NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

ANA ISABEL FERREIRA DOS SANTOS

Bachelor of Sciences in Electrical and Computer Engineering

RF-CMOS SWITCHED-CAPACITOR POWER AMPLIFIER FOR NB-IOT RF TRANSCEIVERS

MASTER IN ELECTRICAL AND COMPUTER ENGINEERING

NOVA University Lisbon
March, 2023

NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING

RF-CMOS SWITCHED-CAPACITOR POWER AMPLIFIER FOR NB-IOT RF TRANSCEIVERS

ANA ISABEL FERREIRA DOS SANTOS

Bachelor of Sciences in Electrical and Computer Engineering

Adviser: João Pedro Abreu de Oliveira
Assistant Professor, NOVA University Lisbon

MASTER IN ELECTRICAL AND COMPUTER ENGINEERING

NOVA University Lisbon
March, 2023

RF-CMOS Switched-Capacitor Power Amplifier for NB-IoT RF Transceivers

To Achille

Acknowledgements

I would like to express my gratitude to all the professors who played a vital role in my academic journey. Their knowledge and wisdom have been crucial in shaping my future. A special thanks to the professors from the electronics department for the numerous opportunities they provided me with. I am particularly grateful to my supervisor, Professor João Pedro Oliveira, for his support and patience in guiding me through the challenges I encountered.

Lastly, I want to express my gratitude to my family and closest friends for the endless encouragement and support they have provided me throughout my academic and personal life. I especially want to thank Achille for his unconditional support and invaluable discussions during the development of this project. My appreciation for all that you have done for me cannot be put into words.

Abstract

The increasing market of Narrowband Internet of Things (NB-IoT) applications brings new challenges and constrains in the design of fully integrated transmission architectures, capable of delivering the desired output power with the highest efficiency and linearity, ensuring the longest battery lifetime of the devices. This work is focused on the study and implementation of the most power consuming block within the transmission chain: the Power Amplifier (PA). In this regard, a Switched Capacitor Power Amplifier (SCPA) is designed to operate at a frequency of 0.9 GHz and aiming the maximum output power allowed by the standard of 23 dBm.

The final architecture includes a matching network that connects to eight unit PA cells through an LC filter. Each unit PA cell is made of a cascoded class-D PA, two drivers, a level shifter and two selection logic blocks. All the blocks were developed using RF components from a UMC 130nm CMOS process with a 1.2V/2.4V supply voltage. The results show that the architecture is able to produce a maximum output power of 15.61 dBm with a maximum Power Added Efficiency (PAE) of 26.52% and a Total Harmonic Distortion (THD) of 0.68%. In the same conditions, the measured HD2 and HD3 are of -70.23dBc and -43.41dBc, respectively.

Additionally, a modulation stage was implemented in VerilogA in order to evaluate the impact of sending different symbols in the SCPA performance. The block, designed for a 16 QAM modulation, is responsible for generating both the number of unit PA cells to be selected and the phase of the clock connected to each PA cell, depending on the amplitude and phase of the constellation points being transmitted.

Keywords: SCPA, Matching Network, Cascoded Class-D, PAE, THD, CMOS

Resumo

O mercado crescente de aplicações *IoT* de largura de banda estreita coloca novos desafios e restrições no desenvolvimento de arquiteturas de transmissão totalmente integradas, capazes de produzir a potência desejada com o máximo de eficiência e linearidade possíveis, de forma a garantir o maior tempo de vida de bateria dos dispositivos. Este trabalho foca-se no estudo e implementação do bloco da cadeia de transmissão que mais consome: o amplificador de potência. Neste sentido, um amplificador de potência de condensadores comutados é desenhado para operar à frequência de 0.9GHz com o objetivo de produzir à sua saída o valor de potência máxima permitida pelo standard de 23dBm.

A arquitetura inclui uma malha de adaptação que liga a oito *PA*s unitários através de um filtro LC. Cada *PA* unitário consiste num amplificador de potencia *class-D cascoded*, dois *drivers*, um *level shifter* e dois blocos de lógica de seleção. Todos estes blocos foram desenvolvidos usando componentes RF da tecnologia CMOS 130nm da UMC com uma tensão de alimentação de 1.2V/2.4V. Os resultados mostram que a arquitetura é capaz de produzir uma potência à saída de 15.61dBm, com uma PAE de 26.52% e uma distorção harmónica total de 0.68%. Nas mesmas condições, os valores medidos da HD2 e HD3 são de -70.23dBc e -43.41dBc, respetivamente.

Adicionalmente, um andar de modulação foi implementado em VerilogA, de forma a avaliar o impacto de enviar diferentes símbolos na performance do amplificador. Este bloco, desenvolvido para uma modulacao 16QAM, é responsável por gerar o número de unidades de *PA* a serem selecionados e o relógio de fase que liga a cada *PA* unitário, dependendo da amplitude e fase dos pontos da constelação a serem transmitidos.

Palavras-chave: SCPA, Malha de adaptação, Classe D *Cascoded*, PAE, THD, CMOS

Contents

List of Figures

List of Tables

Acronyms

ACLR Adjacent Channel Leakage Ratio 10, 11

CDCM Class-D Current-Mode 16, 17
CDVM Class-D Voltage-Mode 16, 17
CMOS Complementary Metal-Oxide-Semiconductor 1, 2, 3, 16, 19, 25, 28, 29, 30

DAC Digital-to-Analog Converter 5, 26
DC Direct Current 11
DPA Digital Power Amplifier 7, 26, 27, 28
DSP Digital Signal Processor 5, 24

EER Envelope Elimination and Restoration 24, 25, 26
ET Envelope Tracking 25, 26
EVM Error Vector Magnitude 9

GaAs Gallium Arsenide 29
GaN Gallium Nitride 20, 29

IF Intermediate Frequency 5, 6
IoT Internet of Things 1

LDMOS Laterally Diffused Metal Oxide Semiconductor 29
LO Local Oscillator 5, 6, 24
LPWAN Low Power Wide-Area Network 1

MOSFET Metal-Oxide-Semiconductor Field-Effect Transistor 27

NB-IoT Narrowband Internet of Things xi, 1, 2, 11, 96

PA Power Amplifier xi, 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 34, 35, 89, 95, 97

PAE Power Added Efficiency xi, xiii, 11, 12, 28, 87, 95, 96

PVT Process Voltage and Temperature xxiii, xxvi, 85, 86

RF Radio Frequency 5, 6, 8, 11, 12, 22, 24, 26, 36, 38

SCPA Switched Capacitor Power Amplifier xi, xiii, xxii, 2, 3, 27, 28, 31, 34, 35, 47, 57, 59, 64, 65, 67, 68, 72, 78, 79, 80, 83, 85, 89, 95, 96, 97

SiGe Silicon Germanium 29, 30

SoC System-on-Chip 2

THD Total Harmonic Distortion xi, xiii, 45, 47, 49, 59, 61, 65, 68, 69, 70, 72, 74, 75, 79, 80, 82, 85, 86, 87, 96

ZCS Zero Current Switching 19

ZVS Zero Voltage Switching 17, 18, 19

1

Introduction

The chapter herein aims to provide a context for the PA to be implemented in a 130nm Complementary Metal-Oxide-Semiconductor (CMOS) technology with a 2.4V/1.2V supply voltage. It will be put forward what drives this project and what it intends to achieve.

1.1 Context and Motivation

The rise of connected devices represents the ever-increasing market share of Internet of Things (IoT). NB-IoT, in particular, is a Low Power Wide-Area Network (LPWAN) standard that aims to unleash a whole new set of IoT devices and services, such as smart cities, smart homes and buildings, asset tracking, healthcare monitoring and environmental monitoring. Furthermore, these require small bandwidth and little need to transmit, sending small amounts of data on an occasional basis. The services deploying this technology strive for plenty of coverage, little power consumption (the less, the better), great efficiency and the ability to connect with many devices. Often, due to the nature of NB-IoT applications, the designers look forward to endow their devices with a battery lifetime large enough to withstand over a decade of use. From this perspective, since the general efficiency of the radio system largely hinges on the efficiency of the PA, aiming at the highest efficiency possible is paramount to achieve a longer battery lifetime and decrease costs. Its development can be challenging because it needs to generate a high output power with great efficiency and enough linearity.

Many of these radio systems have been implemented in CMOS technology, even though integrating the PA in this kind of systems remains a burdensome task. The reason for this has to do with the fact that the amplitude of the output power stems from the values of both voltage and current. Because the supply voltage is constant and low in value, only the current can be tuned. In order to get a greater current, the transistor must be increased in size, making the chip grow with it, hence hindering the goal of integrating one inside the other. As a result, most of these radio systems are not fully integrated, having their matching network (that will be further discussed in this document) implemented off-chip. Besides, a larger PA entails that more power will get dissipated, thus decreasing

efficiency. Such dissipated power, released in the form of heat, causes the temperature in the chip to increase, also harming the PA's performance.

Nevertheless, due to the low production cost and the possibility of integrating the whole radio system in one chip - System-on-Chip (SoC) - it becomes compelling to integrate these PAs in a single CMOS radio chip. For this to happen, it is essential to minimize the power consumption of this block in the transmission chain, so as to obtain a SoC design with plenty of efficiency for little cost. The SCPA, theme of this thesis, seeks to take advantage of the scaling of CMOS technology while maintaining sufficient linearity and high efficiency.

1.1.1 NB-IoT Specifications and Project Goal

In order to implement the PA, it is important to first understand what are the specifications and frequency bands allowed for NB-IoT applications. Table 1.1 presents the operation mode, bandwidth, modulation, access and maximum transmitted power requirements, specified in 3GPP releases 13 and 14. Table 1.2 summarises the different frequency bands allocated to NB-IoT in Europe.

Table 1.1: NB-IoT specifications [29].

	CAT-NB1	**CAT-NB2**
3GPP Release	Release 13	Release 14
Operation Mode	FDD (Half Duplex)	
Bandwidth	180 KHz	
Modulation	Uplink: BPSK / QPSK Downlink: QPSK	
Access	Uplink: Single Tone (15KHz / 3.75KHz) SC-FDMA (15KHz) Downlink: OFDMA (15KHz)	
Maximum Transmit Power	20 / 23 dBm	14 / 20 / 23 dBm

Table 1.2: NB-IoT frequency bands in Europe [29].

NB-IoT Band	**Uplink (MHz)**	**Downlink (MHz)**	**Bandwidth (MHz)**
B3	1720 - 1785	1805 - 1880	75
B8	880 - 915	925 - 960	25
B20	832 - 862	791 - 821	30

Bearing in mind the above-mentioned NB-IoT requirements, the goal of this project is to implement a PA able to produce a maximum ideal output power of 23dBm. Moreover, the adopted carrier frequency will be 0.9GHz, which falls within the B8 frequency band allocated for uplink.

1.2 Main Contributions

A paper of the presented work title "Design of an RF-CMOS Switched Capacitor Power Amplifier for NB-IoT RF Transceivers" has been submitted and accepted for the 7th edition of the International Young Engineers Forum on Electrical and Computer Engineering – YEF-ECE 2023.

1.3 Thesis Outline

This document is organised into four chapters: Introduction, Literature Review, Design and Implementation in CMOS technology and Conclusions and Future Work.

The first and current chapter contextualises the SCPA to be implemented and its necessity. It briefly explains the specifications that need to be satisfied for NB-IoT applications and the goal of the project.

The second chapter, Literature Review, presents the study of the different classes of PAs and the major issues and limitations in their design. It also presents the different metrics used to evaluate PAs performances, as well as some efficiency and linearization enhancement techniques used to overcome PAs limitations. Various examples of implementations are provided and compared in this chapter. Moreover, the performance of some state-of-the-art PAs is shown for different technologies.

The third chapter, Design and Implementation in CMOS technology, presents the design and implementation of the SCPA. It is divided into nine sections. In the first two sections, the ideal behaviour of the SCPA is analysed and the adopted architecture is briefly explained. In the next six sections, a detailed analysis of each block that integrates the SCPA (namely, the matching network, the cascoded class-D PA, the driver, the level shifter and the selection logic) and the respective results are presented and discussed. In particular, the impact of adding a new block in the architecture is carefully analysed. In the eighth section, a final optimization is performed, where the different blocks are reviewed and the system is simulated across different PVT corners and additional Monte Carlo analyses are performed. Finally, in the ninth section, an additional modulation stage is presented in order to analyse the PA's behaviour when a sequence of symbols is transmitted.

The last Chapter, Conclusions, summarises the most relevant results for the designed architecture. It addresses the aspects that could be improved or done differently in the project.

Lastly, the Appendix, includes the designed schematic of the SCPA, the verilogA code implemented in the modulation stage and the final simulation of the whole system when 16 symbols are transmitted.

2

Literature Review

2.1 Transmitter Architectures

Before diving into the study of the power amplifier, it is of paramount importance to have an understanding of what types of transmission architectures exist, where the amplifier fits within these architectures and what components are involved in the processing of the signal to be transmitted.

Two main groups of transmitter architectures are to be considered: heterodyne transmitters and direct upconversion transmitters. The key difference between these architectures lies in the use of an Intermediate Frequency (IF) by the former before converting the baseband signal to Radio Frequency (RF), whereas the latter converts the baseband signal directly to RF. In both architectures, the transmitter is responsible for performing modulation, upconversion and power amplification. Therefore, they incorporate a Digital Signal Processor (DSP), Digital-to-Analog Converter (DAC), Mixers, Local Oscillator (LO) and a PA.

The current section aims to briefly analyse the aforementioned groups of transmitter architectures.

2.1.1 Heterodyne Transmitter

Currently, the heterodyne transmitter, whose architecture is depicted in Fig.2.1, is the most popular transmitter architecture.

In this transmission scheme, the digital signal processor and the DAC blocks generate, from the input bits stream, a baseband signal split into in-phase (I) and quadrature (Q) components. These processed I and Q signals are modulated in quadrature to their respective intermediate frequencies and are summed up to produce the IF modulated signal. One should observe that it is easier to perform the quadrature modulation at lower frequencies since the quadrature output matching is better than it is for higher frequencies. In order to eliminate the IF harmonics and reduce the transmitted noise, the signal is filtered through an IF band-pass filter. Then, the resulting IF modulated signal is upconverted (to the carrier frequency), filtered through an RF band-pass filter to remove

the unwanted mirrored sideband, sent to the power amplifier where it is amplified and further transmitted by the antenna.

Nonetheless, the RF band-pass filter is particularly hard to implement since it has to reject the sideband by a large factor (typically 50 to 60 dB). This filter is usually passive and built with expensive off-chip components. Due to the off-chip passive components in both RF and IF filters, this architecture does not allow full integration [9, 36].

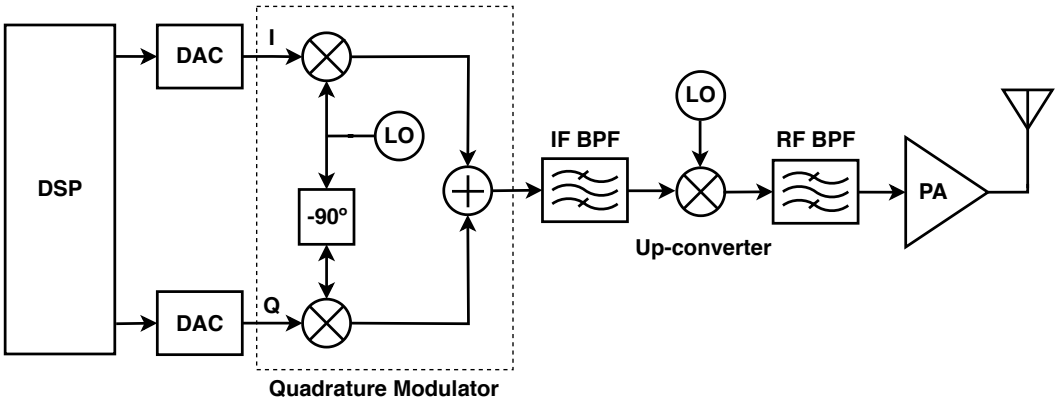


Figure 2.1: Block diagram of the heterodyne transmitter.

2.1.2 Direct Upconversion Transmitter

In the direct upconversion transmitter, whose basic block diagram is illustrated in Fig2.2, the quadrature modulator converts the baseband signal directly to RF, hence the name. In this architecture, the RF carrier frequency is given by the LO frequency at the mixer input.

Because this architecture does not require an intermediate frequency, no filter needs to be implemented to eliminate the mirror sideband that would have resulted from the IF upconversion. For this reason, and unlike the previously mentioned transmission architecture, the direct upconversion architecture can be integrated effortlessly.

Nevertheless, this transmitter architecture is not flawless: the "injection pulling" or "injection locking" of the LO by the PA's output signal is a major issue. Triggered by the fact that the spectrum of the PA's output signal is centered in the same frequency as the signal the LO generates, this phenomenon can be defined as the shift of the frequency of the LO's signal towards the frequency of the noise being undesirably injected by the PA. As a result, any existing leakage from the PA may corrupt the spectrum of the LO's output signal. In order to avoid "injection pulling", it is necessary to guarantee a reverse isolation higher than 60 dB [32].

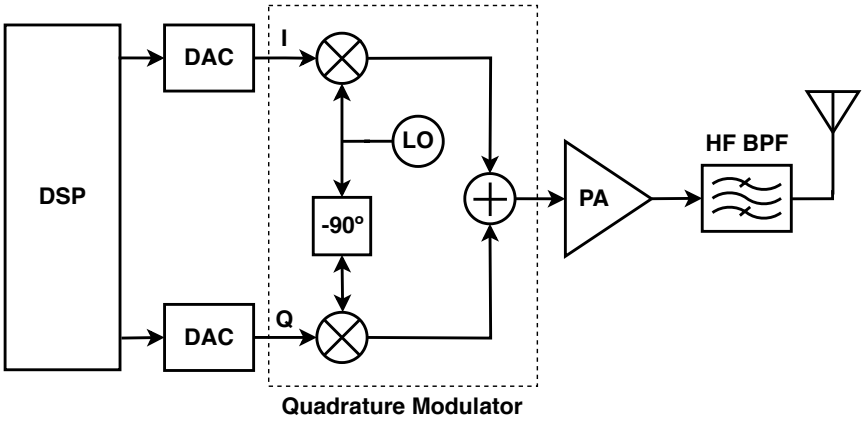


Figure 2.2: Block diagram of the direct upconversion transmitter.

