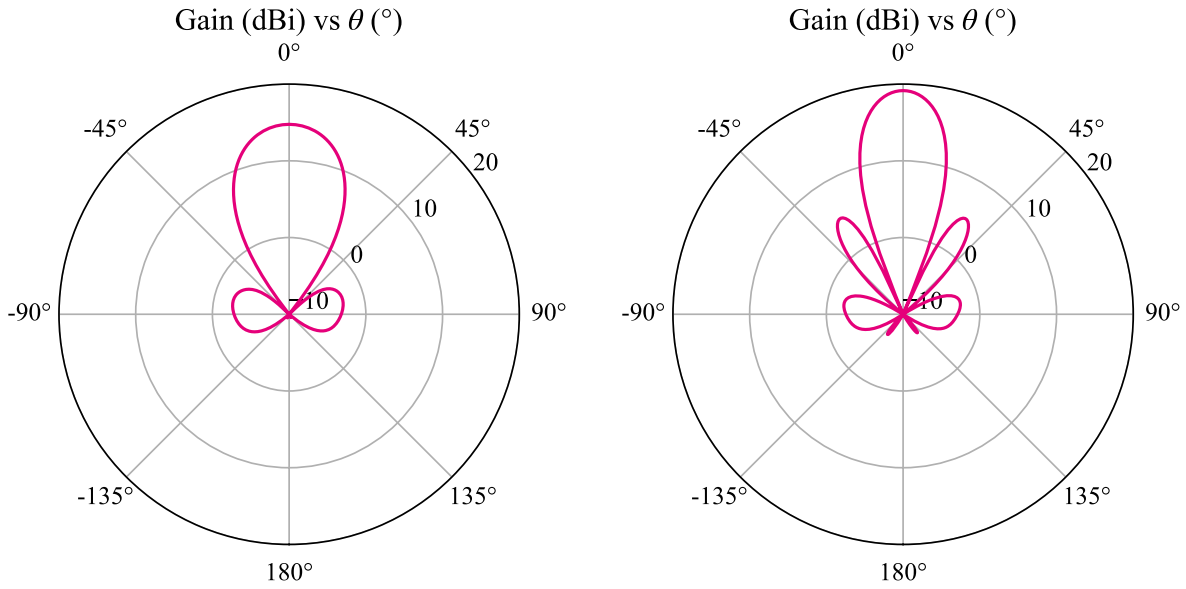


Figure 2.2. Designed 9 (a) and 25 (b) element arrays gain in dB pattern according to θ

within the 2D arrays, they are considered a matrix of antennas with coordinates coded (i, j) with i the x-axis coordinate and j the y-axis coordinate.

2.1.4.1 Phase recombination mechanism

Before digging into the array simulation with complex sources, the coupling mechanism can be modeled with simple elements shown in figure 2.3 to understand its behavior: Let's take an $N-$array fed by ideal and identical, phased-controlled, sinusoidal sources. An element m within the array is considered 'victim' and is fed by a source m and controlled by a phase shift ϕ_m. In a perfect array, this source sees directly the antenna's impedance and, if perfectly matched, no returning waves. In reality, the incident signal delivered by the source goes through an RF path m that is represented by an attenuation $\alpha_{path\,m}$ and a delay/phase $\phi_{path\,m}$. Some assumptions are made: We assume that the source is perfectly matched to its antenna; hence the return loss is null $V_{RL} = 0$. We also consider, in the first place at least, that the $N-1$ aggressor sources are 'off'. Hence no signal traveling from the aggressors to the victim. The victim see no return waves, so $v_{back} = 0$. In these conditions, the source m experiences no disturbances, and the power transfer is complete from source to load.

From now on, the aggressors are no longer passives. The array, coupling the element between themselves, induces an incoming wave $v_{back} \neq 0$ composed of the sum of the

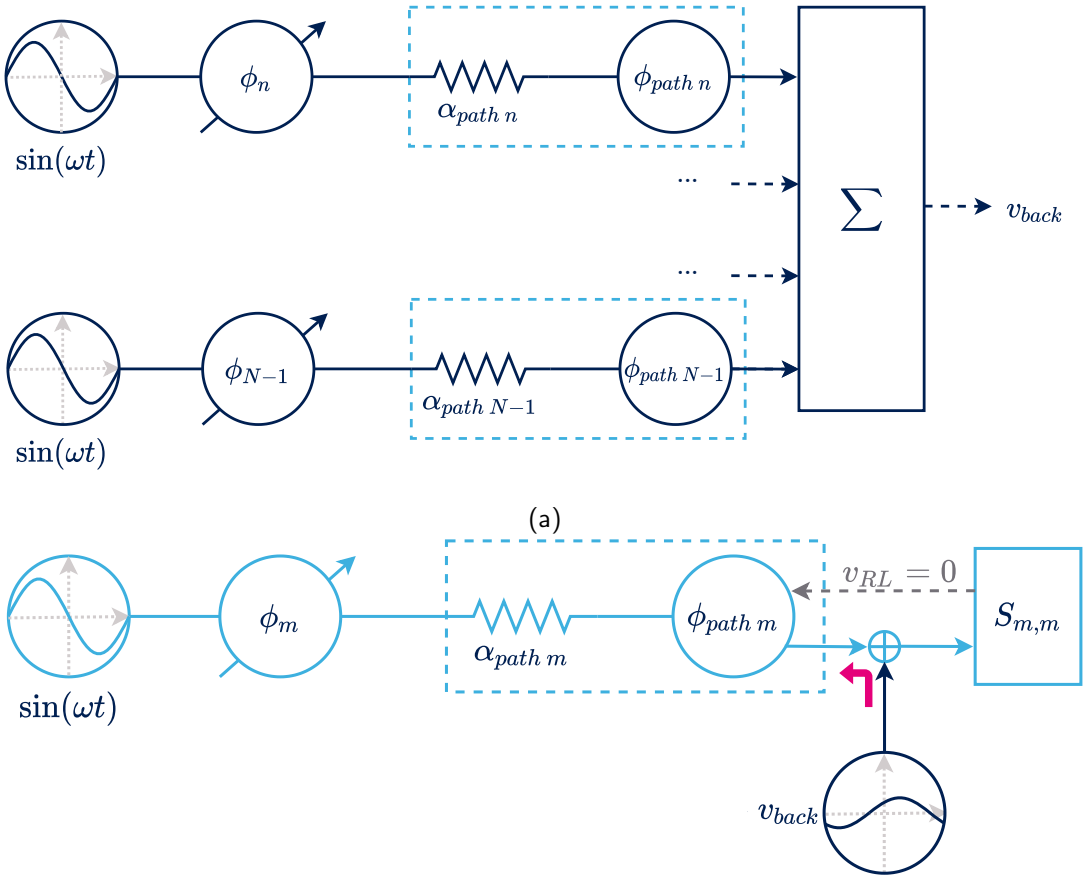


Figure 2.3. Signal recombination through different phase shifts and attenuation between $N-1$ aggressors (a) and on victim elements (b)

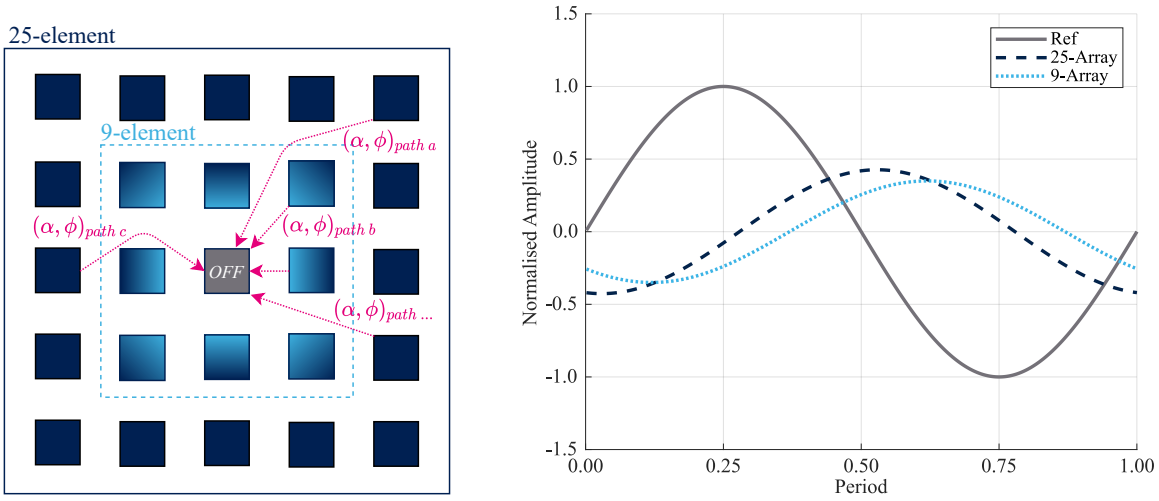
contributions of each aggressor and their RF path S_{nm} defined by α_n and ϕ_n for module and phase. It can be expressed as:

$$v_{back}(t) = \sum_{n=1}^{N-1} \alpha_{path\,n} \sin\left(\omega t + \phi_n + \phi_{path\,n}\right) \tag{2.2}$$

v_{back} is then composed of the controlled phase shifts ϕ_n, of the delay $\phi_{path\,n}$ and the attenuation $a_{path\,n}$ that is basically $|S_{mn}|$

The array simulations performed allow us to illustrate the shape in amplitude and phase of the signal v_{back}: In both arrays, the central element (at coordinates $(2,2)$ in the 9-element array and $(3,3)$ in the 25 element array) is passive while all others are actives, representing the victim vs. aggressor scenario. First, the aggressors are fed with a 1 V peak amplitude synchronous signal ($\phi_n = 0$). The only phase shift that composes the coupled signal is then

the sum of all $\phi_{path\,n}$. The shape of v_{back} is observed in both cases at the output of the victim source m:



(a) (b)

Figure 2.4. 9 and 25-element arrays fed by a unitary peak amplitude synchronized signal (a) and the different signals coupled to the central passive element (b)

Based on the S-parameters obtained at 28 GHz, figure 2.4b shows the reflected waves v_{back} compared to a reference incident signal feeding the central element of each array. Both 9 and 25-element arrays coupled signals present a slight difference of amplitude of 0.08 V but especially a significant phase shift relative to the reference ($\Delta\varphi_{25-elts}$ = 100° and $\Delta\varphi_{9-elts}$ = 133°) that is different according to the array size. Typically if the element observed is set 'on' and delivers the reference signal, the coupled signals will add themselves to the reference signal and change its amplitude that is translated in a certain amount of VSWR. As the $9-$element array coupled signal is practically in phase opposition to the reference signal, the reference will undergo a higher VSWR than the $25-$element array.

As noted, the reflected waves present a similar amplitude between the 9 and 25-element cases, this phenomenon highlights that the second row of antennas that exists only in the 25-element array presents a coupled signal to the central element that does not impact the A-VSWR drastically.

Unlike we could have anticipated, having a higher number of neighbors in an antenna array is not necessarily a synonym of higher ALM. The multiple phase recombinations determine and dictate the amount of A-VSWR endured.

Within a population of radiating elements, the central antenna benefits from a symmetrical environment that fosters destructive phase recombinations of the coupled signals. On the

other hand, the elements on the edges of the array suffer from a strong asymmetry of their environment. Notwithstanding the ALM effect that is our main concern in this analysis, it can be noted that the passive mismatch suffered from the surrounding antennas is usually treated with dummy elements placed all around the array [Hol13] and passively contribute to reducing the A-VSWR.

Among the antenna array feeding techniques, windowing or weighting [DOE16] is commonly implemented to improve the antenna's radiation pattern: the function assigns different weights to each array antenna that control the amplitude of their feeding source. As the elements on the edges are major contributors to the generation of side lobes, especially when scanning wide beam angles [Bal05], the weighting function, most of the time, reduces the excitation amplitude of the elements in the periphery of the array.

2.2 Characterization of interconnection losses

The study of the mutual coupling effect has shown that it depends on three critical parameters: The feeding phase ϕ_n, the RF path attenuation α_{path}, and delay ϕ_{path}. To refine the previous results, PA-to-antenna interconnections must be taken into account in the RF path S_{mn}. The amount of losses separating the radiating elements and their source of power is dictated by the type of front-end architecture and the nature of the physical connections.

2.2.1 Origin of interconnection losses

The architecture of the system defines the mutual coupling path. In consequence, it may be very various according to the technology choices. Cheaper substrate, conductors, and chip to PCB interconnections represent non-negligible losses that play an essential role in the evaluation of the ALM. Working at frequencies above 60 GHz, the small wavelength size allows the designers to implement single on-silicon and on-packaging antennas, limiting the length of interconnections between their active circuit that would be a source of parasitics. At frequencies around 30 GHz, the wavelength is still large enough to coexist integrated circuits and PCB antenna. The chain of interconnections presented on the path of the signal coming from the output of the PA and going to the antenna's input shows more important losses and parasitic effects. Consequently, the power reaching the antenna is dissipated and reduced by a few dB, and these losses must be considered while designing the PA, or they may lead the designer to go back to the drawing board to ensure that the specified EIRP of the system is respected. On the other hand, as these losses increase, the PA gets more and more isolated

from the antenna, and this isolation benefits to the reduction of the active VSWR presented by the antenna:

Considering an incident wave a_1 coming from the source s, going toward the interconnection element that is represented by an attenuation coefficient α and finally feeding the load Z_L:

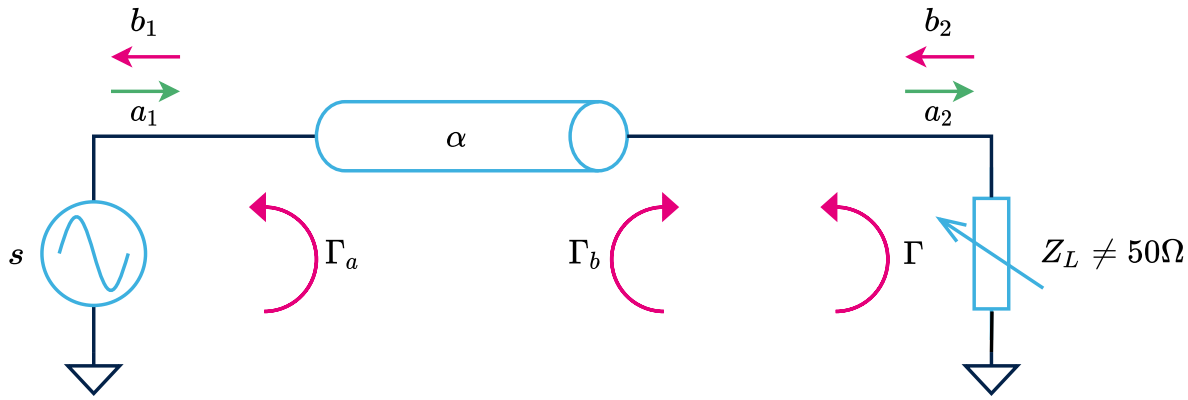


Figure 2.5. Reflection interconnection path under load mismatch

The wave a_2 received at the load is given by:

$$a_2 = \alpha.a_1 - \Gamma_a.a_1 \tag{2.3}$$

In case of mismatch at the load, the reflected wave b_2 is given by:

$$b_2 = \Gamma.a_2 \tag{2.4}$$

The reflection is submitted again to the losses and come back as b_1 to the source:

$$b_1 = \alpha.b_2 - \Gamma_b b_2 \tag{2.5}$$

The reflected wave at the source is then: (considering mismatch at the interconnection)

$$b_1 = \alpha.\Gamma(\alpha.a_1 - \Gamma_a.a_1) - \Gamma_b(\Gamma(\alpha.a_1 - \Gamma_a.a_1)) \tag{2.6}$$

Considering $\Gamma \gg \Gamma_a$ and $\Gamma \gg \Gamma_b$:

$$b_1 \approx \alpha^2.\Gamma.a_1 \tag{2.7}$$

$$B_1 \approx A_1 + 20\log(\alpha) + 10\log(\Gamma) \tag{2.8}$$

with $B_1 = \log(b_1)$ and $A_1 = \log(a_1)$

To summarize, the reflected wave coming back to the output of the PA responsible for the A-VSWR is a portion of the main signal given by the amplifier affected twice by the losses of the interconnections. Therefore, it is crucial to have a good idea of the elements that affect the signal between the antenna and the PA to estimate the A-VSWR as it can be significantly reduced in some front-end configurations.

2.2.2 Front End topologies

As seen the section 1.1.2.2, the classic and most widespread 5G front-end topology is based on switched Tx and Rx paths [Li+21] on a single probe-fed patch antenna. The connections between the integrated circuit and the PCB are composed of a flip-chip ball grid array (BGA) die and vias through the PCB. It can be estimated that at 28 GHz, the TRx switch [LH18] [LCK19] represents itself around 1.5 dB of insertion loss. Working at mmW frequencies, the connections between the chip and the PCB cannot be realized with wire bounding as it is too lossy. Flip-chip BGA connections ensure a good RF path that is estimated by [UC01] and [LZZ16] to represent only 0.2 dB. The remaining losses concerning IC and PCB connections and antenna feeding have been simulated with HFSS and estimated to an average of 0.8 dB.

In the switched topology (figure 2.6a) we find a variable gain amplifier (VGA) amplifier followed by the PA on the Tx path. Assuming that the switch is integrated on the same chip as the PA/LNA, we find the loss of the TRx switch, then followed by interconnections due to the IC to PCB transition, and finally the PCB structures with the transmission lines and diverse vias to feed the antenna. Giving a total amount of losses of 2.5 dB,

Another type of front-end, less widespread, nevertheless very interesting for its performance, is called 'common-leg' [KR19] (figure 2.6b). The transmission and receiving path are, this time, switched before the amplifiers. The PA/LNA paths are connected to the antenna with cross-polarization techniques; for instance, reception is performed in horizontal polarization and reception in vertical polarization. Cross polarization naturally shows isolation between the two channels. Before the two amplifiers, the common-leg is composed of the $T_{x\,IN}$ and the $R_{x\,OUT}$ coming to a dual switched branch with the phase shifter and the VGA. With no switch between the antenna and the circuit, the common-leg topology benefits from lower output losses, giving a total loss of 1 dB. However, this type of front end multiplies by two the amount of IC-to-antenna connections in terms of integration.

On the one hand, the losses decrease the system's efficiency. Still, concerning ALM, the phenomenon will be less visible to the amplifiers in higher loss systems as the interconnections isolate the PA from the antenna. This is why it is imperative to take into account the

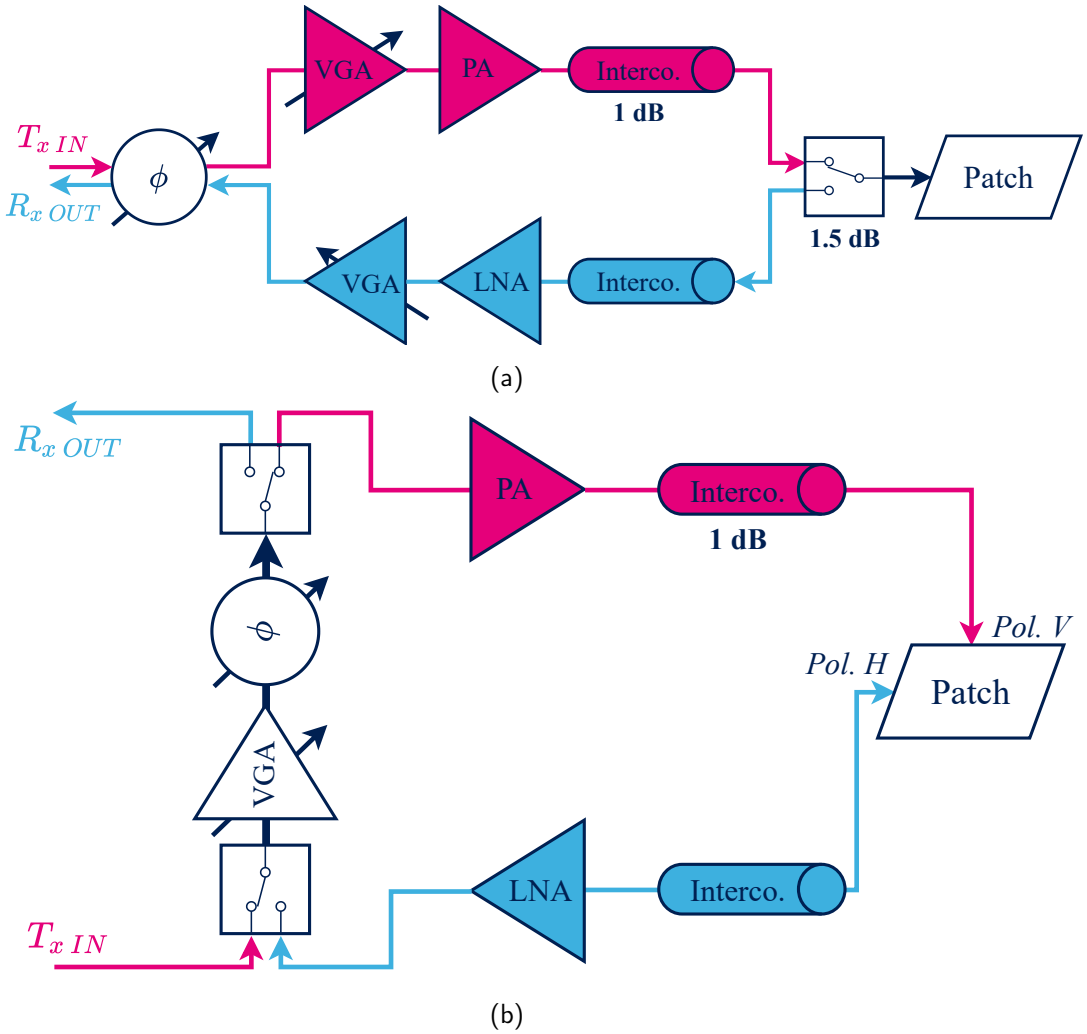


Figure 2.6. Front-end classic switched topology (a) and common-leg topology (b)

environment of the amplifier and the technology choices made to fully evaluate the amount of variation brought back by the mutual coupling. In the design process, this is not always easy for the IC designers to precisely know what will be plugged behind the device. In the best case, they are provided a specific VSWR immunity but not much; however, the next section will show that from simple designs, it is possible to derive a precise evaluation of the behavior of A-VSWR.

2.3 Beam steering load variation effect

The following analysis shows the results of the ALM produced by the BF. In this regard, co-simulation is required to simulate the 3D EM array model with the Golden Gate simulation of the ST CMOS065SOIMMW PA. The first part of this analysis is performed only using extracted S-parameters from the array. The principle of co-simulation required for EM radiation results consists of several steps of simulation that will be detailed later. In each case, the simulation bench for large arrays may be pretty tedious to set up. There are several variables per element to describe in equations and as many probes to place to monitor P_{in}, P_{out}, and P_{dc}.

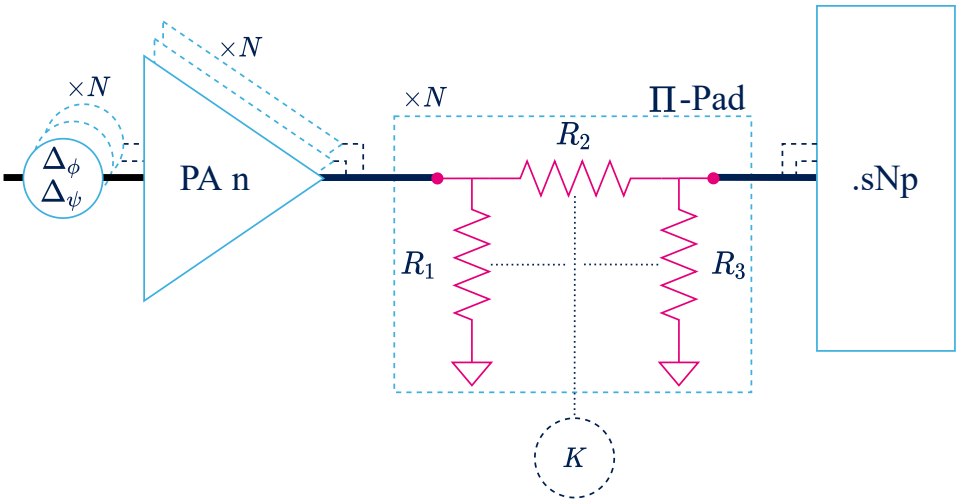


Figure 2.7. Simulation schematic of N PA with Π-pad attenuators loaded by the array S-parameters file

To take into account the interconnection losses, a Π-pad attenuator composed of three resistors R_1, R_2, and R_3 is inserted at the output of each amplifier and controlled via the attenuation factor K. The resistors are defined by:

$$R_1, R_3 = Z\left(\frac{K+1}{K-1}\right) \tag{2.9}$$

$$R_2 = Z\left(\frac{K^2}{2K}\right) \tag{2.10}$$

Within this case, the PA and antenna are matched to 50 Ω, $Z = Z_S = Z_L = 50\Omega$.
The amount of losses that each FE topology represent give:

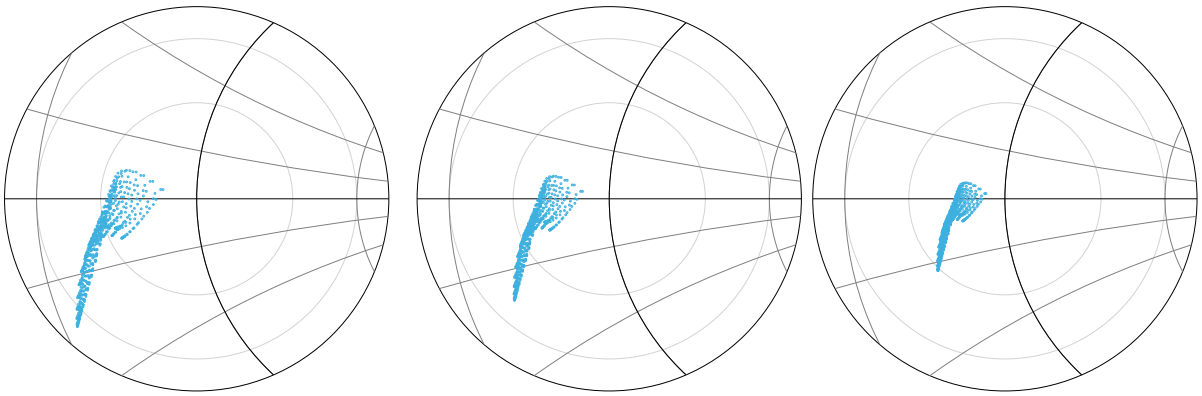
$$K_{\alpha=1dB} \approx 1.122 \tag{2.11}$$

$$K_{\alpha=2.5dB} \approx 1.334 \tag{2.12}$$

Before continuing, it is important to keep in mind that the performances shown in this section are intrinsic to the PAs, hence taken directly on the input and output of the device. The computed output power and efficiency do not consider the losses introduced.

2.3.1 Single element active VSWR

A study of an antenna among the array is proposed. The active impedance of this element is monitored while parameters such as the 2D steering angle and the interconnection losses are swept.



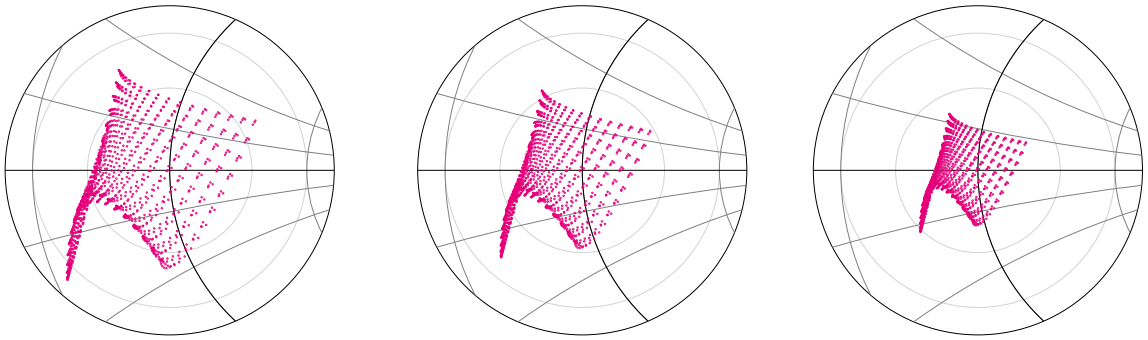
(a) (b) (c)

Figure 2.8. Load variation seen by the PA of the element (3,3) across a ± 60 dB 2D beam scan with $\alpha = 0$ dB (a), $\alpha = 1$ dB (b) and $\alpha = 2.5$ dB (c) in smith chart ($\rho_{max} = 0.4$) with 1.5:1 and 2.0:1 VSWR circles

To have a visible hint of the impedance variation, the active impedance is plotted in the smith chart. Figure 2.8 shows this load variation pattern of the central element of the 25-array $(3, 3)$ for different amounts of losses with a maximum beam steering angle of 60°. Each point observable in the chart corresponds to an iteration of a beam scan sweep with a step of 2°. The A-VSWR obtained is visible thanks to the VSWR circles for 1.5:1 and 2:1 traced. The first example (figure 2.8a) constitutes a worst-case example that is not a realistic case but interesting to use as a point of comparison. In this case, the load evolves within an A-VSWR range comprised between 1.2:1 and 2.1:1, with an average of 1.6:1. The higher A-VSWR obtained is for a steering angle of $(\theta, \phi) = (0°, 0°)$, the beam steering causes the impedance

to shift closer to 50 Ω as the most extreme angles show lower A-VSWR. With additional losses, we can observe the pattern area significantly shrinks and gets closer to the center of the smith chart. For instance, with 1 dB of losses (figure 2.8b), the variation range remains within the VSWR circle 2.0:1 with a maximum VSWR of 1.81, a minimum of 1.11, and an average A-VSWR of 1.43.

Finally, for the classic front-end topology losses (visible in figure 2.8c) ($\alpha = 2.5$ dB) the load variation pattern stays confined within the 1.5:1 VSWR circle with an average A-VSWR of 1.29.