2.2 Power Amplifier

So as to implement a PA, it is worth getting a grasp over the performance parameters associated with PAs and getting to know the main PA class types. Because the goal of a PA is to amplify an input signal in a way that it can be transmitted via antenna, linearity arises as a factor to be taken into account to ensure that the output signal is an exact amplified replica of the input signal, without distortions of any kind. Likewise, since the PA is the most power-consuming block in the transmission architecture, maximizing its efficiency is a matter of the utmost importance. Bearing this in mind, in the present section, both of these parameters will be subject to further analysis. In addition to this, some metrics widely used to measure the non-linearities of the PA, which are typically reflected in terms of spectral regrowth and AM-AM/AM-PM distortions, will be addressed.

PAs can be implemented according to different topologies and configurations, so they are split into classes (usually organised in two groups: linear mode and switch mode), where different PA classes present different trade-offs between linearity and efficiency. These classes will be studied in detail.

Broadly speaking, deploying just one amplifier is not sufficient to obtain the set-out linearity and efficiency. For this reason, different linearity and efficiency improvement techniques are generally implemented so as to ensure the best performance of the PA and the transmission system as a whole. In like manner, these techniques will be addressed thoroughly. Within these techniques, the RF-based-DAC will be analysed in more detail. These amplifiers are typically made up of smaller unit PA cells, connected in parallel. A digital codeword that is built according to the envelope of the input signal is then used to individually activate these amplifiers in order to produce the output signal. In reality, the PA that the current work aims to further develop falls within the category of Digital Power Amplifier (DPA)s, making them the project's primary focus. Finally, the last section of this chapter presents a comparison of the state-of-the-art *PA*s for different technologies.

2.2.1 Linearity

Power amplifier linearity is of major concern as it guarantees that the amplified output signal holds the information carried by the input signal unspoiled. Its study becomes more important as the amplifier operates with larger input signals, causing nonlinear effects to appear. These nonlinear effects lead to the amplitude distortion of the output signal, which is reflected in two ways: the compression effect, when a sinusoidal signal is applied at the input (single-tone test), or the intermodulation (IM) products, when two sinusoidal signals close in frequency are applied at the input (two-tone test). In this section, both these effects are studied.

The 1dB compression point is an important parameter to analyse the compression effect and can be described as the theoretical point at which the real output power decreases 1dB from the ideal output power. From this point on, the real output power no longer increases linearly with the input power and rapidly reaches its maximum value (P_{sat}), leading to power gain compression. This behaviour is illustrated in Fig.2.3. For this reason, the 1dB compression point defines the limit between the linear and the nonlinear operation regions of a RF PA. It is desirable to operate near the 1dB compression point, where the RF PA reaches its maximum efficiency with acceptable linearity.

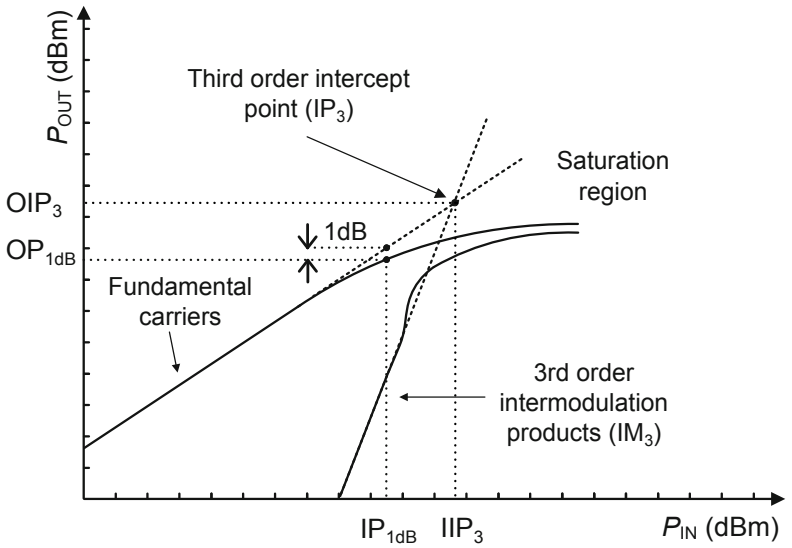


Figure 2.3: 1dB compression point and third-order intercept point. Adapted from [39], Chapter 2.

The above-mentioned intermodulation products are odd order mixed products of the two input frequencies. These products are depicted in Fig.2.4, where ω_1 and ω_2 correspond to the input signals' frequencies, $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$ correspond to third-order intermodulation products and $3\omega_2 - 2\omega_1$ and $3\omega_1 - 2\omega_2$ correspond to fifth order intermodulation products. Due to their proximity with the main frequencies, the third-order IM products cannot be filtered out so they will interfere with the desired signals. In order to analyse their impact, the third-order intercept point is used. This point, illustrated in

Fig.2.3, is obtained when the extrapolated curve of the third-order IM products intercepts the extrapolated curve of the fundamental carrier and indicates how large the PA input signals can be before IM products appear. Typically, its output power value, OIP_3, is 10 dB higher than the 1dB compression point output power of the fundamental carrier, OP_{1dB}. Notwithstanding, this approach is not practical, as this point is placed in the nonlinear PA operation region.

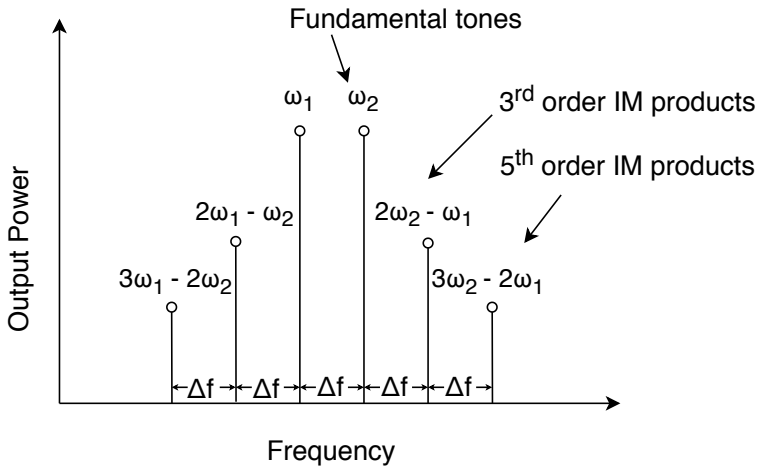


Figure 2.4: Spectrum of two-tone IM products. Adapted from [39], Chapter 2.

2.2.2 Error Vector Magnitude

The non-linearities of the PAs are often characterized using AM-AM and AM-PM distortions. The AM-AM distortion describes the effect of the compression or saturation of the output power with the increase of the input power, analysed in the previous section. The AM-PM distortion, in its turn, describes the phase difference between the input and output signals, which is a consequence of the nonlinear capacitors within the PA.

Typically, when the PA operates with non-constant envelope signals, such as quadrature amplitude modulation (QAM) signals, these non-linearities can strongly impact the constellation of the transmitted signals. In fact, the AM-AM and AM-PM distortions will cause the constellation points to deviate from the ideal constellation points in both amplitude and phase. As a result, the constellation points are more dispersed and rotated when compared to the constellation points of the input signal. An example of this effect is depicted in Fig.2.5 for a 16-QAM signal.

The Error Vector Magnitude (EVM) arises as a metric to evaluate the accuracy of the modulated transmitted signal, by measuring the difference between the ideal constellation point and the transmitted one (the bigger the distance between these two points, the worse the EVM will be). Likewise, the higher the AM-AM and AM-PM distortions are, the worse the EVM will be.

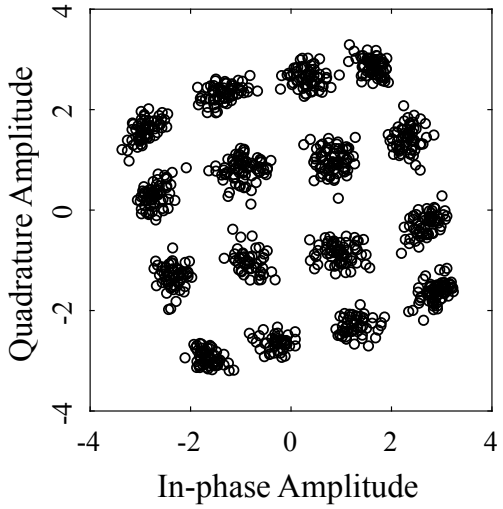


Figure 2.5: Effect of the AM-AM and AM-PM distortions in a 16 QAM constellation.

Another interesting phenomenon that might appear in PAs is the memory effect, in which previous symbols affect the current transmitted symbols. It usually occurs due to the biasing networks and the non-idealities of the devices [50], Chapters 1, 2 and 3. If the output node is not properly discharged at the moment the second symbol is being transmitted, the previous symbol might impact the current transmission. Although in this work the constellation of the transmitted symbols is not going to be analysed, a sequence of different symbols will be transmitted in order to evaluate this impact.

2.2.3 Spectral Regrowth

As discussed in section 2.2.1, unwanted intermodulation products may appear in the output spectrum of the PA when two input signals are applied. These harmonics may lead to what is known as spectral regrowth of the transmitted signal. Spectral regrowth is described as the extension of the band of the output signal, that can go up to three times the original band, in the case of the third-order IM products, or five times the original band, in the case of the fifth-order IM products [39]. When it happens, the spectrum of the adjacent channels might be compromised. Two metrics commonly used to limit the spectral regrowth are the Adjacent Channel Leakage Ratio (ACLR) and the Spectrum Emission Mask.

2.2.3.1 Adjacent Channel Leakage Ratio

In order to limit the amount of power that leaks into the adjacent channels and pollute their spectrum, the ACLR standard metric has to be satisfied. ACLR is defined as the ratio of the integrated power in the adjacent channel over the integrated power in the band, as shown in reference [50], Chapter 7. Fig. 2.6 presents an example of the ACLR.

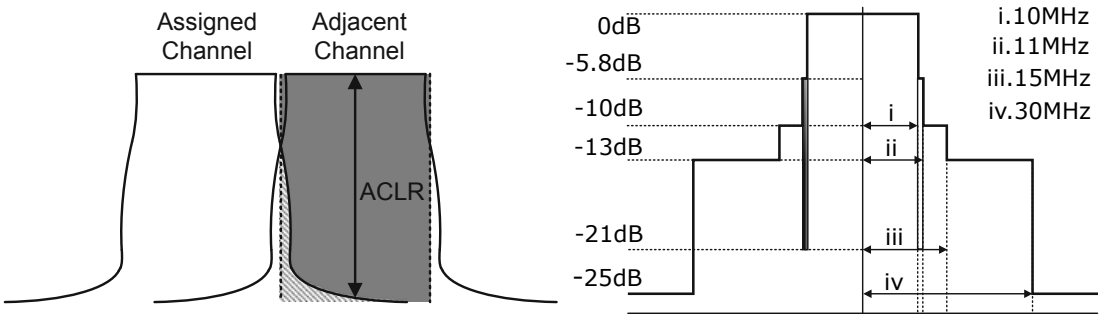


Figure 2.6: Example of Adjacent Channel Leakage Ratio (left) and Spectrum Emission Mask of a 20 MHz channel for 3GPP LTE Standard (right). Adapted from [39].

2.2.3.2 Spectrum Emission Mask

Additionally to ACLR, the Spectrum Emission Mask requirement has to be taken into consideration. The idea behind it is to define different masks of power levels that cannot be surpassed. These maximum power levels are set comparatively to the power in the centre of the channel and will vary according to the standard used [39]. Fig. 2.6 presents an example of the Spectrum Emission Mask.

2.2.4 Efficiency Metrics

Efficiency is one of the most important performance parameters in RF power amplifiers, as they are often the dominant energy dissipators in a transceiver. It quantifies the amount of consumed Direct Current (DC) power that is converted into useful RF power. Due to the PA high power consumption, high efficiency is crucial to improve the battery lifetime of NB-IoT applications. Two traditional metrics commonly used for evaluating efficiency are presented below. In this work, the PAE is used as a metric to evaluate the efficiency.

2.2.4.1 Drain Efficiency

The drain efficiency gives the ratio of the RF output power to the input DC power required to bias the amplifier:

$$\eta_d(\%) = 100 \times \frac{P_{out}}{P_{DC}}. \tag{2.1}$$

Although this metric is good for evaluating the impact the load network and the biasing may have, it is limited since it does not measure the effects of the input side of the PA in terms of efficiency.

2.2.4.2 Power-Added Efficiency

The PAE is similar to the drain efficiency, except the RF input power is taken into account. In literature, different approaches are considered while defining the PAE, as the RF input

power can be removed from the RF output power or be added to the DC power supply. Nevertheless, the former is the most common and is given by:

$$PAE(\%) = 100 \times \frac{P_{out} - P_{in}}{P_{DC}} = \eta_d \left(1 - \frac{1}{G_{ain}}\right). \tag{2.2}$$

Note that the PAE can also be rearranged in terms of drain efficiency and RF power gain G_{ain} (P_{out}/P_{in}). This metric makes it possible to analyse what is happening at the input side and thus estimate more accurately the energy consumed by the PA.

However, it is important to clarify that for some projects, namely this one, it is not possible to consider the impact of the input power, since the amplifier itself generates the output voltage. In this circumstance, both the PAE and the drain efficiency will measure the same.

2.2.5 Classes of Power Amplifiers

Power amplifiers are traditionally arranged in different classes according to the circuit's configuration and the conduction angle. The classes of PAs can be lumped into two main groups: Linear Mode and Switch Mode PAs. In the linear mode group (classes A, AB, B and C), the PA uses a transistor as a voltage-controlled current source, whereas in the switch mode group (classes D, E, F, etc), the PA uses a transistor as a switch. In the following sections, these two groups will be described and compared in some detail.

2.2.5.1 Linear Mode Power Amplifiers

Linear mode power amplifiers are endowed with high linearity, despite presenting a reduced efficiency (always lower than 100% even in the ideal case) when compared to switch mode power amplifiers. Thereby, this group aims to amplify non-constant envelope signals such as amplitude-modulated signals. Every linear mode power amplifier class follows the topology depicted in Fig.2.7, where the C_{BLOCK} refrains the DC current from passing through to the load and the parallel L_0C_0 forms a filter tuned to the fundamental frequency. The different classes are distinguished by the conduction angle, which is chosen according to the bias voltage. This conduction angle can vary between 0 and 360º (or 2π) and defines the duration for which the transistor is conducting (the current is flowing through the load).

As demonstrated in [39], Chapter 3, the maximum drain efficiency of any linear mode PA can be calculated by applying the following expression:

$$\eta_{d(max)}(\%) = 100 \times \frac{\alpha - \sin(\alpha)}{4\left(\sin(\frac{\alpha}{2}) - \frac{\alpha}{2}\cos(\frac{\alpha}{2})\right)}, \tag{2.3}$$

where α corresponds to the conduction angle of the PA. From the equation, one can easily see that decreasing the conduction angle allows us to achieve higher efficiency. Bearing this in mind, let us now discuss in more detail the classes in this group.

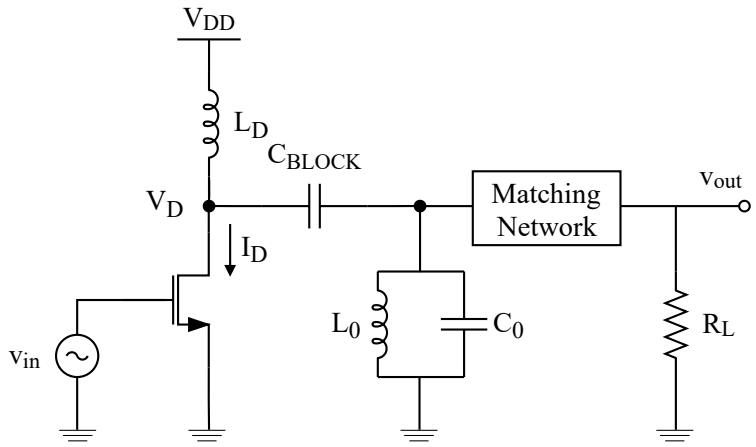


Figure 2.7: Scheme of linear mode power amplifiers.

i. Class-A Power Amplifier

In a class-A power amplifier, the transistor is biased so the quiescent bias current[1] is half the maximum drain current, allowing the device to remain in the saturation region for an entire cycle of the input signal. As a consequence, the device conducts current along its 360º (or 2π) conduction angle, producing a waveform similar to the input waveform. Class-A's biasing point, ideal drain current and drain-to-source voltage waveforms are presented in Fig. 2.8, where theoretical $V_{DS(max)}$ is given by $2V_{DD}$ and $I_{D(max)}$ corresponds to $2V_{DD}/R_L$.

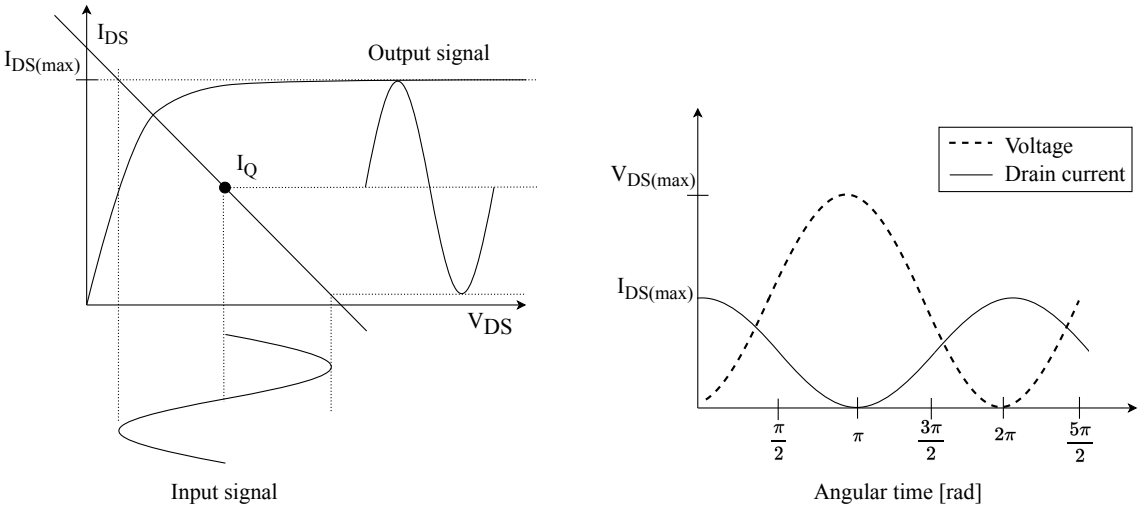


Figure 2.8: Class-A's biasing and ideal drain voltage and current waveforms.

[1] The quiescent bias current corresponds to the amount of current consumed by the circuit when no input drive is applied.

Due to its wide conduction angle, this class presents the best linearity among all classes. However, higher linearity comes at the expense of lower efficiency, as the transistor is conducting at all times. Recalling the drain efficiency expression 2.1, one can see that the maximum efficiency for this class is given by:

$$\eta_{d(max)}(\%) = 100 \times \frac{1}{2}\left(\frac{V_{out(max)}}{V_{DD}}\right)^2, \tag{2.4}$$

where the maximum absolute value of the output signal $V_{out(max)}$ is V_{DD}, leading to a theoretical maximum drain efficiency of 50%. The same result can be obtained by using the generic expression of the efficiency for linear mode power amplifiers presented in Eq.2.3 when the conduction angle is 360º. Nevertheless, because the transistor requires a certain voltage in order to enter the saturation region, the output voltage can never reach V_{DD}. The same applies to drain current and source-to-drain voltage, as they cannot go up to their theoretical maximum values either.

ii. Class-B Power Amplifier

In a class-B power amplifier, the transistor is biased at the threshold voltage, where there is no quiescent bias current, allowing the device to easily change its operation between the saturation and cut-off regions. In this case, the transistor remains in the saturation region for just half the input signal, meaning the device conducts current for a conduction angle of just 180º (or π), thus producing a sinusoidal waveform in the first half of the period and a zero signal in the second half. Class-B's biasing point, ideal drain current and drain-to-source waveforms are illustrated in Fig.2.9, where $V_{DS(MAX)}$ and $I_{D(MAX)}$ are the same as for class-A amplifiers.

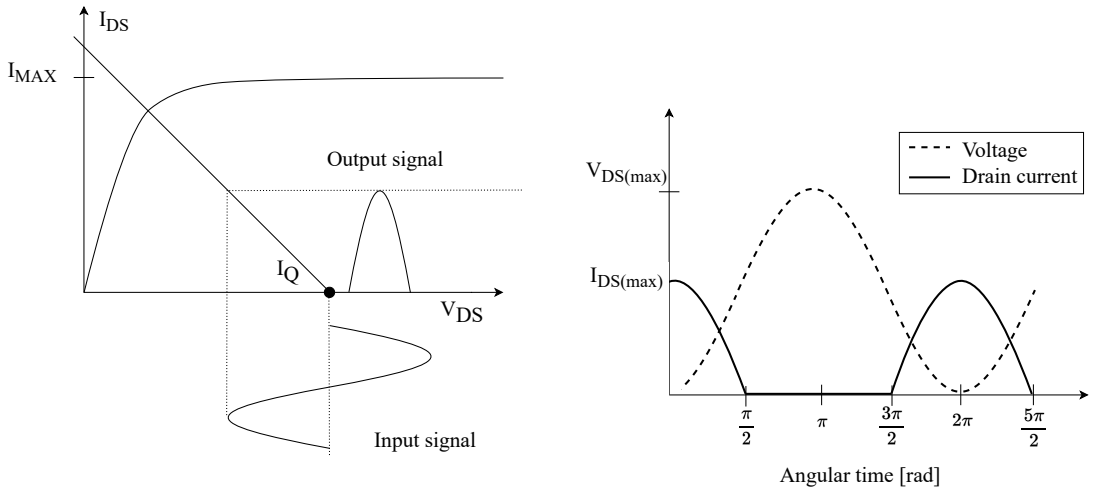


Figure 2.9: Class-B's biasing and ideal drain voltage and current waveforms.

Moreover, having a shorter conduction angle than its class-A counterpart, class-B power amplifiers make it possible to achieve a higher drain efficiency. One can obtain the maximum drain efficiency using the expression 2.3, when the conduction angle, α, is 180º. This results in a theoretical maximum efficiency of 78.5% (once more, this value cannot be reached since the output voltage is always lower than V_{DD}). Although the reduction of the conduction angle implies an increase in the maximum drain efficiency, having the transistor be shut off creates undesired distortion products that affect its linearity, which is already lower than that of class-A.

iii. Class-AB Power Amplifier

Class-AB power amplifiers are placed between class-A and class-B operation modes, so their conduction angle varies between 180º (or π) and 360º (or 2π). Consequently, as the conduction angle increases, the time the transistor is driving increases, causing the maximum drain efficiency to decrease from 78.5% to 50%. Despite this, there is an increase in linearity, suggesting a trade-off between linearity and efficiency should be considered. In fact, by varying the biasing point, it is possible to choose the conduction angle that best suits the circuit's requirements. Class-AB's biasing point, ideal drain current and drain-to-source voltage waveforms are represented in Fig. 2.10.

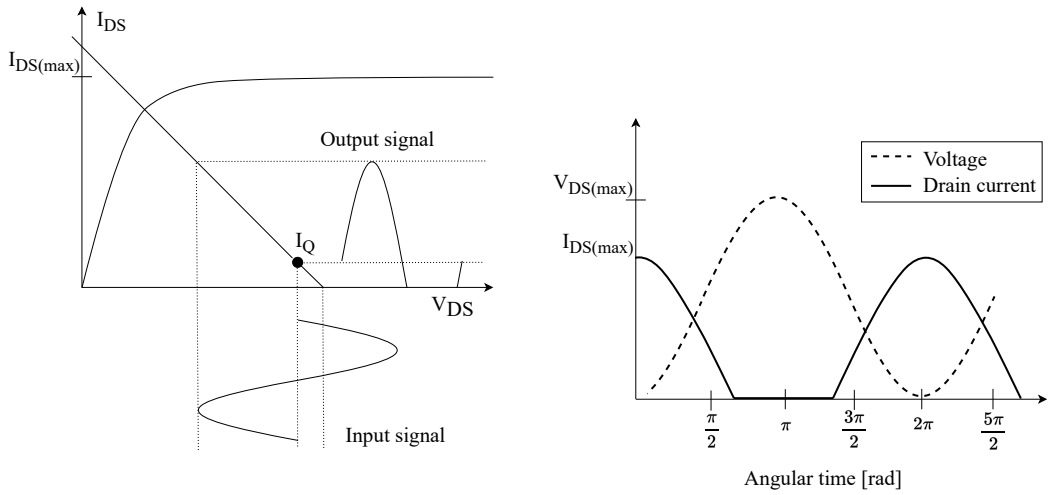


Figure 2.10: Class-AB's ideal drain current and voltage waveforms.

iv. Class-C Power Amplifier

In a class-C power amplifier, the transistor is biased below the threshold voltage. Thereby, the transistor remains in the saturation region for less than half a cycle of the input signal, meaning its conduction angle is below 180º (or pi). Class-C's biasing point, ideal drain current and drain-to-source voltage waveforms are represented in Fig.2.11. The maximum drain efficiency of a class-C PA can be calculated using expression 3.35, by

varying the conduction angle, α, from 0º up to 180º. By reducing the conduction angle towards zero, the theoretical maximum drain efficiency increases from 78.5% to 100%. However, its linearity plummets as well as its maximum output power. Due to their poor linearity, class-C power amplifiers are best suited to amplify constant envelope signals.

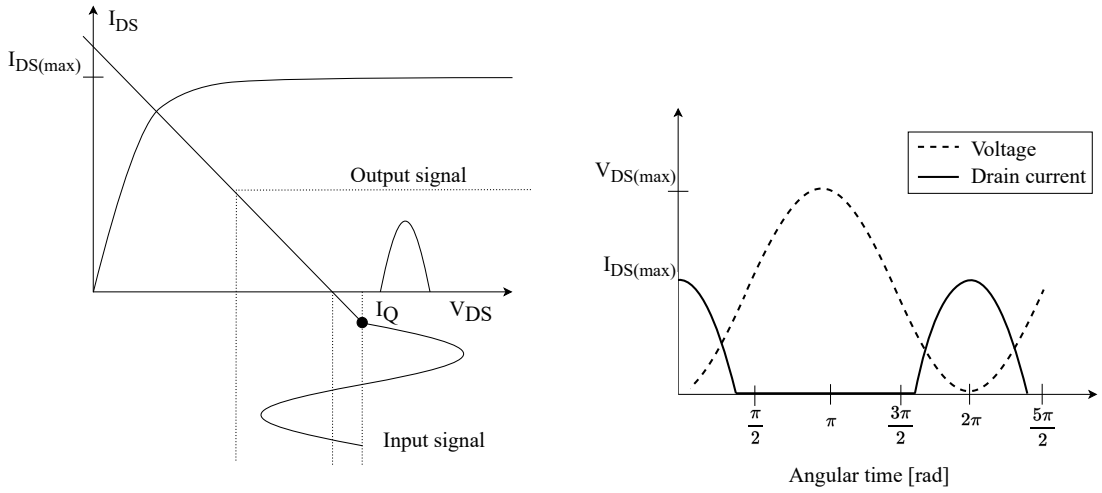


Figure 2.11: Class-C's ideal drain current and voltage waveforms.

2.2.5.2 Switch Mode Power Amplifiers

Unlike linear mode power amplifiers, switch mode power amplifiers are characterized by their high efficiency but poor linearity. The reason behind this is that the transistors are operated as switches, instead of controlled current sources. Posited that the transistors act as ideal switches, no power losses occur and the maximum theoretical drain efficiency is 100%. On that account, switch mode PAs are commonly used to amplify constant envelope signals.

i. Class-D Power Amplifier

A class-D power amplifier uses two transistors in a push-pull configuration and can be either implemented in voltage, Class-D Voltage-Mode (CDVM), or current, Class-D Current-Mode (CDCM). Voltage-mode PAs are often called class-D PAs, whereas current-mode PAs are known as inverse class-D PAs. Fig. 2.12 presents the class-D power amplifier topologies for each of these operation modes. Both these architectures share a similar behaviour with the CMOS inverter, in which each transistor is driven with opposite square waves to ensure that only one is selected at a time.

By analysing the drain voltage and the drain currents of the transistors marked in the CDVM topology, one can obtain the waveforms depicted in the first column of Fig. 2.13. Looking into the figure, it is possible to see that when the input voltage is low and the PMOS is selected, the V_D assumes a high value close to V_{DD}. At the same time, the

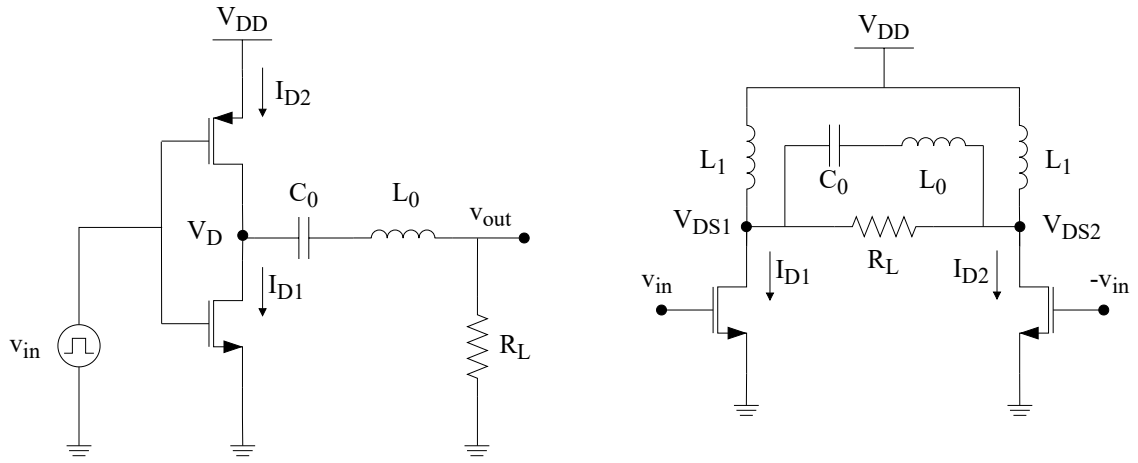


Figure 2.12: Voltage mode Class-D PA and Current mode Class-D PA architectures, respectively

current I_{D2} will flow from the PMOS through the load. Because the NMOS is at high impedance, no current will flow through its branch. The same reasoning can be applied when the input voltage is high and NMOS is selected. In this way, a square waveform is generated at the node V_D. The L_0C_0 filter, tuned to the carrier frequency, removes the harmonics of the square wave, producing a sinusoidal signal.

This conventional class-D PA presents some issues, especially when it comes to operating at higher frequencies. In the first place, the capacitances of the transistors that are charged and discharged during the switch cycle lead to inevitable losses. These losses tend to increase as the frequency goes up. Secondly, since this architecture uses both NMOS and PMOS transistors, it is necessary to design the PMOS larger in size than the NMOS, so the on-resistance in the former equals the on-resistance in the latter. However, increasing the area of the transistor leads to the increase of the input capacitance and, ultimately, the switching time [39, 52]. These topics will be further analysed in this document.

In turn, by analysing the drain voltages and currents marked in the CDCM topology, one can obtain the waveforms presented in the second column of Fig.2.13. In this configuration, depending on the applied input voltage, either M1 or M2 transistors are selected and the DC current is switched between one branch and the other. At the same time, the drain voltage of the selected NMOS is connected to ground, while the drain node of the other NMOS will start charging, hence the transitory state. Taking into account that both transistors are connected to ground, during the switch time there is no voltage across them (and consequently insignificant discharge losses). This phenomenon is known as Zero Voltage Switching (ZVS) and allows the CDCM to achieve higher efficiencies than the CDVM.

Nonetheless, this type of amplifier shows poor linearity and, despite the theoretical

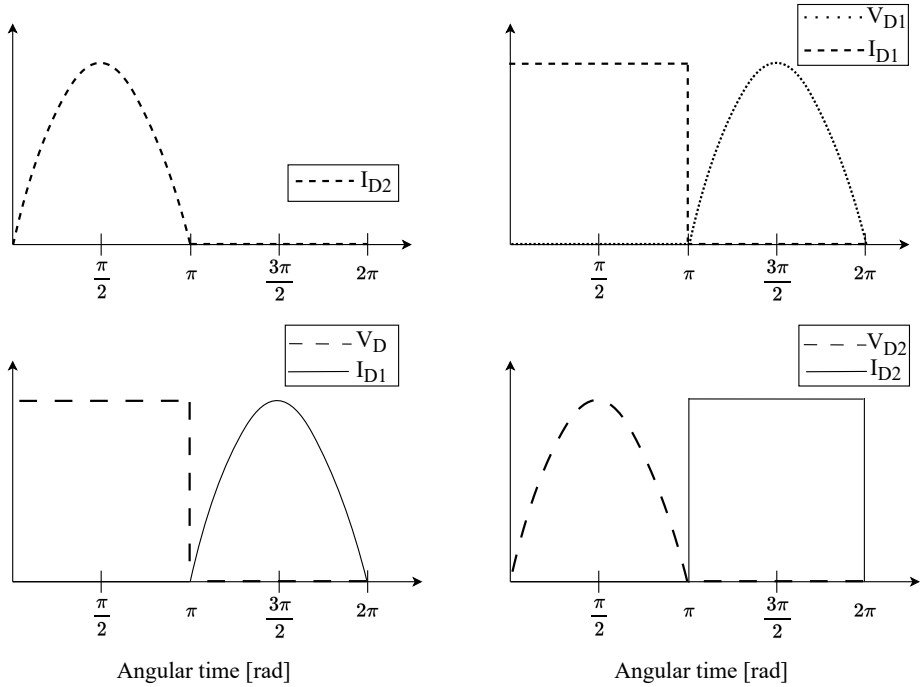


Figure 2.13: Ideal voltage and current waveforms for a voltage-mode class-D and a current-mode class-D, respectively.

drain efficiency being 100%, in reality this value is way lower. For this reason, class-D PAs are usually implemented along with linearization and efficiency enhancement techniques. Besides allowing to achieve better efficiency and linearity, some of these techniques (such as Polar and Outphasing) also allow the operation of class-D PAs at higher frequencies. References [6, 12] are examples of a class-D PA implemented with Polar and Outphasing techniques at GHz frequencies. These techniques will be discussed in the next section.

ii. Class-E Power Amplifier

In the class-E switched power amplifier, the switch operation is performed using a single transistor. Two types of class-E PAs are to be considered: zero-voltage switching (ZVS) and zero-current switching (ZCS). Let us focus on the more basic topology, class E ZVS, depicted in Fig. 2.14.

To understand the principle of the class-E PA, let us observe Fig. 2.14. One can see that when the switch is turned on and the current flows through the drain branch, the L_1 inductor is charged. By the time the switch is turned off, all this current flows through the C_1 capacitor and the voltage at the V_D node increases. The L_0C_0 filter forces this voltage to return to zero at the time the switch is turned on again. As a result, unlike class-D, class-E PAs allow for a less sudden decrease of voltage due to the C_1 capacitor (that includes the inherent parasitic capacitance of the transistor). Besides, in view of the fact that the voltage is zero when the switch is off, there are no switching losses [39]. This effect, also studied in the previous section, is the ZVS.

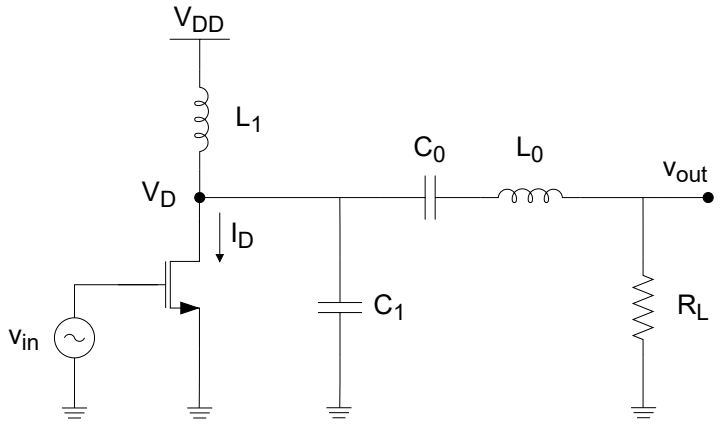


Figure 2.14: Class-E PA.

The class-E PA can also be designed to operate as Zero Current Switching (ZCS), where the switch is turned off at zero current. In such case, its topology does not include the C_1 capacitor. As a result, when the switch is selected, there are losses due to the output parasitic capacitance. For this reason, class-E ZCS show less efficiency than class-E ZVS when operating at higher frequencies. [27]

Similarly to other switch modes PAs, the theoretical efficiency of the class-E PA is 100%. However, any stress of the device provoked by a large peak of drain voltage leads the efficiency to decrease. Besides, it is necessary to ensure that the switch operation is fast enough so that the requirements imposed by the drain current are accomplished [39]. Since the transistor capacitance is absorbed by the matching network, losses no longer occur due to the switching operation. As a consequence, the class-E PA is widely used in high-frequency applications.