(a) (b) (c)

Figure 2.9. Load variation seen by the PA of the element (3,3) across a 90° beam scan with $\alpha = 0$ dB (a), $\alpha = 1$ dB (b) and $\alpha = 2.5$ dB (c) in smith chart ($\rho_{max} = 0.4$) with 1.5:1 and 2.0:1 VSWR circles

To push the analysis further, a simulation of BS is tested to $\theta_{max} = 90°$. This situation is not likely to happen in normal conditions due to the high sidelobe content, but it can represent a scenario of "defect" if, for some reason, the system drifts away from its normal behavior. An interesting behavior is visible in figure 2.9: The additional range of variation appears directly on the smith chart and covers a different area of impedance that is vastly enlarged. However, the new impedances reached are still confined around the center of the smith chart. Hence, under those circumstances, there are no significant changes in A-VSWR in the different cases: no matter the amount of losses, the peak A-VSWR reached remains identical to 60-degree beam steering. The additional impedance values due to $\pm 45\,°$ beam steering reach no higher A-VSWR than those from 60 °beam steering. As the variation pattern is "folded" around 50 Ω we even notice a small decrease of the average A-VSWR from 1.6 to 1.5:1 for $\alpha = 0$ dB and from 1.3 to 1.2 for $\alpha = 2.5$ dB. Also, a minimum A-VSWR of 1.0 is now achieved as the 90° pattern covers the center of the smith chart.

	90°Beam Scan			60°Beam Scan		
α	Min	Avg	Max	Min	Avg	Max
0 dB	1.0	1.5	2.1	1.1	1.6	2.1
1 dB	1.0	1.4	1.8	1.1	1.4	1.8
2.5 dB	1.0	1.2	1.5	1.1	1.3	1.5

Table 2.1. Summary of the A-VSWR achieved on the element $(3, 3)$ within a 25-element array

A summary of the results obtained is shown in table 2.1. It offers a low global increase of the minimum A-VSWR between 90 and 60° beam scan.

2.3.2 Full array active VSWR

After focusing on a single element of the array, we now look at the whole array. The load patterns are plotted for each patch of the 25-array, in figure 2.10. Following the same method as in the previous section, a beam scan of 60° is first performed, and $\alpha = 1dB$ of losses (FE worst case) are considered. The patterns come with various shapes that are more or less scattered within the smith chart.

A vertical symmetry of the array can be found in the pattern shapes. However, the asymmetrical antenna feeding is the cause of the slight horizontal asymmetry in the patterns (between rows 1 and 5, 2 and 4). This asymmetry manifests itself with a more important A-VSWR (noted AV on the figures for convenience) on the lower section of the array, probably due to the slight tilt of the radiating pattern caused by the feeding, resulting in more near field coupling. We can distinguish different types of patterns: the 9 antennas constituting the center of the array show load patterns that are pretty alike, having a relatively identical EM environment. With a bit of extrapolation, we can expect this pattern to repeat itself on most elements of a vast array. The edges are populated with more diverse patterns. The four corners are alike, and so for the element between them. A more critical spreading of impedance values is observed on the patterns of antennas on the periphery of the array. However, they do not necessarily show the highest A-VSWR. While each pattern is different, the impedance first tends to shift away from the center of the chart with the beam scan and then tends to orbit around it for large angle values. Some patterns also show good matching for $(\theta, \phi) \neq 0°$. During a 60° beam scan, the average VSWR simulated on the whole array is 1.76:1.

Figure 2.11 shows a significant spread of the load variation pattern. For most of the elements, the extra points of load variation orbit in a limited area close to the center of the

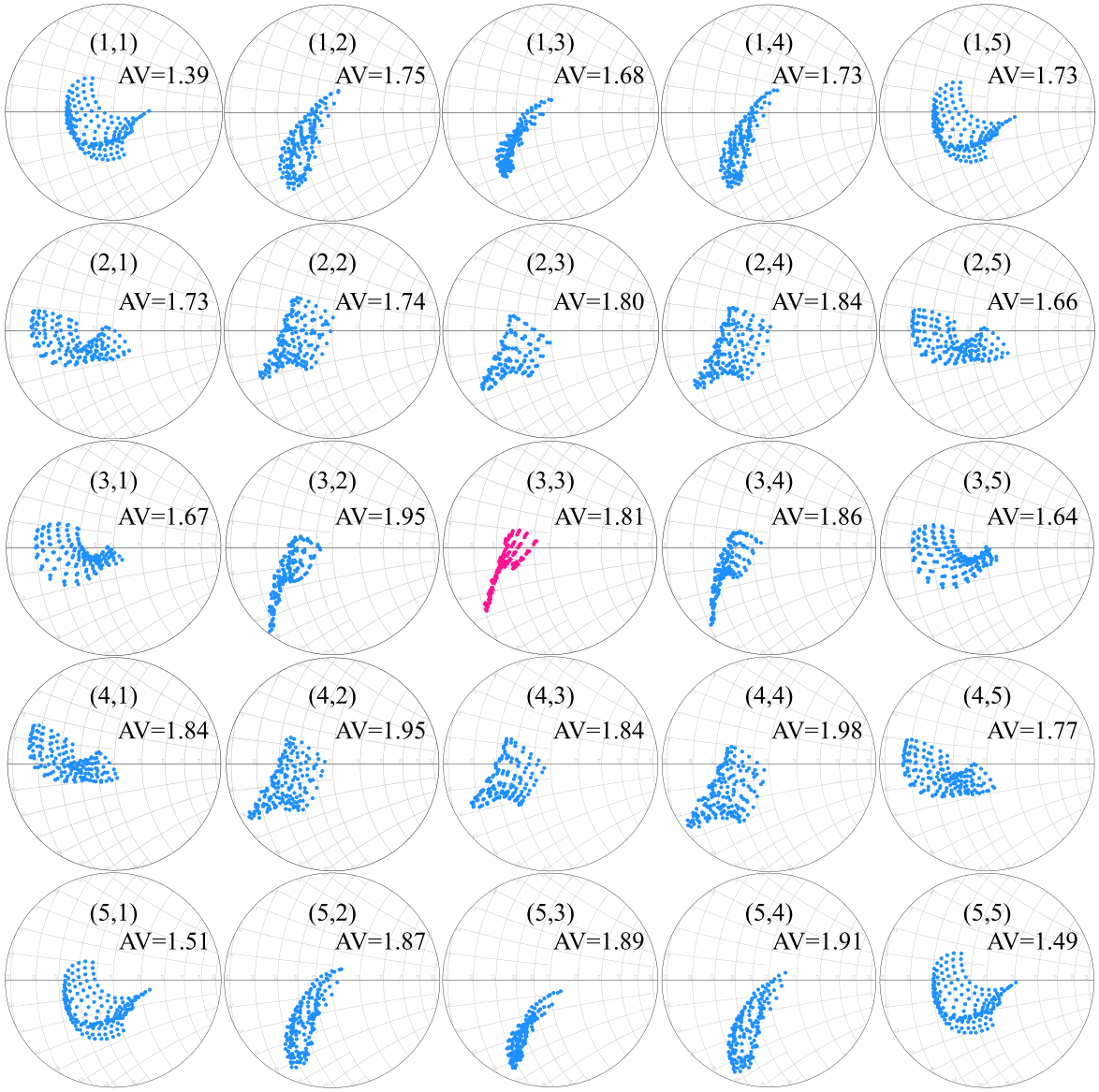


Figure 2.10. 25-element array load variation for a $\pm 30\,^{\circ}$ beam scan in smith chart with $\alpha = 1\,\text{dB}$ ($\rho_{max} = 0.35$) and the maximum A-VSWR reached noted AV

smith chart. Consequently, the increase of A-VSWR is again not significant and stays below 2:1.

These results show another interesting point: One of the characteristics of this load variation is that the patterns of load variations are intrinsic to the antenna array designed. Whatever the source is behind the array (as long as it does not show strong AM/PM distortion), it will see the exact impedance for each angle steered. In case of a substantial impact on the source of the ALM, we can imagine using Chireix type architecture that can track the different

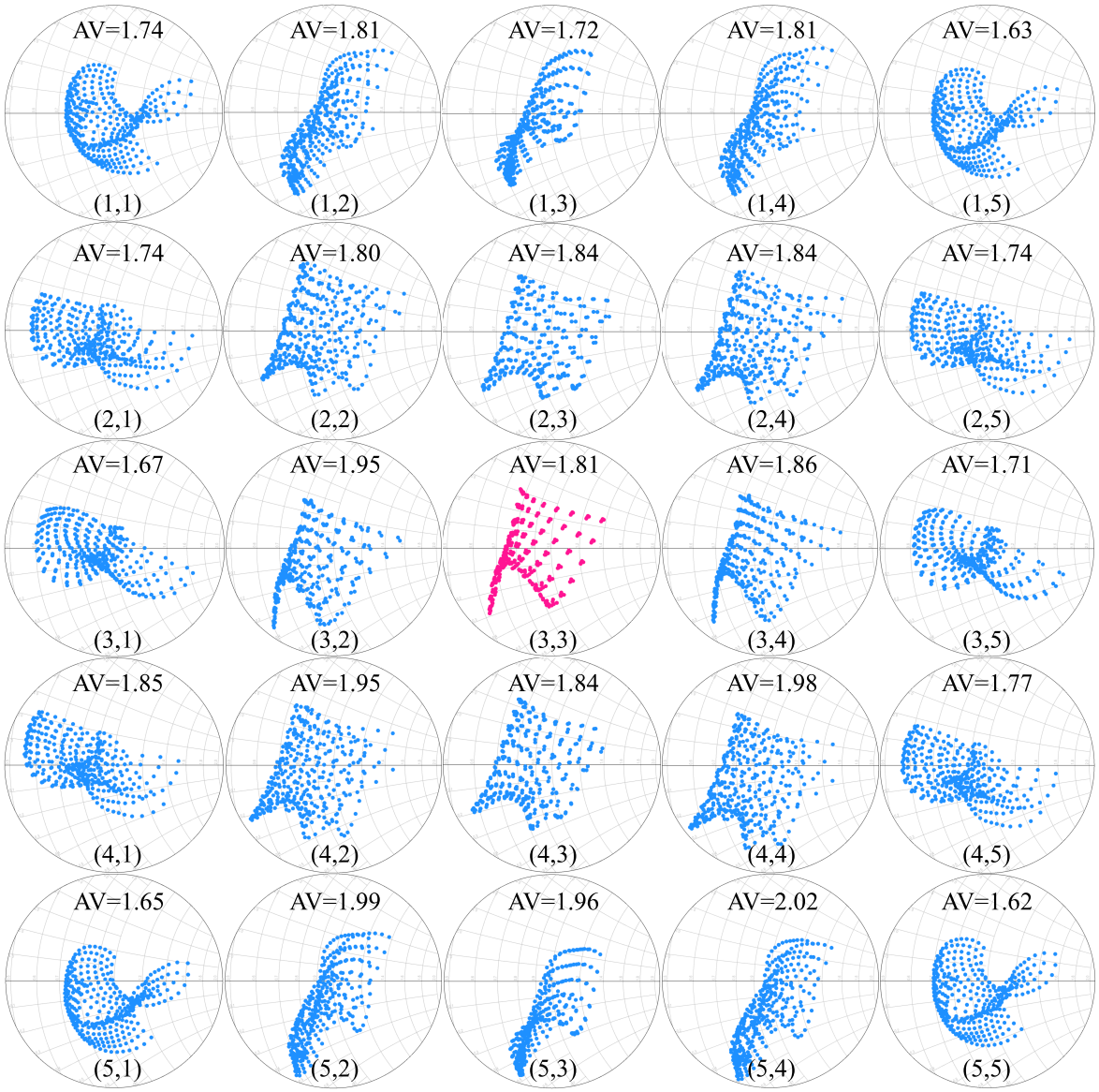


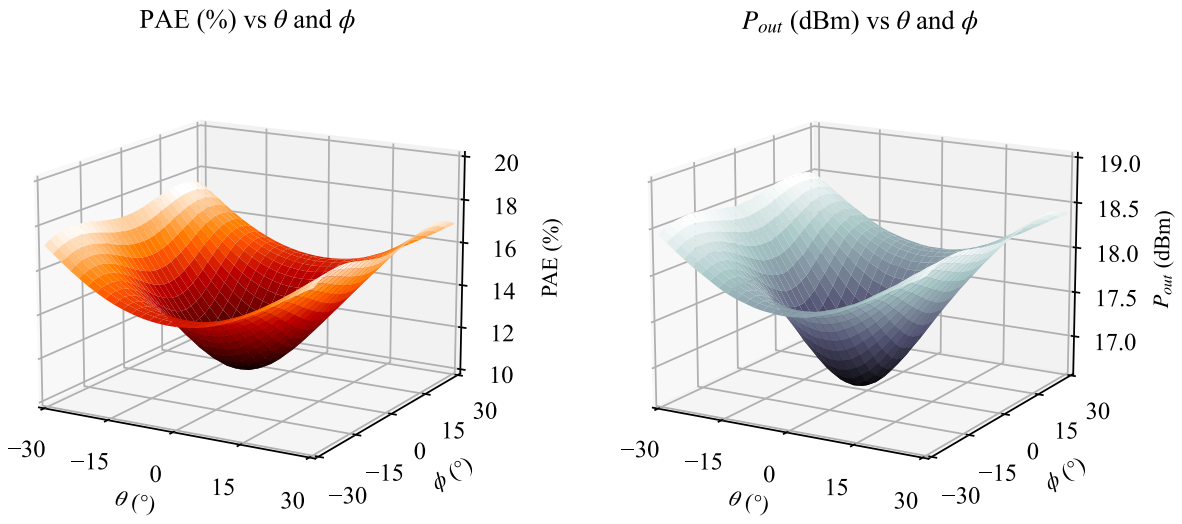
Figure 2.11. 25-element array load variation for a 90° beam scan in smith chart with $\alpha = 1\,\text{dB}$ ($\rho_{max} = 0.35$)

patterns of impedance, or eventually design PAs matched to a different impedance than $50\,\Omega$ by taking into account the ALM. Still, this impedance would be difficult to determined as it depends on the antenna topology.

2.3.2.1 Parasitic Active load modulation on power amplifiers

To derive the impact of the observed A-VSWR on PAs, a simple power amplifier architecture is designed in ST CMOS065SOIMMW technology with as specification an OCP1 of 18 dBm.

The topology chosen is a differential cascode amplifier to maximize gain and input/output (S_{12}) isolation. The transistor's current density is chosen for high transition frequency f_T. It is biased to exhibit good linearity, giving a gain of 14.8 dB, an OCP1 of 18.4 dBm, and a PAE of 17 %. The design is simulated with a full EM extraction of the top and the transistor core's post-layout simulation (PLS).



(a) (b)

Figure 2.12. PAE (a) and P_{out}(b) of the element $(2, 3)$ across a 60° 2D beam-scan with $\alpha = 1$ dB

A series of comparisons of PAE and P_{out} of elements that are the most and the least subject to high A-VSWR is proposed. The beam scan is again performed in 2D with 60°, and the two front-end configurations are compared with their respective losses. To provide visual data to the performances variations, 3D plots are shown with the 2D space of beam-steering represented on the (x, y) plane and the data of interest on the z-axis. The simulation's angle step is 1° on both axes allowing smooth shapes at the expanse of harmonic balance (HB) simulation time, which can be significant with advanced parasitic extraction of the PA design. From the simulation, we, without surprise, note that the two elements showing higher and lower A-VSWR respectively show lower and higher performances.

In the first case, we consider $\alpha = 1$ dB; during the full-beam scan, the PAE fluctuates for the element $(2, 3)$ between 10.8 and 17.6 %. The symmetry of the EM environment of the element located within the center of the array reflects itself in the pattern of performance shown in figure 2.12. We have four maximums that occur for the largest angles with a constant decrease of PAE and P_{out} when the beam is heading toward $(0°, 0°)$, for which the PA exhibits

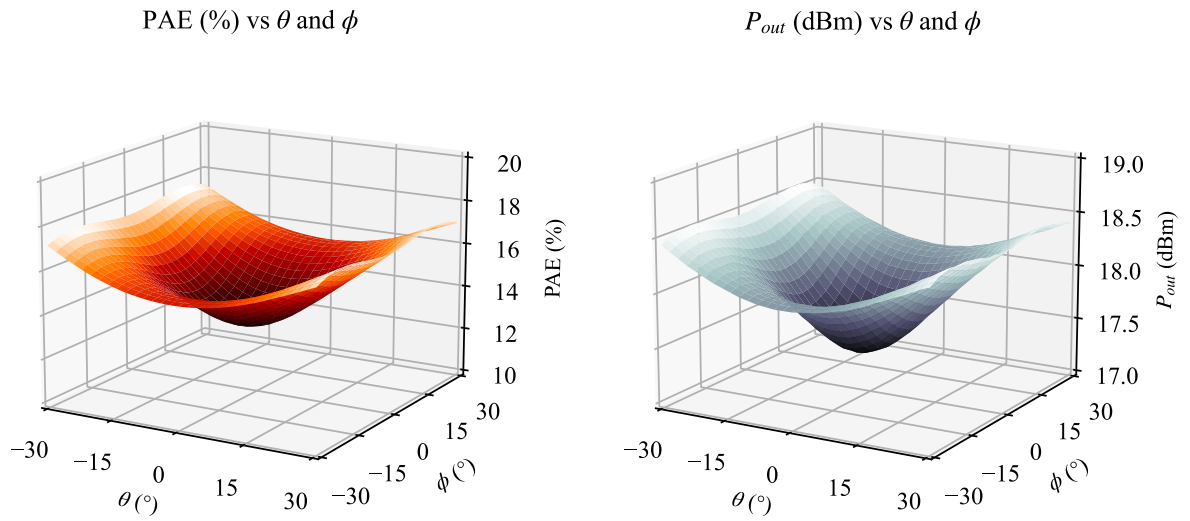


Figure 2.13. PAE (a) and P_{out}(b) of the element $(2,3)$ across a 60° 2D beam-scan with $\alpha = 2.5$ dB

its lower performances. The output power variation is between 18.6 and 16.6 dBm for an average of 17.8 dBm on the beam scan. With $\alpha = 2.5$ dB, the patterns of PAE and P_{out} from figure 2.12 adopt a flatter shape in figure 2.13 with minor variation from 12.9 % to 17.5 %, keeping an average of 15.5 % efficiency, and from 17.3 dBm to 18.5 dBm keeping an average of 18.1 dBm of output power.

We can also notice that the element keeping good performances is $(1,1)$ in the top corner of the array; it shows a much flatter pattern preserving the PAE between 17.6 and 18.8 dBm with an average of 18.1 dBm (within the specifications).

2.3.2.2 High peak power added efficiency PA and ALM

A simulation of the beam steering ALM has been performed on a second power amplifier design optimized for efficiency. Without discussing Doherty architecture, another differential cascode is sized for higher efficiency. The transistor size has been re-scaled and gate biasing has been reduced (class AB), pushing the OCP1 closer to P_{sat} and the peak of PAE. With the lower biasing, the PA shows re-polarization before its compression point, creating gain expansion that increases the PAE. However, the efficiency improvement is at the expanse of gain and linearity. The new PA then exhibits an OCP1 of 18.9 dBm with a PAE of 41 % and a gain of 9.03 dB with an expansion of 0.84 dB. The PA design simulation contains PLS of the MOS transistor and EM simulations of the baluns and interconnections.

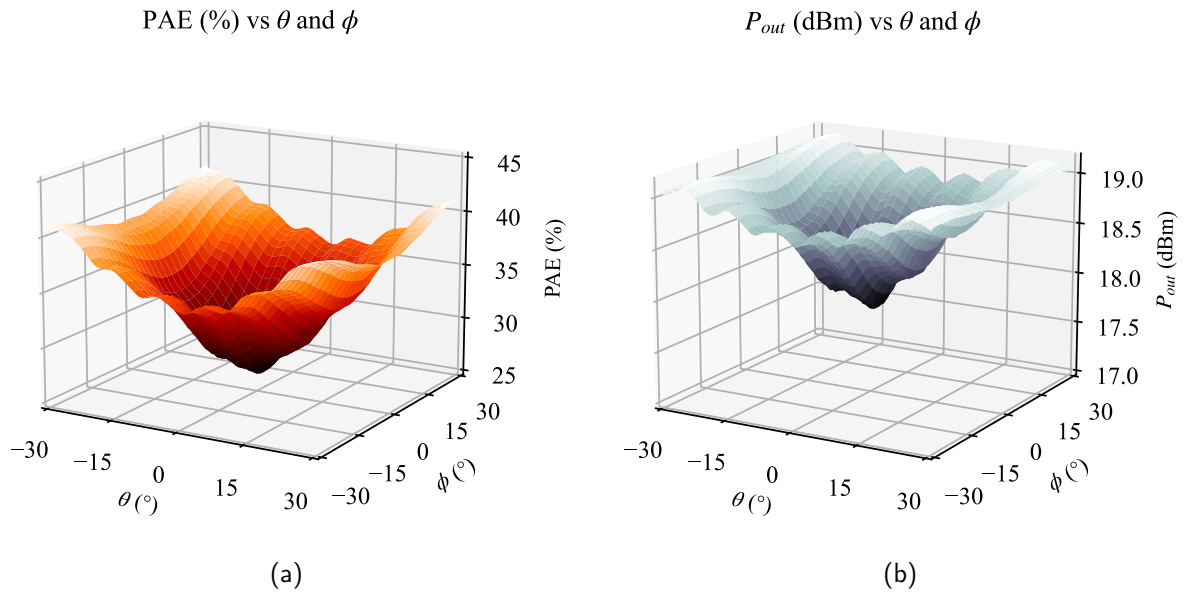
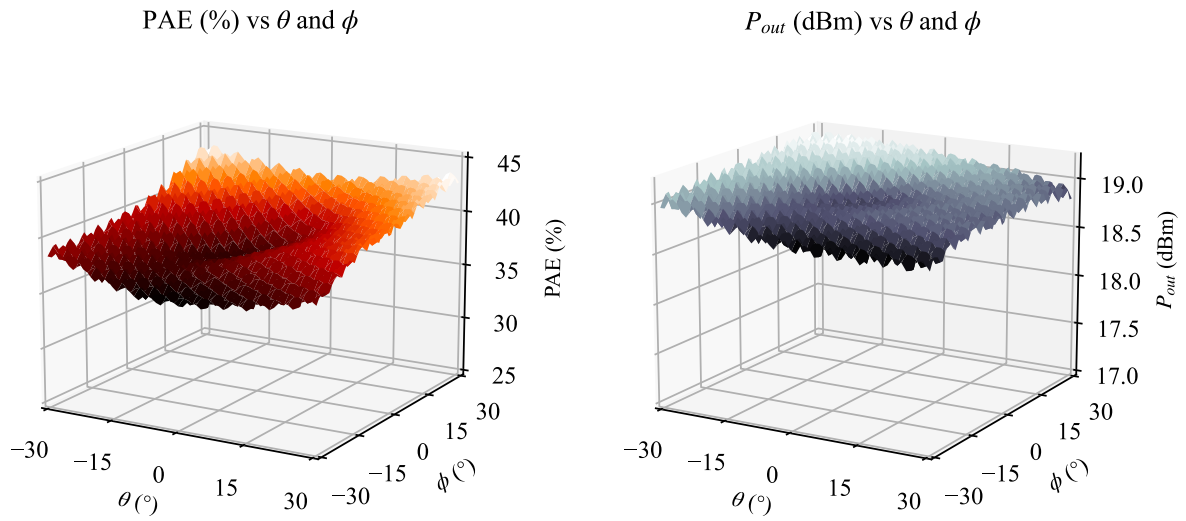


Figure 2.14. PAE (a) and P_{out}(b) of the element $(2,3)$ across a 60° 2D beam-scan with $\alpha = 1$ dB

Once included in the arrays, we can first notice that the new PA design shows a more pronounced variation of performances. These extreme values occur for the element $(2,3)$, the figure 2.14 shows that this particular element's PAE and P_{out} drop when the beam is located around $(\theta,\phi) = (0°,0°)$ due to a VSWR of 2:1. Concerning the output power, the impact of the beam steering is reduced from 18.9 dBm to 17.5 dBm at its lowest. It represents a drop of almost 17 %. Across the beam scan, the average output power delivered is about 18.5 dBm. In the same BS conditions, the PAE drops down to 15.2 % from 40.7 to 25.5 %, with an average of 34.3 %. The maximum of PAE and Pout occurs for wide angles $(\theta,\phi) = (\pm 30°, \pm 30°)$.

On the other hand, the PA the least affected is the one feeding the antenna $(1,1)$ on the top corner of the array; it has a different shape of performance variation, as shown in figure 2.15 with only 8.4 % of variation with a PAE_{min} of 34.6 %, keeping an average efficiency of 38.3 %. The P_{out} follows the same pattern with a low variation of 0.6 dBm between 19 dBm (due to the fact that the load variation goes through an impedance that is exactly the Z_{opt} and giving an extra 0.1 dBm) and 18.4 dBm around an average of 18.7 dBm. The maximum and minimum output power and efficiency occur respectively for $(\theta,\phi) = (-30°, 30°)$ and $(\theta,\phi) = (30°, -30°)$

Here, as the size of the array is limited, each element appears to have a different profile of response to A-VSWR as they do not have precisely the same radiating environment. In much wider arrays, we can expect to see more uniformity in the performance pattern from



(a) (b)

Figure 2.15. PAE (a) and Pout (b) of the element $(1, 1)$ across a 60° 2D beam-scan with $\alpha = 1\,\text{dB}$

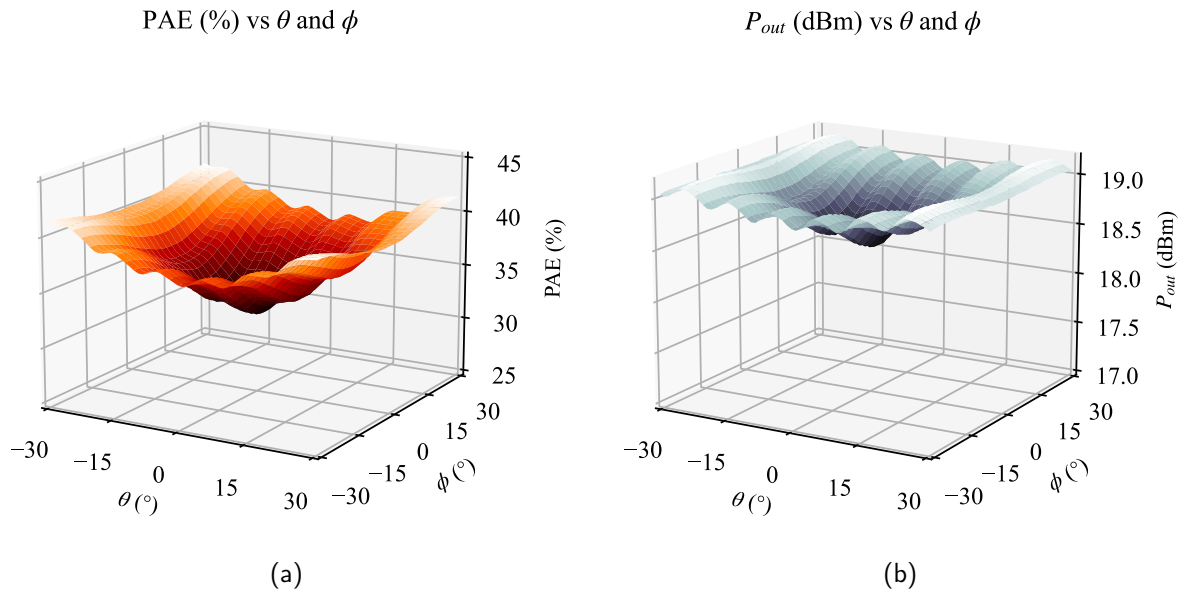
one element to another with disparities approaching the edges. One of the main contributions to the A-VSWR here is the active impedance when no beam steering is performed. It drives the impedance seen by the PA far from $50\,\Omega$, we can observe that the beam steering at an important angle tends to bring the impedance back to the center of the smith chart, hence, restoring the PAE to its original value.

A simulation with 2.5 dB of PA-to-antenna interconnection losses (switched FE topology) is performed, the range of PA performances variation is much more limited. The most affected element sees its PAE vary between 40.9 % and 31 %, which represents 5.3 % less variation compared to a front-end with 1 dB of losses. The output power variation stays contained between 18.9 dBm and 18.1 dBm with an average power of 18.7 dBm. Table 2.2 summarizes the results obtained for better understanding.

	Classic front-end ($\alpha = 2.5\,dB$)						**Common-leg front-end** ($\alpha = 1\,dB$)					
Case	P_{out} **(dBm)**			PAE **(%)**			P_{out} **(dBm)**			PAE **(%)**		
	Min	*Avg*	*Max*	*Min*	*Avg*	*Max*	*Min*	*Avg*	*Max*	*Min*	*Avg*	*Max*
Best	18.6	18.8	19	36.9	39.5	42.8	18.4	18.7	19	34.6	38.3	43
Worst	18.1	18.7	18.9	31	36.8	41	17.5	18.5	18.9	25.5	34.3	40.7

Table 2.2. Summary of the P_{out} and PAE variations among a 25-element array during a 60° beam scan for two front-end architectures

From the data gathered, we notice that the high peak PAE power amplifier suffers from



(a) (b)

Figure 2.16. PAE (a) and Pout (b) of the element $(2,3)$ across a 60° 2D beam-scan with $\alpha = 2.5\,\text{dB}$

more significant drops of performance (up to 15.2 % of PAE) than the more linear PA. Again, the most common topology of FE shows that the important losses due to interconnections significantly reduce the load variation seen by the PA down to 9.9 %. Moreover, it is important to note that the average performances of the power amplifier stay quite close to the nominal values. Meaning that the minimum efficiency is achieved for very few angle values among the full-beam scan reducing the severity of the parasitic active load modulation again.

While the most worrying drops of efficiency that we have identified in worst cases (around 7 %), the reality is even less concerning: Indeed, we expect to have much larger phased arrays working in constant dynamic beam steering; hence, an average data the whole beam scan and on the whole array is more reflective of reality. For a $25-$element array, with 18.9 dBm OCP1 and 41 % PAE PAs in two different front end topologies, the array average performances are the following (Table 2.3):

Average data	Reference PA	$\alpha = 1\,\text{dB}$	$\alpha = 2.5\,\text{dB}$
A-VSWR	1.0	1.4	1.3
Pout (dBm)	18.9	18.8 (-0.1)	19.0
PAE (%)	41	36.6 (-4.4)	38.6 (-2.4)

Table 2.3. Summary of the average performances in power and efficiency under different A-VSWR

The results show that the active load pulling pushes the power amplifiers to work under

VSWR of different intensities according to their location in the array. Furthermore, the choice of the phased-array antenna design and the 5G front-end architecture is also crucial to determine the magnitude of load variation that amplifiers must endure in beam steering conditions.