Different implementations can be found in the literature. Reference [16] shows a differential class-E PA implemented in 0.18μm CMOS technology for a 2.4GHz frequency. They were able to obtain an output power of 28.5dBm with a drain efficiency of 45%. References [1, 2] are also examples of cascode and differential implementations of this amplifier.

ii. Class-F Power Amplifier

The class-F power amplifier takes advantage of the load network to resonate both the fundamental frequency and its harmonics. The idea behind this technique is to eliminate the undesired harmonics of the signal by controlling the way they are terminated. Fig. 2.15 shows the generic circuit of a class F PA. Here, the switching operation is done by using a single transistor and an adjustable load network is designed to tune the harmonics of the input signal at the fundamental frequency.

Looking into the picture, it is possible to see that if a very high impedance is provided to the odd harmonics (which results in an open circuit) and a very low impedance is provided to the even harmonics (which results in a short circuit), a square wave voltage

is obtained at the output of the resonator blocks. Note that for this to happen, the PA needs to be biased at the threshold voltage, similar to the class B PA, so the current is a half-rectified sine wave. Naturally, the higher the number of resonators, the better the filtering of the harmonics and the closer the output power is to a perfect square wave. Notwithstanding, the limited quality factor of the reactive components typically limits the maximum efficiency of the PA and, consequently, the number of resonator blocks used. Since a higher number of resonator blocks imply a higher number of inductors, the number of blocks is usually limited to an optimum value that allows achieving high efficiency.

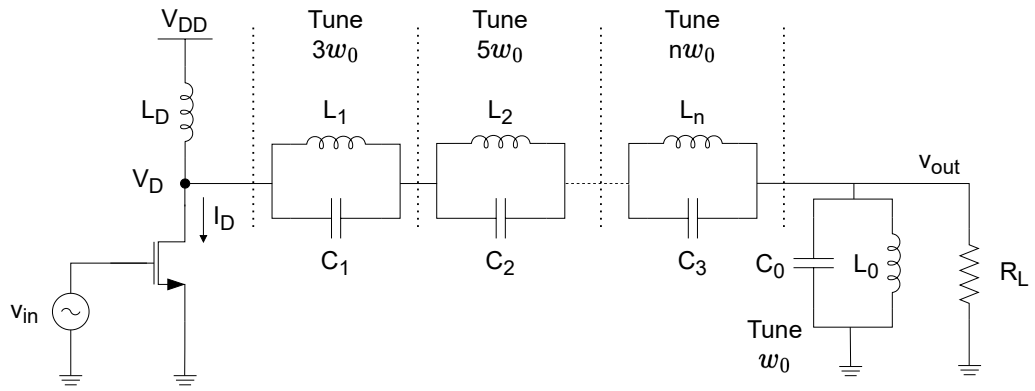


Figure 2.15: Class-F PA.

These type of PAs can also be designed to provide an open circuit to the even harmonics and a short circuit to the odd harmonics. In this case, the PA is biased like in a class-A and square wave current is obtained at the output. This configuration is named inverse class F. In this ideal configuration, where the load network uses ideal components and the intrinsic parasitics of the transistor are not considered, no power dissipation occurs, leading to a maximum theoretical efficiency of 100%. However, in a real implementation, this theoretical value is never achieved.

Reference [17] presents an implementation of a CMOS class-F PA designed to operate at 2.4GHz. In this implementation, the load network was designed to filter up to the third harmonic. They were able to obtain a maximum output power of 19.2dBm with a maximum PAE of 26%. Reference [11] presents an implementation in Gallium Nitride (GaN) of a class-F and inverse class-F dual bands PA, to operate at 1.4GHz and 2.1GHz. The results show the PA achieves a drain efficiency higher than 72% and 69% in the two operation bands.

2.2.5.3 Comparison between Power Amplifier Classes

This section aims to summarize the performance parameters of the different PA classes studied in this section. Table 2.1 presents the trade-off between linearity and efficiency among all class modes.

Table 2.1: Theoretical drain efficiency and linearity in RF PA classes.

Class	Mode	Conduction angle [%]	Max. drain efficiency [%]	Linearity
A	Linear	100	50	Good
AB	Linear	50-100	50-78.5	Good
B	Linear	50	78.5	Moderate
C	Linear	<50	78.5-100	Poor
D	Switch	50	100	Poor
E	Switch	50	100	Poor
F	Switch	50	100	Poor

2.2.6 Efficiency and Linearization Enhancement Techniques

As mentioned before, in order to surpass the degradation in efficiency at the output power back-off [2] and boost the linear performance of a single PA, different efficiency enhancement and linearization techniques have been proposed. In the section herein, those techniques and architectures will be examined, along with their advantages and drawbacks. Note: Some of these techniques already integrate the transmitter architecture analysed in chapter 2.1.

2.2.6.1 Doherty Technique

One way to enhance the efficiency of a linear power amplifier is to use the Doherty technique. Its traditional architecture, illustrated in Fig.2.16, combines two linear power amplifiers in parallel: the main or carrier power amplifier and the auxiliary or peaking power amplifier. The main PA, usually biased in class A, B or AB, operates over the whole power range, while the auxiliary PA, biased in class C, only operates when a higher output power level, usually 6dB below saturation power, is required and the main PA begins to saturate [41, 39].

The idea behind this technique is to have only the main PA operating in linear mode at the maximum efficiency for low-input drives. When the input drive increases and the output power level saturates to a point where the voltage gain starts to decrease, the auxiliary PA begins to supply current to the load, increasing the output power level and the impedance seen at the output node. By using a quarter-wavelength transmission line (impedance inverter), it is ensured that the main PA output impedance decreases up to $2R_L$. To do so, another quarter-wavelength transmission line is necessary at the input of the Auxiliary PA to compensate for the 90º phase shift introduced by the latter [39]. Therefore, by adjusting the amount of auxiliary PA output current, it is possible to guarantee that the output power level is such that the efficiency of the main PA is maximum for the whole input power range. However, quarter-wavelength transmission lines are not

[2] As power amplifiers operate near the nonlinear region to maximize their efficiency, every slight increase in power leads to output power saturation. The power backoff corresponds to the amount of power that has to be subtracted to avoid this effect.

practical when it comes to the use of RF signals, since they occupy a very large area. As a result, many implementations have been replacing the transmission lines with lumped elements or SCT transformers [39].

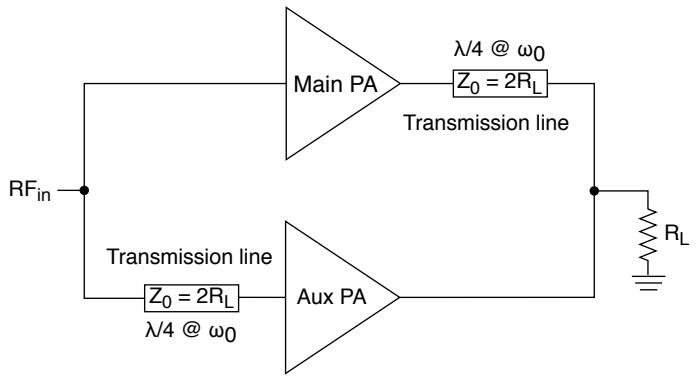


Figure 2.16: Doherty PA architecture. Adapted from [39], Chapter 5.

Reference [28] presents an example of a conventional Doherty PA implemented in 0.18μm CMOS technology. References [26, 3] are examples of transformer-based Doherty PAs. Reference [56] shows a "cascode-cascade" Doherty PA implemented with lumped elements (LC π networks).

2.2.6.2 Outphasing

Outphasing, also known as linear amplification with nonlinear components (LINC), is a technique where the varying envelope input signal is decomposed through a signal component separator into two signals with different phases and constant envelopes. Each of these signals is amplified and then recombined together at the output (with a passive power combiner) to reconstruct the desired signal. This procedure is depicted in Fig.2.17.

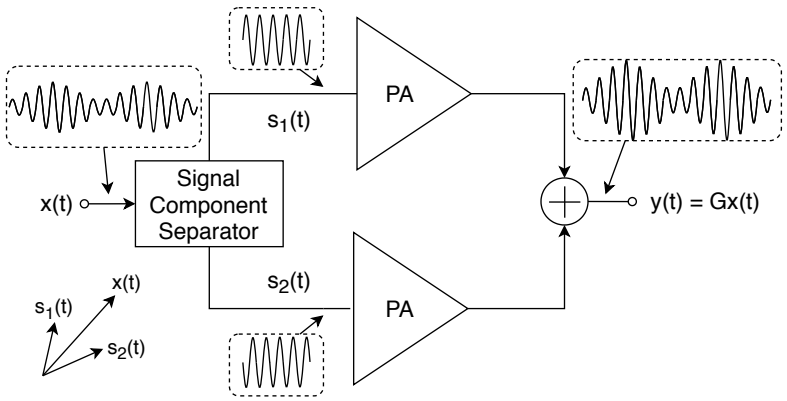


Figure 2.17: Outphasing/LINC PA architecture.

Observing Fig.2.17, one can see that when signals $s_1(t)$ and $s_2(t)$ are in phase, the resulting signal has a larger envelope, whereas when the signals are in antiphase, the

amplitude of the combined signals is cancelled out, producing a low envelope output signal. Note that for this technique to work properly, a delay matching between the two paths is required.

Since PAs amplify constant envelope signals, switching PAs can be used. This allows for very high efficiency without spoiling linearity. Notwithstanding, the power combiner used to obtain the modulated output signal is highly consuming, leading to power waste.

References [4, 20, 18, 15] are examples of PAs implemented with outphasing techniques.

2.2.6.3 Predistorted PAs

A commonly used technique to improve the linearity of the PA is the Predistortion. In such technique, instead of applying the input signal directly in the PA, the signal is first predistorted. By doing this, the non-idealities of the distorted signal are cancelled out during the amplification process.

The predistortion can be either analogue or digital. Typically, in an analogue implementation, a circuit is connected in parallel to the transistor bias. This circuit is used to cancel out the distortions caused by the nonidealities of the amplifier (in particular, the intrinsic parasitic capacitances). However, this kind of predistortion usually offers small improvements and it is not sensitive to any variations of the conditions at which the PA operates [39].

On the other hand, in a digital implementation, the input signal is typically used by a Lookup table (LUT) that controls the predistortion circuit. Usually, this solution offers better results when compared to the analog counterpart, although it is also not able to adapt to the variations in the PA conditions. In order to surpass this issue, closed-loop solutions can be implemented, where the lookup table receives an attenuated version of the output signal in baseband. In this way, the predistortion circuit is also controlled by the output signal of the PA. However, this solution increases the complexity of the architecture which is translated into higher costs [39].

In some works, the combination of the analogue and digital predistortions can also be found. In reference [8], two digital signals are predistorted digitally and modulated using different carriers. The two signals are applied to an analogue predistorter and a single RF signal is sent to the PA. The results show an improvement in the output spectrum, where the intermodulation products are clearly attenuated. In this solution, the lookup table concept is not employed. Reference [37], on the other hand, implements a narrowband dithering technique in feedback channel quantization to improve the effect of a closed-loop digital predistorter. Here, the idea is to inject a dithering signal before the quantization to reduce the non-linear effects and be filtered by the narrow-band digital filter. This combined solution allowed to improve the performance of the overall system by more than 4dBm.

2.2.6.4 Polar or Envelope Elimination and Restoration Technique

In the Envelope Elimination and Restoration (EER) technique, whose architecture is presented in Fig.2.18, the RF input signal is processed along two different paths. In the first path, the signal is passed through a limiter, where its envelope is eliminated, resulting in a constant phase-modulated signal that is used to drive the input of the PA. The phase-modulated signal can then be amplified. In the meanwhile, in the second path, the signal goes through an envelope detector where its envelope is obtained and further sent to the PA supply modulator (Low Frequency Amplifier). Depending on the envelope information, the modulator varies the supply voltage going to the RF PA, so the input envelope is restored. Hence, the PA combines the phase information received at the input with the envelope information, to restore the signal at the output.

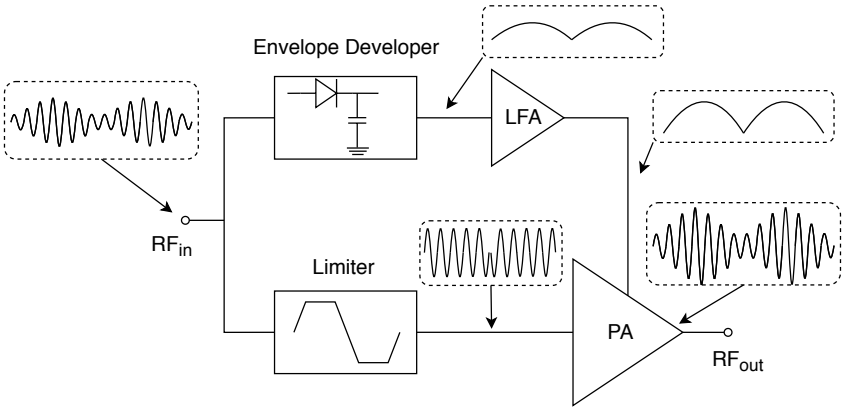


Figure 2.18: EER PA architecture. Adapted from [39], Chapter 5.

A similar approach can be made where the separation of the input signal into phase and amplitude components is made in baseband instead of RF. Thereby, a DSP is used to separate the baseband signal into both components. Hence, the constant phase-modulated signal is upconverted to the carrier frequency provided by the LO and sent to the PA. For this reason, the limiter and the envelope detector are dismissed from this architecture. This technique is also known as Polar technique and its block diagram is illustrated in Fig. 2.19.

Since the PA's input signal has a constant envelope, switch-type PAs are the most used in these techniques. For this reason, EER and Polar architectures can achieve high efficiency and provide good linearity.

Regardless of its favourable characteristics, this architecture poses some challenges. Firstly, in order to efficiently restore the input signal, it is necessary to guarantee that the envelope and the RF paths are time aligned. Secondly, bearing in mind that the overall efficiency depends on both the supply modulator and the PA, a high-efficiency modulator is needed. Nevertheless, the efficiency of this kind of modulator is affected when the input signal bandwidth increases. Lastly, considering that the phase-modulated signal

entails a higher bandwidth than the original input signal, this technique is confined to narrowband applications. [39]

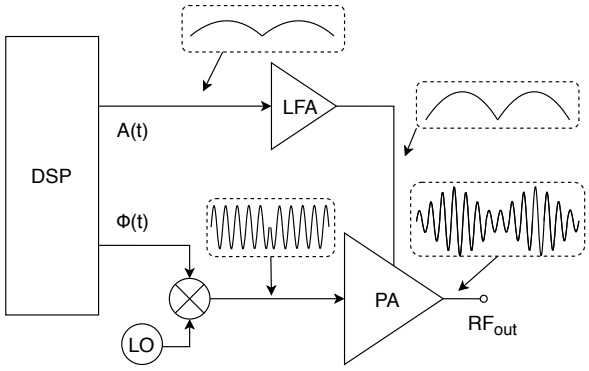


Figure 2.19: Polar/EER PA architecture. Adapted from [39], Chapter 5.

.

Reference [45] presents an EER architecture that uses a class-E PA and a class G supply modulator in 130nm CMOS. The class G linear supply modulator operates with two power supply voltages. The first power supply voltage is used for as long as the input signal amplitude is smaller than a certain threshold. Above this point, the second power supply voltage is used instead. The dynamic nature of the supply modulator allows to achieve higher efficiency. References [30, 38] show PAs that use EER and EER/Polar techniques.

2.2.6.5 Envelope Tracking Technique

The Envelope Tracking (ET) emerges as an alternative technique to EER, where the limiter is removed from the architecture, as shown in Fig. 2.20.

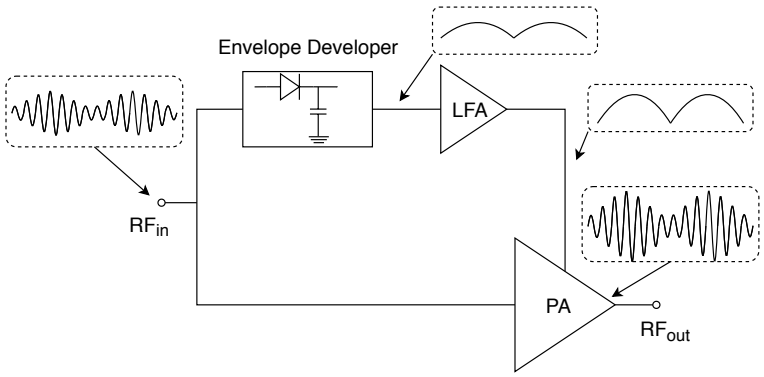


Figure 2.20: ET architecture. Adapted from [39], Chapter 5.

Observing Fig. 2.20, one can see that, by removing the limiter, the PA's input signal includes both phase and amplitude information, while with the EER technique, the PA's

input signal contains phase information only. As a consequence, the envelope signal that goes through the supply modulator is now only used to provide the required voltage to the PA. The modulator supplies the PA with just enough voltage (that varies according to the envelope of the modulated input signal) so that the input signal can be amplified.

In view of the fact that both the amplitude and phase information are provided at the input of the PA, the requirements of phase matching between the RF and the envelope paths are softened. Moreover, because the PA's input signal and the baseband signal bandwidths are the same, this architecture makes it possible to work with signals with larger bandwidth than EER architectures. However, unlike EER, the PA's input signal has a varying envelope, so this technique aims to enhance the efficiency of linear PAs.

Reference [48] presents an ET PA implemented in 0.18μm siGe BiCMOS technology. In references [24] and [22] ET PAs are implemented in 0.18μm CMOS technology. The ET PA architecture may also be combined with other above-mentioned techniques. References [49, 23] are examples of an ET PA technique combined with EER and Doherty techniques, respectively.

2.2.6.6 RF-DAC based PA

Every digital-based power amplifier architecture springs from the EER architecture, studied in the previous section. Here, the PA and the supply modulator are replaced by a power RF DAC composed of several units of PA cells. For this reason, DPAs are also known as RF-Power-DACs. In this architecture, the unit PA cells are connected in parallel and selected according to a digital code word that depends on the envelope of the input signal. Thus, by selecting the unit PA cells to be summed at the output, it is possible to modulate the output signal [50, 47, 25]. Fig.2.21 illustrates this procedure for a polar configuration.