2.3.3 Active load modulation on radiation patterns

We have constituted a consequent folder that exonerates the antenna from charge for high impact on power amplifiers. We can, however, determine with the help of our 25-element array the variation of output power has the potential to deteriorate the beam radiated by changing the phase and amplitude of the signal at the output of each PAs. Up to now, the precedent analyses were based on HB with S-parameters files extracted from EM simulations. Obtaining radiation results with circuit excitation is not that trivial, especially when complex integrated technology designs are involved. Therefore, a co-simulation tool is necessary. Thanks to Keysight and their tool suite, we managed to set up a bench capable of deriving the radiated electric field (E-field) of the designed antenna array when excited by our ST SOIMMW PA matrix.

2.3.3.1 Circuit-Antenna co-simulation flow

The principle of the co-simulation flow shown in figure 2.17, first of all, is based on the three conventional results that have been used extensively in this work. The HB results from the circuit, EM simulation of the array, and the extracted S-parameters from the array.

To successfully obtain the wanted results, the flow must be derived in this way: First, we must identify the impedance that matches the circuit (PA) and the radiating structure (antenna), then design both devices on their respective tools. Similar to the setup in section 2.1, the radiating element is simulated with the chosen solver (see section 2.1.1) under RFPro. The S-parameters of each port of excitation are exported with an .SnP file. The circuit under Virtuoso/Golden Gate is loaded by an S-parameters box containing the antenna simulation results. The HB simulation is performed with Golden Gate and generates a dataset file containing the excitation that feeds the S-parameters box. Back to RFPro, the EM model of the antenna fed with a feature called "circuit excitation", which is defined as the dataset file generated by Golden Gate. All the simulation solutions deriving from the eventual parameter sweep under Golden Gate are selectable to feed the EM model. The radiation results such as directivity, E-field, and gain are finally plotted.

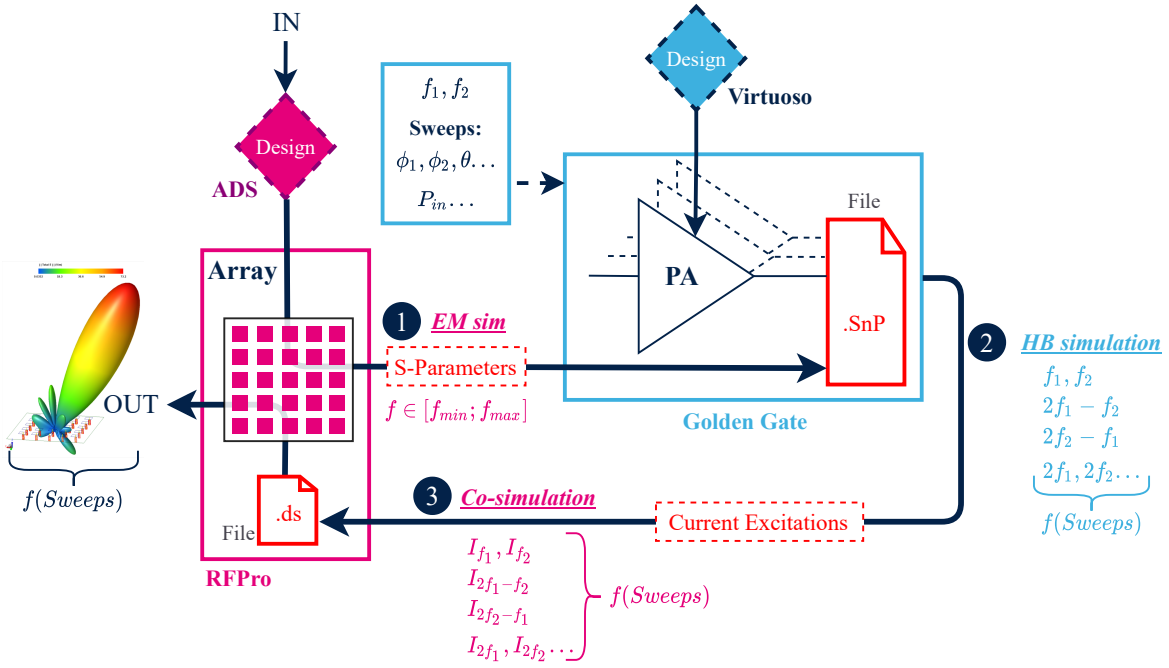


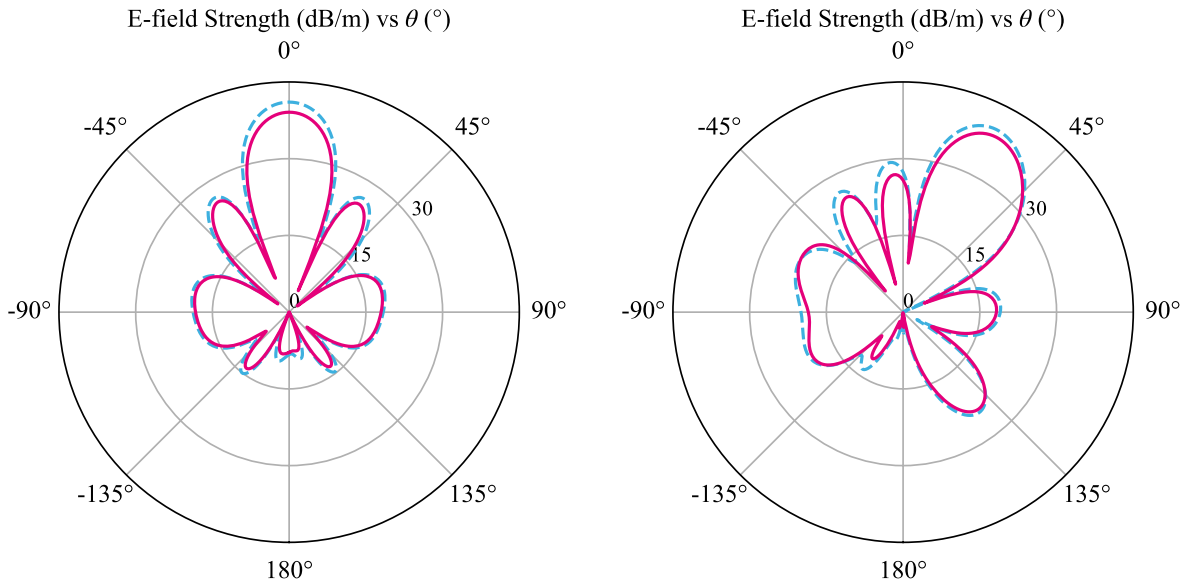
Figure 2.17. Three steps general co-simulation flow for radiation and circuit characterization

2.3.3.2 Radiation results

With the bench ready to co-simulate, a comparison between two cases is proposed: The perfect case is simulated with 25 ideal sources delivering the same output power as the nominal power of our designed PA at OCP1. The mutual coupling symbolized by the S_{nm} terms present within the S-parameters file are forced to 0. On the other hand, the real case reuses the setup from previous analyses with 25 designed PAs plugged at the back of the S-parameters file, including mutual coupling. The beam generated in both cases is simulated for two angles $\theta = (0, 30)°$ (we put aside 2D beam steering to simplify the results). With this comparison, we should determine whether or not the mutual coupling and the resulting ALM has a significant impact on the beam shape.

We consider two potential distortions of the beam, the peak E-field strength in dB and a shift of the peak beam angle. Finally, we keep the two different cases of interconnection losses and plot the radiating patterns for $\alpha = 1\,\text{dB}$ and $\alpha = 2.5\,\text{dB}$:

At first glance, we see in figure 2.18 that the radiating pattern has good surviving skills to load modulation. The general shape of the pattern is well preserved; no particular shift of the beam is observed when $\theta = 0\,°$, the peak E-field strength that was initially of 41.1 dB undergoes a drop of $-2\,\text{dB}$. When steered to $\theta_{steer} = +30°$, a light but still noticeable beam



(a) $\theta_{steer} = 0°$ (b) $\theta_{steer} = +30°$

Figure 2.18. E-field Strength pattern according to θ with (line) and without parasitic ALM (dotted) effect for $\alpha = 1$ dB

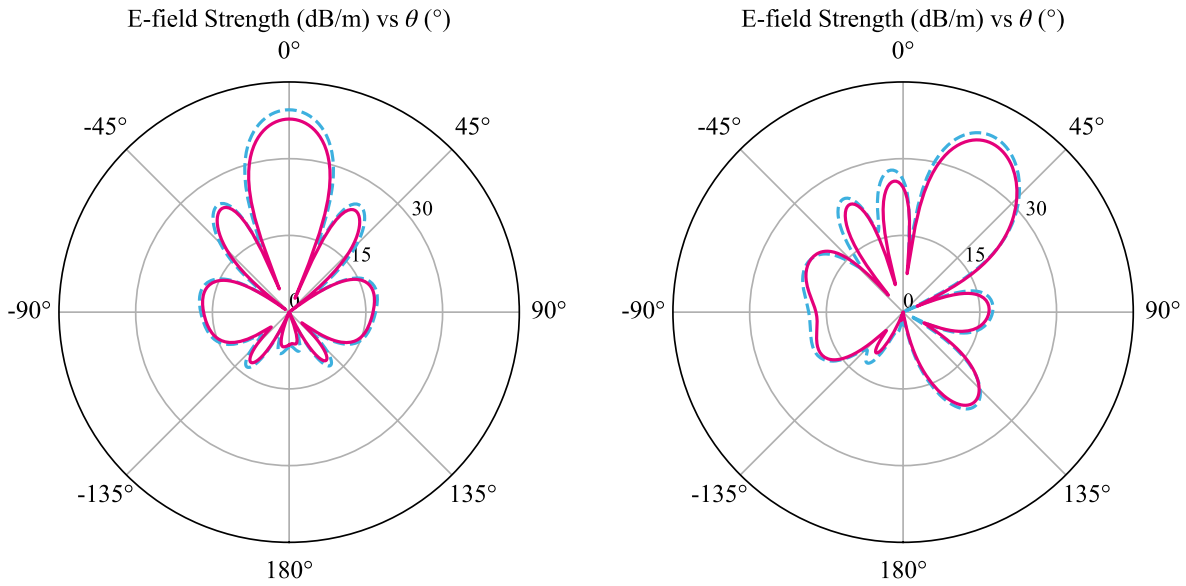
shift exists of around 1° that is within the incertitude of simulation as the step of the plotted data is of 1°, the beam peak amplitude has been reduced by 1.3 dB.

As a reduced impact on PAs performances has been evaluated with $\alpha = 2.5$ dB it is logical to expect a reduced effect on the beam pattern. Indeed, with $\theta_{steer} = 0°$ (figure 2.19a), the amplitude is reduced by −1.8 dB, no angle shift is again noticeable. And finally, we observe a same beam amplitude reduction of 1.3 dB and an angle shift around 1° (within the incertitude of simulation again) when $\theta_{steer} = +30°$ (figure 2.19b).

	α = 1 dB		**α = 2.5 dB**	
Angle θ	**0**	**30**	**0**	**30**
Amplitude shift (dB)	-2.0	-1.3	-1.8	-1.3
Angle shift (°)	<1.0	+1.0	<1.0	+1.0

Table 2.4. Amplitude and angle beam distortion with different PA-antenna losses and steering angle

Table 2.4 summarizes the amplitude and angle variations. We obtain an already known tendency globally among the different amounts of PA to antenna losses without surprise. As both cases represent different amounts of performance variations for the power amplifiers, it shows that we are dealing with the number of disturbances that are negligible in terms of radiation impact. However, the limited size of the array may hide the significance of the



(a) $\theta_{steer} = 0°$ (b) $\theta_{steer} = +30°$

Figure 2.19. E-field Strength pattern according to θ with (line) and without parasitic ALM (dotted) effect for $\alpha = 2.5\,\text{dB}$

effects simulated. Indeed, with larger arrays providing much narrower beams with larger gain, the ALM effect is susceptible to being less discrete. A beam shift of a few degrees may impact the communication link. On the other hand, it is important to remind that the analysis conducted is set up with worst-case excitation. Usually, each source amplitude exciting the array is submitted to weighting functions to optimize PSR. The strategy used in this work is maximizing the main beam amplitude with a uniform amplitude distribution across the array.

2.4 Conclusion

From this analysis, we can draw different conclusions: First, the full-wave simulation of an antenna array is necessary to derive the true interactions between the IC and the radiating elements. Unfortunately, this step is very consuming in terms of time and computing resources, forcing us to work on the reduced-sized array to rely on many assumptions for the rest of the analysis.

Then, despite the concerns with which we started this work on parasitic ALM, this chapter concludes on a positive note: Indeed, the many balanced amplifiers designed to be robust to A-VSWR up to 2:1 were raising a major concern about the need for VSWR-proof amplifiers in 5G PAA. The proposed results show that with a comprehensive analysis involving simulation

with PA and antenna and considering all the RF paths between them, we can achieve direct results of all the PAs behavior. Furthermore, we can derive radiation results from an antenna fed by the designed PAs with co-simulation. On the overall study, the parasitic ALM is finally limited due to the isolation that represents the interconnection losses, and secondly, thanks to the different compensations of phase in the coupled signal RF path. We have evaluated an overall drop of $4\,\%$ of the PAE on the amplifiers within the 25 element array simulated under a 60° beam steering.

The radiation results have been obtained with a new co-simulation method with the help of Keysight. Again it shows very moderate consequences of the parasitic ALM on the beam shape with beam deviation that does not go up to 1° and amplitude variation that does not go beyond 2 dB of amplitude.

With the results obtained on our limited array, we can speculate on what happens in a very large array (≥ 32 elements). We can expect our $4\,\%$ drop of PAE to be even reduced, as a larger sample of elements will have the same EM environment hence the same behavior in terms of load variation. Concerning the strategy to adopt in terms of PA design, we may adopt two approaches: First, we can focus on a design maximizing robustness; it must then comply with the most demanding specifications representing the worst case in which the device will work. However, with the results obtained, we can disagree with this strategy. Indeed, it implies the implementation of complex architectures that most of the time sacrifice efficiency. On a distributed system with a large sample of devices such as PAA, the extra level of robustness will benefit a few elements while potentially costing efficiency to the whole array. This brings us to the second strategy. We can sacrifice robustness for a limited number of elements that will benefit from it and keep the designs simpler for the sake of efficiency for the major part of the array.

The disruptive patterns created by the edge of the array will less contribute to the overall efficiency impact. Antenna techniques and judicious approach of the antenna active impedance matching can also greatly contribute to the impact of the parasitic ALM.

With the very handy co-simulation bench that we have exploited, another problematic phenomenon involving PAs and nonlinearities has been identified and will be treated in the next chapter.

References

[Bal05] Constantine A Balanis. *Antenna theory: analysis and design*. Wiley-Interscience, 2005.

[DOE16] Daniel-Nicuşor Depărăţeanu, Teofil-Cristian Oroian, and Florin Enache. "Side lobes level reduction in planar array antenna using two-dimensional window functions". In: *2016 International Conference on Communications (COMM)*. June 2016, pp. 125–128. DOI: 10.1109/ICComm.2016.7528330.

[Hol13] Eric Holzman. "On the use of dummy elements to match edge elements in transmit arrays". In: *2013 IEEE International Symposium on Phased Array Systems and Technology*. Oct. 2013, pp. 549–552. DOI: 10.1109/ARRAY.2013.6731887.

[KR19] Umut Kodak and Gabriel M. Rebeiz. "A 5G 28-GHz Common-Leg T/R Front-End in 45-nm CMOS SOI With 3.7-dB NF and 30-dBc EVM With 64-QAM/500-MBaud Modulation". In: *IEEE Transactions on Microwave Theory and Techniques* 67.1 (Jan. 2019), pp. 318–331. ISSN: 1557-9670. DOI: 10.1109/TMTT.2018.2873374.

[LCK19] Kyeong-Hyeok Lee, Sunkyu Choi, and Choul-Young Kim. "A 25–30-GHz Asymmetric SPDT Switch for 5G Applications in 65-nm Triple-Well CMOS". In: *IEEE Microwave and Wireless Components Letters* 29.6 (June 2019), pp. 391–393. ISSN: 1558-1764. DOI: 10.1109/LMWC.2019.2914123.

[LH18] Wonho Lee and Songcheol Hong. "Low-loss and Small-size 28 GHz CMOS SPDT Switches using Switched Inductor". In: *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2018, pp. 148–151. DOI: 10.1109/RFIC.2018.8428974.

[Li+21] Zekun Li et al. "A 24–30-GHz TRX Front-End With High Linearity and Load-Variation Insensitivity for mm-Wave 5G in 0.13-µm SiGe BiCMOS". In: *IEEE Transactions on Microwave Theory and Techniques* 69.10 (Oct. 2021), pp. 4561–4575. ISSN: 1557-9670. DOI: 10.1109/TMTT.2021.3101232.

[LZZ16] Hongbin Li, Quanming Zhao, and Panpan Zuo. "Modeling and simulation of multilayer flip-chip package". In: *2016 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*. July 2016, pp. 1–2. DOI: 10.1109/NEMO.2016.7561647.

[UC01] K.B. Unchwaniwala and M.F. Caggiano. "Electrical analysis of IC packaging with emphasis on different ball grid array packages". In: *2001 Proceedings. 51st Electronic Components and Technology Conference (Cat. No.01CH37220)*. May 2001, pp. 1496–1501. DOI: 10.1109/ECTC.2001.928034.

Chapter 3

Using 5G phased array properties to solve third-order intermodulation issues

3.1 Beam Diversion for linearity improvement

Section 1.4 has shown the particular conditions of radiation that affect the IMD3. This chapter will show how these peculiarities can be turned into a system solution to relax PA linearity requirements and will present an implementation on ST CMOS065SOIMMW technology of a PA able to take advantage of this solution. We will first use a co-simulation flow (section 2.3.3.1) to illustrate different intermodulation scenarios and show in which conditions IMD3 radiations are disturbing communications. From the various cases of intermodulation radiation, the first step will consist in the identification of situations where PA non-linearities may cause interference in base station-to-user communication.

3.1.1 IMD3 and dual-beam forming (co-simulation)

3.1.1.1 Upgrade of co-simulation bench

One of the main challenges faced with IMD3 is the frequency proximity of the main tones. IMD3 spectrum exhibits upper and lower frequencies (see section 1.4.2.1). The co-simulation bench needs some upgrades to derive the IMD3 radiation beam. The 25 element array with the 25 PA exhibiting 18 dBm of output power is reused. A bi-tone HB simulation is set up. As a 2D beam steering is not needed in this chapter, and to speed up the simulations, ψ parameter has not been used. The setup becomes:

-Two frequencies f_1 and f_2 for each user;

-Constants needed for beam steering (d, k, λ);

-Two angles θ_1 and θ_2 correspond to the two-beam angular location;

-Two matrices P_1 and P_2 of 25 phases each, respectively derived from the values of θ_1 and θ_2 ;

-A fixed input power P_{in} (split on the two tones) is applied to each PA;

-A fixed biasing (V_G, V_{GC}, V_{DS}) for each amplifier.

Roughly 50 parameters are computed to run the HB simulation for any couple of angles θ_1 and θ_2. A sweep of the angles would imply a re-computation of those parameters for each iteration. The previous analyses were performed with single tone HB simulation where seven harmonics were calculated to converge; in this bi-tone case, seven harmonics per tone are computed, which is considerably impacting the computing time and resources. The final results are again obtained with the same co-simulation flow used previously with RFPro.

3.1.1.2 Case study of IMD3 beamforming

In former wireless communication standards, relying on a single antenna, the intermodulation was radiated in every direction the antenna was showing gain. With 5G, a highly directive beam is formed and oriented. Equation 1.31 has demonstrated that with multiple beams steered in the same direction, the IMD3 are oriented in the same direction. An example is shown in figure 3.1 with two beams steered at $(20; 20)°$.

Situation reminder: UE_1 and UE_2 communicate with the same base station, each with a tone and a dedicated beam. The antenna elements are driven by the sum of the two tones and the IMD3 power generated by the power amplifiers. The tones are separated in frequency by $\Delta f = 200\,\text{MHz}$ at 27.9 GHz and 28.1 GHz.

We remember from the first chapter that the model given by [Hem02] allows computing the beam angle with which the IMD3s are radiated (equations 1.32 and 1.33).

$$\theta_{IMD3,\alpha} = \arcsin\left\{\frac{2f_{UE_1}\sin\theta_{UE_1} - f_{UE_2}\sin\theta_{UE_2}}{2f_{UE_1} - f_{UE_2}}\right\} \tag{3.1}$$

$$\theta_{IMD3,\beta} = \arcsin\left\{\frac{2f_{UE_2}\sin\theta_{UE_2} - f_{UE_1}\sin\theta_{UE_1}}{2f_{UE_2} - f_{UE_1}}\right\} \tag{3.2}$$

It is difficult to find any rule of thumb to describe the behavior of the IMD beam. Firstly, we face a limitation of the equations 1.32 and 1.33: The inverse sine function is defined within the boundaries $(-1 : 1)$, so when the delta between angle θ_1 and θ_2 makes the term inside the arcsine exceed these boundaries, the IMD3 beam angles are no longer defined. Secondly, the maximum angle of steering θ_{max} causes the IMD3 beam to fade and then to re-appear

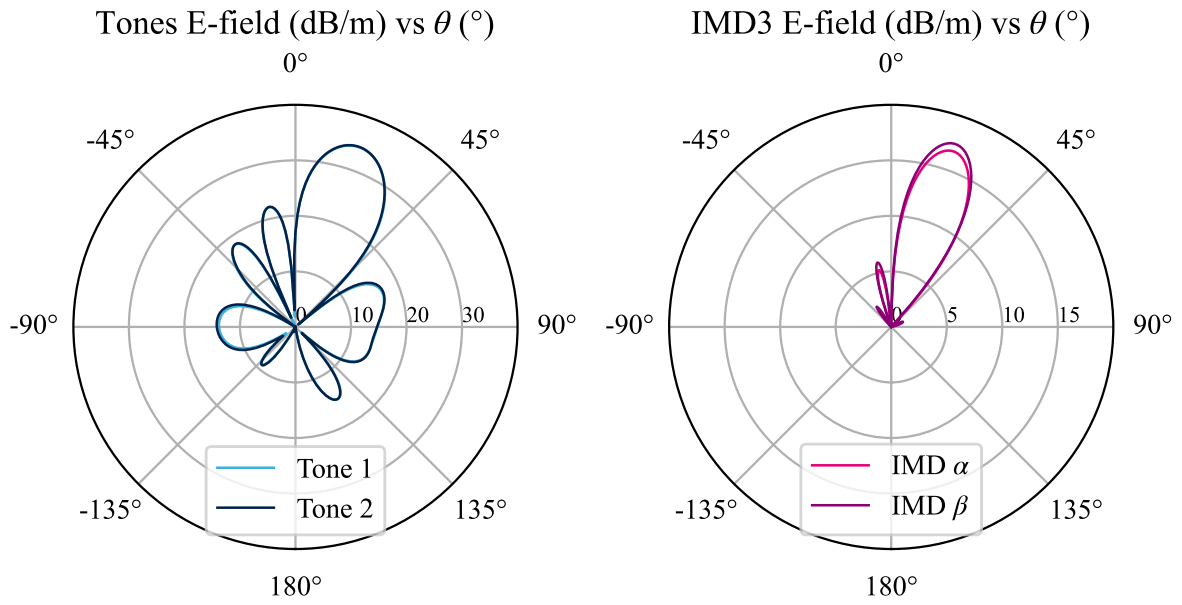


Figure 3.1. Radiated tones for $(\theta_{UE_1}; \theta_{UE_2}) = (20; 20)°$ and resulting IMD3 beams in mono-directional beam steering

with an opposite sign. Again, this phenomenon is not represented by the equation given; the bench of simulation, which works for any case, is used to derive the IMD3 radiation angle. An analysis of several particular cases is proposed.

3.1.1.2.1 Case 1: IMD3 outside of users' range

The two users are located on opposite angles $(\theta_{UE_1}; \theta_{UE_2}) = (10; -10)°$. The terms inside the arcsine function are within the range of definitions of the equations 1.32 and 1.33 and give the two IMD3 $(\theta_{IMD3,\alpha}; \theta_{IMD3,\beta}) = (31.7; -31.1)°$.

We observe, thanks to the co-simulation bench, that in this case (figure 3.2) UE_1 and UE_2 are "IMD3 free", but at the expanse of the users located around. Indeed, both IMD3 are radiated in different directions that can disturb other devices independent of our base station's communication. In this case, we also observe that simulation results match equations predictions.

3.1.1.2.2 Case 2: IMD3 inside of users' range

UE_1 and UE_2 are located in opposite directions, with the larger angle difference achievable. Hence, at the maximum angle reached by the array on $(\theta_{UE_1}; \theta_{UE_2}) = (30; -30)°$. With $f_1 = 27.9$ GHz and $f_2 = 28.1$ GHz the current setup we have:

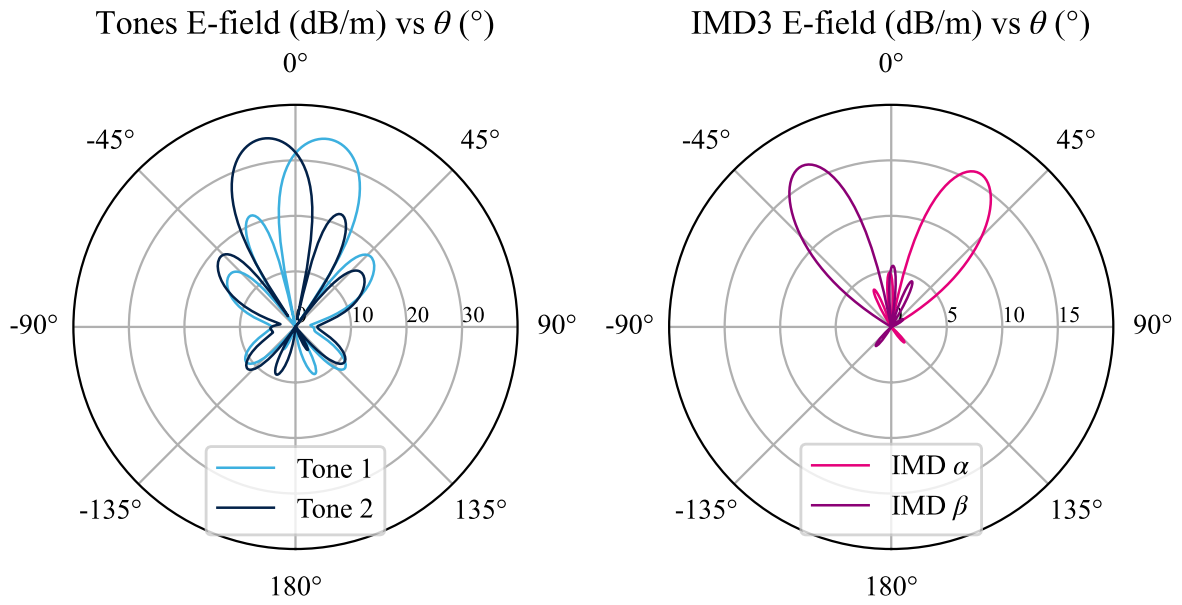


Figure 3.2. Radiated tones for $(\theta_{UE_1}; \theta_{UE_2}) = (-10; 10))^\circ$ and resulting IMD3 beams outside users' range

$$\frac{2f_{UE_{1,2}} \sin\theta_{UE_{1,2}} - f_{UE_{2,1}} \sin\theta_{UE_{2,1}}}{2f_{UE_{1,2}} - f_{UE_{2,1}}} = \pm 1.5 \tag{3.3}$$

As $\sin(\theta_{IMD3,\alpha,\beta}) \notin (-1:1)$, $\theta_{IMD3,\alpha,\beta}$ are then not defined. However, the co-simulation shows in figure 3.3 that two beams are radiated with $(\theta_{IMD3,\alpha}; \theta_{IMD3,\beta}) = (-30; 30)^\circ$.

In this case, the two users receive each an IMD3 component. This represents a worst-case scenario for two users at two different locations as they have a reduced SNR due to IMD3.

3.1.1.2.3 Case 3: IMD3 cancellation

Another particular case occurs when the users are located at $(\theta_{UE_1}; \theta_{UE_2}) = (0; 30)^\circ$. By computing the IMD3 angles, we find that $(\theta_{IMD3,\alpha}; \theta_{IMD3,\beta}) = (-30.5; 83.2)^\circ$, and we immediately notice that $\theta_{IMD3,\beta} > \theta_{max}$, that is the maximum angle of steering of the array defined by the relationship of equation 1.12

For UE_1 and UE_2, it is a very favorable scenario; The IMD_β is out of the radiation range of the array and consequently is not radiated with high directivity. Furthermore, the IMD_α is emitted out of the scope of the users. However, other users located between $\theta = (-30; 0)^\circ$ can receive the IMD_α created by the signals dedicated for UE_1 and UE_2.

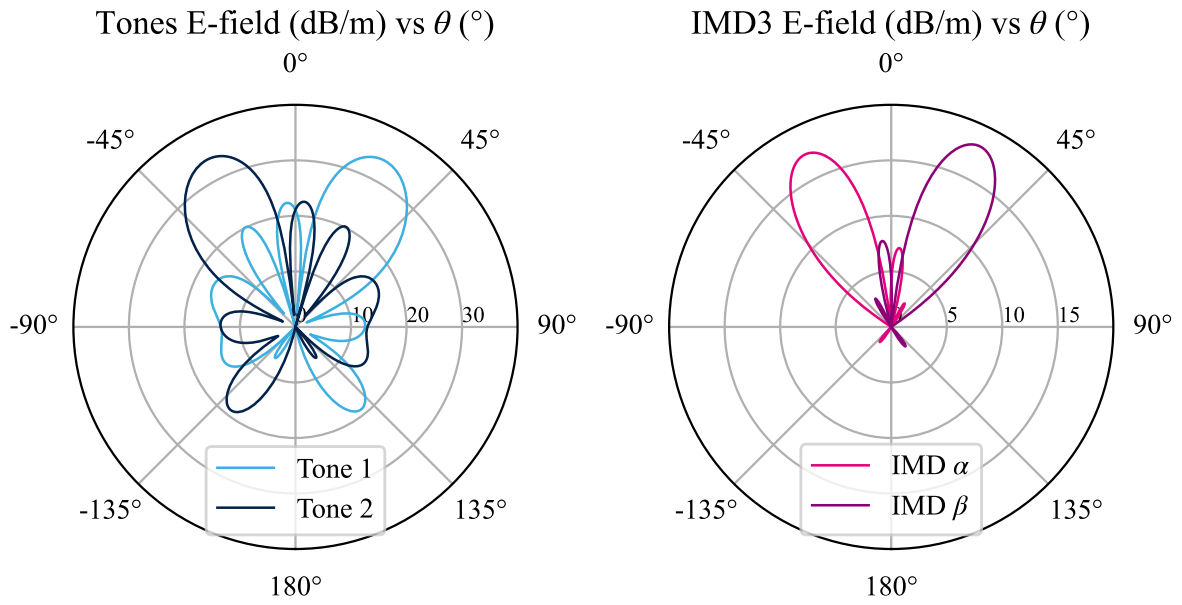


Figure 3.3. Radiated tones for $(\theta_{UE_1}; \theta_{UE_2}) = (-30; 30)°$ and resulting IMD3 beams inside users's range

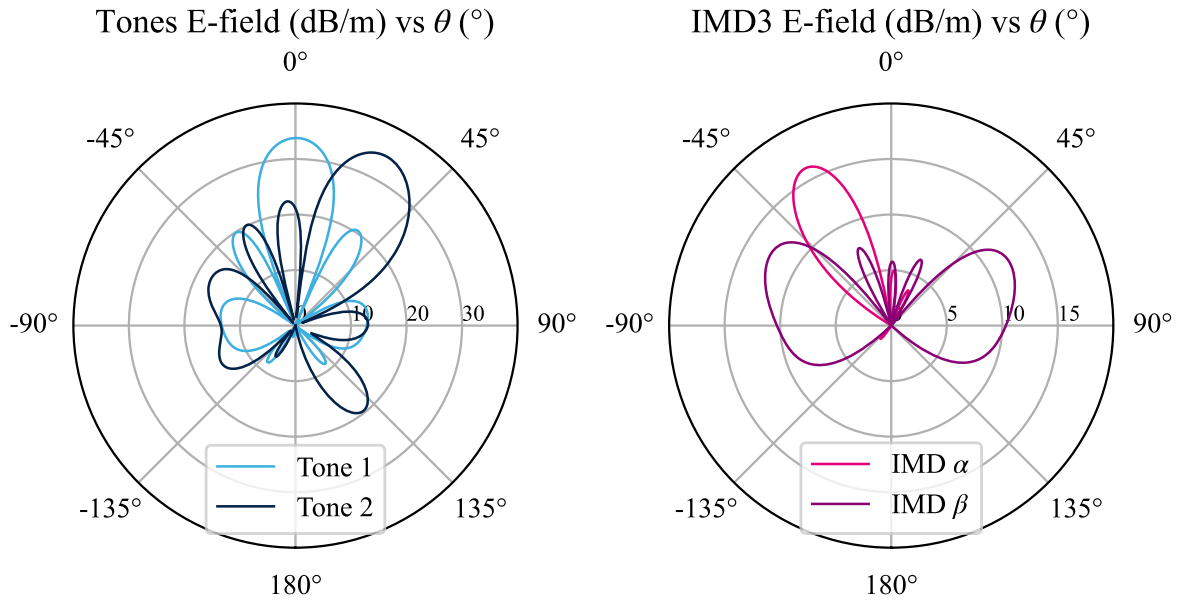


Figure 3.4. Radiated tones for $(\theta_{UE_1}; \theta_{UE_2}) = (0; 30)°$ and resulting IMD3 beams

3.1.1.3 IMD3 beam diversion principle

The case study has shown that IMD3 radiation can occur where a potential user is. It is not allowed to radiate power above a certain power level outside of your target angle to allow any user to communicate with the base station. The usual way to comply with this constraint is to use a very linear PA so that the amount of IMD3 before beamforming is low enough. Of course, this solution costs power consumption (need for PA linearization schemes, higher bias current, etc...). We propose an idea based on beam-steering for linearity relaxation: the concept resides in the diversion of the IMD3 radiation beam; when the IMD3 lobe is formed toward a user, we propose to divert it out of the user range or ideally outside of the maximum range of beamforming capability θ_{max}. We must first evaluate the linearity relaxation that we can expect from IMD3 beam diversion. In this regard, we look at a beamformed in a given direction θ for a given array size. It has a certain width, and with an additional angle shift θ_{sft} we obtain a rapid drop of power emitted in the original direction.

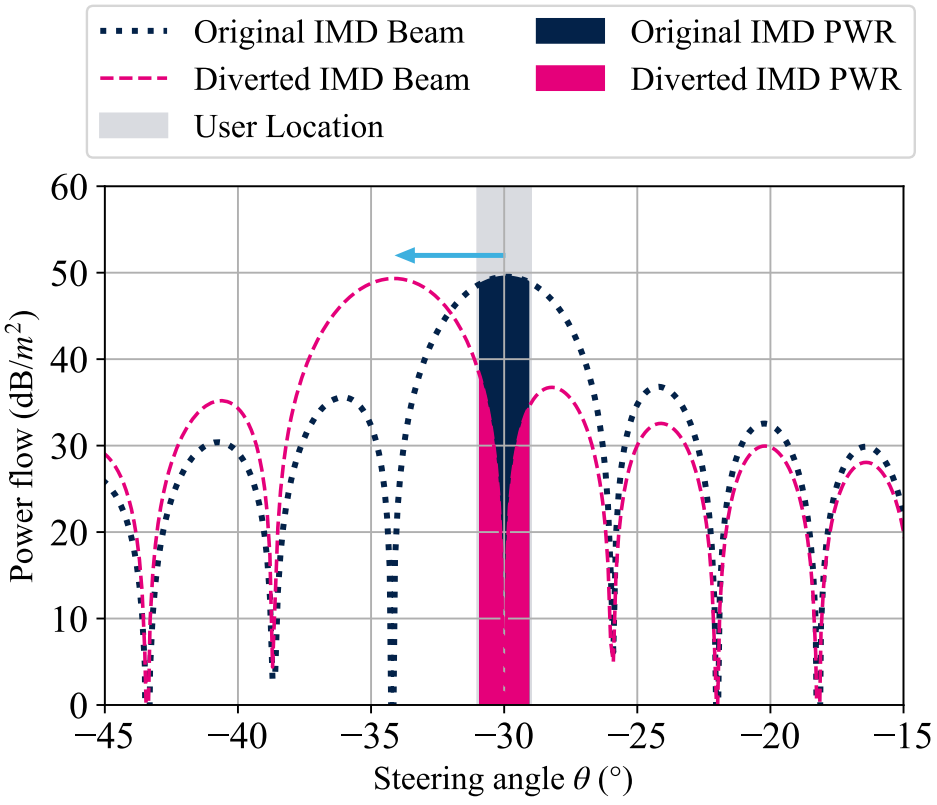


Figure 3.5. 30°beam and 34°deviated beam power flow vs. angle θ in a 1024-array

Let's reuse the example of figure 3.3 from section 3.1.1.2.2. Two users are located at $\pm 30°$;

the previous analysis showed that the IMD3 beams are heading toward the two users. Then, we want to apply beam diversion to the user located in a 2° wide sector at $-30°$. As can be observed in figure 3.5, the amount of IMD3 energy received by the sector for a given phase shift will be higher for a beam with strong directivity. This means that the diversion technique will be more suited for large arrays requiring less shift for the same linearity relaxation. Since we cannot simulate a large array with the co-simulation flow, we can obtain quick results with the help of the array factor [Bal05]. Software such as HFSS can provide a pre-computed radiation pattern of an N-array from only a single element simulation. Based on those results, we can now emulate a 1024 element array. With the graph shown in figure 3.5, we see that starting with an original IMD3 beam located at the user location; we can minimize the IMD3 energy received by the user's sector if we can coincide the user location with a zero of radiation of the IMD3 pattern. These nulls occur for θ_n [Bal05] given in equation 3.4.

$$\theta_n = \arccos\left[\frac{\lambda}{2\pi d}\left(-\Delta\phi \pm \frac{2n}{N}\pi\right)\right] \tag{3.4}$$

With λ the wavelength, d the element spacing, n the null index, and N the number of radiating elements on the axis perpendicular to the radiation direction. With a $\lambda/2$ spaced $1024-$array and a beam steered at $-30°$ ($\Delta\phi = -90°$), the four first nulls given by equation 3.4 are located around the main beam at $(-38.7, -34.2, -25.9, -22.0)°$. To steer the first null of the IMD3 beam toward the user, a shift of angular position of $\theta_{shift} = 5\,°$ is targeted, with as small margin taken. The IMD3 $\Delta\phi$ must be changed by at least $\Delta\phi = kd\sin\theta_{shift} = 15.7\,°$ relatively to the main beam excitation to reach the first null and send it toward the user. By integrating the power flow on a 2° sector around the user's location, a full-beam would transmit 65 dBpm^2. With the first null aligned, the power flow transmitted on this sector is reduced by 17 dB giving the relaxation that can be obtained with beam deviation.

3.2 IMD3 and second harmonic injection

A mean of control is needed to implement beam diversion on IMD3 by changing the phase of intermodulations with at least $\pm 15°$ without perturbing the main signal phase and amplitude. To achieve such results, the analysis of IMD3 behavior under 2nd harmonic injection is studied.

3.2.1 Bi-tone signals and non-linear devices

The generation of IMD3 is shown in this section with a simulation of a simple nonlinear device in moderate signal: A capacitor $\widetilde{C}$ is chosen as it is representative of the non-linear output impedance of a transistor (see section 1.4.1.1). This capacitance is dependent on the $N-$order function of the voltage across its terminals:

$$\widetilde{C}(v_{\widetilde{C}}) = C_0 + \sum_{n=1}^{N} C_n v_{\widetilde{C}}^n \tag{3.5}$$

With C_0 the linear capacitance and C_n the capacitances of each order.

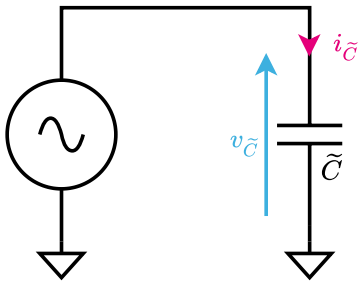


Figure 3.6. Non-linear capacitance fed by a voltage source

The ratio of charges $Q_{\widetilde{C}}$ in the capacitor $\widetilde{C}$ is derived from:

$$Q_{\widetilde{C}} = \widetilde{C} \cdot v_{\widetilde{C}} = C_0 \cdot v_{\widetilde{C}} + C_n v_{\widetilde{C}}^2 + ... + C_n v_{\widetilde{C}}^{N+1} \tag{3.6}$$

The current being the variation of charges within the capacitor, is expressed as:

$$i_{\widetilde{C}} = \frac{dQ_{\widetilde{C}}}{dt} \tag{3.7}$$

The second order of $i_{\widetilde{C}}$ corresponds to $C_n v_{\widetilde{C}}^2$ which will create an $(N+1)^{th}$ order harmonic of the fundamental frequency component in Q and thus in $i_{\widetilde{C}}$.

To observe the components generated on $i_{\widetilde{C}}$ the capacitor is simulated with a voltage source delivering the voltage $v_{\widetilde{C}}$ across its terminal, as shown in figure 3.6.

3.2.1.1 1st order non-linear device

First, we consider a capacitor model limited to the order $N = 1$. The capacitor is then defined by

$$\widetilde{C}(v_{\widetilde{C}}) = C_0 + C_1 v_{\widetilde{C}}(t) \tag{3.8}$$

The harmonic content generated is observable with a simple simulation: The capacitor $\widetilde{C}(v_{\widetilde{C}})$ is fed with two series voltage sources generating two bi-tones signals. A main source delivering $v_{\widetilde{C}}/2$ on both frequencies $f_1 = 27.9\,\text{GHz}$, $f_2 = 28.1\,\text{GHz}$, and an auxiliary source delivering a variable voltage on frequencies $2f_1 = 55.8\,\text{GHz}$, $2f_2 = 56.2\,\text{GHz}$. The linear capacitance C_0 is arbitrarily chosen to be 1 pF and the coefficient $C_1 = 0.5\,\text{pF}$ is kept small enough to avoid too strong harmonic content that might prevent the simulation from converging.

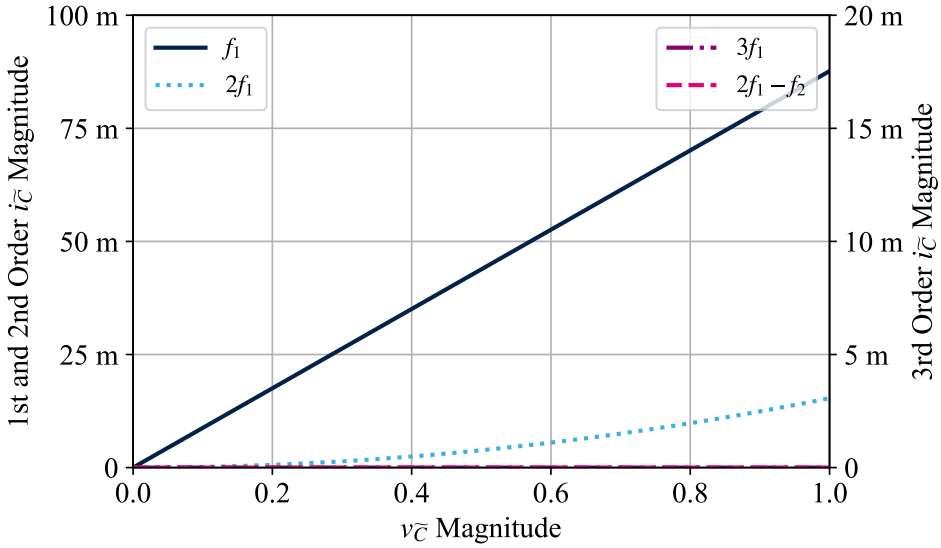


Figure 3.7. Fundamental, IMD3, 2nd, and 3rd harmonics current magnitude vs. $v_{\widetilde{C}}$ magnitude

As the first step, the auxiliary source providing $2f_{1,2}$ is off. As the capacitance is not dependent on the second-order of $v_{\widetilde{C}}$, by increasing the voltage amplitude of the source $f_{1,2}$ (figure 3.7), only 2nd harmonic current is generated; in contrast, the 3rd harmonic and 3rd order intermodulations stay null.

We now enable the auxiliary source that feeds the capacitor with a constant second harmonic bi-tone voltage, $v_{\widetilde{C}}$ is again swept. Figure 3.8 shows the second harmonic current that is generated, the increase due to the natural 2nd harmonic current generated by the non-linear capacitor is barely visible. Furthermore, we notice that now third order components are present: Although the fundamental current amplitude did not change, we have an increase of 3rd harmonic. More importantly, the second harmonic injection allows the generation of

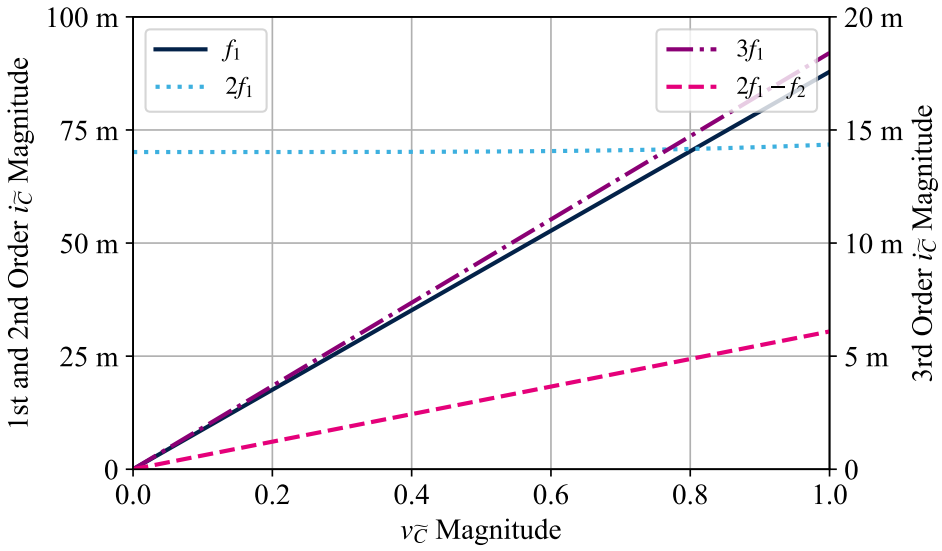


Figure 3.8. Fundamental, IMD3, 2nd and 3rd harmonics current magnitude vs. $v_{\widetilde{C}}$ magnitude under 2nd harmonic injection

IMD3 thanks to the mixing of the 1st order capacitance nonlinearity with the 2nd harmonic content injected.

3.2.1.2 2nd order non-linear device

In the previous section, IMD3 was generated by a combination of fundamental and second harmonic injection. To have IMD3 generated from fundamental itself, let's now add an additional order of nonlinearities to the capacitance model ($N = 2$); $\widetilde{C}$ becomes:

$$\widetilde{C}(v_{\widetilde{C}}) = C_0 + C_1 v_{\widetilde{C}}(t) + C_2 v_{\widetilde{C}}(t)^2 \tag{3.9}$$

The auxiliary source is turned off, and $v_{\widetilde{C}}$ is swept. The nonlinearities generated are shown in figure 3.9. We observe that the non-linear capacitance created fundamental, 2nd , 3rd harmonic currents and IMD3 due to the intrinsic nonlinearity of the capacitance.

By turning on the auxiliary second harmonic source, we will introduce additional 3rd order nonlinearities that will be added to the natural 3rd order nonlinearities generated by the capacitance. However, we can achieve a mix of injected phased-controlled 2nd harmonic with the 1st order nonlinearity of $\widetilde{C}$ to synthesize artificial IMD3. Then, we can expect to influence the phase and the magnitude of the IMD3 naturally generated by $\widetilde{C}$ with the synthetic IMD content generated by the injection.

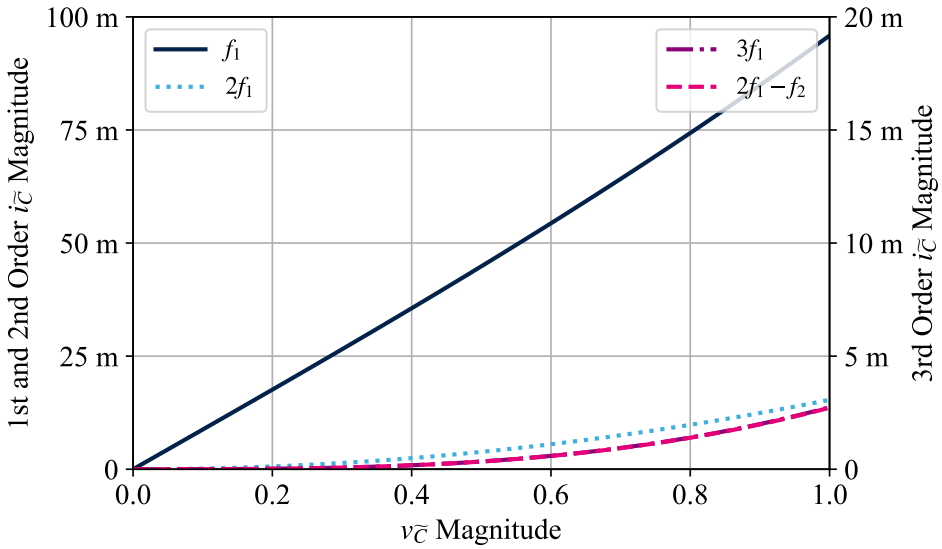


Figure 3.9. Fundamental, IMD3, 2nd, and 3rd harmonics current magnitude vs. $v_{\widetilde{C}}$ magnitude

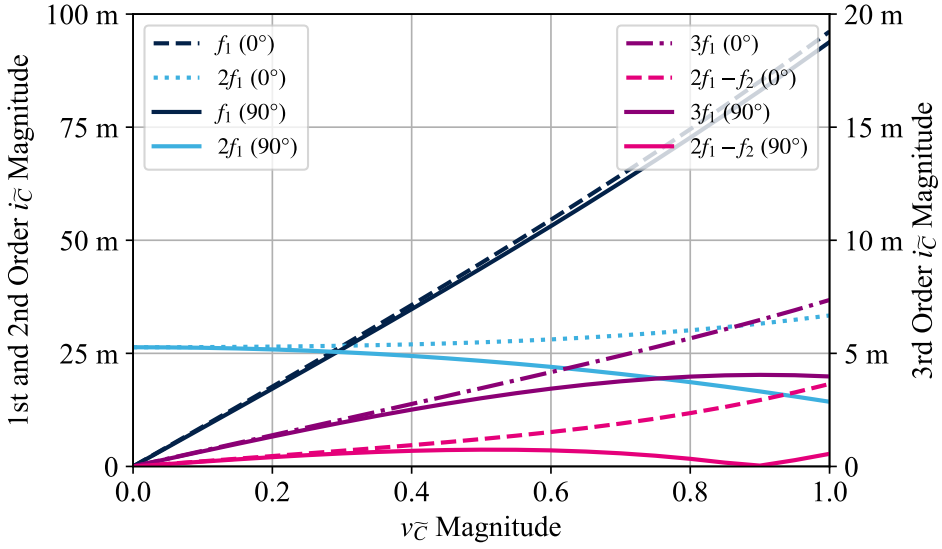


Figure 3.10. Fundamental, IMD3, 2nd, and 3rd harmonics current magnitude vs. $v_{\widetilde{C}}$ magnitude for $\phi = (0; 90)°$

A phase shift ϕ is then introduced in the auxiliary source, changing the phase of the second harmonic injection. To keep the results readable, the current composition is plotted for two phase-shifts values: $\phi = 0\,°$ as a point of comparison and $\phi = 90\,°$, where the most amplitude variations of the nonlinearities occur. Figure 3.10 shows several interesting facts: The fundamental current shows low amplitude variation between the two angles of injection

(around 2 %). The second harmonic content composed of the constant injected component plus the natural component generated by the device shows a negative slope with the increase of $v_{\widetilde{C}}$, reaching a drop of 57 % of amplitude at $v_{\widetilde{C}_{max}}$. The third harmonic increase is softened with the rise of $v_{\widetilde{C}}$, getting to a level of 45.9 % lower than with $\phi = 0\,°$. Last but not least, our IMD3 level is kept very low and is even nullified at $v_{\widetilde{C}} = 0.9\,\mathrm{V}$.

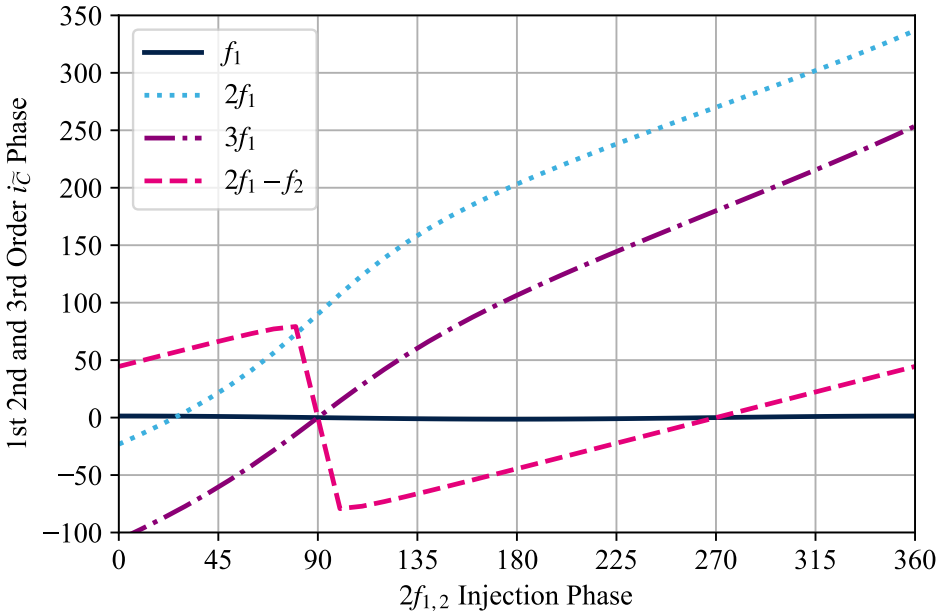


Figure 3.11. Fundamental, IMD3, 2nd, and 3rd harmonics current phase vs. 2nd harmonic injection phase ϕ

This first step of simulations demonstrates that, by injecting $2f_1$ and $2f_2$ signals with the correct phase and amplitude at the output of a non-linear system, we can tune the value of IMD3 components ($2f_1 - f_2$ and $2f_2 - f_1$) both in-phase and amplitude. The next step will help us to understand how this property can be used to tune the IMD3 beams in a phased array context.

$v_{\widetilde{C}}$ is fixed to 0.9 V, ϕ is swept between 0 and 360°. The phase of the 1st, 2nd, and 3rd component is monitored and shown in figure 3.11. As in amplitude, the fundamental phase is independent of the injection phase, which is essential not to change the properties of the proper signal. The 2nd harmonic current is quasi-linear due to recombination of the variable controlled current injection with the natural second harmonic current generated by the nonlinearities of $\widetilde{C}$. On the other hand, the 3rd harmonic shows a somewhat similar pattern to $2f_1$. Concerning the IMD3, we directly observe a discontinuity at 90° that corresponds to

the phase sweet spot where the IMD3 is canceled (figure 3.11). Furthermore, we see that the IMD3 phase has a phase variation range of around 158.4°.

This section has shown that through 2nd order non-linear mechanisms, we can achieve control of the third-order intermodulation distortion as well in phase and amplitude with 2nd harmonic injection. The phase modulation is then the brick required to achieve the beam diversion targeted. The next step will be to implement this technique inside a power amplifier.

3.3 IMD3 Phase Shifting Power Amplifier implementation

3.3.1 ST CMOS065SOIMMW technology

The CMOS SOIMMW technology of STMicroelectronics is a PD-SOI technology in early development during this work. It is derived from a 65 nm SOI node that originally had transistors using a thick gate oxide and a 1.2 V nominal operating voltage. During the thesis, the technology has been updated with a thinner gate oxide that handle a 1 V nominal voltage and 40 nm gate length transistors, for higher frequency applications. Two back-end-of-lines (BEOLs) are available: 5 metal layers, one with a single thick metal and a 7 metal layers BEOL with two thick metals. As its name suggests, the emergence of such technology follows the need for RF front-end for mmW frequencies. As mentioned in section 1.4.1.1, SOI substrates benefit from passive device performances; high resistivity substrate below the buried oxide combined with a trap-rich layer allows low-loss passives by getting rid of the lossy silicon bulk from bulk technologies. The high resistivity SOI substrate is also a significant drive for space application. The insulator properties allow radiations hardening, which is critical for the IC to face the harsh environment of space.

3.3.2 Circuit Design

3.3.2.1 Base Power Amplifier Design

The power amplifier designed targets an OCP1 of 18 dBm according to the specifications for 5G AA power amplifiers (between 17 and 20 dBm) [Par+16] [IR17]. A differential architecture with cascodes is chosen.

First, cascodes are widely used for millimeter-wave amplifiers [Par+19] [ÇR19]. They show several advantages [Raz11]; cascodes allow a larger voltage swing balanced between the two transistors, limiting their stress. The cascode topology also allows high gain, good isolation, and reduces the miller effect. The latter consists of the multiplication of C_{dg} capacitance

(see section 1.4.1) by the voltage gain A_v. As the C_{dg} feedback increase with frequency, the bandwidth of an amplifier is reduced by the Miller effect. A common-gate (CG) stage is not subject to the Miller effect because of the grounded gate that absorbs the drain signal feedback. On top of that, CG stage presents a low impedance to the common-source (CS) stage ($1/g_{m\,CG}$), making the voltage gain of the CS stage, A_v ($= g_{m\,CS}/g_{m\,CG}$), in the order of magnitude of 1 reducing the miller effect impact drastically. However, this topology's voltage swing suffers from twice the threshold voltage v_{knee} preventing from using the v_{ds} full swing.

Secondly, the differential aspect of the topology present also pertinent advantages [Raz00]. It naturally shows common-modes rejection. Hence, every order eventh harmonics (a potential 2nd harmonic content for instance...) are filtered out by differential structures such as baluns and transformers. Furthermore, it allows large voltage swing (two times the voltage swing of a single-ended structure, with the same supply voltage), high linearity, and simpler biasing that is enabled with the use of baluns. However, the baluns represent a large area and, special care must be given to the cold spots for proper common-mode rejection and avoid stability issues.

3.3.2.1.1 Transistor Sizing

The ST CMOS065SOIMMW technology is proposed with two transistor gate lengths of 65 nm and 40 nm with a 1 V operating voltage. However, as an experiment for reliability concerns, we have exceptionally considered $V_{DD} = 1.2\,\text{V}$ for 65 nm transistor counting for relaxation with L.

A metric to determine the potential of a technology node for power amplifier design is to use the frequency transition f_T determined by (3.10).

$$f_T = \frac{f}{\text{Im}\left\{\frac{1}{h_{21}}\right\}} = \frac{g_m}{2\pi(C_{gs} + C_{gd})} \tag{3.10}$$

With h_{21} a term of the hybrid matrix represents the current gain of the two-port network with its output shorted, and f the extraction frequency. f_T is the frequency for which the current gain of a transistor is null. A rule of thumb for proper technology choice is to verify that the maximum f_T reach at least five times the frequency of operation. The transition frequency is extracted and shown in table 3.1 for both technology nodes on a unitary reference cell of 10 fingers of 1 µm width. From these results, we see that the 40 nm transistor shows a $f_{T\,max} \gg 5f_{operation}$ and that the 65 nm node shows a $f_{T\,max} > 5f_{operation}$. The 65 nm node is then sufficient and also allows a larger voltage swing with a higher V_{DD} while the smaller

l	$N_f \times W$	V_{DD}	$f_{T\,max}$
40 nm	$10 \times 1\,\mu m$	1 V	293 GHz
65 nm		1.2 V	193 GHz

Table 3.1. Transition frequencies for 40 and 65 nm of ST CMOS065SOIMMW technologies

transistor is more suitable for applications around 60 GHz.