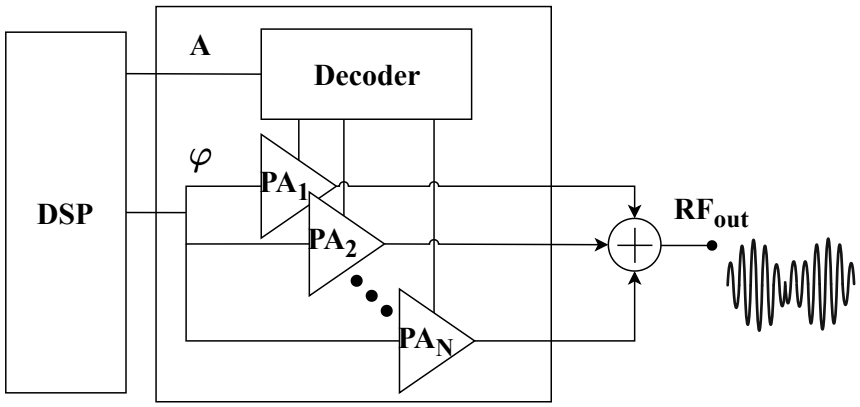
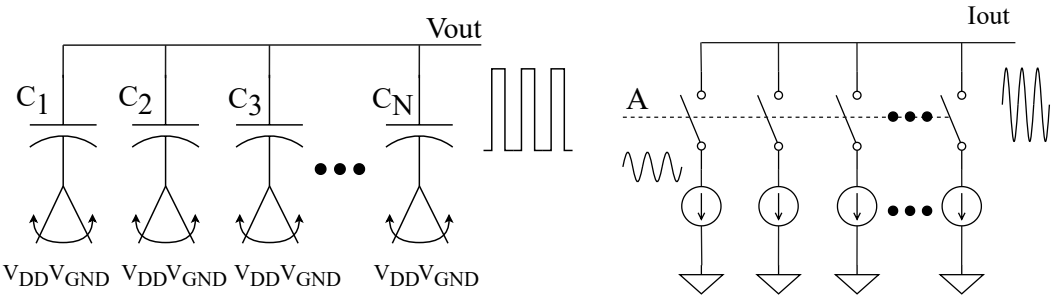


Figure 2.21: Block diagram of the DPA composed of unit PA cells and a decoder. Adapted from [50], Chapter 8.

Broadly speaking, DPAs can be sorted into current mode and voltage mode. When operating in current mode, unit current cells are switched at the carrier frequency. Each

current cell is made of a transistor operating as a current source and the output current is obtained by enabling or disabling these current cells [50, 45]. Fig. 2.22(b) illustrates this current switching structure. DPAs of this kind are inherently nonlinear due to the finite output impedance of the transistor when it operates as a current source. As a result, it can be very difficult to obtain a high output power and guarantee both high efficiency and linearity at the same time. For this reason, it is of major concern to design each current cell with high-output impedance to boost linearity. Typically, most of these conventional DPAs are implemented with additional bits of resolution (more unit PA cells) along with linearization techniques (such as digital predistortion) to enhance the overall linearity [50, 45]. References [46, 13, 43] show examples of current-mode DPAs implemented in polar configuration. The last two are specifically designed for NB-IoT applications. Reference [10] presents a quadrature dual-band transmitter with a current-mode DPA.



(a) Voltage Mode digital based PA

(b) Current Mode digital based PA

Figure 2.22: DPA technique: (a) based on switch capacitors (b) based on switch current sources. Adapted from [50], Chapter 8.

In turn, in voltage mode DPAs (also named SCPAs), the transistor of each PA cell acts as a switch and controls the amount of charge that is distributed to an array of capacitors [50]. As shown in Fig 2.22(a), depending on the envelope of the input signal, the bottom plates of the capacitors are either switched, at the carrier frequency, between ground or the supply voltage or held at the ground. Unlike current mode DPAs, SCPAs do not suffer from the abovementioned nonlinearities. However, due to the device's capacitances and the nonlinear behaviour of the Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) switches, the linearity is slightly affected. Additionally, this kind of digital amplifier poses a limitation: the maximum power output signal is restricted to the supply voltage. Despite this, SCPAs can easily yield a high output power with high efficiency and good linearity. Furthermore, this technique has the advantage of scaling very easily with CMOS technology with weakened parasitics: firstly, because it uses capacitors that are area efficient and, secondly, because it uses fast CMOS switches. All things considered, voltage-mode DPAs and, in particular, SCPAs, are usually the preferred choice within digital PAs.

Reference [47] presents a fully integrated SCPA for polar transmitters. The circuit was implemented in a 0.90μm CMOS process and achieves peak P_{out} and PAE values of 25.2dBm and 45% respectively. Other voltage-mode DPA configurations have also been implemented: references [55, 21] present a SCPA in quadrature configuration and reference [5] in multiphase configuration. References [53, 44] are examples of Class-G SCPA implementations that achieve high power efficiency at the output power backoff by using two power supply voltages (and thus having the supply voltage greater than V_{DD}). This allows to work around the problem that typical SCPA implementations face of having their output voltage limited by the supply voltage.

2.2.6.7 Comparison of Techniques

Table 2.2 presents a brief overview of the major advantages and drawbacks related to each efficiency and linearization technique studied in the previous section.

Table 2.2: Comparison of efficiency and linearization techniques.

Technique	Advantages	Drawbacks
Doherty	Possible use of lumped elements or SCT transformers; Easy implementation;	Large overall area; Less efficient (Linear PAs); Narrowbandwidth; High synchronization;
Outphasing	Very high efficiency; Easy implementation;	Power combiner; High synchronization; Narrowbandwidth;
Predistortion	Increases linearity; Increases efficiency	Digital implementations increase complexity and costs; Requires high bandwidth;
EER/Polar	Very high efficiency with good linearity;	Degradation of the modulator efficiency for wideband signals; Narrowbandwidth; High synchronization;
RF-based-PA	Very high efficiency with good linearity; No modulator used;	Narrowbandwidth;
ET	Low Synchronization; Lower modulator bandwidth needed compared to EER;	Less efficient (Linear PAs);

2.2.7 Comparison of Published PAs

A vast number of works have been published in the last years using different technologies and targeting several applications. Most of these works have been collected, examined and can be found in the survey presented in [33]. So as to have a better understanding of how the PA's performance varies according to the technology, some of these works were extracted from this survey and are presented in this section.

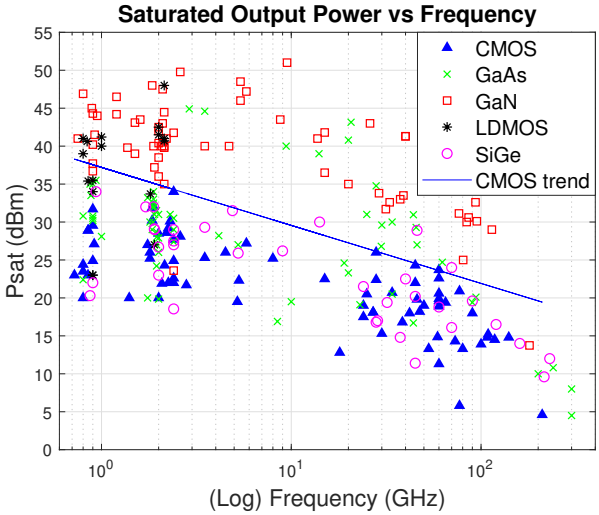


Figure 2.23: Saturated Output Power and Frequency of the SoA Power amplifiers.

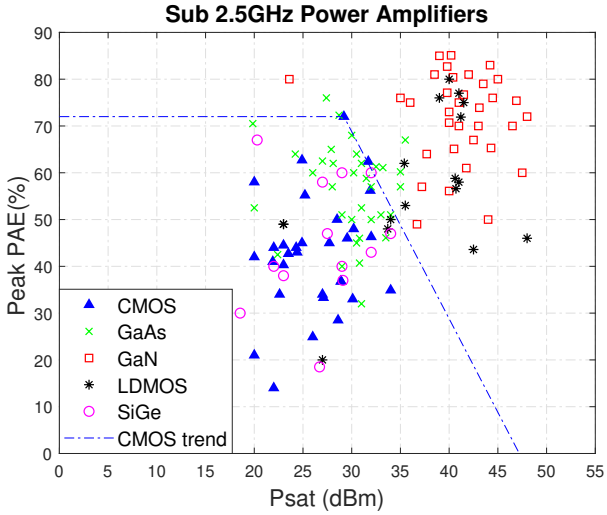


Figure 2.24: Peak Power Added Efficiency and Saturated Output Power of the SoA sub-2.5GHz Power amplifiers.

Fig. 2.23 shows the saturated output power of the PA as a function of the frequency for CMOS, Gallium Arsenide (GaAs), GaN, Laterally Diffused Metal Oxide Semiconductor (LDMOS) and Silicon Germanium (SiGe) technologies. Fig. 2.24 depicts the peak PAE

as a function of the saturated output power for the same abovementioned technologies, when the frequency application is below 2.5GHz.

From Fig.2.23 it is possible to observe that, regardless of the technology used, increasing the frequency application limits the maximum output power. Besides this, it is also possible to conclude that CMOS and SiGe technologies present the worst performance among all. This result is also clear when looking into the data presented in Fig.2.24. This suggests that there is a limitation in the design of PAs when implemented with CMOS technology (marked in the figures with a blue line). These limitations are mainly due to the limited voltage supply (that can go up to $2V_{DD}$), along with breakdown phenomena (oxide breakdown) and hot carrier degradation, that arise when the technology is scaled down and the electrical field in the oxide layer increases. Moreover, the limited voltage swing due to the threshold voltage, the losses of integrated inductors and transformers and the losses of the transistors also play an important role in the CMOS poor performance [39].

3

Design and Implementation in CMOS technology

3.1 Switched Capacitor RF Power Amplifier

Broadly speaking, a SCPA consists of an array of capacitors that are connected to the load through a matching network, as represented in figure 3.1. The array of capacitors is driven by a square wave generated by the local oscillator at the RF carrier frequency.

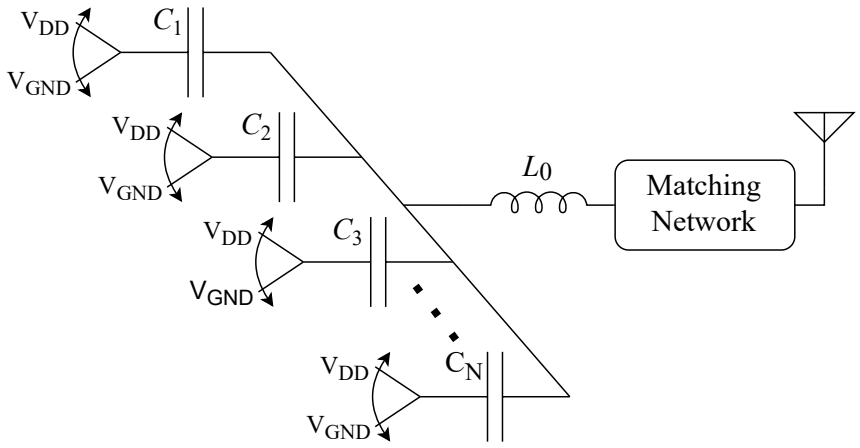


Figure 3.1: Ideal model of a single-ended SCPA.

In such architecture, the bottom plate of the capacitors can be either dynamically switched between the supply voltage and ground or simply connected to ground, depending on the digital input code, proportional to the envelope of the input signal. Assuming N as the total number of capacitors in the SCPA, the array can be divided between the n on capacitors, which are switched, and the $N - n$ off capacitors, which are connected to ground. Typically, all the capacitors in the array have the same capacitance value. Note that, because the top plates are never switched, the total capacitance connected to the matching network is always constant, regardless of the number of capacitors which are effectively switched. Furthermore, by using a lossless inductor connected between the

array of capacitors and the matching network, one can create a simple filter that will tune the output voltage at the fundamental frequency. Similarly, the matching network acts as a second filter and has the additional role of transforming the load resistance into the optimum resistance (this topic will be discussed in more detail in section 3.3.4). The described behaviour is depicted in figure 3.2.

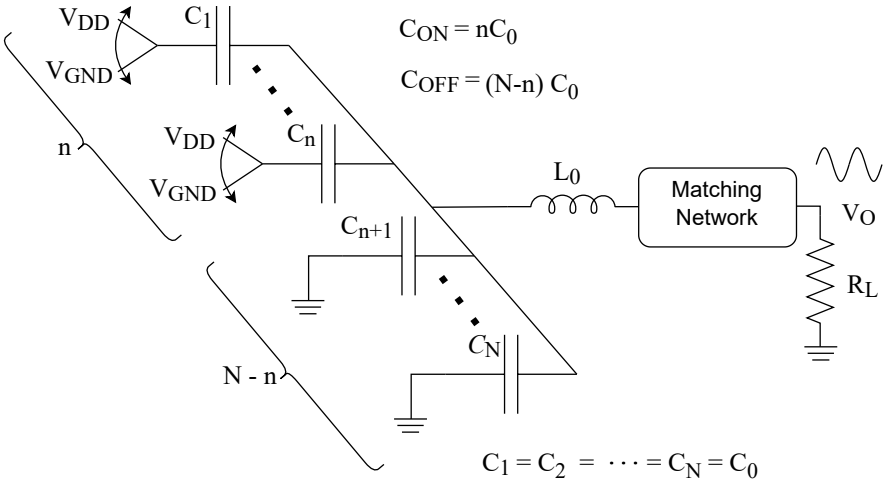


Figure 3.2: Example of the SCPA operation mode when n capacitors are switched between V_{DD} and ground and (N-n) capacitors are connected to ground.

As a result, the output voltage of the SCPA will be proportional to the number of switched on capacitors, $V_o \propto \left(\frac{\sum C_{ON}}{C}\right) V_{DD}$, and its maximum value is achieved when all the capacitors are on, as expressed in the following equation:

$$V_o = \frac{2}{\pi}\left(\frac{\sum C_{ON}}{C}\right) V_{DD} = \frac{2}{\pi}\left(\frac{n}{N}\right) V_{DD}, \tag{3.1}$$

where V_{DD} corresponds to the maximum voltage applied at the bottom plates of the capacitors. Note that, the maximum voltage is multiplied by a factor of $2/\pi$. This factor, that will be analysed in more detail in chapter 3.4.1, corresponds to the first coefficient of the fourier series and is the result of the filtering of the square waveform. In this way, a sinusoidal waveform is obtained at the desired frequency. The output power on the optimum resistance will increase with the output voltage according to the equation 3.2.

$$P_o = \frac{V_o^2}{R_{OPT}} = \frac{2}{\pi^2}\left(\frac{n}{N}\right)^2 \frac{V_{DD}^2}{R_{OPT}}. \tag{3.2}$$

Let us now calculate, for this ideal analysis, the efficiency of the SCPA. Assuming that the switching action at the bottom plate of the capacitors is modelled by ideal switches, then no resistive nor capacitive losses occur. In view of this, the only loss mechanism

is related to the power required to charge and discharge the capacitor array. In order to calculate this power loss, let us take a look at figure 3.3.

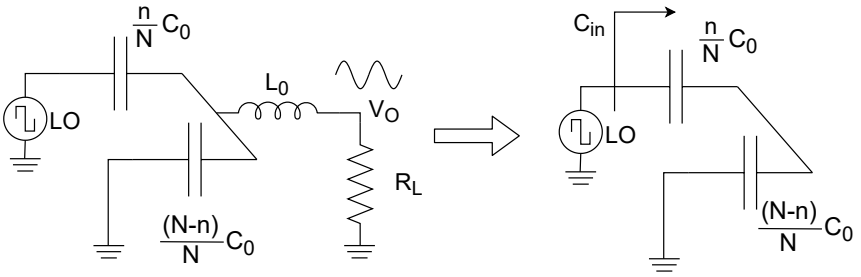


Figure 3.3: Ideal simplified model of the SCPA and equivalent circuit use to calculate the P_{SC}.

Looking at figure 3.3, it is possible to perceive that all the n switched and $N-n$ non-switched capacitors can be reassembled into two capacitors whose capacitance is multiplied by n and $N-n$, respectively. If one assumes that the switching square waveform that drives the on capacitors has sufficiently sharp edges (fast switching transitions), then it is possible to approximate the inductor as a constant current source. In these conditions, the inductor can be seen as an open circuit and the effective input capacitance C_{in} driven by the selected switches can be calculated as the series of the two capacitances:

$$C_{in} = \frac{n(N-n)}{N^2} C_0, \tag{3.3}$$

where C_0 is the capacitance value of each individual capacitor of the array. Thus, the power needed to charge and discharge the array P_{SC} can be calculated according to equation 3.4:

$$P_{SC} = C_{in} V_{DD}^2 f. \tag{3.4}$$

Considering this, one can then obtain the ideal power added efficiency of the SCPA given by equation 3.5:

$$PAE_{ideal} = \frac{P_o}{P_o + P_{SC}}. \tag{3.5}$$

It can also be useful to look at efficiency as a function of the quality factor of the network. This parameter evaluates the efficiency of the ideal SCPA by measuring the quality of the resonant circuit formed by the LC filter and it can be estimated using the equation described in 3.6.

$$Q_{loaded} = \frac{2\pi f L_0}{R_{OPT}} = \frac{1}{2\pi f C_0 R_{OPT}}. \tag{3.6}$$

Combining equations 3.6 and 3.3 and rewriting 3.4, one can obtain the efficiency given by equation 3.7:

$$PAE_{ideal} = \frac{4n^2}{4n^2 + \frac{\pi n(N-n)}{Qloaded}}. \tag{3.7}$$

So as to have a better understanding of how the Q_{loaded} and the number of switched-on capacitors affect the efficiency of the SCPA, let us analyse figure 3.4. In this figure, one can see the ideal PAE of an SCPA for $Q_{loaded} = 1$, $Q_{loaded} = 3$ and $Q_{loaded} = 10$, along with the ideal PAE of a conventional current mode DPA and of a classic class-A PA studied in the previous chapter.

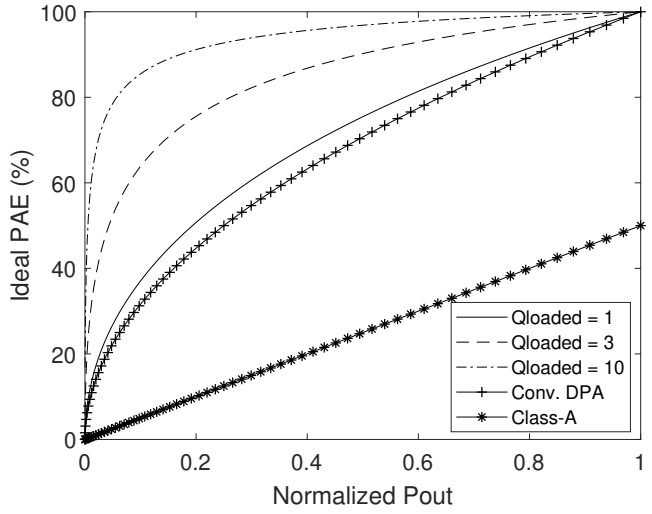


Figure 3.4: Comparison of the PAE efficiency versus the normalized output power for an SCPA with different Qloaded, a current-mode DPA and a classic class-A power amplifier - obtained using equation 3.5, equation 2.4 and the equation presented in [25].