Losses between 1 and 2 dB are expected from the matching circuit. Consequently, the transistors are sized for an OCP1 of 20 dBm.

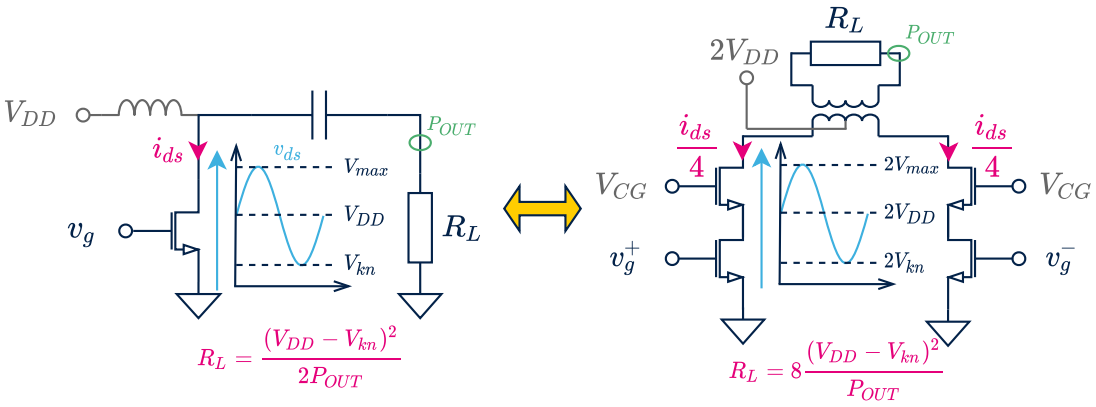


Figure 3.12. Load sizing from a common source amplifier to a differential cascode amplifier

A rough estimation of the PA sizing can be extracted from basic voltage/power considerations. If we look at the common source prototype in figure 3.12, assuming that the voltage waveform is a sine wave, the output power can be expressed as:

$$P_{out} = \frac{V_p^2}{2R_L} \tag{3.11}$$

(2 stands for a peak to root mean square (RMS) conversion).

Assuming that 1 dB compression point is reached when V_{DS} equals V_{kn}, $P_{out\,OCP1}$ is equal to:

$$P_{out\,OCP1} = \frac{(V_{DD} - V_{kn})^2}{2R_L} \tag{3.12}$$

Which can also be written as:

$$R_L = \frac{(V_{DD} - V_{kn})^2}{2P_{out\,OCP1}} \tag{3.13}$$

In other words, the load real part is determined by V_{DD} (technology characteristics related to reliability) V_{kn} (that will depend a little bit on transistor sizing but will be in the order of magnitude of 200 mV), and the output compression point. Moving from a common source to differential cascode structures (figure 3.12), the formula can still be used, considering differential structure as two branches delivering $P_{out}/2$ each on $R_L/2$, and the cascode as a device capable of sustaining twice the voltage, with twice the V_{kn}. Finally:

$$\frac{R_L}{2} = \frac{2(V_{DD} - V_{kn})^2}{\frac{2P_{out\,OCP1}}{2}} \tag{3.14}$$

And thus:

$$R_L = 8\frac{(V_{DD} - V_{kn})^2}{P_{out\,OCP1}} \tag{3.15}$$

Using our hypothesis ($V_{DD} = 1\,\text{V}$, $V_{kn} = 200\,\text{mV}$, $P_{out\,OCP1} = 0.1\,\text{W}$) bring us to a load resistance of 51.2 Ω.

Of course, this value is a starting point for the design and has to be refined with circuit simulations

The differential cascode architecture allows us to synthesize a reasonable load R_L. If we had chosen the common source topology, R_L would represent only 3.2 Ω.

A unitary cell with 10 fingers of 2.5 µm is used to determine the transistor size. The biasing current is chosen from this cell and then scaled to the size of the final transistor to deliver the correct output power. We set I_{DS} so that the transistor exhibits a maximum transition frequency f_T; furthermore, a margin of 10 % is taken relative to the peak to anticipate further process/temperature (figure 3.13).

The final transistor core is composed of a pair of 350 µm transistors with 140 fingers of 2.5 µm width. The current per finger is about 180 µA with $V_{GS} = 640\,\text{mV}$.

3.3.2.1.2 Balun Design

We have designed a test structure to enable on-wafer measurement of the circuit with Ground-Signal-Ground RF probes. The differential circuit is matched to the single 50 Ω impedance of the measurement probes with the help of balun transformers. The ST CMOS065SOIMMW BEOL stack comprises 7 metal layers, including 2 thick layers; this feature allows the design of

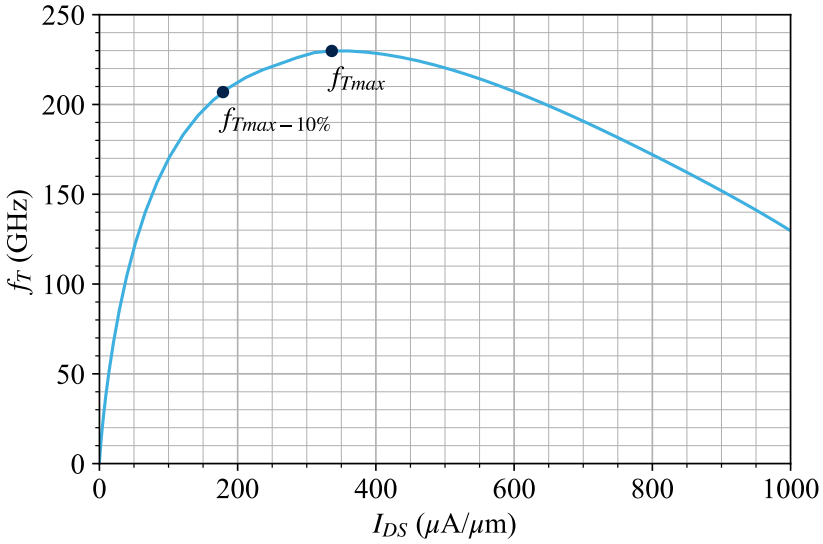


Figure 3.13. Transition frequency versus current density

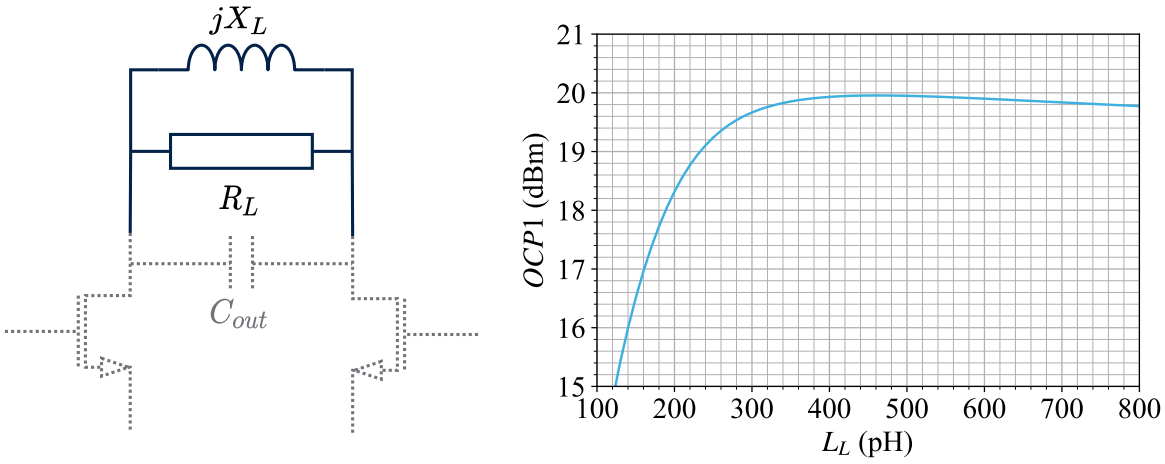
baluns with primaries and secondaries on two different layers for high coupling coefficient and low losses. Furthermore, the fringe capacitors from the design kit using metal layer 2 (M2) to metal layer 4 (M4) are used to adjust the impedance transformation. Particular attention is given to the size of the capacitances; they must be kept small to preserve their high-quality factor and minimize the losses of the matching circuit.

The input balun has to match the 50 Ω single probes to the gate parallel differential 110 Ω impedance. The primary is designed in metal layer 7 (M7) and the secondary with metal layer 6 (M6), the gate polarization V_{G1} is connected to the secondary center tap. Given the impedance transformation ratio, an octagonal 1 turn to 1 turn shape is enough to obtain the correct matching.

The output balun is designed to present on its primary a complex differential load Z_L (3.16) composed of a real part R_L (??) in parallel with an imaginary inductive part $X_L = jL_L\omega$.

$$Z_L = \frac{1}{R_L} + \frac{1}{X_L} \tag{3.16}$$

L_L is swept (figure 3.14b) and chosen to ensure a complete transfer of output power by fully compensating the imaginary capacitive part of the output of the transistors. As a result, the output capacitance is fully compensated for $L_L = 460$ pH.



(a) Output impedance representation

(b) OCP1 vs. compensating inductor L_L

Figure 3.14. Output impedance (a) and L_L sweep for optimal output power (b)

We have a parallel impedance composed of $R_L = 51.2\,\Omega$ and $L_L = 460\,\text{pH}$ which transformed in series impedance gives:

$$R_{L\,series} = \frac{R_L}{1+Q^2} \tag{3.17}$$

And:

$$L_{L\,series} = L_L \frac{Q^2}{1+Q^2} \tag{3.18}$$

With:

$$Q = \frac{R_L}{L_L.\omega} \tag{3.19}$$

With ω at 28 GHz that finally gives:

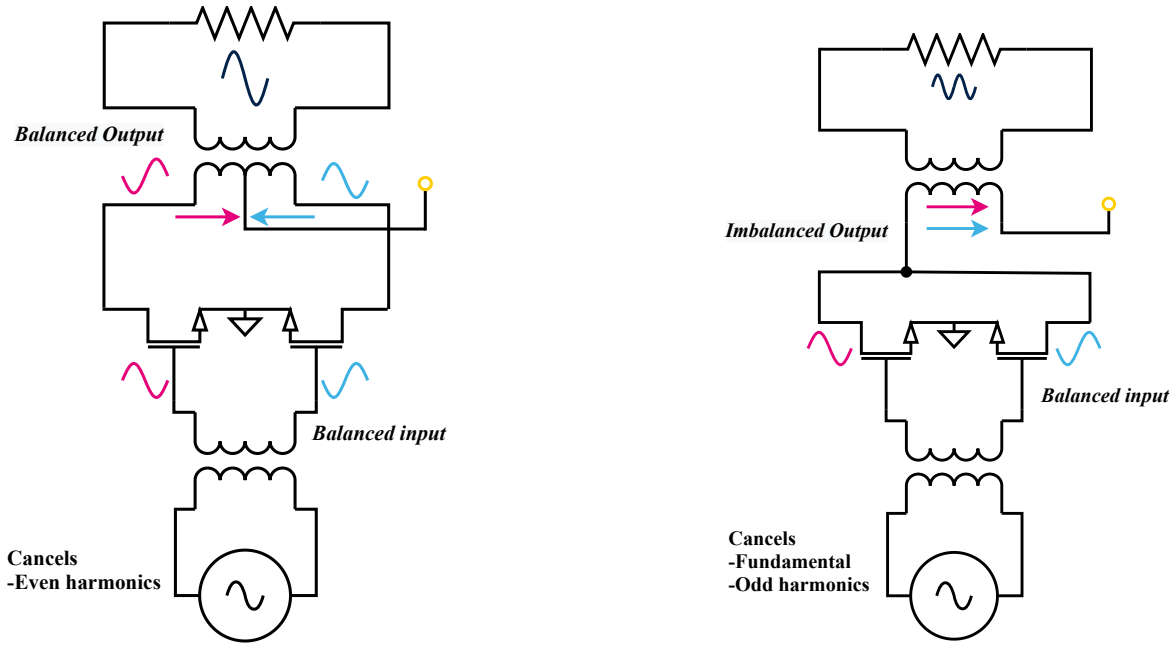
$$Z_{L\,series} = R_{L\,series} + j\omega L_{L\,series} = 36.6 + j23.1\,\Omega \tag{3.20}$$

$Z_{L\,series}$ is then transformed to the 50 Ω output probe impedance. A rectangle shape is chosen; the primary is made of a single turn of M6 with the drain DC polarization V_{DD} access connected to its center tap. Another single turn of M7 is used for the secondary.

With an EM simulation with Keysight Momentum, each input and output balun represent around 0.7 dB of losses.

3.3.2.2 Second harmonic source design

Among the solutions to generate a second harmonic, one that allows good integration with a power amplifier is the "Push-Push" amplifier. It is primarily known in oscillator literature for its frequency multiplication property. Its principle can be compared to the differential amplifier as shown in figure 3.15: The differential amplifier (figure 3.15a) is a differential architecture with a balanced input and output. Each transistor alternately supplies current to a connected load, and, as already said, it naturally cancels the even harmonics at its output. On the other hand, the push-push (figure 3.15b) consists of the same pair of Gm cells but with balanced input and unbalanced output with both drains connected to the same node. When a differential signal feed the transistors, the even odd-order harmonics (including fundamental) being 180° out of phase cancels when even order harmonics (especially h2), being in phase, are summed. The push-push amplifier is usually characterized by its conversion gain.



(a) Differential (b) Push-Push

Figure 3.15. differential and Push-Push amplifier architectures

Due to its unbalanced output, the push-push amplifier (PP) is doubled, and the two outputs are connected to each drain of our differential PA. It allows direct second harmonic injection through current recombination in the non-linear impedance represented by the PA.

3.3.2.2.1 Transistor sizing

The PP is designed to have a maximum conversion gain, delivering the maximum 2nd harmonic current while consuming the minimum DC current to minimize the degradation of the circuit efficiency. To do so, the transistor size is chosen to generate the maximum IMD3 control with consumption that remains under 25 mA; representing around 16 % of the PA consumption. As a result, the final push-push architecture is realized with 2 pairs of 37.5 µm wide transistors and shows a maximum conversion gain of −10 dB for $P_{in} = 1$ dBm.

3.3.2.2.2 Balun Design

As we are dealing with much smaller transistors, the impedance to match is larger (small C_{gs}), making the balun more challenging to design due to the high impedance transformation ratio. Indeed, 1 kΩ must be matched to 50 Ω. To help with the balun design, a parallel resistor of 477 Ω between positive and negative input is added to reduce the impedance seen to around 323 Ω. The transformation ratio is finally achieved with a two-turn balun.

3.3.2.3 Biasing circuit design

The bias current in the two PA sections are generated using current mirrors whereas the cascode voltage (V_GC) is generated by the current crossing two diode-connected MOS transistors. Reference current inputs are connected to pads and were generated by lab requirements. The principle of the current mirror (figure 3.16) resides in a diode-mounted transistor coupled to another MOS gate. We can express the currents flowing into the drains of transistors M_1 and M_2 as (3.21) and (3.22).

$$I_{REF} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_1 (V_{GS} - V_{TH})^2 \tag{3.21}$$

$$I_{OUT} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_2 (V_{GS} - V_{TH})^2 \tag{3.22}$$

With C_{ox} the oxide capacitance and μ_n the electron mobility. As both transistors share the same V_{GS} and the same die, they undergo the same process and temperature variations (affecting C_{ox}, μ_n, and V_{TH}). Hence the process, voltage, temperature (PVT) effect is canceled.

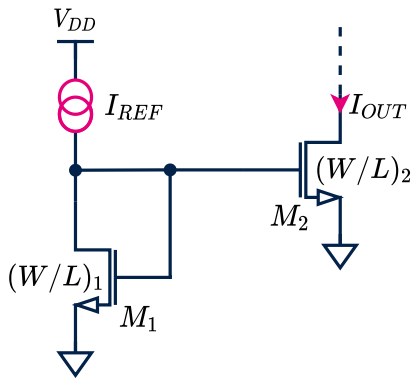


Figure 3.16. Current Mirror principle

With M_1 and M_2 of different sizes, the current I_{OUT} can be expressed as (3.23):

$$I_{OUT} = K.I_{REF} = \frac{(W/L)_1}{(W/L)_2} I_{REF} \tag{3.23}$$

For the CS bias of the PA and PP (see figure 3.17), the size of transistors M_1 and M_2 (W/L) are identical and, their current I_{OUT} is adjusted with I_{REF}. The ratios for the MOS size is are shown in table 3.2.

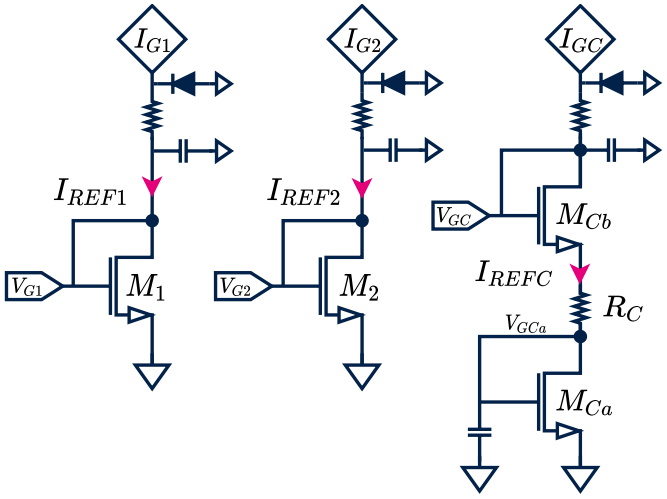


Figure 3.17. Current Mirror biasing for PP and PA

The desired cascode common gate voltage is $V_{GC} = 1.95\,\text{V}$ (chosen to have a balanced v_{ds} swing between the CS and CG stages). Hence, the current mirror cannot simply be made with two identical stacked diode-mounted transistors as $V_{GCa} \neq V_{GCb}$. To achieve the V_{GC} targeted, V_{GCb} must be equal to $V_{GC} - V_{GCa} = 1.95 - 0.7 = 1.25\,\text{V}$. It could

be obtained by reducing the ratio W/L of M_{Cb}; however, it results high stress of the MOS (with $V_{Cb} > VDD$!). A resistor R_C is then added to the source of M_{Cb} to lower V_{GCb}. A compromised have been found for: $R_C = 40\,\Omega$ and $M_{Cb}(W/L) = 125$. In addition, as the DC

	PA CS	PP CS
K	28	6

Table 3.2. Current mirrors' MOS size ratio

probes providing the biasing currents are directly connected to transistor gates, protection is implemented with a diode and a resistor against any electrostatic discharge (ESD) that might occur during measurements. A general schematic with each piece put together is shown in figure 3.18.

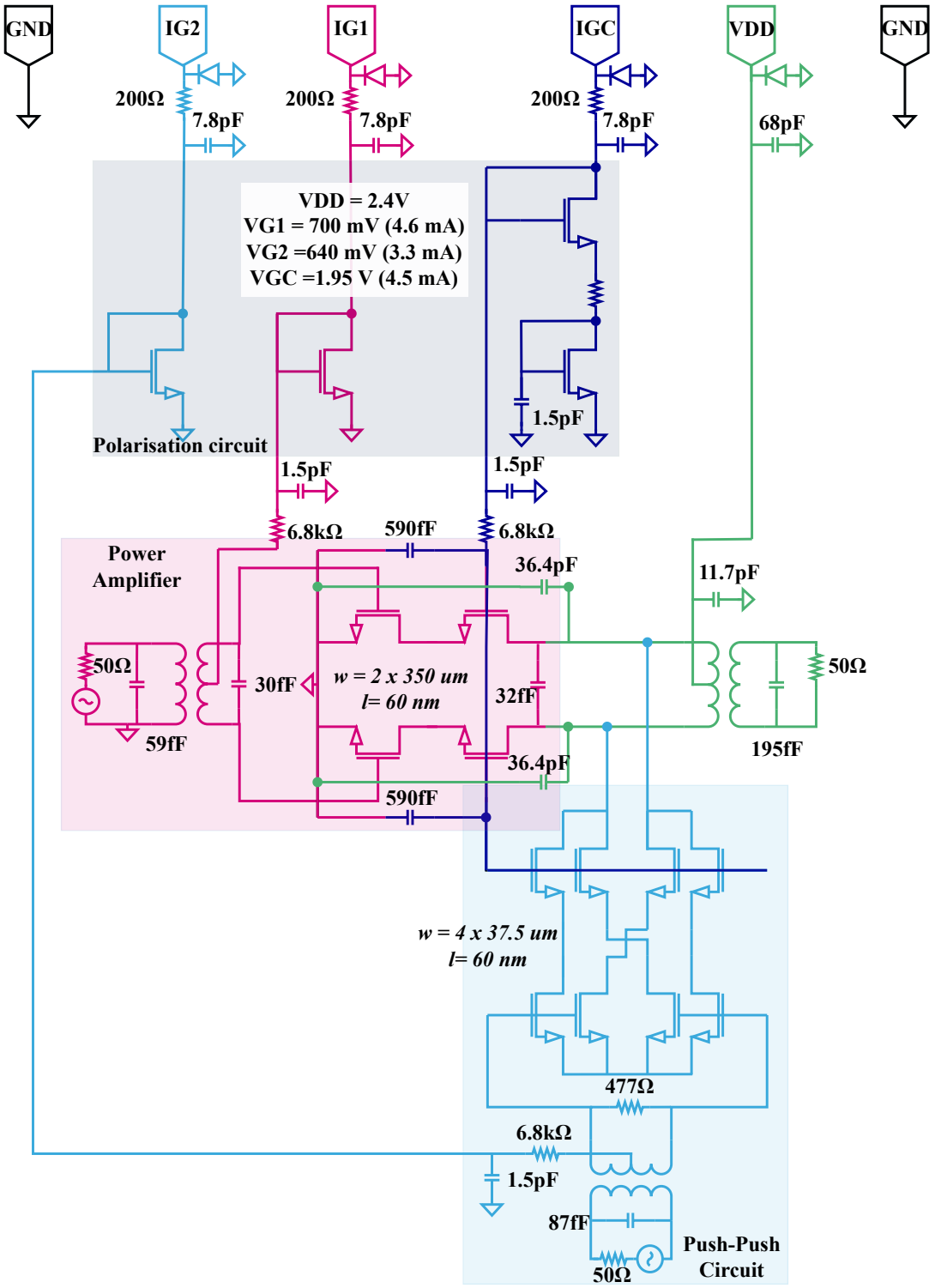


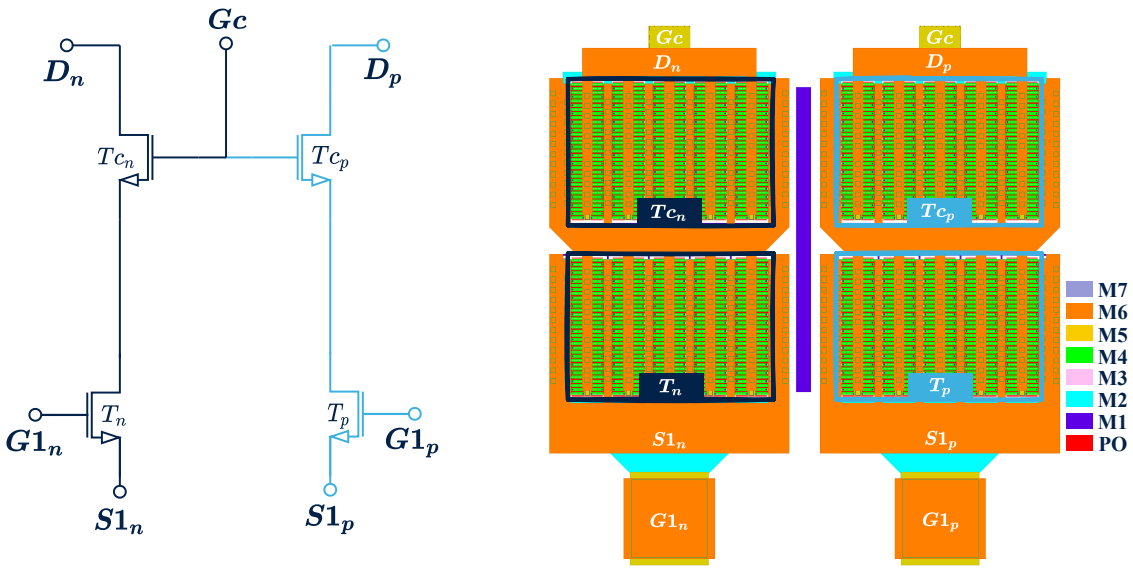
Figure 3.18. Full IMD3PSPA schematic

3.3.3 Circuit Layout

As the circuit architecture is composed of two inputs and one output, the strategy is to preserve the PA input and output on the same axis to minimize the length of the interconnections on the RF path that carry the power.

3.3.3.1 Transistor cores

For each core (PA and PP) the transistor's drain and source fingers are connected with stacked layers and vias from metal layer 1 (M1) to M2, reducing the contact resistance but at the cost of higher C_{DS} capacitance.



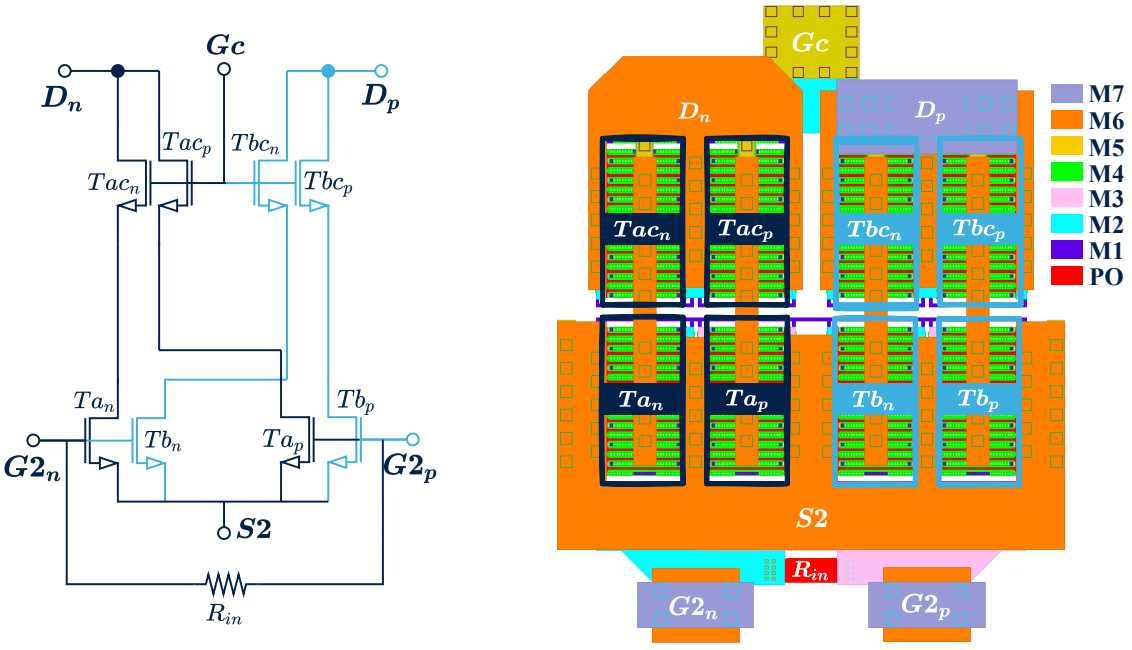
(a) (b)

Figure 3.19. Schematic (a) and layout (b) disposition of differential PA

Multiple aspects have been carefully taken into account in the design of the transistors' access layout. First, each transistor has been subdivided into five sections to maintain a small x/y ratio to avoid a critical dispersion of the active core that will result in long connections and high losses. The final structure measures $25 \times 38\,\mu m$. Each finger is connected to a stack of thin metal layers from M1 to M4 to connect other thick layers. Concerning the common-source transistors, their sources and drains are linked with the thick M6 to the ground and the common-gate MOS source. Their gates are connected to the input matching through layers M1 and M2 that join the matching capacitor with a whole stack of vias reaching the M6 input balun secondary. On the other hand, the common-gate transistors have their drain connected

to the primary output balun through the M6 layer. Their gate is connected to the biasing through the medium metal layer 5 (M5) layer.

For reliability and electromigration concerns, a certain number of vias are required by the design rules manual (DRM) for connections that must flow significant currents. According to the DRM, the design can sustain an operating temperature of 125 °C with the amount of current and the number of vias implemented.



(a) (b)

Figure 3.20. Schematic (a) and layout (b) disposition of the double push-push architecture

The PP core is arranged to simplify all the crossed connections according to the schematic from figure 3.20a: The input path crosses to keep the output drains straight for fewer interconnections, as shown in figure 3.20b. The common-sources gates are connected with M2 for the negative branch of the differential architecture and with metal layer 3 (M3) for the positive branch, allowing a cross-over. The source and drain connections are insured with M6 links. The output drains are connected on two different thick layers to enable cross-connections to the PA and output balun nodes.

3.3.3.2 Input 1 and output layout

The input of the power amplifier (figure 3.21) is located on the west part of the circuit and is composed of the designed balun (a) associated with its two matching capacitors (c_1) and

(c2). The transformer's primary is connected to the ground-signal-ground (GSG) pad (b) and the secondary to the input of the transistor (f). The biasing circuit (d) (see section 3.3.2.3) is situated at the north of the input balun creating the polarisations VG_1, VG_2, and VGC are respectively supplied by the M5 (d1), (d2) and (d3) access. VG_1 is connected to the center tap of the balun's secondary while VG_2 is crossing through the balun to reach the south input. Both voltages are decoupled with capacitors (e1) and (e2), respectively.

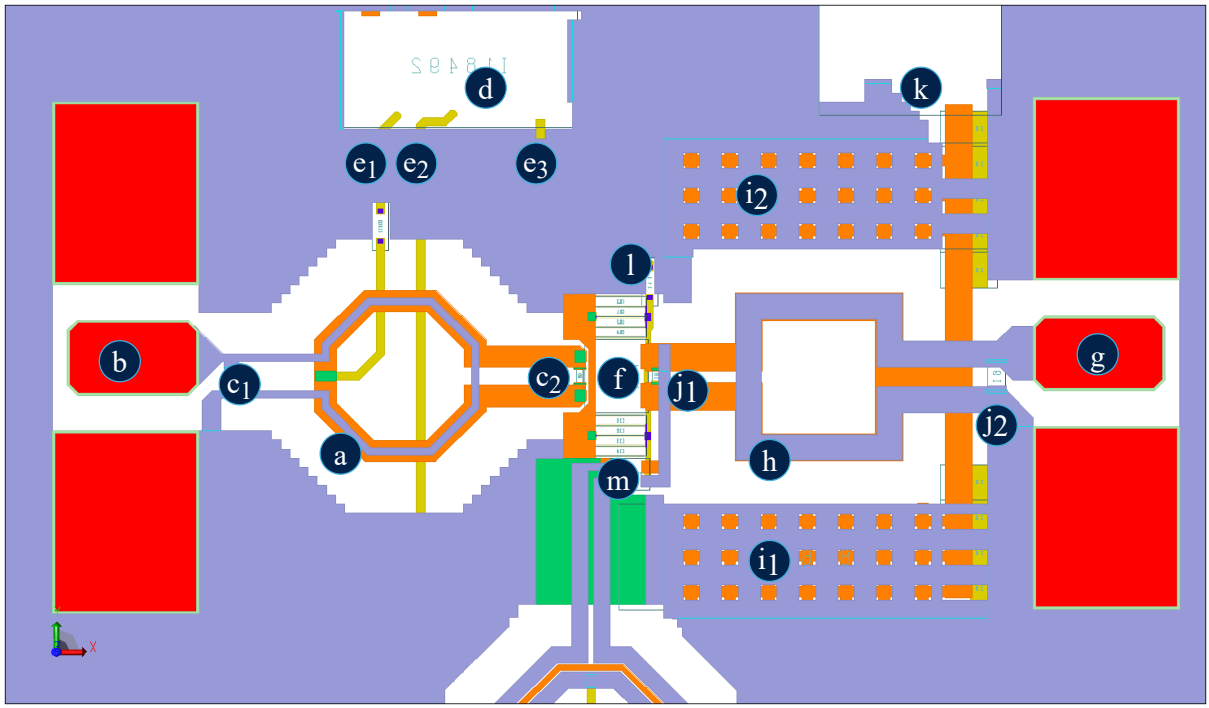


Figure 3.21. Input and output layout of the PA

On the east section of the circuit is located the output composed of the output pad (g) and the transformer (h) associated with its two matching capacitors (j1) and (j2).

Two large banks of capacitor (i1) and (i2) are placed along the balun on the north and south sides to decouple the V_{DD} supply coming from (k). The banks are composed of multiple elementary blocks that each contains a large fringe capacitor with a high-quality factor dealing with the higher frequency decoupling and four small poly-silicon capacitor with lower Q that deals with lower frequency decoupling.

The drains of the transistor (f) and of the double PP circuit (m) are connected to each branch of the balun's primary. Finally, the common-gate polarization is distributed to the two active cores in M5 from (d3) to (l).

3.3.3.3 Input 2 layout

The second input feeding the second harmonic generation is located south of the circuit. Again, we find the balun (n) for the input matching of the push-push amplifier with its capacitor (p). The secondary center tap is connected to the polarization VG_2 through M5 and a decoupling capacitor (r), and then each branch of the transformer is connected to the double PP. In order to prevent any space issues between the RF probe body that will be used for measurements, a 50 Ω coplanar waveguide (CPW) line (q) using M7 stacked with alucap layer (AP) is used between the primary of the balun and the RF pad (o).

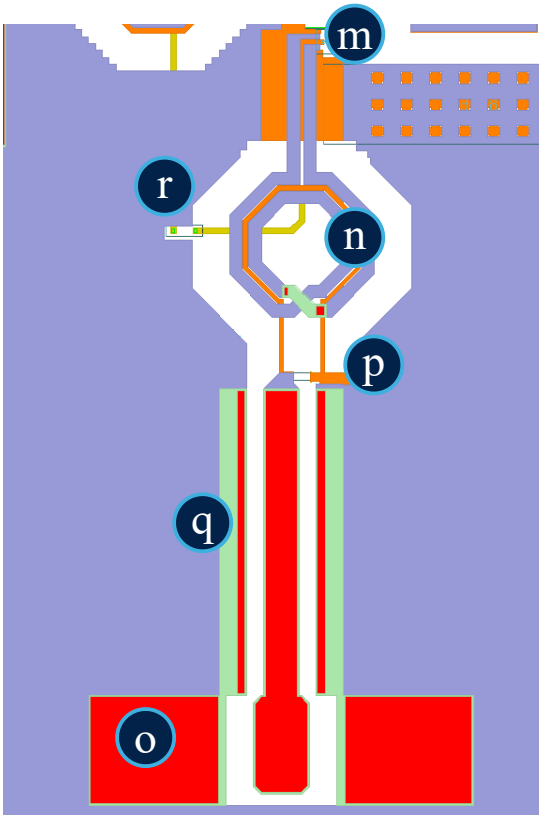


Figure 3.22. Second input layout

3.4 IMD3 PSPA simulation and measurements

An analysis via simulation has shown that the 65 nm transistor length version is sufficient for 28 GHz applications. It allows large enough transition frequency f_T higher DC voltage enabling larger v_{ds} voltage swing, which is crucial for power amplifiers. To confirm the results of this

analysis, the IMD3 PSPA has been designed with 40 and in 65 nm transistors. A standalone version of the power amplifiers (without PP) has also been implemented.

3.4.1 Design and simulation setup

The circuit is simulated and designed in Cadence environment. The different simulations, DC, S-parameters for small-signal, and HB for large and non-linear signals, are performed with Keysight Golden Gate simulator. Furthermore, the top of the circuit layout (including PLS extractions, design kit (DK) components, passive and plane ground filling elements) is netlisted with RFPro which is used as a cockpit that allows a circuit partitioning between PCells. RFPro uses model provided by the design kit, small area blocks, such as MOS transistor matrix, that are modeled using PLS flow, and long connections that are extracted by Momentum EM tool.

3.4.2 Measurement setup

The measurements are performed directly with GSG probes on a 300 mm wafer.

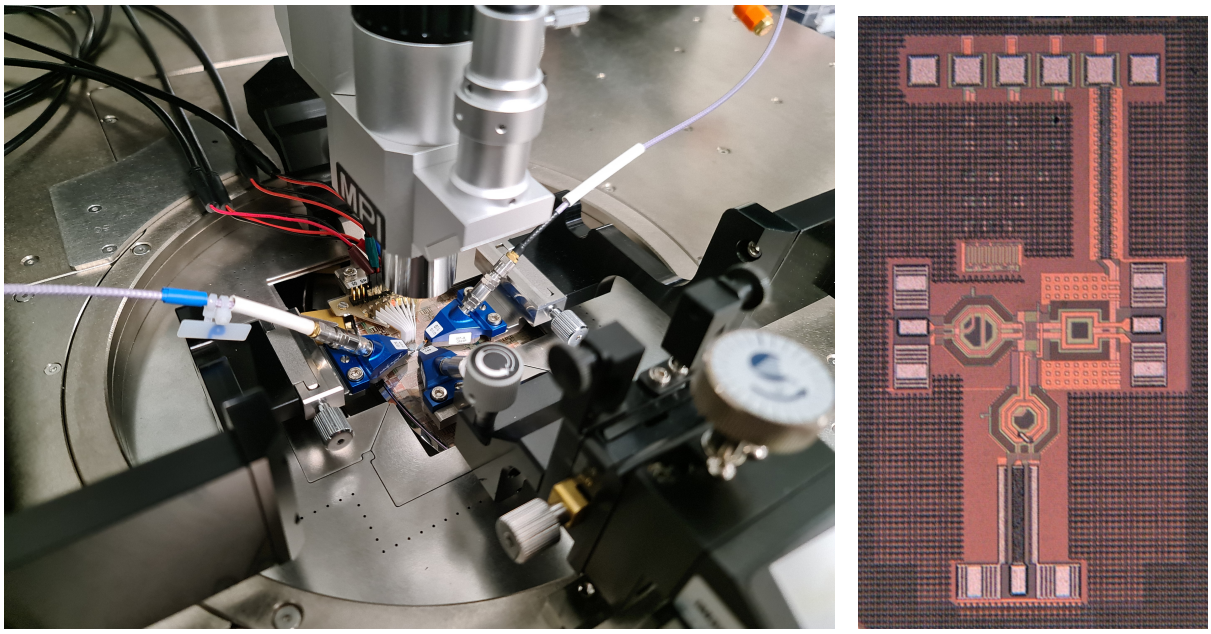
3.4.2.1 S-parameters

The S-parameters have been measured on a 67 GHz Keysight N5247A PNA-X network analyzer. The results have been obtained in two phases; first, the PA path has been measured in a two-port configuration, while the second input has been measured with a single port calibration. The input of the amplifier is fed by port 1, the push-push by port 2, and the output of the device under test (DUT) is measured by port 3.

The analysis is performed under 50 Ω between 0 and 67 GHz. The performances of the power amplifier within the DUT are evaluated with a sweep of power in single tone mode on input 1. The second input is loaded by 50 Ω. The output power is measured on port 3, and the DC power consumption is monitored.

3.4.2.2 Power sweep

The sweep of power has been performed with a Rohde and Schwartz 67 GHz ZVA, an amplifier ZVA-0.5W303G+ has been added at the input of the device to obtain sufficient power to cover the suited P_{in} sweep. The probes used are GS/SG 67 GHz |Z| probes from Cascade.



(a) (b)

Figure 3.23. DUT on two-port measurement with input 2 loaded (a) IMD PSPA die photograph (b)

3.4.3 S-parameter and power sweep simulation and measurement results

3.4.3.1 S-parameters

An S-Parameter simulation between 1 GHz and 67 GHz is performed of the 3-port circuit, and the essential parameters are plotted in figure 3.24.

The input of the PA is on port 1, the input of the PP is on port 2, and the output is on port 3. The simulation shows a small-signal gain S_{31} of 15.0 dB, a PA input matching S_{11} of 28.9 dB, and a PP input matching S_{22} of 22.6 dB. The output return loss S_{33} is relatively low (2 dB) as the output matching has been designed to present a load suited for power matching and not for small-signal matching. The PA also shows good isolation (S_{13}) with 33.2 dB.

The measurements results are plotted along with the simulation in figure 3.24. We first notice an overall shift of −2 GHz (around 7 %) of all the parameters toward lower frequencies. Furthermore, a more critical frequency shift is visible on the input matching S_{11} from 28 to 23 GHz (around 18 %). Consequently, the circuit shows at 28 GHz in measurement a lower gain with 10.8 dB, a PA input matching S_{11} of 5.8 dB, a PP input matching S_{22} of 22 dB, an output return loss S_{33} of 3.1 dB, and isolation (S_{13}) of 30.8 dB. Before continuing to

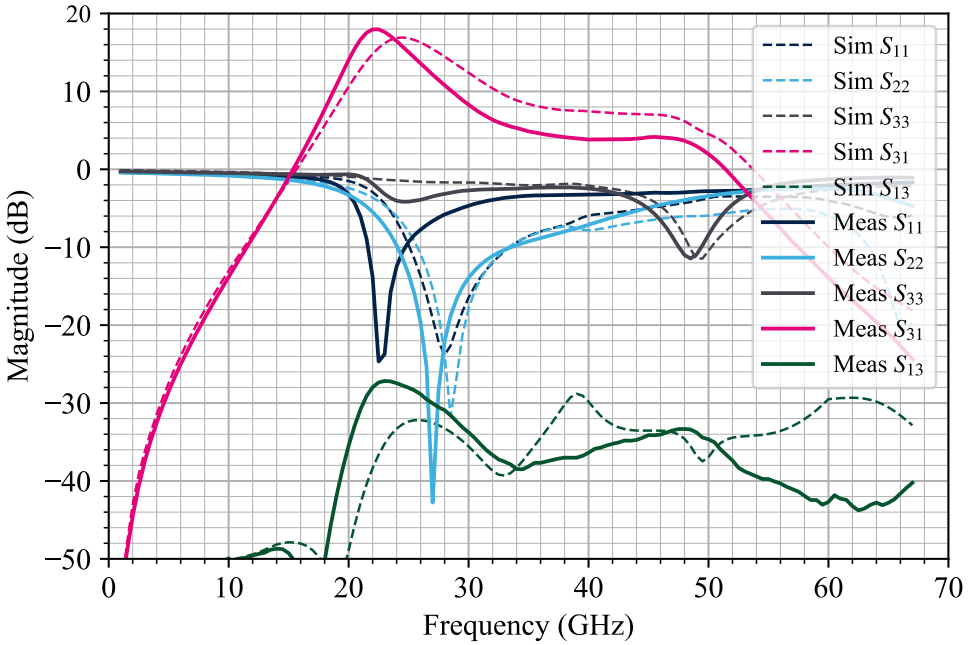


Figure 3.24. S-Parameters of the 3-ports circuit in simulation (dashed) and measurement (line)

power simulation and measurements, the mismatch between simulation and measurements is investigated to explain the different phenomenons observed.

3.4.3.2 Power analysis

The simulation results and the measurements of the power amplifier with $V_{DD} = 2.4\,\text{V}$,$I_{DD} = 159\,\text{mA}$, at 28 GHz with the push-push biased with $I_{G2} = 0$ (disabled) are shown in figure 3.25.

Small signal gain exhibits the same behavior as shown with S-parameters measurements (11 dB in measurements vs. 15 in simulation). This is consistent with shifted max gain frequency observed in S-parameters. OCP1 is slightly lower in measurements (17.7 dBm) compared to simulation (18.5 dBm). Lower measured gain results in lower PAE (12.6 % in measurement vs. 16.7 in simulation). As the OCP1 is close to the target, we can conclude that PA output load, presented by the output balun is close to the target (what is consistent with the good agreement between measured and simulated S_{22} at 28 GHz).

Different power sweep measurements has been performed at 23 and 25 GHz. Even though they show a higher gain due to the frequency shift, we have chosen to keep working at 28 GHz

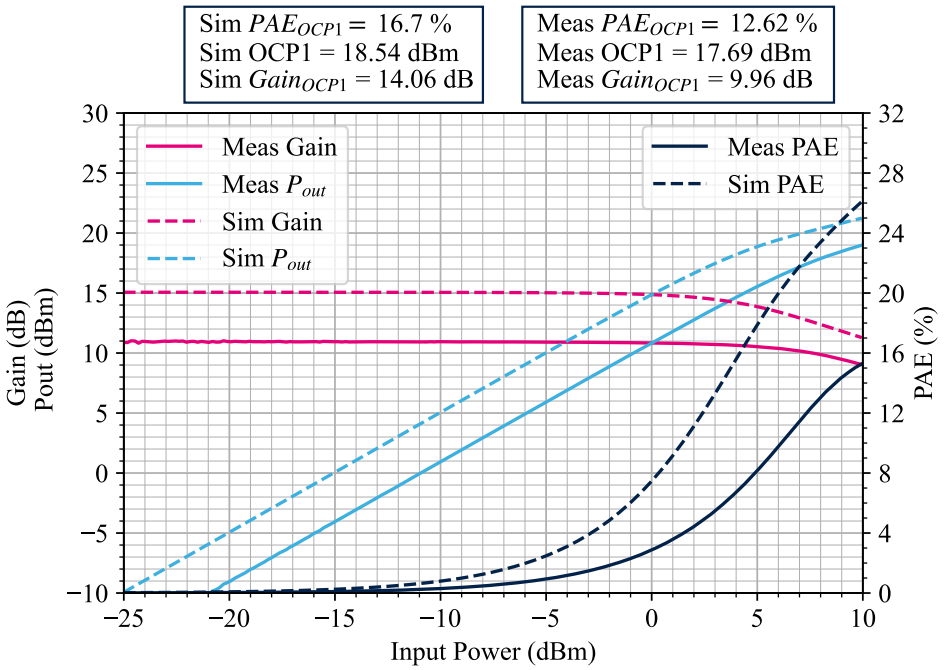


Figure 3.25. Gain, Pout, and PAE vs. input power

as it respects the only specification given (OCP1) and better coherence with the rest of our work.

Anyway, the small and large signal results show some flaws in our simulation bench. Therefore, an analysis is proposed before proceeding to IMD results.

3.4.3.3 Simulation-measurements mismatch

As said earlier, it is essential to remember that the technology used is in its early stages of maturity. First of all, the contributions that have been measured from the same multi-project wafer (MPW): The other designs have shown an overall shift of -7 % in frequency of S-parameters. It is visible here through the 2 GHz shift in our contribution. To identify the contributors to the mismatch observed, we proceeded by identifying the potential mistakes made in the setup of the EM simulation, then by verifying the elements provided by the design kit, BEOL, and the MOS models were aligned with the process of fabrication.

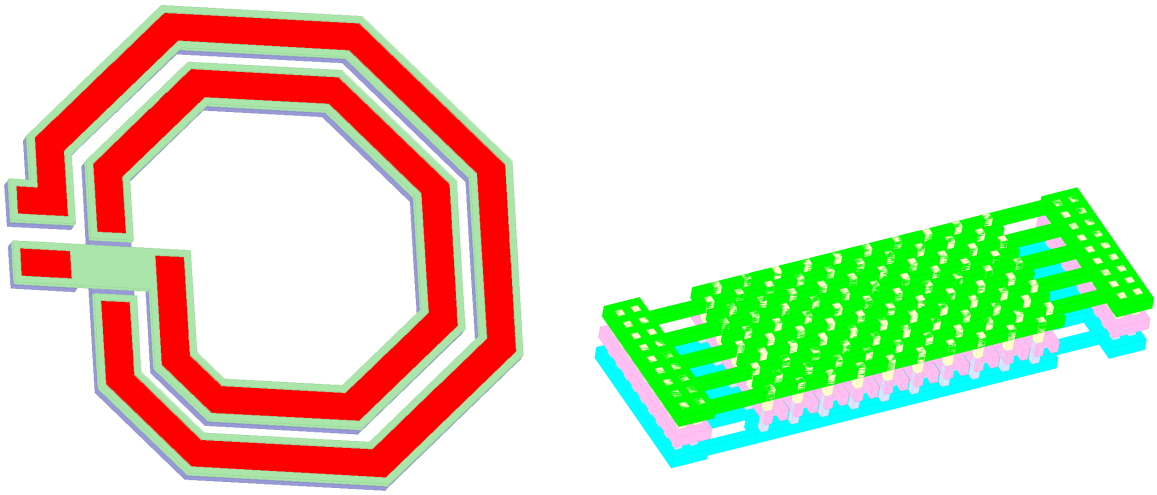
3.4.3.3.1 Debug: Simulation setup

The simulation has been verified with other designers and set up according to Keysight engineers' recommendations. A few parameters must be chosen carefully: First, the frequency sweep must contain a point at 0 Hz for the convergence of DC simulation. An adaptive sweep must

be performed between 0 and a maximum frequency covering at least 3 harmonics. A linear sweep is added around 28 GHz with 100 MHz steps for IMD simulations. On top of that, the Momentum matrix solver must be in direct dense mode for faster convergence, with a mesh density of 50 cells per wavelength and a mesh of 1 µm width is forced around the edges. As Momentum is a "2.5D" solver, the BEOL can be described with 2D or 3D models for thick conductors and vias; the conductors are chosen to be modeled with a 3D physical model and the vias by 2D models with an exception for the via layer between the top tick metal layer M7 and the aluminum AP layer. The simulation bench used for the design have been validated at this stage; it confirms that the EM simulation methodology is correct and that we can eventually reuse it for retro-simulations.

3.4.3.3.2 Debug: Back end of the line

To confirm the BEOL description that is used for the EM simulation, an analysis has been performed on known devices such as an inductor (figure 3.26a) and a metal-oxide-metal (MOM) fringe capacitor (figure 3.26b) derived from the design kit.

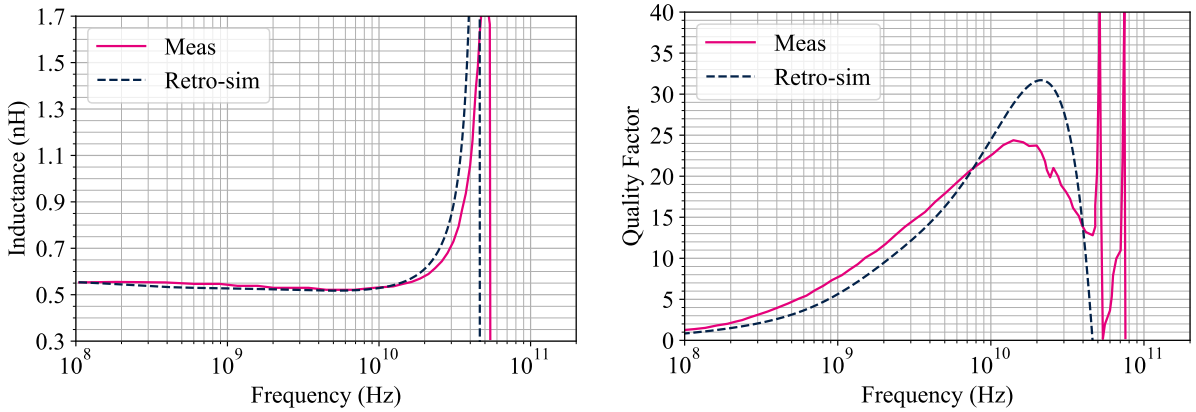


(a) (b)

Figure 3.26. Two turn inductor (a) and MOM fringe capacitor (b)

Concerning the inductor, some measurements have been performed on structures that have been embedded on a previous MPW. A measured A two-turn 500 pH coil of 12 µm width and 80 µm diameter has been retro-simulated with the same EM setup used for our circuit design. By comparing the measurement and the retro simulation of the inductor in value L

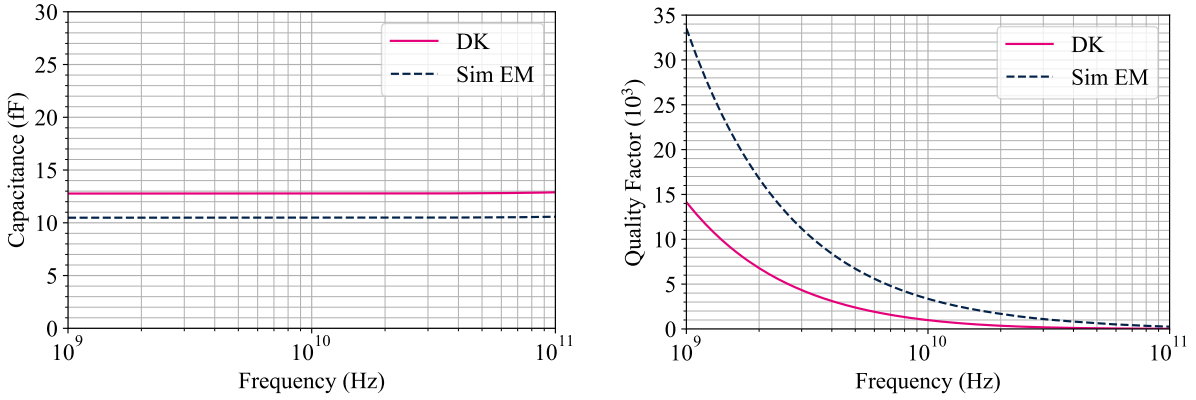
(figure 3.27a) and Q (figure 3.27b), we see that we have a good fit between both data, which validate the EM simulation bench and the correlation of the BEOL model with the process of fabrication.



(a) (b)

Figure 3.27. Measured and retro-simulated Inductance (a) and Q factor (b)

As no standalone MOM capacitors are available for measurement, the capacitor that composes the input matching is studied by comparing their model (issued from the design kit) with an EM simulation performed with the setup that the inductor study has validated. Due to the complex geometry of fringe capacitors, a small capacitor is studied in order to limit the EM simulation time and improve the simulation accuracy. It is made of M2, M3, and M4 metal layers and, according to the DK model, has a capacitance of 12.8 fF.



(a) (b)

Figure 3.28. Model and retro-simulated Capacitance (a) and Q factor (b)

The results show in figure 3.28 that as the capacitance is small with a small inductive part, it presents low variation and no resonance up to 100 GHz. We observe that the model is

around 17.9 % higher in terms of capacitance (figure 3.28a) than the simulation that shows 10.5 fF instead of the 12.8 fF expected. Also, the DK model is somewhat pessimistic in terms of quality factor (figure 3.28b) as it considers the potential process variations on the thin metal layers.

As we are trying to explain a shift in frequency of a matching network of about 18 %, we need to put a scale on the variations observed on the element composing the matching network: With a network consisting of inductive and capacitive components, let's consider, based on what we have seen, that the inductive part of the balun is well aligned with the measurement. An 18 % variation on the resonant frequency would represent almost 50 % of the take on the capacitance value. Then, it is safe to say that the potential variation observed on the BEOL and the capacitors are slight in front of the frequency shift at stake.

3.4.3.3.3 Debug: Front end of the line

Before diving into the MOS transistor analysis, a bit of chronological elements needs to be mentioned to understand our approach. As the ST CMOS065SOIMMW technology is under development, its design kit elements get regular updates over time. The circuit had been designed with a DK version that we call "A". By the time the circuit was processed and the first investigations started on the measurement results, a version "B" of the design kit has been released with some updates affecting only the active devices. With this update, the S-Parameters are re-simulated and shown in figure 3.29. The new simulation shows an overall closer fit with the measurements. It confirms that we had a good model of the BEOL but some maturity was yet to be gained on the modelization of the FEOL. Indeed, we have a better alignment of the small-signal gain and of the second input S_{22} matching around 26 GHz, also a slight improvement of S_{33}. Finally, the input matching S_{11} shows a much closer resonance compared to measurements around 24 GHz

While the S-parameters seem to have caught up a bit with the measurement, the fit is far from perfect, especially for S_{11}. We also notice some differences in the peak gain amplitude and on the output matching that need to be studied. An investigation has been conducted in Appendix A and revealed important changes, especially on the transistor model's input capacitance C_{gs}, which is responsible for the changes in the input matching circuit.

With this new design kit, a new sweep of power is shown in figure 3.30. Similarly, with the S-parameters, we have a better correspondence of the measurements with the simulation. The mismatch of the small-signal gain is noticeable with 2.17 dB less in measurement, lowering the PAE from 15.7 % to 12.6 %. Furthermore, we have a closer fit of the OCP1 with a difference of 0.46 dBm with measurements.