By inspection, it is possible to observe that the quality factor of the network has a significant impact on the efficiency of the SCPA. This means that, by assuring a high Q_{loaded}, one can guarantee a higher efficiency. However, the value of this parameter is highly limited by the quality factor of the reactive components on chip (in particular the inductor), which also limits the efficiency. What is more, the ideal efficiency of 100% is achieved when all the capacitors are switched on (for higher output power) and its value degrades when the number of switched-on capacitors decrease (for lower output power).

Another important aspect to note is that the ideal efficiency of the SCPA is always better than the one obtained using a current mode DPA or a class-A PA, regardless of the Q_{loaded} value of the network and the number of switched on capacitors. This difference increases when a higher Q_{loaded} is ensured.

Nevertheless, in a practical implementation, there are several sources of power loss that must be taken into account and that will strongly degrade the maximum efficiency. In order to study this matter, let us first analyse the selected architecture.

3.2 Proposed Architecture

The proposed architecture of the PA is presented in figure 3.5 and it follows the structure of the first SCPA presented in reference [47]. This architecture includes a modulation stage (represented at the bottom of the figure) that is responsible for controlling the N parallel unit cells that are connected to the antenna through the matching network.

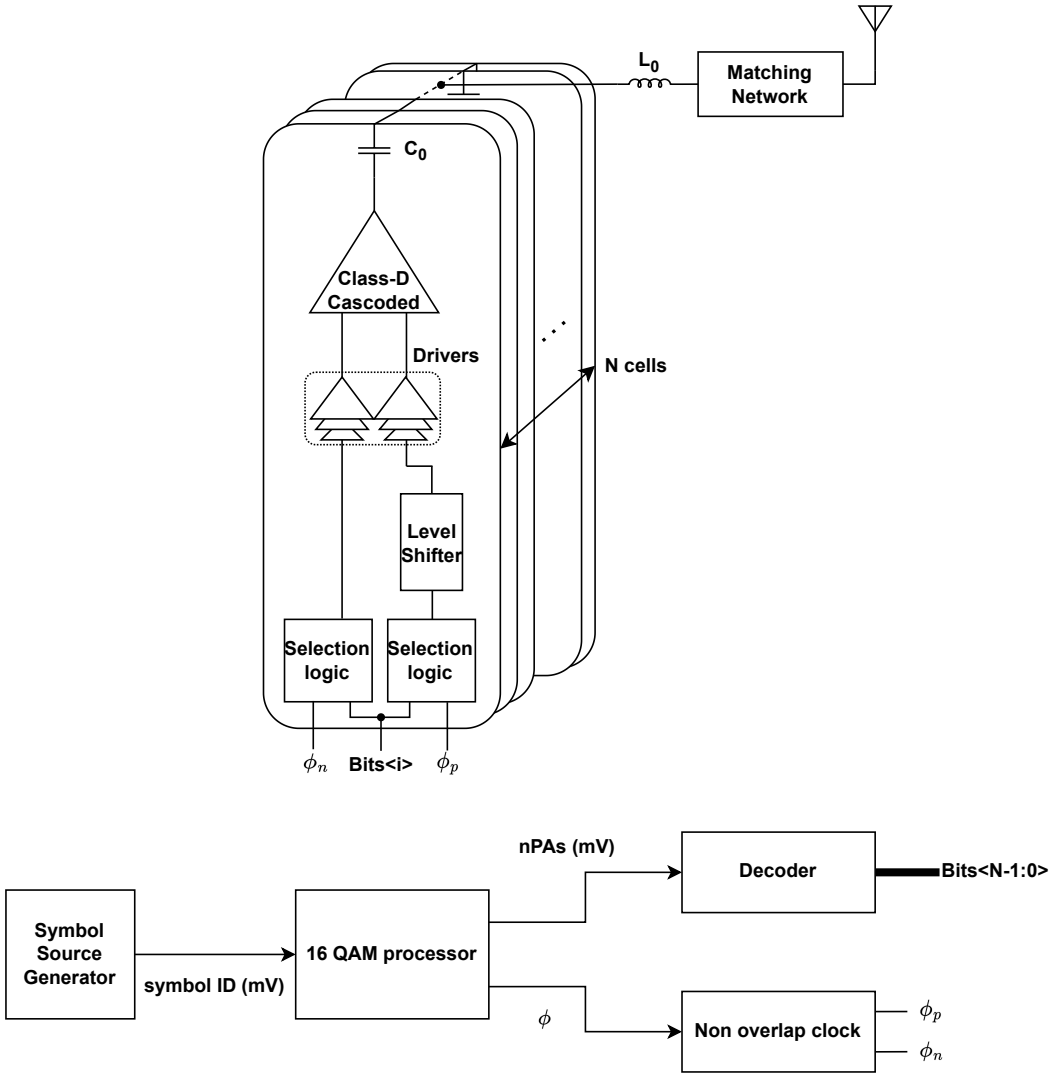


Figure 3.5: Proposed architecture of the SCPA.

The modulation stage, designed for a 16 QAM modulation, is composed by four blocks: the symbol source generator, the 16 QAM processor, the decoder and the non-overlap clock. Together, the blocks are responsible for deciding the number of PAs and the phase clocks that need to be propagated to the different N parallel unit cells (through the signals Bit<N-1:0>, Φ_n and Φ_p), according to the symbol that is being transmitted. Although in this figure two non-overlap clocks are being represented, in reality, a single clock will be employed. This topic will be further analysed in this chapter.

The output signals of the modulation stage are then sent to the N parallel unit PA cells and used as an input in the two different selection logic blocks (one, that will feed the PMOS of the class-D cascoded switch and the other that will feed the NMOS of the same block). Depending on the value of the *Bits* signal, each unit PA cell is either dynamically switched according to the phase clock or tied to ground. In this work, two selection blocks will be studied: the AND gate and the transmission gate. Besides this, an additional block, the level shifter, needs to be added in the PMOS path of the cascoded class-D in order to shift the phase clock by 1.2V. After the selection logic and the level shifter blocks, a driver needs to be added in the architecture in order to drive the capacitances of the cascoded class-D PA. This block consists of a stage of inverters connected in a cascade configuration. In this work, the impact of the number of inverters on the performance of the architecture will be carefully analysed. Finally, following the driver, the main switches of the architecture, made of cascoded class-D PAs, are implemented. The resulting signal at the output of the filter of the N unit PA cells will then be transformed by the matching network and sent to the antenna.

3.3 Matching Network

In order to obtain the desired output power, it is necessary to define an optimal working impedance, which does not correspond to the typical impedance value of 50 Ω of the on-chip load. In fact, since the aim of a RF PA is to deliver the maximum output power, it is desirable to have a very small resistive impedance when compared to the resistive impedance of the antenna. Bearing this in mind, it is of the utmost importance to guarantee that the load impedance is transformed into the optimal impedance. This means that, when looking into the antenna direction from the optimal impedance perspective, the characteristic impedance seen has to be the optimal impedance and, when looking into the circuit direction from the load impedance perspective, the characteristic impedance seen has to be the load impedance. This impedance-matching process is done by using a matching network.

3.3.1 Impedance Matching Concept

The idea of impedance matching is more complex than simply transforming the load impedance into the desired optimum impedance. In fact, for some applications where the impedance matching is fundamental, the circuit is not even designed to have a specific optimum impedance. For this reason, before presenting the chosen matching network, let us first analyse in more detail the concept of impedance matching.

For circuits operating at very high frequencies, the period time of a sinusoidal signal or the period time of a digital square wave becomes smaller in such a manner that it is comparable to the propagation time of the circuit. When this happens, the signal varies inside the conductor itself and the analysis of the circuit must take into consideration

distributed parameters. In this analysis, the conductor is seen as a transmission line where the propagated waves represent the input signal. Note that, these waves are a function of time and space, which means that, at a certain instant of time, the voltage across the line varies according to the distance to the source. In transmission lines, there are always two waves being propagated at the same time, the incident wave and the reflected wave, regardless of how the line is terminated. For instance, if the line is terminated in an open circuit, there will be always an incident wave as a result of the size of the line and, because there is no load to absorb the power of the transmitted wave, all the energy has to return to the source, creating a reflected wave.

So as to have a better understanding, let us consider V_i as the amplitude of the incident wave, V_r as the amplitude of the reflected wave and ρ as the reflection factor, defined as the ratio between the two waves, $\rho = V_r/V_i$. As V_r is the reflected wave, it is fair to assume that $|V_r| < |V_i|$. What is more, considering that the reflected wave can present the same phase or the opposite phase of the incident wave, ρ can vary in the following way $-1 \leq \rho \leq 1$. For a reflection factor of -1 (opposite phase) the line is terminated in a short circuit and for a reflection factor of 1 (same phase) the line is terminated in an open circuit. Ideally, the load should absorb the maximum power possible, which means that no reflected wave should be propagated and $\rho = 0$. The point at which this condition is verified depends not only on the load, but also on the impedance characteristic of the line. This impedance is a parameter of the line and can be measured, at any point, as the ratio of the voltage and current. When $\rho = 0$, the load impedance and the characteristic impedance of the line are matched. However, the impedance of the load that represents the circuit not always equals the characteristic impedance of the line. For this reason, it is necessary to find a way to adapt the circuit by using a matching network.

3.3.2 Cascaded L-Shape Matching Network

There are different types of matching networks, as they can be implemented using either lumped elements, which involve passive components as capacitors and inductors, or distributed elements, that employ lines and stubs. Typically, the former is mainly used for circuits operating at a few GHz, whereas the latter is preferred when the circuits operate at very high frequencies. Regardless of the elements used, a matching network can be designed to maximize the power transfer, to maximize the gain or to minimize the noise and it can be projected to be narrowband or wideband, depending on the application purpose. For this project, a lumped-element narrowband network will be designed to maximize the power transfer. Let us now discuss some aspects to consider in its design.

Firstly, the switching action of the SCPA generates a lot of higher harmonics at the output that need to be attenuated to an acceptable level or filtered out. To achieve this, an ideal bandpass matching network or low pass matching network must be employed in order to eliminate all the spurious content. From this perspective, it is desirable to select a network with inductors connected in series with the output in order to block

the harmonics. Note that, this suppression or attenuation is intrinsically related to the quality factor of the matching network, which is limited by the losses of its reactive components, in particular, the inductor (this topic will be further addressed in chapter 3.3.3). Naturally, a higher quality factor implies better filtering. Secondly, and in line with what was mentioned in the first point, it is imperative that the components' sizes are significantly small, in order to ensure its intrinsic resistive losses are reduced. In this way, the power losses are minimized and more power is transferred to the load. Moreover, smaller components mean less area occupied, which makes the integration of the whole RF PA on chip easier.

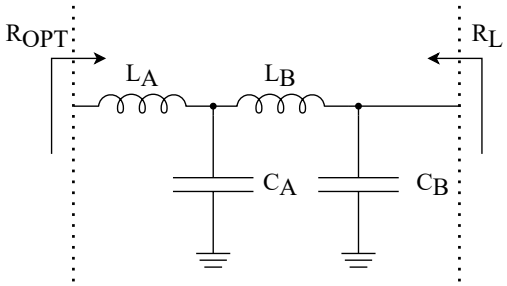


Figure 3.6: Ideal proposed cascaded L-shape matching network.

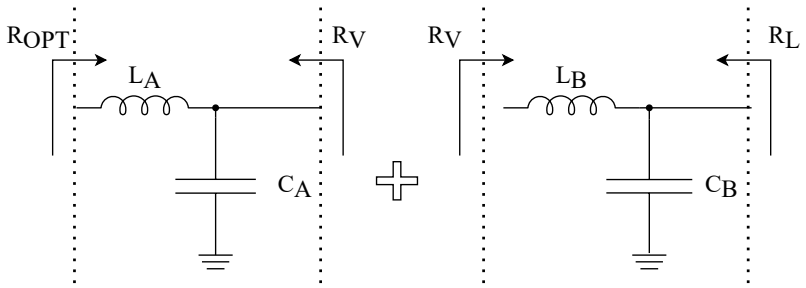


Figure 3.7: Equivalent simplified circuit of the cascaded L-shape matching network.

As shown in figure 3.6, the adopted matching network consists in a cascaded L-shape matching network, that can be seen as the sum of two L-shape matching networks connected through a virtual resistance, as depicted in figure 3.7. In this configuration, the circuit can be analysed as two simple L-shape matching networks, where $R_{OPT} < R_V$ and $R_V < R_L$ (R_V being the virtual resistance, R_{OPT} being the optimal resistance and R_L being the load resistance). Let us now understand how to dimension the reactive components of the network.

Looking into the figure 3.7, it is possible to see that the quality factor of each simple L-shape matching network can be estimated as a function of the input and the output impedance, as expressed by equations 3.8 and 3.9:

$$Q_A = \sqrt{\frac{R_V}{R_{OPT}} - 1}, \tag{3.8}$$

$$Q_B = \sqrt{\frac{R_L}{R_V} - 1}. \tag{3.9}$$

By assuming that both halves have the same quality factor ($\mid Q_A \mid = \mid Q_B \mid$), one can easily obtain the virtual resistance expression $R_V = \sqrt{R_{OPT} R_L}$. Considering the previous results, it is now possible to obtain the expressions that allow estimating the components' values of the matching network in 3.10:

$$L_A = \frac{R_{OPT} Q_A}{w_0}, \tag{3.10a}$$

$$C_A = \frac{Q_A}{w_0 R_V}, \tag{3.10b}$$

$$L_B = \frac{R_V Q_B}{w_0}, \tag{3.10c}$$

$$C_B = \frac{Q_B}{w_0 R_L}, \tag{3.10d}$$

where w_0 is the angular frequency (rad/s).

3.3.3 Impact of Non Idealities

All the analyses done so far assumed that the reactive components are ideal. However, in a real implementation, the non-idealities of these components strongly impact the impedance matching, especially when considering that the circuit will operate at a considerably high frequency. Figure 3.8 shows a generic model that represents the circuit of a real inductor and a real capacitor. Note that, these generic models do not represent the components usually integrated on-chip. However, they describe a similar behaviour.

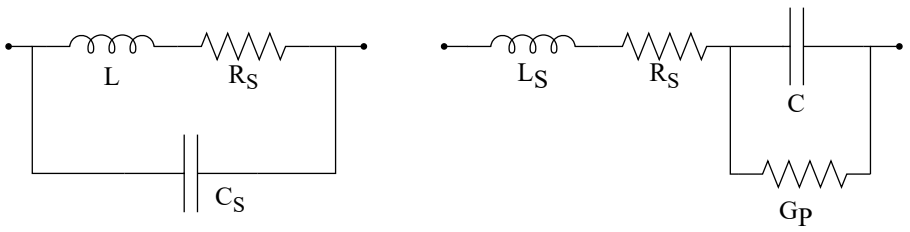


Figure 3.8: Model of a real inductor (left) and a real capacitor (right). Adapted from [31].

As shown in figure 3.8, a real inductor is essentially the parallel of a parasitic capacitance and the series of the inductor and its intrinsic parasitic resistance. On the other hand, a real capacitor can be seen as the series of a parasitic inductor, a parasitic resistance and the parallel of the capacitor and the parasitic conductance.

The impedance curves of a real inductor and a real capacitor as a function of the frequency can be found in figure 3.9. In an ideal case, the inductor impedance varies linearly with the frequency. However, when considering the real model, the impedance value

starts increasing exponentially until it reaches its maximum value at the self-resonant frequency. Starting from this point, the inductor starts behaving as a capacitor and its value continually decreases in a exponential manner. In the same way, in an ideal case, the capacitor impedance is inversely proportional to the frequency. Nonetheless, when considering losses, its impedance value starts decreasing almost exponentially until it reaches its minimum value at the self-resonant frequency, the value at which it starts behaving as an inductor.

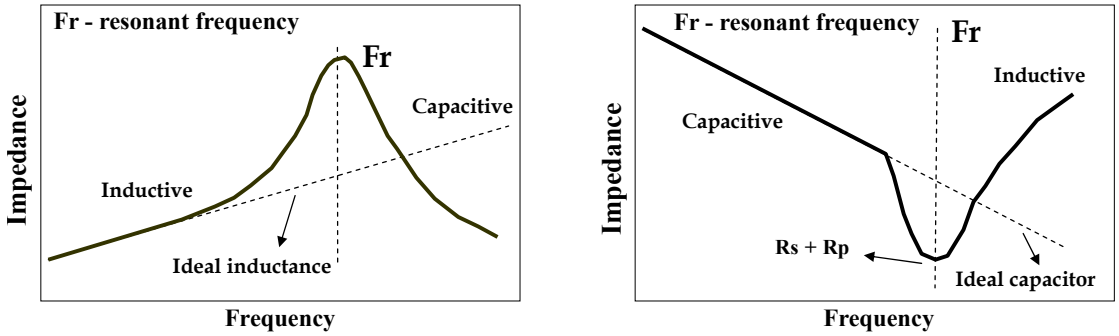


Figure 3.9: Impedance as a function of the frequency for a real inductor (left) and for a real capacitor (right). Adapted from [31].

In brief, the self-resonant frequency is the frequency at which the ideal component resonates with its intrinsic parasitics. At this frequency, the impedance value of the reactive component is strongly affected, resulting in an extremely high impedance in the inductor and in an extremely low impedance in the capacitor. So as to avoid the self-resonant of the component, the operating frequency should be considerably lower when comparing to the self-resonant frequency of the reactive elements. Note that it can be very difficult to control the self-resonant frequency of the reactive components as it is technology dependent. All things considered, it is possible to perceive that this effect strongly impacts the impedance matching, where a slight change in the reactive components affects the quality factor of the impedance matching.

For simplicity, in the sizing of the matching network the real inductor is modelled as an ideal inductor in series with a parasitic resistance and the real capacitance is simply modelled as an ideal capacitance. The parasitic resistance of the inductor can be obtained using equation 3.11:

$$R_s = \frac{w_0 L}{Q_L}, \tag{3.11}$$

where w_0 is the angular frequency of the circuit (rad/s), L is the inductor value and Q_L is the quality factor of the inductor.

3.3.4 Modelling

The first step towards the implementation of the SCPA is the design of the matching network discussed in section 3.3.2. Considering the impact of its non-idealities in the impedance matching, it is worth analysing a more realistic model of the SCPA and its matching network, as depicted in figure 3.10. The model includes the losses of the inductors previously studied, as well as the non-idealities of the switches and it follows a similar approach to the one referenced in [34].