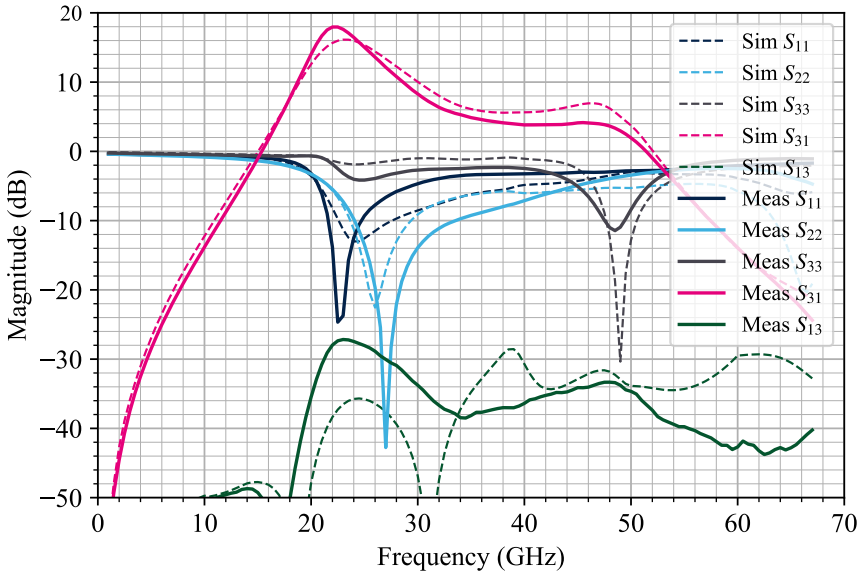


Figure 3.29. S-Parameters of the 3-ports circuit in simulation with DK B (dashed) and measurement (line)

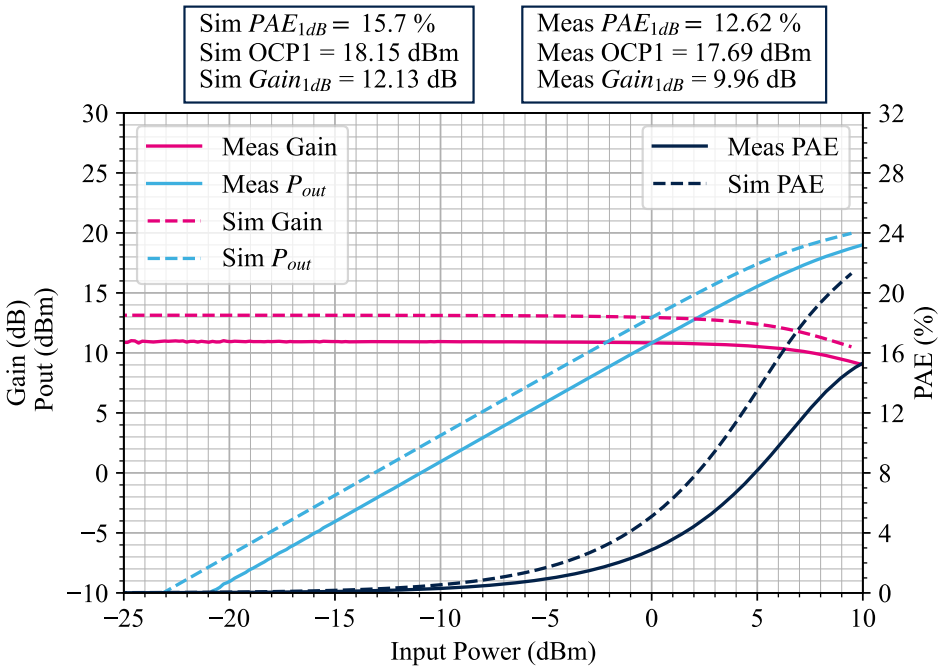


Figure 3.30. Gain, Pout, and PAE vs. input power in measurement and updated DK B

3.4.4 3rd order intermodulations phase shift simulation and measurement results

3.4.4.1 Measurement setup

One of the main challenges of this work resides in the measurement of the IMD3 PSPA. First, the measure of nonlinearities in phase requires non-linear vector network analyzer (NVNA) capabilities that are usually not embedded in standard network analyzers without extra options. A device able to provide a bi-tone signal is also needed. The setup proposed is based on the 2-port Rohde & Schwartz ZVA 67 GHz. As every network analyzer, it is able for each port

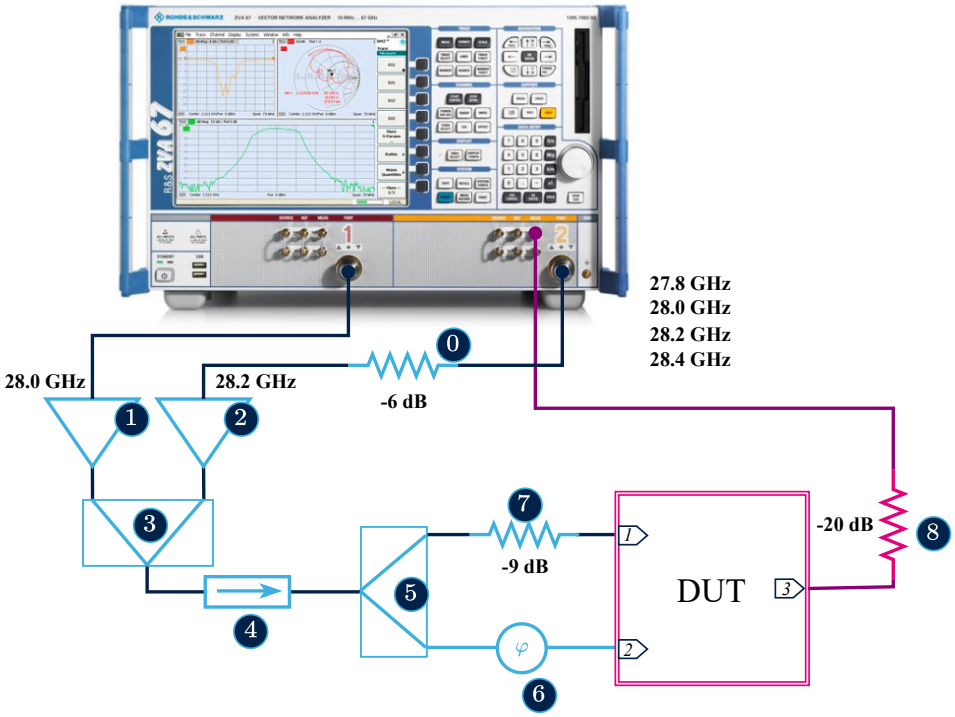


Figure 3.31. Measurement setup of the IMD3 PSPA with an R&S ZVA 67 GHz

n to generate and measure waves a_n and b_n. For our purpose, both port sources a_1 and a_2 are delivering the two tones we need while the port 2 receiver measures the DUT outgoing wave b_2. Both sources are coherent in phase between themselves to derive the phase at each frequency measured; then, an image of the spectrum is visible in magnitude and phase. The measurement setup is shown in figure 3.31 and assembled in figure 3.32: As the ZVA power sources are limited and do not meet our requirements in terms of level of power, each tone is amplified by two amplifiers (BNZ3498 BZY ① and ZVA-0.5W303G+ ②). The two amplifiers being different, they do not show the same gain with 24 dB for the BNZ3498 and 30 dB for

the ZVA-0.5W303G+, then a 6 dB attenuator ⓪ is added before ② to balance the amplitude of the tones. Each tone must be amplified independently to prevent any generation of IMD3 by the setup of measurement. Next, the tones are combined (ZN2PD-K44+ ③) and then split (ZN2PD-K44+ ⑤) to reach our device's two inputs. An isolator (MTC K5156VFF④) is inserted between the splitter/combiner to prevent any returning wave due to imperfect matching of the following devices in the chain that would recreate IMD by returning to the PAs. In order to ease the independent control of the power of input 2 (second harmonic generation) the output power of the ZVA is kept constant: A fixed attenuator ⑦ (actually composed of 3 dB and 6 dB attenuators) is placed on input 1 so that the amplifier is working at constantly at 7 dB of back-off (BO) (P_{avg}). On the second input is placed the phase shifter (PE82P1000 ⑥) for IMD3 control. An attenuator ⑦ ensures the control of the level of power of the injection. The latter was originally intended to be replaced by a variable attenuator RFVAT0040A30 for a more flexible injection power control. Unfortunately, its linearity was not sufficient and was polluting the DUT input signal with strong IMD content forcing us to work with constant injection power, limiting the capabilities of the bench. The original plan was to measure the phase control with different input amplitudes. This will not be possible with this setup. Finally, the output of the DUT is measured on port 2 of the ZVA through a 20 dB attenuator to avoid hitting the maximum power tolerated by its receiver or creating any nonlinearity (compression, IMDs) that can interfere with the PA non-linearities. Ideally, a selective filter would have been used to remove the harmonics generated by the amplifiers in the input chain. Still, unfortunately, the bench is already facing limitations in terms of power level that would be even more important with the additional losses due to the filter. More illustration and an extended list of equipment are shown in appendix B. To summarize, the bench as it is now will enable us to do the measurements. In an ideal world, we would have chosen a source with a higher power to have more options to add filters/attenuators and variable attenuators to have full flexibility for input power selection (on both inputs).

In terms of phase calibration, the DUT is first measured under a bi-tone excitation without injection by the ZVA, the spectrum with the tone and IMD3 are saved. Then, the spectrum, amplitude, and phase of fundamental tones and IMD3 products will be reused as a reference when the injection is active. All the values presented in the following sections will be relative values (delta amplitude and delta phase) to this reference.

3.4.4.2 Measurement results

With the setup described in section 3.4.4.1 put in practice, the effect of second-harmonic injection on IMD3 is simulated and measured. The phase shifter used is characterized manually

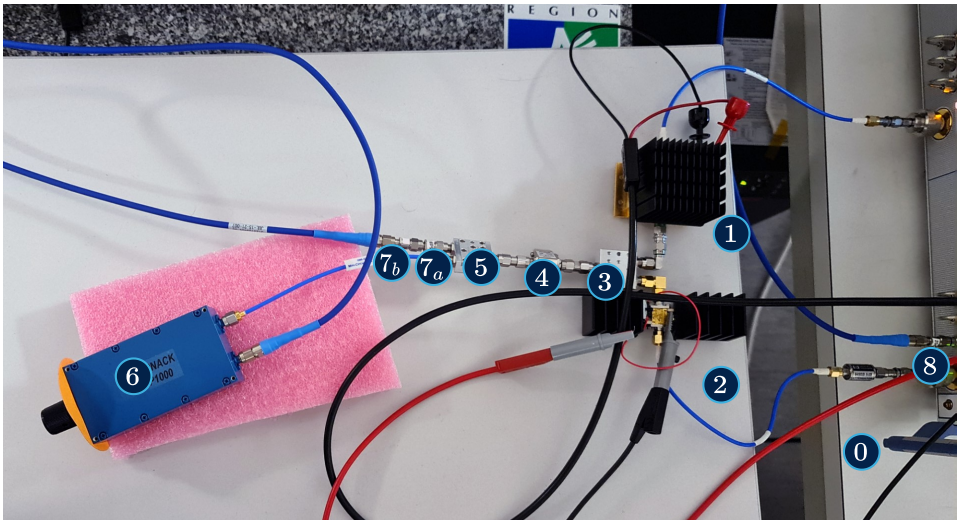
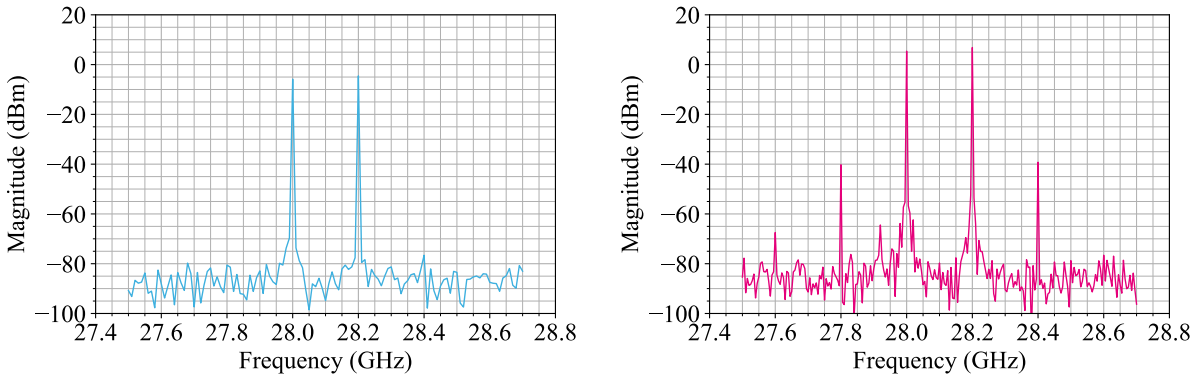


Figure 3.32. Photograph of the input recombined and split path

to correlate the command position and the phase shift obtained.



(a) (b)

Figure 3.33. PA Input (a) and output (b) bi-tone signal spectrum

The input power delivered to the PA input is −4 dBm per tone, while the power at the input of the PP is -2 dBm per tone. The two tones generated are located at 28 and 28.2 GHz with an amplitude of 5.5 dBm each. The input of the power amplifier is first measured to verify that the bi-tone signal is "clean", hence with no IMD visible on the spectrum, as seen in figure 3.33a. The output signal of the DUT shown in figure 3.33b is first measured without injection (with a 50-Ohm termination on the input 2 probes) and saved as a reference as it represents the spectrum generated naturally by the amplifier.

The second input is enabled and fed with the phase-controlled bi-tone (−2 dBm per tone) signal. A manual sweep on the phase shifter is performed; the spectrum of the output signal

is saved for each value. In order to show the variations created by the injection, the tones and the IMD3 magnitude are normalized relative to the reference signal from figure 3.33b. The phases of the different components are normalized to the initial phase before the phase shifter sweep hence, starting with zero. For convenience, the IMD3 $2f_1 - f_2$ is noted α et $2f_2 - f_1$ is noted β.

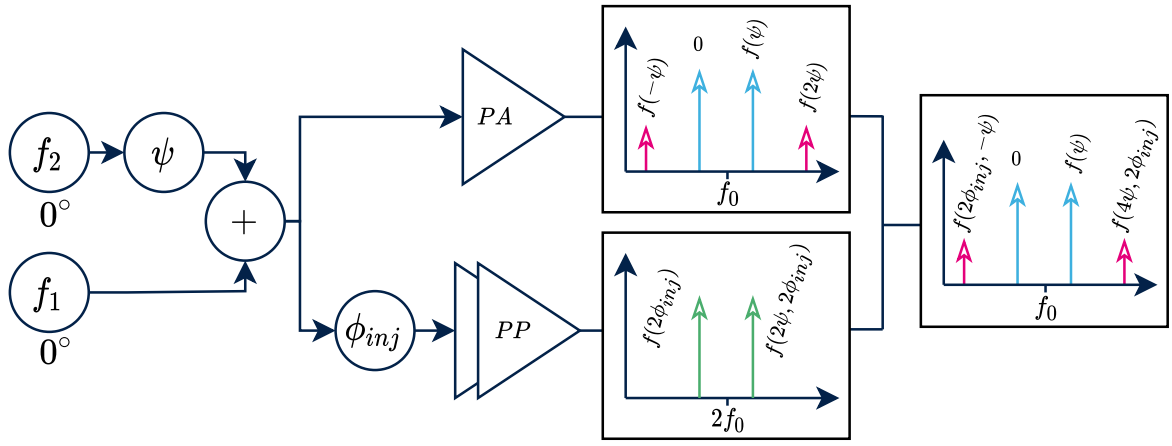


Figure 3.34. Mechanism of phase variation in the circuit

The phase relations at different system stages are illustrated in figure 3.34. In the simulation, the phase at t_0 is null for each tone, and the passive elements (splitter/phase shifter) are considered perfect. In measurement, each passive element shows slight impedance variations between f_1 and f_2. Also, the two sources are coherent in phase; we can imagine at t_0 that f_1 is referenced at 0° and f_2 is constantly phase shifted by an arbitrary value ψ. The tones are recombined and amplified by the PA. Due to its nonlinearities, the first and second-order are mixed to create IMD3. As ψ propagates to second harmonic $2f_2$ with a factor 2, it propagates to IMD3. On the other path, f_1 and f_2 first undergo the controlled phase shift ϕ_{inj}, then are doubled in frequency and are a function of 2ψ for f_0 and $2\psi + 2\phi_{inj}$ for f_2. Both spectra are mixed in the non-linear impedance of the PA. The tones remain unchanged while the IMD are dependent on ψ and $2\phi_{inj}$. However, the phase shift range is complex to derive as it depends on amplitudes and AM/PM distortion effects.

First, an eye is kept on the tones f_1 and f_2 magnitude and phase to check their immunity to phase injection. Figure 3.35a shows that with the injection, the amplitude of the main tones remains unchanged in measurements. It shows a variation between 0 and 0.44 dB (with 0.2 dB within the ZVA measurement tolerance). Small variations below 2° are visible on the

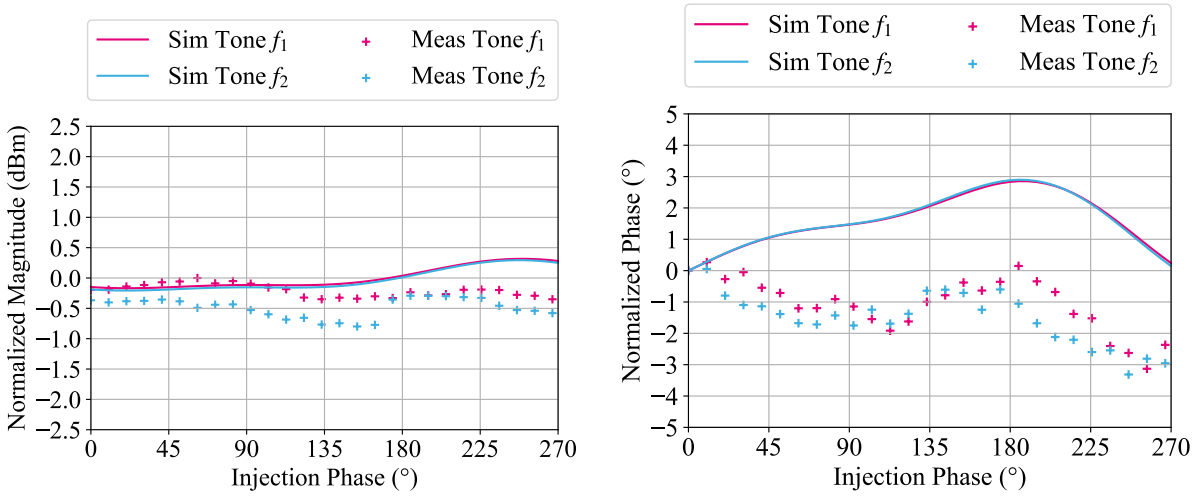


Figure 3.35. Phase (a) and magnitude (b) of tones f_1 and f_2 vs. input 2 phase injection at 7 dB back-off

phase on the range of injection of interest. However, it does not show a clear correlation with the injection and the variations are comprised within the tolerance (2°) of the ZVA.

The amplitude and phase of the IMD3 according to the injection phase in figure 3.36 compared to simulation at the same back-off region (with measurement-simulation discrepancies in mind). The phase sweep is performed up to 270° to see at least a period of variation (180° at the second-order) and its periodic characteristic. The magnitude of the IMD3 in figure 3.36a

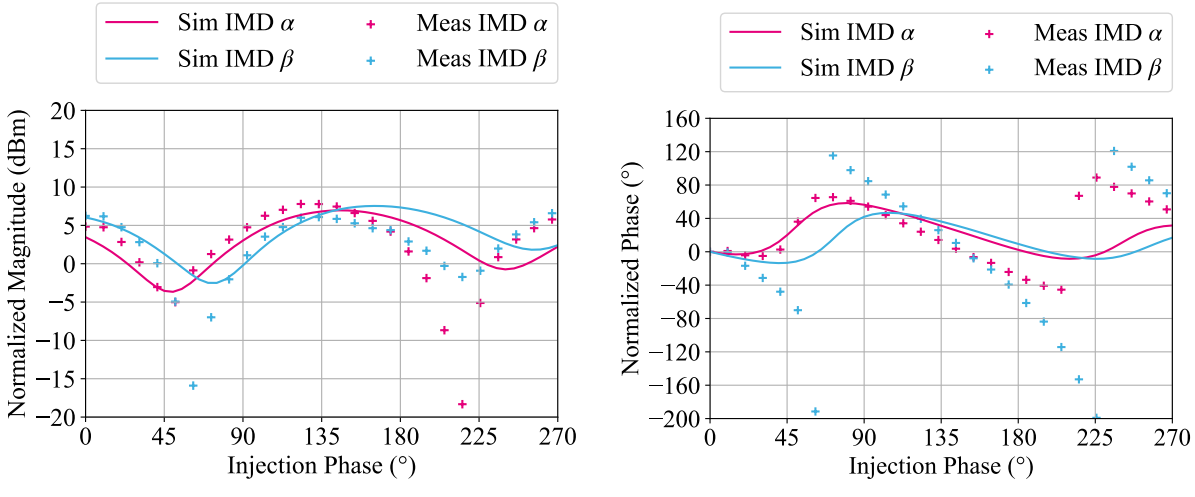
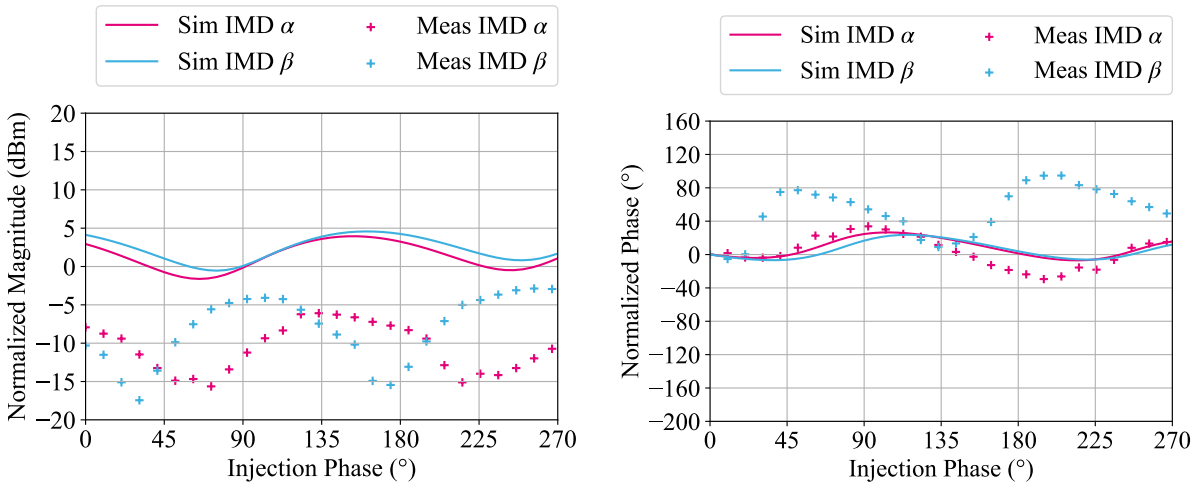


Figure 3.36. Phase (a) and magnitude (b) of IMD3 difference product vs. input 2 phase injection at 7 dB back-off

shows the dependency with the phase injection. Both IMD products have their maximum and minimum amplitude for the same injection phase (around 50° and 135°). It is interesting to note that the injection enables the decrease of at least one IMD3 of 15 dB. In the simulation, we observe an amplitude modulation of 10 dB that shows a reduction of -4 dB. We also keep a phase modulation of the intermodulation products that occur in measurement periodically with a range of 153° for IMD α and that is wrapped and covers 360° for the IMD β. The simulation shows a phase variation of 65° for both IMD3s.

If we compare measurements and simulations, we do not get a 100% match. The general behavior is caught by simulations. Nevertheless, there is some mismatch. This is something that was expected since main PA measurements (3.4.3) are already not aligned with simulations. We observe that IMD3 α and β (both amplitude and phase) are aligned in simulation and phase shifted in measurements. This is probably due to AM/PM distortion not being properly modeled in simulation [AKJ12]. We also observe that the phase variations are more important in measurements compared to simulation. This is a good news because this demonstrates that measured PA will be able to provide targeted phase shift and thus enough beam diversion.

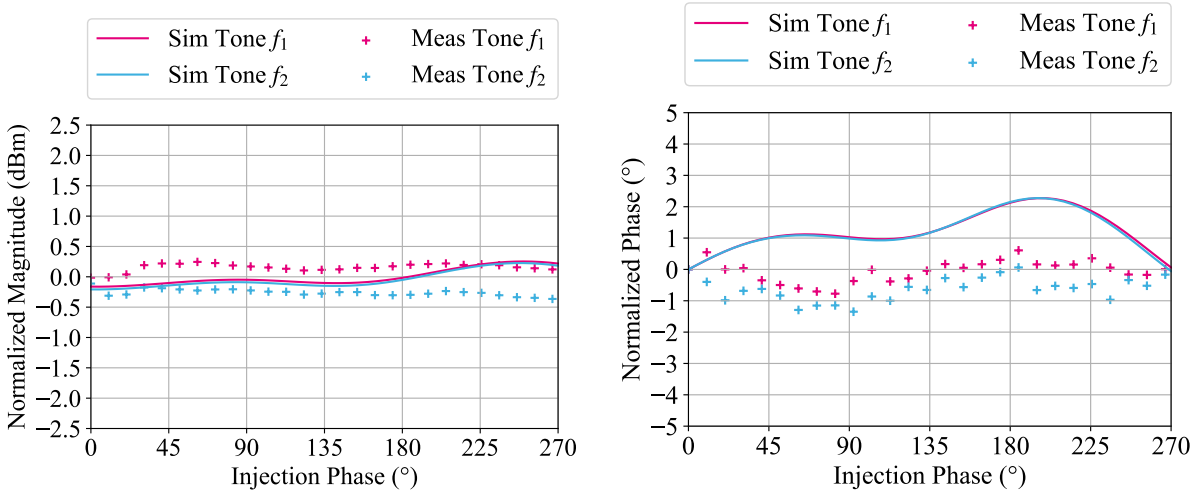


(a) (b)

Figure 3.37. Phase (a) and magnitude (b) of IMD3 difference product vs. input 2 phase injection at 4 dB back-off

To get another set of results, the injection has been measured with 4 dB back-off (instead of 7) by removing one of the attenuators at PA input (figure 3.37b). A lower control range is expected as we get closer to the compression point, large-signal nonlinearities are starting to occur, and the natural amount of IMD3 generated becomes important (as when P_{in} increases by 3 dB, the IMD3 level increases by 9 dB). Results are reported in figure 3.38.

Indeed, the range of magnitude and phase variation is reduced with a range of -15 and −6 dBm in magnitude and 60° of phase range for the IMD3 α. The second IMD3 β shows a magnitude range between -17 and −4 dB and a range of phase variation of 94°. The simulation shows a similar effect as the range of control has been reduced by 5 dB and by 35°.



(a) (b)

Figure 3.38. Phase (a) and magnitude (b) of tones f_1 and f_2 vs. input 2 phase injection at 4 dB back-off

Once again, the integrity of the tones has been preserved with phase variations contained between -0.8 and 0.6° for f_1 and 0 and −1.3° (within the tolerance of the ZVA in phase measurement). The magnitude variation remains under 0.25 dB for f_1 and 0.36 dB for f_2.

With those results, we confirm that the phase shift target of 30° for maximum relaxation on the PA is achievable with the architecture of the PA proposed. Furthermore, as this experiment is in its first stage of development, the device could enable a reduction of IMD3.

There is other techniques that can provide such relaxations. Digital pre-distortion [PMP19] has demonstrated equivalent improvement (15 dB) with even 15 % power consumption reduction, when counting digital consumption. Nevertheless, digital pre-distortion is unpractical in antenna arrays context because it would require that all PAs non-linearities are matched. Analog pre-distortion/post-distortion can be a solution as well. [HB17] demonstrated 15 dB (ACLR) improvement with 10 % PAE degradation. This solution looks promising but complex and has been demonstrated in a multi-board discrete implementation and may be difficult to integrate into a single chip. The novel technique proposed in this paper can achieve similar IMD3 relaxation with compact and low circuit complexity fitting an already dense 5G phased array architecture with a minimized impact on the power consumption.

3.5 Conclusion

In this chapter, we have proposed a novel technique called Beam Diversion. Based on the observation that IMD3 nonlinearities are beam-formed, Beam Diversion gives IMD3 beam the proper orientation so that it does not interfere with any user communication. Beam diversion has been patented during the PhD. It has been established that using Beam diversion in a 1024 elements array can relax the PA linearity requirements (IMD3) with 17 dB, with a meaningful positive impact on current consumption. To exercise Beam diversion, control has to be added to the system to modify the phase of the IMD3 applied to the antenna. A novel PA architecture has been proposed, using a phased controlled 2^{nd} harmonic injection to control the IMD3 phase. A circuit with this PA has been designed in SOImmW technology and characterized. Results demonstrate the validity of the approach, even if measurement results do not 100% match simulations. Even if the feasibility of beam diversion using second harmonic injection has been demonstrated with the PA prototype in the lab, with sine wave signals, some steps are still needed to get a turn key solution. First, IMD3 control has to be demonstrated with modulated signals, representatives of 5G transmissions. We do not anticipate any blocking point, but a demonstration is still needed. The second axis is calibration. Measurements reveal that IMD3 angle is not a linear function of ϕ_{inj}. On top of that, the $IMD3_{phase}(\phi_{inj})$ may vary with output power, temperature, process, and thus, a methodology to select the right ϕ_{inj}, for a desired IMD3 angle is needed.

References

[AKJ12] Kashi Aazar Saadaat, Mahmoud Kamarei, and Mohsen Javadi. "Estimation of Intermodulation Rejection Value as a Function of Frequency in Power Amplifier Using AM-AM and AM-PM Diagrams Based on Power Series Analysis". In: *Circuits and System* 3 (June 2012), pp. 282–287. DOI: 10.4236/cs.2012.33039.

[Bal05] Constantine A Balanis. *Antenna theory: analysis and design*. Wiley-Interscience, 2005.

[ÇR19] Umut Çelik and Patrick Reynaert. "An E-Band Compact Power Amplifier for Future Array-Based Backhaul Networks in 22nm FD-SOI". In: *2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2019, pp. 187–190. DOI: 10.1109/RFIC.2019.8701866.

[HB17] Yushi Hu and Slim Boumaiza. "Joint RF pre-distortion and post-distortion linearization of small cell power amplifiers". In: *2017 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*. 2017, pp. 66–69. DOI: 10.1109/PAWR.2017.7875575.

[Hem02] C. Hemmi. "Pattern characteristics of harmonic and intermodulation products in broadband active transmit arrays". In: *IEEE Transactions on Antennas and Propagation* 50.6 (June 2002), pp. 858–865. ISSN: 1558-2221. DOI: 10.1109/TAP.2002.1017668.

[IR17] Paramartha Indirayanti and Patrick Reynaert. "A 32 GHz 20 dBm-PSAT transformer-based Doherty power amplifier for multi-Gb/s 5G applications in 28 nm bulk CMOS". In: *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2017, pp. 45–48. DOI: 10.1109/RFIC.2017.7969013.

[Par+16] Byungjoon Park et al. "Highly Linear mm-Wave CMOS Power Amplifier". In: *IEEE Transactions on Microwave Theory and Techniques* 64.12 (Dec. 2016), pp. 4535–4544. ISSN: 1557-9670. DOI: 10.1109/TMTT.2016.2623706.

[Par+19] Hyun-chul Park et al. "A High Efficiency 39GHz CMOS Cascode Power Amplifier for 5G Applications". In: *2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2019, pp. 179–182. DOI: 10.1109/RFIC.2019.8701809.

[PMP19] Saroj K Patro, Rabindra K Mishra, and Ajit K Panda. "A Comparative DC Power Budget Analysis of RF Linearization Architectures for 1 Watt Power Amplifier for Low Power Wireless Base Station". In: *2019 4th International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT)*. 2019, pp. 1389–1393. DOI: 10.1109/RTEICT46194.2019.9016810.

[Raz00] Behzad Razavi. *Design of Analog CMOS Integrated Circuits*. 1st ed. USA: McGraw-Hill, Inc., 2000. ISBN: 0072380322.

[Raz11] Behzad Razavi. *RF Microelectronics (2nd Edition) (Prentice Hall Communications Engineering and Emerging Technologies Series)*. 2nd. USA: Prentice Hall Press, 2011. ISBN: 0137134738.

Conclusion

To conclude...