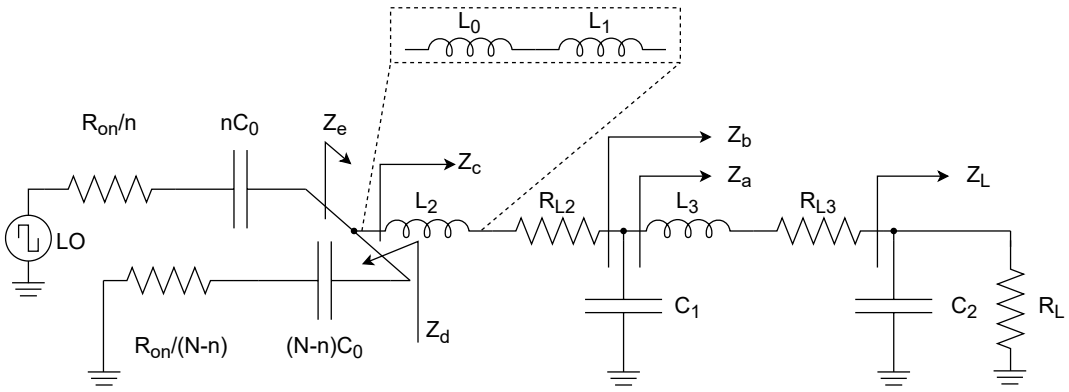


Figure 3.10: Simplified equivalent model of an SCPA with an L shaping matching network with resistive losses.

In a real implementation, the switches of the array of capacitors are made with cascoded class-D PAs. These PAs are implemented using MOSFETs, that have a linear conduction resistance and parasitic capacitances, that lead to power losses. Minimizing the conduction resistance implies having bigger transistors, while minimizing the parasitic capacitances implies having smaller transistors. Considering this trade-off, the transistors are typically optimized to have an optimum resistance that allows to achieve the desired output power, while assuring the losses of the parasitic capacitances are minimized. This topic will be addressed in more detail in chapter 3.4. For now, let us simply introduce in the model the impact of the on-resistance R_{on} of the non-ideal switches. In this way, the n on switched capacitors can be grouped into one branch, with the total R_{on}/n on-resistance in series with the total nC_0 capacitance, whereas the $N-n$ off capacitors can be grouped into another branch, with the total $R_{on}/(N-n)$ on-resistance in series with the total $(N-n)C_0$ capacitors tied to ground. These two branches are then connected to the load through the matching network, where each inductor is modelled with its intrinsic parasitic resistance (as discussed in chapter 3.3.3).

Let us now estimate the most important expressions of the model. In order to do this, let us first extract all the impedances of the circuit, starting from the one in the far right (Z_L) all the way to the one in the far left (Z_e), as presented in figure 3.10. The expressions of the impedances Z_L, Z_a, Z_b, Z_c, Z_d and Z_e are described in equations 3.12 to 3.17, respectively:

$$Z_L = \frac{\frac{1}{j\omega C_2} \cdot R_L}{\frac{1}{j\omega C_2} + R_L}, \tag{3.12}$$

$$Z_a = j\omega L_3 + R_{L3} + Z_L, \tag{3.13}$$

$$Z_b = \frac{\frac{1}{j\omega C_1} \cdot Z_a}{\frac{1}{j\omega C_1} + Z_a}, \tag{3.14}$$

$$Z_c = j\omega L_2 + R_{L2} + Z_b, \tag{3.15}$$

$$Z_d = \frac{R_{on}}{N-n} + \frac{1}{j\omega(N-n)C_0}, \tag{3.16}$$

$$Z_e = \frac{Z_c \cdot Z_d}{Z_c + Z_d}. \tag{3.17}$$

Bearing in mind the impedance expressions previously obtained, one can easily estimate the voltage across the intrinsic parasitics of the inductors L_2 (which is given by the losses of inductor L_0 used in the filter and the losses of the inductor L_1 used in the matching network) and L_3. The voltages are obtained by applying a voltage divider in R_2 and R_3, as presented in 3.18 and 3.19:

$$V_{R2} = \frac{2V_{dd}}{\pi} \cdot \frac{Z_e}{Z_e + \frac{R_{on}}{n} + \frac{1}{j\omega n C_0}} \cdot \frac{R_2}{Z_c}, \tag{3.18}$$

$$V_{R3} = \frac{2V_{dd}}{\pi} \cdot \frac{Z_e}{Z_e + \frac{R_{on}}{n} + \frac{1}{j\omega n C_0}} \cdot \frac{Z_b}{Z_c} \cdot \frac{R_3}{Z_a}. \tag{3.19}$$

Using the same approach, the voltage across the on-resistance of the on capacitors $V_{Ron,1}$ and the off capacitors $V_{Ron,2}$ can be estimated in 3.20 and 3.21.

$$V_{Ron,1} = \frac{2V_{dd}}{\pi} \cdot \frac{\frac{R_{on}}{n}}{\frac{R_{on}}{n} + \frac{1}{j\omega n C_0} + Z_e} \tag{3.20}$$

$$V_{Ron,2} = \frac{2V_{dd}}{\pi} \cdot \frac{Z_e}{Z_e + \frac{R_{on}}{n} + \frac{1}{j\omega n C_0}} \cdot \frac{\frac{R_{on}}{N-n}}{\frac{R_{on}}{N-n} + \frac{1}{j\omega(N-n)C_0}} \tag{3.21}$$

Thus, the power losses of the inductors and the on-resistances of the switches can be obtained by using equations 3.22, 3.23 and 3.24. Note that, because the expressions of the voltages are complex, it is necessary to use the module in the calculations.

$$P_{RL2} = \frac{1}{2} \mid V_{RL2} \mid^2 \frac{1}{R_{L2}} \tag{3.22}$$

$$P_{RL3} = \frac{1}{2} \mid V_{RL3} \mid^2 \frac{1}{R_{L3}} \tag{3.23}$$

$$P_{Ron} = \frac{1}{2} \mid V_{Ron,1} \mid^2 \frac{n}{R_{on}} + \frac{1}{2} \mid V_{Ron,2} \mid^2 \frac{(N-n)}{R_{on}} \tag{3.24}$$

In the same way, the output voltage V_o and the output power P_o can also be derived by using expressions 3.25 and 3.26:

$$V_o = \frac{2V_{dd}}{\pi} \cdot \frac{Z_e}{Z_e + \frac{R_{on}}{n} + \frac{1}{j\omega n C_0}} \cdot \frac{Z_b}{Z_c} \cdot \frac{R_L}{Z_a}, \tag{3.25}$$

$$P_o = \frac{1}{2} \mid V_o \mid^2 \frac{1}{R_L}. \tag{3.26}$$

The Power Added Efficiency, in its turn, is now given by the ratio of the output power P_o and the sum of the output power, the power needed to charge and discharge the array of capacitors P_{SC} (already estimated in 3.4) and all the additional power losses of the conduction resistance of the switches P_{Ron} and of the inductors P_{RL2} and P_{RL3}, as shown in 3.27.

$$PAE_{model,1} = \frac{P_o}{P_o + P_{Ron} + P_{RL2} + P_{RL3} + P_{SC}} \tag{3.27}$$

3.3.4.1 Initial Sizing

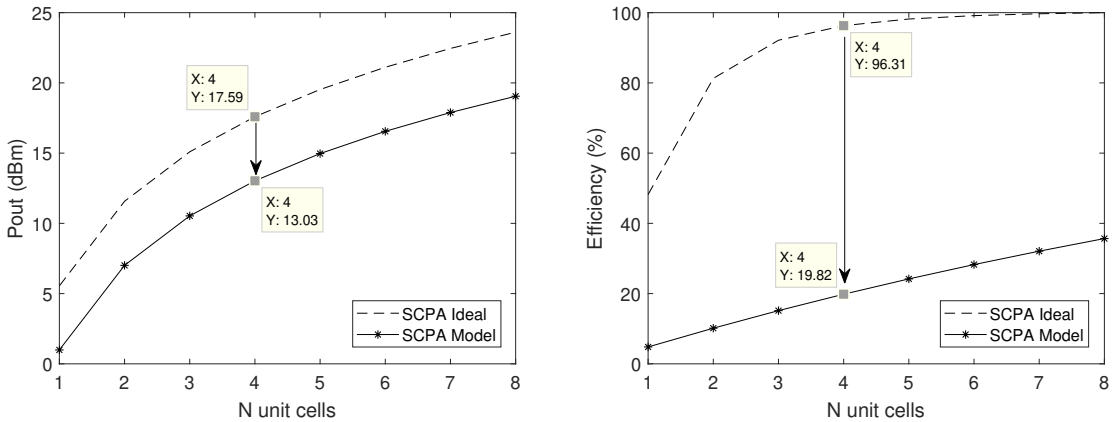
In this project, an optimal resistance of 5Ω was estimated in order to guarantee an ideal output power of 23dBm. Considering that the optimal resistance is given by the losses of both the switches and the inductors that integrate the filter and the matching network, each switch was sized with an on-resistance of 2.2Ω, giving a margin for the losses of the inductors. Since 8 PAs are used, the total on-resistance R_{on} of 17.6Ω is estimated. Besides, in order to dimension the filter, an arbitrary capacitance value of $6.8fF$ was chosen for C_0. Consequently, L_0 was sized with an inductance value of 0.58nH. Based on the losses of the on-resistance and the inductor that integrate the filter, the matching network was determined in order to transform the 50Ω into 2.5Ω. The dimensions of the above-mentioned components are presented in table 3.1.

Bearing in mind the components' dimensions, one can now obtain both the ideal and the modelled curves of the SCPA, for both the output power and the efficiency as a function of the number of cells, as depicted in Fig.3.11. Looking into the curves and considering the situation where 4 unit PAs are used, it is possible to see that output power

Table 3.1: Components' dimensions of the simplified equivalent model of an SCPA.

Inductors	Size [nH]
L0	0.58
L1	0.83
L2	1.40
L3	3.69
Capacitors	**Size [pF]**
C0	6.80
C1	29.30
C2	6.57
Resistors	**Size [Ω]**
R_{on}	17.6
R_{Load}	50.0
R_{L0}	0.33
R_{L1}	0.47
R_{L2}	2.09

drops almost 4.6dBm and the efficiency is around 76% lower, when compared to the ideal case. Besides this, for a maximum number of 8 unit PA cells, the maximum output power is around 19dBm for a maximum PAE of 36.43%. Although these results may still look promising, it is important to remember that some non-idealities, not covered in this model, will lead to inevitable degradation in the results.



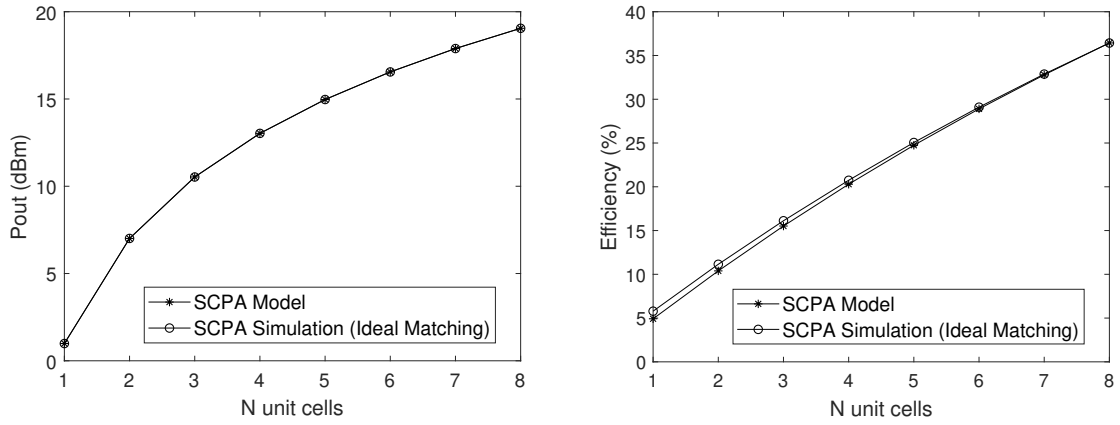
(a) Output Power

(b) Power Added Efficiency

Figure 3.11: Modelled and ideal values of the output power and efficiency per unit cell, considering the resistive losses of the components.

3.3.5 Simulation Results

The section herein presents the results of the real SCPA model illustrated in figure 3.10. The main goal is to present the study conducted to optimize the sizing of the matching network when using real components. As an initial step, all the components used are ideal with the resistive losses modelled. The values used are presented in the table 3.1.



(a) Output Power

(b) Power Added Efficiency

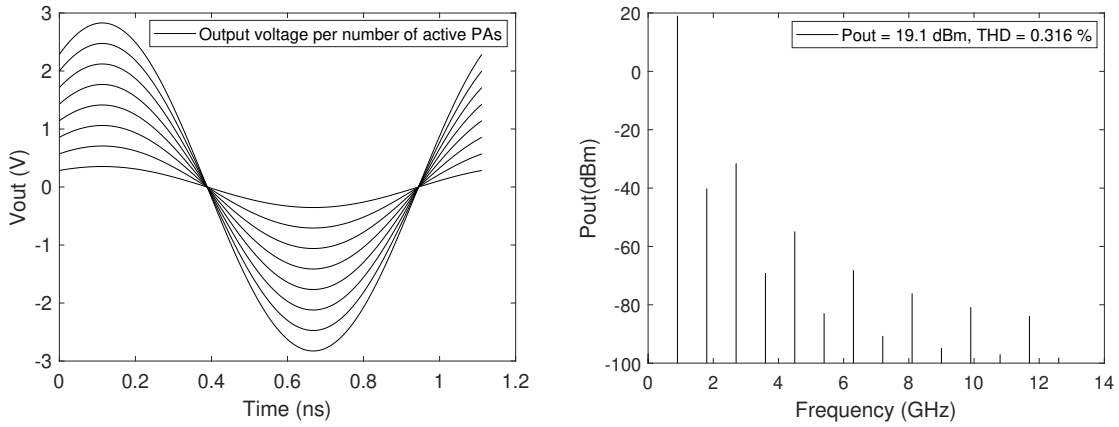
Figure 3.12: modelled and simulated values of the output power and efficiency per unit cell when using an ideal matching network with modelled losses.

The output power and the power added efficiency per unit PA cell simulated and modelled are presented in figure 3.12. By inspection, it is possible to observe that both graphics present approximately the same behaviour. This is an important result, as it validates the model.

Figure 3.13 illustrates the simulated results of both the output voltage per unit cell as a function of time (the higher the number of active PAs, the higher the amplitude of the output voltage) and the spectrum of the output power. As shown in figure 3.13(a), the output voltage is perfectly filtered, regardless of the number of cells used, without significant distortion. The spectrum of the output power, shown in figure 3.13(b) shows that the harmonics close to fundamental frequency are well attenuated, as reflected in the value of the THD.

All the results presented at this point were considering ideal inductors and capacitors with modelled losses. However, as discussed in chapter 3.3, real inductors and capacitors will be used in the final architecture. Bearing this in mind, a second analysis was performed to evaluate the impact on both efficiency and output power when real components are used. For this case, the results show a drop in the efficiency of 17.22% (from 36.43% to 19.21%) and in the output power of 4.48 dBm (from 19.05 dBm to 14.57 dBm). As already discussed in section 3.4.2, these results can be easily explained when thinking about the real model of the inductor and the capacitor. Since the matching network was

designed for ideal components, every mismatch in the impedance of the network will lead to inevitable losses.



(a) Output Voltage

(b) Output Power Spectrum

Figure 3.13: Simulated Output Voltage per unit cell and Output Spectrum when using an ideal matching network.

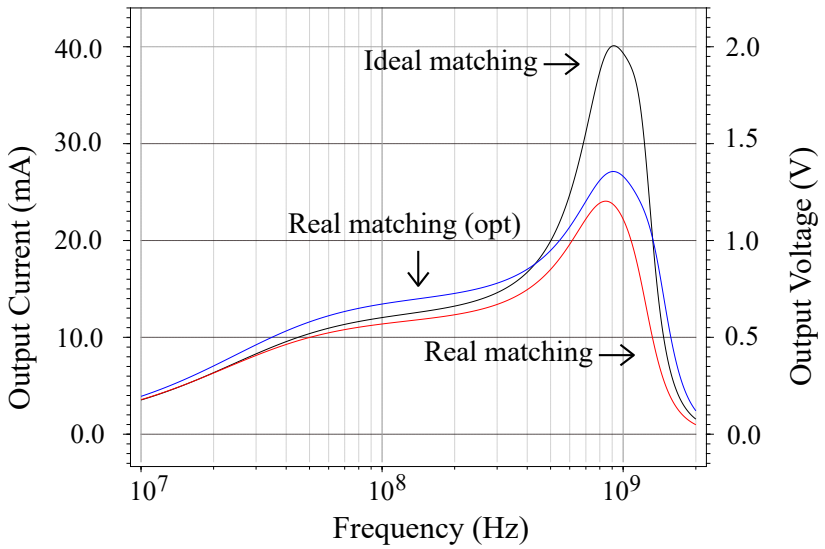
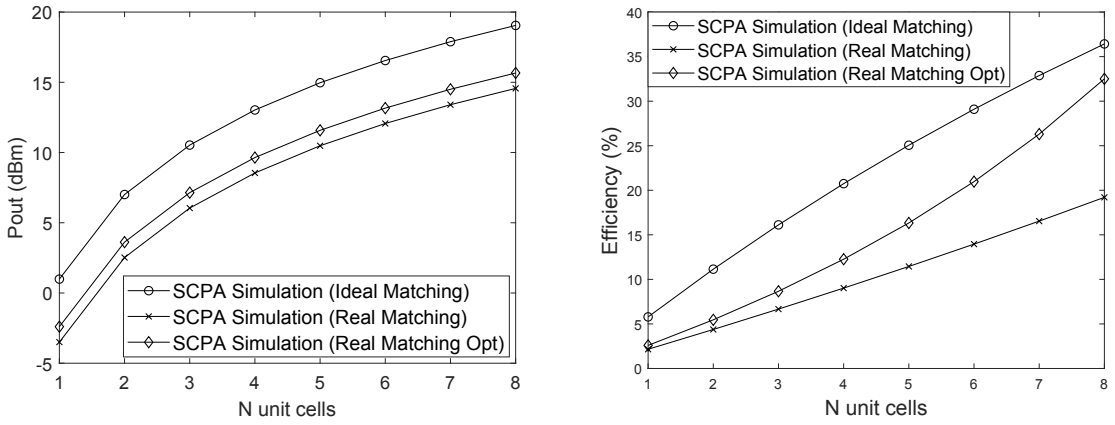


Figure 3.14: Matching network optimization: current and voltage as function of the frequency for the ideal matching network, the real matching network and the real matching network optimized. Note: The current and voltage curves are overlapped.