This work aimed to explore the impacts of Antenna Arrays, such as the ones of 5G mmW base stations, on PA and PA design. Indeed, using Antenna Arrays for cellular communications brings new challenges to the PA designer, such as being able to cope with a large number of PAs on the same board while maintaining functionality, especially the accuracy of the beam angle and power efficiency. The usual way to approach this constraint is to design PAs not sensitive to the environment (balanced PA as an example), but this comes with an efficiency penalty. We were convinced that an environment-aware PA design would bring better results. In a first step, we focused on being able to model the antenna array environment and thus simulate the PA in the AA environment accurately. A co-simulation flow has been developed, mixing electrical and electromagnetic simulations. Based on Keysight tools, this flow can transform electrical stimuli into electromagnetic radiations, counting for PA circuit behavior, antenna, coupling between the antenna, PCB effects. This simulation bench revealed the impact of active load modulation, coming from antenna to antenna coupling, on the PAs. This effect is difficult to catch because it depends on many parameters, including antenna pitch, substrate characteristics, number of antennas. Results have been somewhat surprising. The active load-pull, which was supposed to be a big deal for antenna arrays, finally does not significantly affect the beam orientation. The efficiency impact on the overall array is very limited (only the PAs on the periphery of the array are impacted). This study was published at EuCAP 2021 conference. In a second step, we decided to use the co-simulation bench to explore another aspect of the antenna arrays. The beamforming principle relies on the phase relations between the signals applied at each antenna element. If it is easy to control the phase of a signal generated on purpose, the phase of spurious frequencies can be more unpredictable. We worked with a use case where the antenna is generating 2 beams with different frequencies for two users. The co-simulation bench revealed that IMD3 products created by PA non-linearities are beamformed in a direction that can interfere with the communication of a third user. Even if this issue can

be fought by using a linear enough PA (that will cost in terms of efficiency); we proposed to use the properties of the Antenna Array to get rid of it. We proposed a technique called "Beam Diversion" (a patent has been filed) that consists of changing the IMD3 beam orientation to avoid any user. The beam diversion principle is very simple: As IMD3 beam direction depends on the IMD3 phase, we proposed to add an IMD3 phase control on the PA. We estimated that with 4° of IMD3 beam diversion, which corresponds to 17° IMD3 phase shift, we could relax the PA IMD3 constraint with 17 dB. To implement this IMD3 phase control on the PA, the principle of second harmonic injection has been explored. A PA based on a cascode differential amplifier and two Push-Push stages have been designed in ST CMOS065SOIMMW technology. Measurements confirmed, even if the measurement does not 100% match the simulations, that this PA can control the phases of the IMD3 and thus can enable beam diversion without degrading the fundamental signal. The measurement results associated with the beam diversion principle were submitted in T-CAS II / ISICAS 2022 conference when the PA design is about to be in ESSCIRC.

4.1 Perspectives

One of the challenges that still need to be fulfilled is the characterization of the active load-pull effect through measurements. To obtain this type of result, the complexity of the millimeter-wave 5G phased array demands a high integration level that is close to the realization of an actual product. Indeed, the base system would require the design of an AA with $N \times N$ elements on a multi-layer substrate, and $N \times N/4$ integrated circuits with 4 PAs need to be designed and packaged. The beam-forming circuit would also be directly embedded on chips to avoid having complex source management. To get good results to validate the simulation, multiple probe points would be inserted to obtain the load variation on different locations in the array. One can imagine a setup that gets the complex current and voltage at the output of others PAs feeding the antennas to deduce the impedance seen and the output power delivered. Furthermore an independent DC monitoring network is required to observe the efficiency variations of each amplifier. However, a few 'tricks' can be imagined to simplify this process: To ease the integration constraint, the measurement can occur at a lower frequency, with larger patch antennas and wider spaces between them, enabling a more 'lab friendly' setup. Also, partial results can be evaluated with simple S-parameters measurements. With some data processing using active S-parameters, the load variation effect on power amplifiers must be quantified with simulations.

The IMD3 phase shifter power amplifier prototype realized is still at its early stages of development and needs deeper analysis for improvement and tests: Further investigations need to be done on the correlation between the simulation and the measurements. First, simulation must be performed with more mature transistor models for a better fit. Furthermore, the setup used at the input of the circuit is quite complex. A deeper analysis must be performed on the devices such as splitters that have imperfect matching across frequency. Especially with different Δf between tones, this mismatch is susceptible to affect the tone phase on the different paths (PA or PP input). The device has also shown IMD3 amplitude reduction capabilities that need to be studied and put in equations to, eventually allow implementation on systems that do not necessarily use phased antenna arrays.

With further investigations on optimal 2^{nd} harmonic generation, the push-push amplifier sizing can be refined to improve the overall efficiency. The IC layout was voluntarily focused on prioritizing the symmetry (for reduced common-mode and stability) and minimal losses of the power amplifier. The floorplan did not let many options for the PP layout integration. Furthermore, the small size of the PP transistors imposed an important impedance transformation ratio that could hardly be feasible with a balun transformer. To overcome this issue, resistive elements needed to be added to make the impedance transformation possible, decreasing the performance of 2^{nd} harmonic generation. Also, a study of the other component generated by the injection ($f_2 - f_1$, high orders...) must be conducted.

Also, a calibration algorithm is needed to determine the $\delta\phi_{inj}$ that must be given to the IMD3. The algorithm have to constantly track the optimal linearity relaxation according to the beams generated position and also according the different PVT variations that have an important effect on the IMD3 phase/amplitude.

For now, all the simulations and measurements have been performed for bi-tone sinusoidal signals only. It is crucial in the future to perform tests under OFDMA signal in simulation and measurement to confirm the same behavior and validate its functionality with modulated signals.

4.2 Oral and written productions

4.2.1 Patent

T. Capelli, P. Cathelin, A. Ghiotto, and N. Deltimple, "Phased Array IMD3 Beam Diversion", patent n°PCT/FR2022/050560, 2022.

4.2.2 Journal

T. Capelli-Mouvand, N. Deltimple, P. Cathelin and A. Ghiotto, "Third Order Intermodulations Spatial Filtering for Power Amplifiers in Multi-beam 5G mmW Phased Array," submitted at 2022 International Symposium on Integrated Circuits and Systems (ISICAS), 2022.

4.2.3 International Publications

I. Alaji, T. Capelli Mouvand, M. El-Chaar, A. Lisboa-de-Souza, F. Podevin and S. Bourdel "Cascaded tunable distributed amplifiers for serial optical links: Some design rules" 2020 18th IEEE International New Circuits and Systems Conference (NEWCAS), 2020, doi: 10.1109/NEWCAS49341.2020.9159841.

T. Capelli-Mouvand, A. Ghiotto, P. Cathelin and N. Deltimple, "5G Millimeter Wave Active Phased-Array Antenna Active Load Pulling Evaluation on Power Amplifiers," 2021 15th European Conference on Antennas and Propagation (EuCAP), 2021, pp. 1-5, doi: 10.23919/EuCAP51087.2021.9410977.

T. Capelli-Mouvand, N. Deltimple, P. Cathelin and A. Ghiotto, "Third Order Intermodulations Spatial Filtering for Power Amplifiers in Multi-beam 5G mmW Phased Array," writing ongoing for 48th European Solid-State Circuits Conference (ESSIRC), 2022.

4.2.4 National Publications

T. Capelli-Mouvand, A. Ghiotto, P. Cathelin, and N. Deltimple, "Evaluation de la variation de charge active dans les réseaux phasés d'antennes 5G et son effet sur les amplificateurs de puissance", XXIIèmes Journées Nationales Microondes, 8-10th of June 2022, Limoges

4.2.5 Presentations

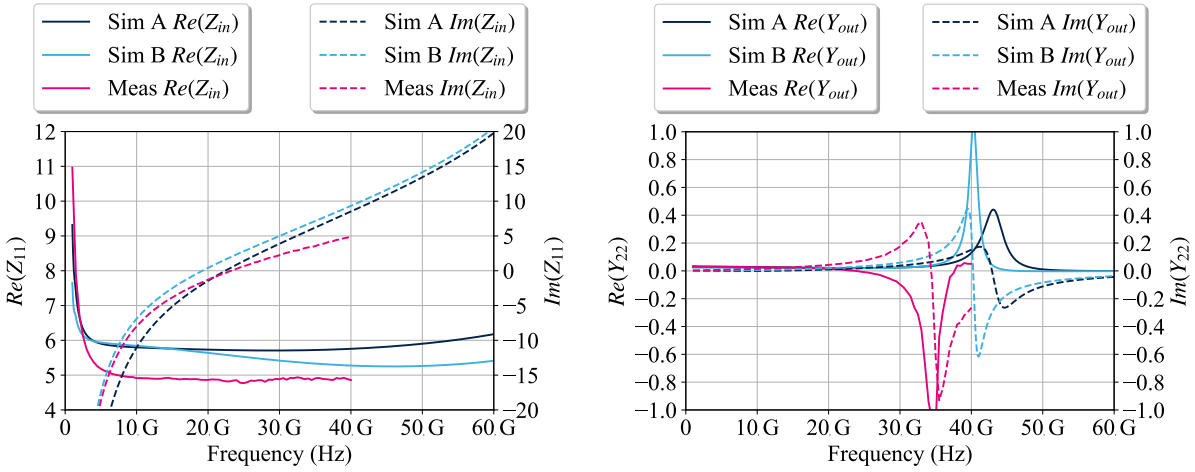
Thomas Capelli-Mouvand, "Predicting coupling between Phased Array and 5G mmW Power Amplifier", Keysight Technologies Tech Talks, Monday 18th and Tuesday 19th of January 2020

Thomas Capelli-Mouvand, "5G NR Antenna arrays : PA designers' nightmare or opportunity to implement new solutions ?", ST-IMS common laboratory, 24th of November 2020

Appendix A

MOS transistor measurement/simulation discrepancies study

As the matching network of the PA is composing an impedance transfer between 50 Ω with the capacitive input/ output of the differential pair of MOS, we use the measurement structure of a single MOS transistor that has been embedded with the main circuit, with probing to monitor its polarization. Unfortunately, no de-embedding structure has been designed, an interpretation of the variations observed have been proposed using the impedance and admittance parameters is proposed. To do so, the real and imaginary part of input impedance Z_{11} and output admittance Y_{22} are plotted on figure A.1 with the measurements performed compared to the top layout simulations with the models of the DK A and B.



(a) (b)

Figure A.1. Embedded input impedance (a) and output admittance (b) of a measured MOS transistor vs. DK A and DK B simulation

The output admittance (figure A.1b) is showing a resonance between the output capacitance and the inductive part of the measurement structure. Assuming that the inductive part is constant between simulations and measurements, the resonance indicates a higher capacitive effect from simulation A to B and even higher in measurements as the resonant frequency is decreasing.

Concerning the input, we see that the real part of Z_{11} especially reflect the gate resistance R_g, while its imaginary part represents the capacitance C_{gs} with the equivalent capacitance due to the Miller effect C_{gmi}. From the results shown on figure A.1a we observe that the image of R_g is lower for DK B compared to DK A at 28 GHz showing a closer fit with the measurement that shows around 5 Ohm. On the other hand, the imaginary part of the input impedance undergoes low variations between design kits A and B which seems counter-intuitive given the strong variations of the results obtained with the PA.

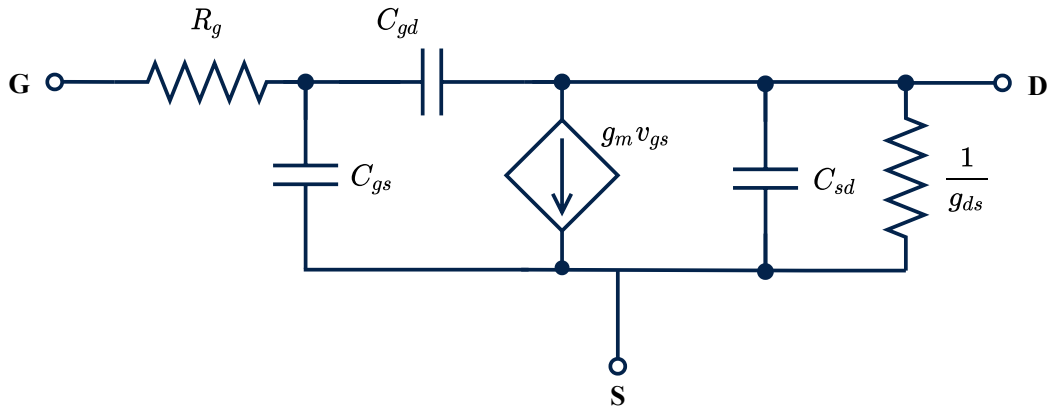


Figure A.2. 8-element MOS transistor model with non-reciprocal capacitances

With the lack of deem-bedding structure and by considering the effect of the measurement structure invariant between the different DK simulation and measurement, we can expect that the variations observed on Z and Y parameters are only due to the MOS device. Then to quantify the changes between the design kits an extraction of the main parasitic elements of both models of the transistor, is proposed at 28 GHz. Thus C_{dg}, C_{gg}, C_{gs}, C_{dg}, C_{sd}, G_m, G_{ds} and R_g are extracted from the S-parameters of a single MOS used in the differential pair according to the method given by [Pas+17].

The extraction of the capacitances (figure A.3) of both models highlights major changes in the MOS definition. We see that we have low variations of capacitances C_{dg}. On the other hand, there is an important re-evaluation of the source-drain capacitance from 74.5 fF to 105 fF.

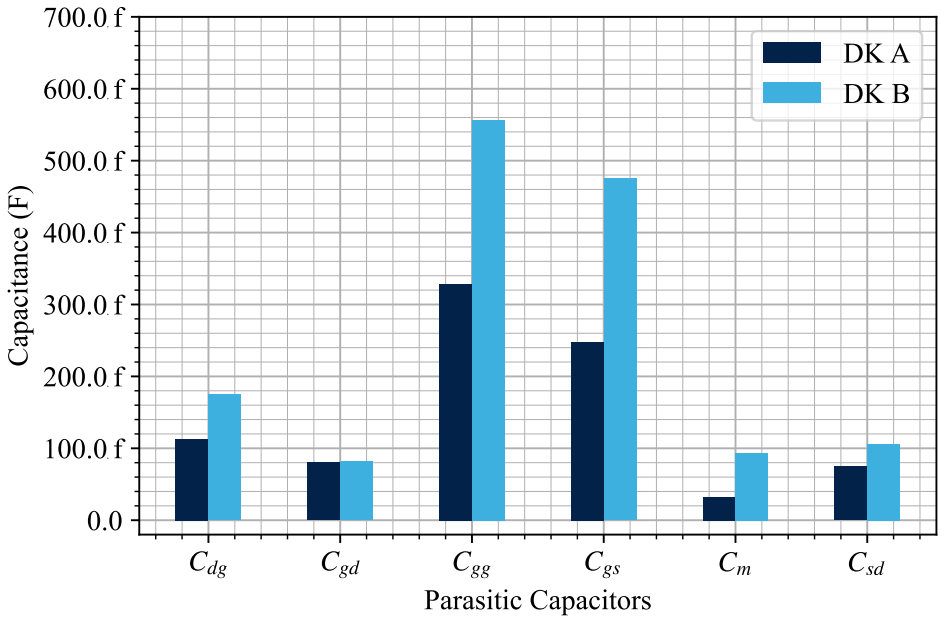
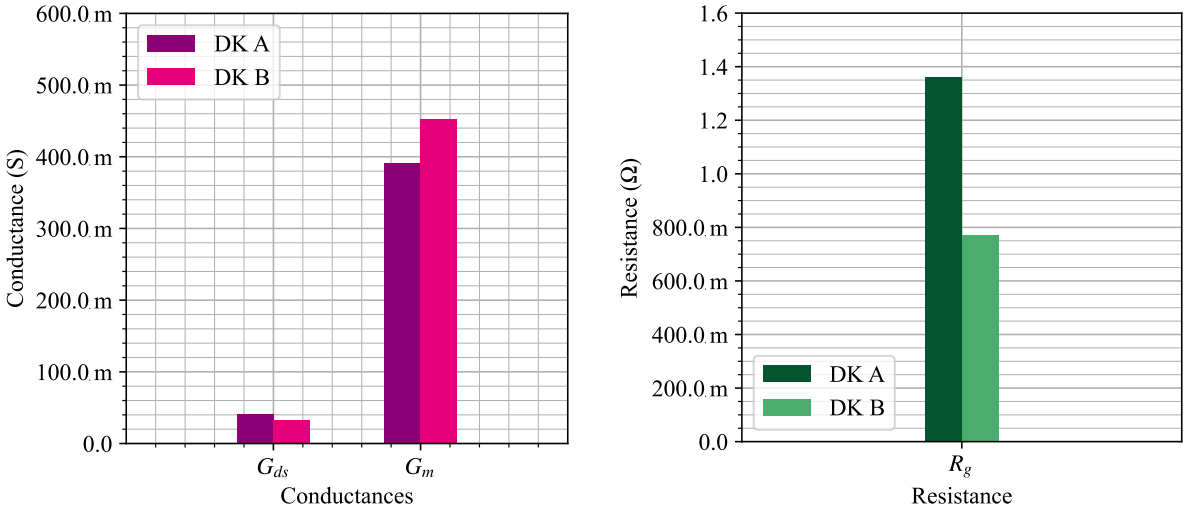


Figure A.3. Parasitic MOS capacitance extraction from S-parameters of the MOS transistor models



(a) (b)

Figure A.4. Conductances (a) and gate resistance (b) extraction from S-parameters of the MOS transistor models

Relating to the output admittance we can use the resonant frequency on Y_{22} to determine the inductive part value L_{ext} of the measurement structure: We assume that

$$Im_{Y_{22}} = jC_{sd}\omega + jC_{pad}\omega + jC_{gd}\omega + \frac{1}{jL_{ext}\omega} \tag{A.1}$$

With $C_{pad} \approx 20\,\text{fF}$ the capacitance represented by the measurement pad. Then the inductive part from the resonant frequency f_r is

$$L_{ext} = \frac{1}{4\pi^2 f_r^2 (C_{sd} + C_{pad} + C_{gd})} \tag{A.2}$$

Which using the values extracted from DK B with $f_{rB} = 40\,\text{GHz}$ gives $L_{ext} = 85.3\,\text{pH}$. Then if L_{ext} resonates with C_{pad} and C_{sd} and C_{dg} of the design kit A, we should expect:

$$f_{rA} = \frac{1}{2\pi\sqrt{L_{ext}(C_{sd} + C_{pad} + C_{gd})}} = 43.2\,\text{GHz} \tag{A.3}$$

Which fits with the resonance observed on figure A.1b. From those results it is then possible to estimate the equivalent capacitance $(C_{sd} + C_{gd})_{meas}$ of the measured transistor based on the measured resonant frequency $f_{r\,meas}$:

$$(C_{sd} + C_{gd})_{meas} = \frac{1}{4\pi^2 f_{r\,meas}^2 L_{ext}} - C_{pad} \approx 248.5\,\text{fF} \tag{A.4}$$

To summarize we have:

$$(C_{sd} + C_{gd})_A < (C_{sd} + C_{gd})_B < (C_{sd} + C_{gd})_{meas} \tag{A.5}$$

With $(C_{sd} + C_{gd})_A = 156.2\,\text{fF}$, $(C_{sd} + C_{gd})_B = 185.6\,\text{fF}$ and $(C_{sd} + C_{gd})_{meas} = 248.5\,\text{fF}$

Pursuing with the MOS extraction between the DK, we find that the gate-source capacitance that is estimated to be of 475 fF instead of 248 fF, which represent almost 50 % of variation. The transconductance G_m also increase from 405 mS to 452 mS and R_g is reduced by 42 % from 1.32 Ω to 770 mΩ.

From the parameters extracted and their variations, we can estimate the variations expected on the input impedance on the transistor: With an input impedance expressed under the form

$$Z_{in} = R_{in} + \frac{1}{jC_{in}\omega} \tag{A.6}$$

We can identify from our small-signal model of figure A.2 that

$$R_{in} = R_g \tag{A.7}$$

	C_{gs}	C_{gd}	C_{gmi}	C_{sd}	G_m	R_g	$\mathrm{Im}(Z_{in})$
DK A	247.8 fF	80.6 fF	1.1 pF	74.5 fF	391 mS	1.4 Ω	−4.2 Ω
DK B	556.6 fF	81.7 fF	0.99 pF	105.8 fF	451 mS	0.77 Ω	−3.9 Ω
Δ (%)	124.6 %	1.4 %	−10 %	−42 %	15.4 %	−43.4 %	−7.7 %

Table A.1. Input impedance variations

and that

$$C_{in} = C_{gs} + C_{gmi} \tag{A.8}$$

With C_{gmi} the equivalent capacitance brought by C_{gd} by Miller effect

$$C_{gmi} = C_{gd}.G_m.R_{out} \tag{A.9}$$

and

$$R_{out} = \mathrm{Re}(Z_{out}) = \frac{R_{load}}{1 + (R_{load}C_{out}\omega)^2} \tag{A.10}$$

Which finally gives for Z_{in}:

$$Z_{in} = R_g - j\frac{1}{\omega\left(C_{gs} + C_{gd}.\left(G_m.\frac{R_{load}}{1+(R_{load}C_{sd}\omega)^2}\right)\right)} \tag{A.11}$$

Among the elements that compose $\mathrm{Im}\{Z_{in}\}$, we have the variations between the design kits reminded in the table A.1 which finally translate themselves on $\mathrm{Im}\{Z_{in}\}$ with variations of −7.7 %. In simple common source configuration, the miller effect is showing an important equivalent capacitance C_{gmi} due to C_{gd} and to the gain $G_m.R_{out}$. C_{gmi} is then large enough (1.9 pF) to mask the C_{gs} variations (250 fF) and minimize the variations of the input impedance.

Looking back at the PA design, with the topology used, the cascode architecture has the particularity to have a common gate transistor that shows a low impedance at the drain of the common source transistor to minimize the miller effect. In consequence, the input impedance of the cascode is quasi-independent from the output and is then mostly dependent on C_{gs}. As the design kit shows strong variations of C_{gs} from one version to another, larger variations can be expected on the input impedance of a cascode compared to a common source topology.

As we have seen, the DK B still does not fit the measurements either in output admittance or input impedance. In the light of the transistor parameters extraction, we were able to estimate the variation of the couple $(C_{gd} + C_{gs})$. Given the difference in the resonant frequency of the output admittance and the input impedance between the simulation and the measurement,

we can suspect that C_{gd} that is predominant in the input matching and also contributing to the output admittance needs particular adjustments in the MOS model for a better alignment with measurements.

References

[Pas+17] Francisco Pasadas et al. "Small-Signal Model for 2D-Material Based FETs Targeting Radio-Frequency Applications: The Importance of Considering Nonreciprocal Capacitances". In: *IEEE Transactions on Electron Devices* 64.11 (Nov. 2017), pp. 4715–4723. ISSN: 1557-9646. DOI: 10.1109/TED.2017.2749503.

Appendix B

IMD3 PSPA Measurement Bench

The bench of measurement (B.1) using the ZVA R&S has been set up using short connections between the elements of the input chain to reduce losses and cables toward the three $|Z|$ GSG 67 GHz probes (B.2a) from Cascade and the analyzer. Four power supply Agilent E3631A (B.2b) was needed to power the DUT and two others to power the amplifiers at the input chain.

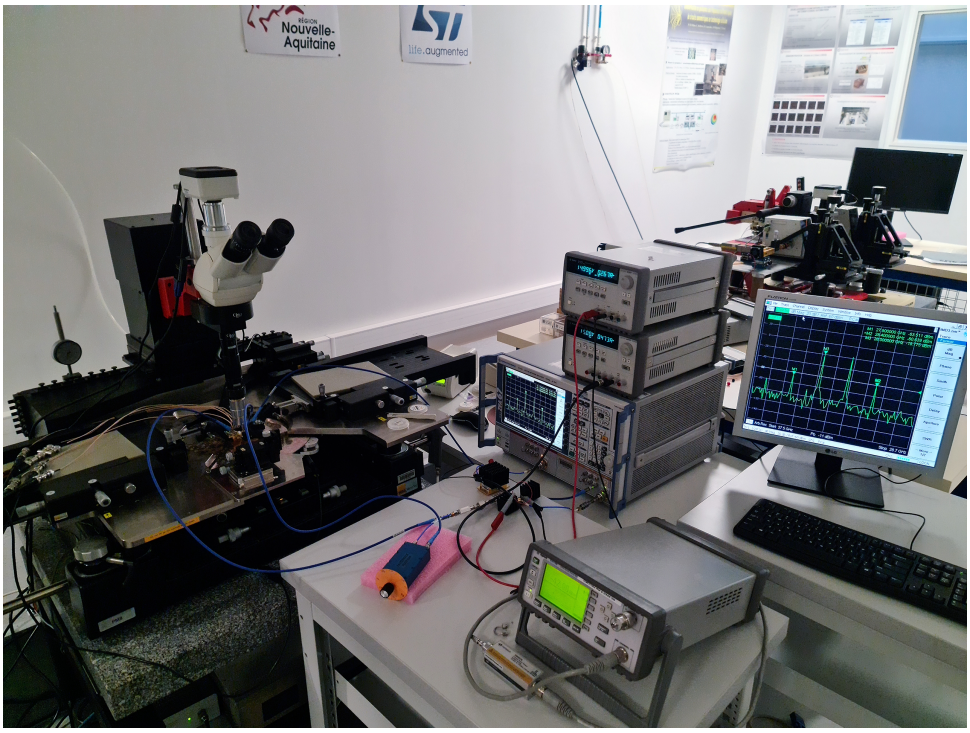
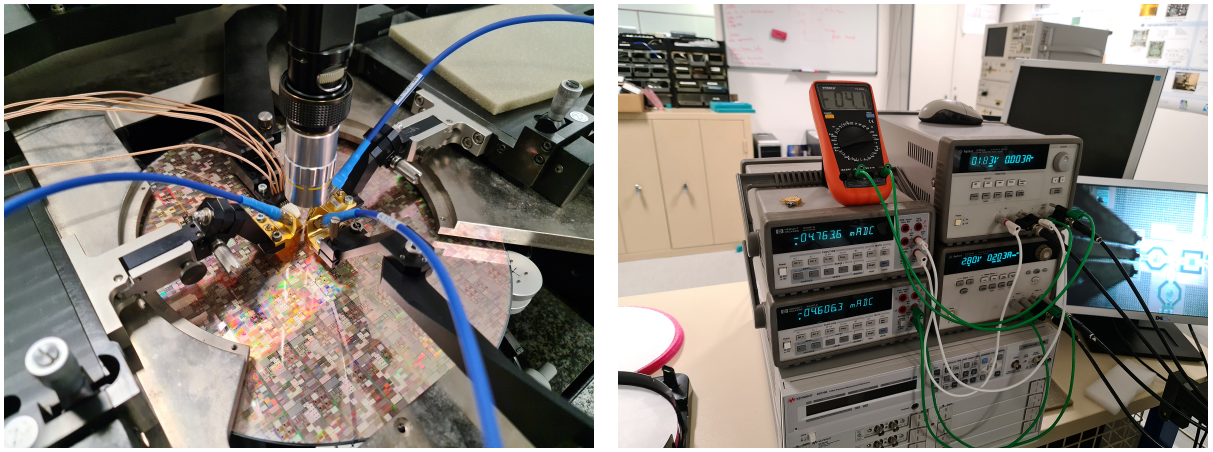


Figure B.1. Bi-tone measurement setup of the DUT with an R&S ZVA 67 GHz

Each current powering the circuit has been monitored with ammeters HP 344001Q and Tenma 72-2590.

A power meter Agilent E4418B has been used to ensure a balanced level of power on each path of the input. The input power calibration of the ZVA has been performed with its



(a) (b)

Figure B.2. 3-port bi-tone measurement of the DUT (a) and DUT DC power supplies and monitoring (b)

power meter tool at the output of the chain in order to take into account the gain of the amplifiers added in the chain and the losses of the other devices. The output calibration has been performed with input and output probes on a through from a calibration kit to remove the losses of the probe and cables from the measurement.

The ZVA is able to provide direct measurements of the spectrum in phase and amplitude thanks to its frequency conversion option that allows versatile configurations of internal and external sources and the internal receivers.

Amplificateur de puissance pour réseaux phasés d'antenne 5G multibande en technologie ST CMOS065SOIMMW

Résumé : Les télécommunications mobiles afin de subvenir à leurs insatiables besoins, trouvent des moyens d'améliorer leurs capacités depuis maintenant plus de 30 ans. En 2019, la 5ème génération est à l'épreuve afin de garantir une connexion non seulement au marché des téléphones mobiles toujours croissant, mais aussi au vaste univers des objets connectés. Afin de pouvoir remplir ses objectifs, la 5G marque une expansion sans précédent en termes de bandes de fréquences utilisées. En effet, des bandes jusqu'à 60 GHz et plus font partie des ambitions du réseau et cela implique de radicaux changements technologiques qui impactent toute l'électronique dédiée. Nouvelles fréquences, niveau d'exigences relevé, les réseaux phasés d'antenne sont introduits et imposent une toute nouvelle architecture système et une interface inédite pour les front-end RF.

Dans ce travail nous proposons donc une analyse de ces réseaux d'antenne phasés et des contraintes qu'ils représentent en particulier pour les amplificateurs de puissance en termes de variation de charge parasite et de non-linéarités. Une évaluation de la variation de charge est proposée ainsi que de son impact sur la performance des amplificateurs. Le comportement des non-linéarités telles que les produits d'intermodulation du troisième ordre est montré dans les réseaux d'antennes. Un concept permettant de relâcher les contraintes de linéarité des amplificateurs 5G est proposé ainsi qu'une implémentation en technologie ST CMOS065SOIMMW à 28 GHz.

Mots-clés : 5G, Réseaux phasés d'antenne, Amplificateur de puissance, Non-linearités, Variation de charge active

---

Power amplifier for multi-band 5G phased arrays in ST CMOS065SOIMMW

Abstract: In order to satisfy their insatiable needs, mobile telecommunications find ways to improve their capabilities for more than thirty years. In 2019, the 5^{th} generation is challenged to guarantee a connection non only to the mobile telecommunications market, but also to the vast universe of the internet of things. To achieve its goals, the 5G presents an unprecedented expansion on the frequency spectrum. Indeed, bands up to 60 GHz and more are targeted. It implies an important turnover for the dedicated electronics. New frequencies, high level of requirements, the phased arrays are introduced and impose a whole new system architecture and a new interface for the RF front ends.

In this work, we propose an analysis of these phased arrays with the constraints that they represent especially to the power amplifiers, in terms of parasitic active load pulling and non-linearity. An evaluation of the active load modulation is proposed with its impact on the power amplifier. The non-linearities behaviors such as the third order intermodulations in the phased array are shown. A concept allowing relaxation of linearity constraints on power amplifier is proposed and possible implementation in ST CMOS065SOIMMW technology is realized at 28 GHz

Keywords: 5G, Phase arrays, Power Amplifier, Non-linearities, Active load pull

---

Unité de recherche

UMR 5218 - IMS - Laboratoire de l'Intégration du Matériau au Système Université de Bordeaux, 351 Cours de la libération, 33405 Talence cedex, France.