In order to surpass this issue and boost the efficiency of the SCPA, an optimization of the real matching network is performed. In this analysis, the output voltage and current are measured for a range of frequencies and the components of the matching network are swept. By comparing the maximum output current and voltage of both the ideal and real implementations as a function of the frequency, it is possible to improve the results. Fig.3.14 presents the values of current and voltage for the ideal matching network, the

real matching network and the real matching network optimized. Looking at the figure, it is possible to see that there is an improvement in the values of both current and voltage for fundamental frequency, after the optimization.

The simulated Output Power and Efficiency per number of active PA cells for the 3 above-mentioned cases are presented in Fig.3.15.



(a) Output Power

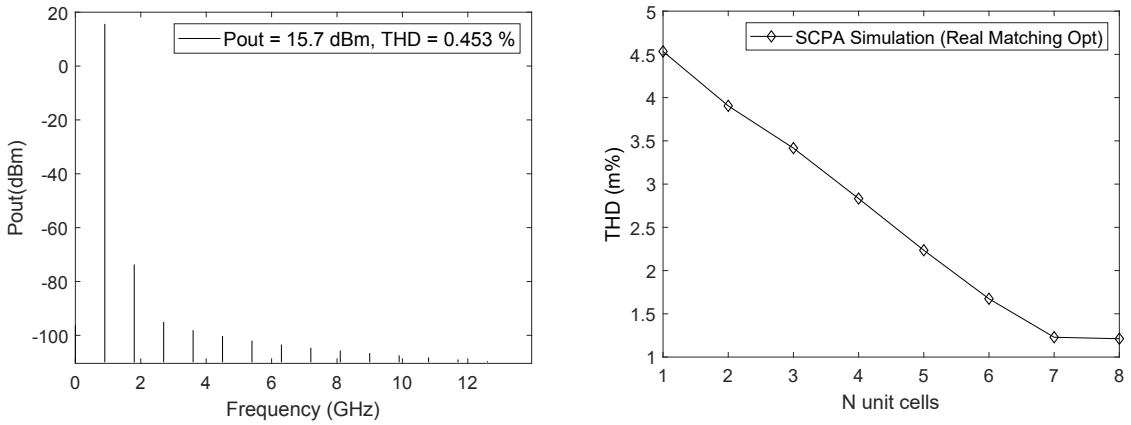
(b) Power Added Efficiency

Figure 3.15: Comparison of the simulated output power and efficiency per unit cell of the SCPA model when using ideal components, real components and real components after optimization.

From inspection, it is possible to see that the curve of the output power obtained when using a real matching network (either with the initial values or with the optimized values) presents a similar behaviour to the one obtained in the ideal case, apart from a gain factor. However, the efficiency curve presents a slightly different behaviour. The non-idealities of the real inductors and real capacitors of the matching network may justify the difference between the curves. Since it is not possible to estimate their equivalent impedance properly, the modelled efficiency cannot represent exactly what happens in the real case. Note that this difference is also visible between the real matching network and the real matching network optimized curves.

Figure 3.16(a) presents the spectrum of the SCPA output power when using the optimized real matching network in full configuration (all eight unit PA cells on). It achieves a maximum output power of 15.7dBm and a THD of 0.45%. Note that for this optimized case, the second harmonic is below -70dBm. Figure 3.16(b) shows the THD of the optimized case per number of active cells. Analysing the figure, one can see that the lower the number of selected PAs, the higher the distortion. This result suggests that this architecture should switch between five to eight PA unit cells in order to guarantee sufficient linearity. Since NB-IoT applications employ a BPSK/QPSK modulation scheme, this topology can simply be used in full configuration or switch between seven and eight PAs.

The simulated output voltage at the load per number of active PA cells is presented in figure 3.17. In this figure, it is possible to see that the signal is perfectly filtered for all the eight cases.



(a) Output Power Spectrum

(b) Total Harmonic Distortion

Figure 3.16: Output power spectrum of the SCPA model when using real optimized components (all PA cells on) and THD per number of PA cells.

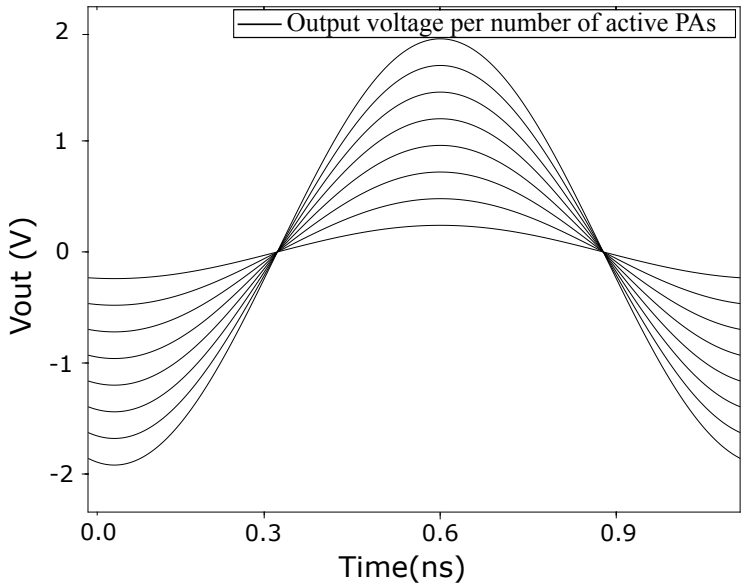


Figure 3.17: Output Voltage per number of active PA cells when using real optimized components.

Table 3.2 presents the results of output power, total harmonic distortion, output voltage, output current and efficiency obtained by the model and obtained in simulation, when all the PA cells are on. The results obtained show that the model provides a good approximation of the behaviour of the circuit when ideal components are used. However, when using real components the non-linear effects impact the estimated output power

and efficiency. The final optimization of the real matching network allows achieving a maximum output power of 15.66dBm with a THD below than 0.45% and maximum efficiency of 32.5%.

Table 3.2: Estimated and simulated results of an ideal SCPA model with modelled losses, a real matching network SCPA model and a real matching network SCPA optimized, when all PA cells are on.

	Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	Efficiency [%]
Estimated	19.05	-	2.84	40.10	36.43
Simulated (ideal)	19.05	0.32	2.01	40.09	36.43
Simulated (real)	14.57	0.39	1.20	23.92	19.21
Simulated (real opt.)	15.66	0.45	1.36	27.14	32.50

The initial and final components dimensions of the matching network are presented in table 3.3. Note that these values will be further adjusted when replacing the ideal components of the circuit with the real ones.

Table 3.3: Components' dimensions of the matching network when using real components after optimization.

Inductors	Initial Size [nH]	Final Size [nH]
L0	0.58	0.58
L1	0.83	1.01
L2	1.41	1.58
L3	3.69	3.69
Capacitors	**Initial Size [pF]**	**Final Size [pF]**
C0	6.80	7.40
C1	29.30	16.35
C2	6.57	5.40

3.4 Cascoded Class-D Power Amplifier

A cascoded Class-D PA consists of a class-D PA with two NMOS and two PMOS transistors connected in a cascoded configuration, as represented in figure 3.18. Such architecture helps to increase the output power by taking leverage of the cascoded structure to achieve a higher supply voltage. This block (that along this project will also be referred to as unit PA cell) is used as a switch in each branch of the SCPA architecture and is designed to have an on-resistance of 17.6Ω.

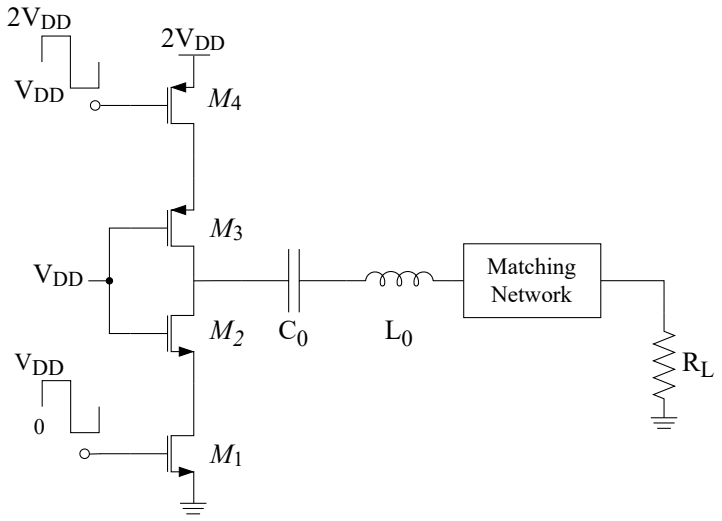


Figure 3.18: Architecture of cascoded Class-D PA connected to a resistive load through a matching network in series with a L_0C_0 filter.

3.4.1 Class-D Ideal Operation

Before proceeding to the analysis and design of the cascoded class-D PA, it is worth first comprehending its ideal behaviour. To simplify the analysis, let us study the case of a simple class-D PA, as depicted in figure 2.12. The input of a class-D stage is a square waveform that switches between the high voltage V_{DD} and ground during its period time, as described in equation 3.28.

$$v(\omega t) = \begin{cases} V_{DD}, & \text{for } 0 \leq \omega t \leq \pi \\ 0, & \text{for } \pi \leq \omega t \leq 2\pi \end{cases} \tag{3.28}$$

Since a square wave is a sum of infinite sinusoidal waves, the previous equation can also be rewritten as the trigonometric fourier series expressed in 3.29.

$$v(\omega t) = V_{DD}\left(\frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin((2n-1)\omega t)}{2n-1}\right) = V_{DD}\left(\frac{1}{2} + \frac{2}{\pi}\sin(\omega t) + \frac{2}{3\pi}\sin(3\omega t) + \cdots\right) \tag{3.29}$$

As already discussed, when the input voltage of the stage is high, the PMOS transistor switches on and the supply voltage is directly connected to the output. For a low input voltage, the NMOS transistor switches on and the output node is tied to ground. In this way, an inverted square waveform is obtained. The L_0C_0 filter will tune the resulting signal at the resonant frequency, transforming the square wave into a sinusoidal wave, as described in equation 3.30,

$$v_o(\omega t) = V\sin(\omega t) = \frac{2}{\pi}V_{DD}\sin(\omega t). \tag{3.30}$$

Note that, this previous result was already used in this work to design the matching network. The output current can now be obtained by simply applying the Ohm law at the load, as expressed in equation 3.31,

$$i_o(\omega t) = I\sin(\omega t) = \frac{2}{\pi}\frac{V_{DD}}{R_L}\sin(\omega t). \tag{3.31}$$

The maximum output power, presented in 3.32, is given by the product of the maximum current to the power of two and the load impedance, divided by 2:

$$P_o = \frac{I^2 R}{2} = \frac{2}{\pi^2}\frac{V_{DD}^2}{R_L}. \tag{3.32}$$

Let us now estimate the ideal power consumption of the topology. By applying an integral during the whole period of the input signal, one can easily obtain the DC current presented in 3.33, which corresponds to the total current that is delivered to the load,

$$I_{dc} = \overline{i_1(\omega t)} = \frac{1}{2\pi}\int_0^{2\pi} i_1(\omega t)d\omega t = \frac{I}{\pi} = \frac{2}{\pi^2}\frac{V_{DD}}{R_L}. \tag{3.33}$$

The DC output power can be estimated as the product of the current previously obtained and the supply voltage, as shown in 3.34,

$$P_{dc} = V_{DD}I_{DC} = \frac{2}{\pi^2}\frac{(V_{DD})^2}{R_L}. \tag{3.34}$$

As a result, an ideal efficiency of 100% is obtained in 3.35,

$$\eta = \frac{P_o}{P_{dc}} = 100\%. \tag{3.35}$$

3.4.2 Impact of Non Idealities

As previously mentioned, the actual MOSFETs used in the cascoded class-D PA have an intrinsic on-resistance and parasitic capacitances. In order to have a better understanding of how these non-idealities affect both the output power and the efficiency of the SCPA, let us first analyse in more detail their nature.

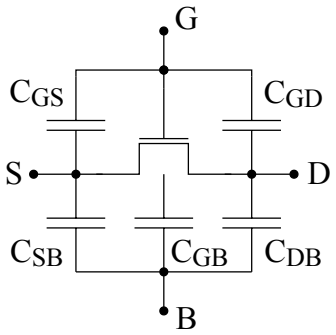


Figure 3.19: Example of the capacitance model of MOSFET.

A generic capacitance model of a MOSFET is illustrated in figure 3.19. Broadly speaking, each MOSFET contains five parasitic capacitances that can be grouped into two different groups: the gate-to-channel capacitances and the junction capacitances. In the first group, composed by the gate-to-source capacitance C_{GCS}, the gate-to-drain capacitance C_{GCD} and the gate-to-bulk capacitance C_{GCB}, the value of each parasitic capacitance varies according to the operation region of the transistor. Typically, these channel capacitances are summed up with a MOS structural capacitance, forming the final capacitances C_{GS}, C_{GD} and C_{GB} presented in the figure. Let us first understand the origin of the former capacitances. Figure 3.20 shows the physical structure of the transistor and the respective gate-to-channel capacitances for three operation regions: cut-off, linear/triode and saturation. By inspection, if the V_{GS} voltage is lower than a certain threshold, the transistor is cut off and no channel is created between the source and the drain terminals. Hence, only the capacitance C_{GCB} exists between the gate and the bulk terminals (as shown in the first case). If the V_{GS} voltage is higher than the threshold voltage and the V_{DS} voltage is such that the transistor is operating between the linear and triode regions, an inversion layer is formed, creating a conductive channel between the source and the drain terminals. As a result, the capacitance C_{GCB} disappears and the capacitances C_{GCS} and C_{GCD} appear (as shown in the second case). If the V_{DS} voltage increases to a point where the channel is pinched-off, the transistor will start operating in saturation mode and only the capacitance C_{GCS} remains (as shown in the third case).

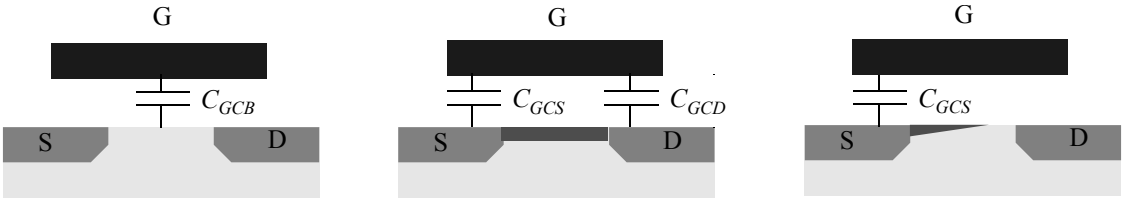


Figure 3.20: Gate-to-channel capacitances depending on the operation mode: 1) transistor in cut-off mode; 2) transistor in linear/triode mode; 3) transistor in saturation mode. Adapted from [35].

Now that the origin of these capacitances was examined, let us understand better how their value changes with the operation mode. Looking into figure 3.21, one can see that the value of these parasitic capacitances changes in a very non-linear way between regions. However, since the transistors operate in a particular region, it is possible to estimate their parasitic capacitance for that region of interest. In this work, because the MOSFETs are used as switches, their operation region will vary between the cut-off (when off) and the linear (when on) regions. In order to evaluate the efficiency of the architecture, only the parasitic capacitances of the linear region (C_{GCD} and C_{GCS}) will be considered. Their values are given by equation 3.36,

$$C_{GCD} = C_{GCS} = \frac{C_{ox}WL}{2} \tag{3.36}$$

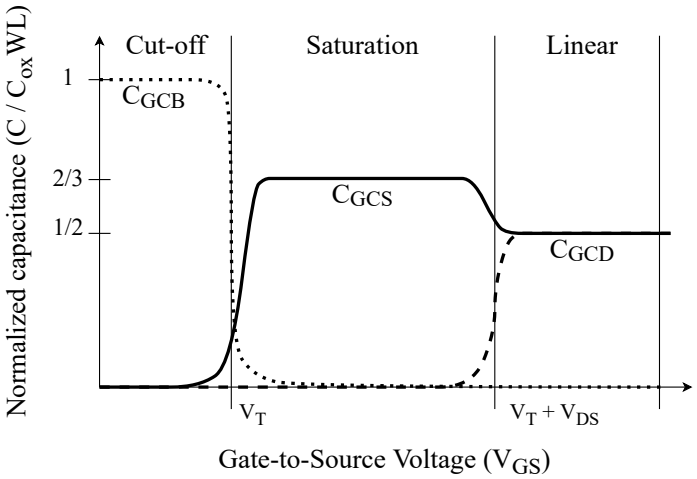


Figure 3.21: Channel-to-gate capacitances C_{GCB}, C_{GCS} and C_{GCD} as function of gate-to-source voltage V_{GS} of a NMOS transistor. The same reasoning can be applied to a PMOS transistor.

where C_{ox} is the the gate oxide capacitance per unit area, W is the width of the transistor and L is the length of the transitor. To properly estimate the gate-to-drain and the gate-to-source capacitances, it is important to add the effect of the structural capacitance previously mentioned. In the physical model of the transistor, both the drain and the source extend their length under the gate oxide for a certain amount x_d, named lateral diffusion. As a result, the effective length of the transistor becomes smaller than the expected length and an overlap capacitance appears. The overlap capacitance can be estimated according to equation 3.37,

$$C_{GDO} = C_{GSO} = C_{ox} x_d W = C_o W. \tag{3.37}$$

where the technology dependent lateral diffusion x_d is usually combined with the oxide capacitance C_{ox} into C_o. In this way, the final gate-to-drain C_{GD} and gate-to-source C_{GS} can then be derived for the linear region in 3.38. From the formula, one can see that these capacitances are proportional to the area of the transistor. Meaning, the bigger the area of the transistor, the bigger C_{GD} and C_{GS} capacitances.

$$C_{GD} = C_{GS} = C_{GCS} + C_{GSO} = \frac{C_{ox} W L}{2} + C_o W. \tag{3.38}$$

Finally, the second group of parasitic capacitances, composed of the junction capacitances C_{SB} and C_{DB}, are a consequence of the reverse-biased source-body and drain-body pn junctions. A detailed explanation of these capacitances can be found in [35]. Due to the complexity in deriving their expressions, the study of these capacitances will not be presented or considered in this work.

Now, let us study the on-resistance of the MOSFET. This resistance represents the connection between the source and the drain and, similarly to what was studied for the

gate-to-channel capacitances, its value varies according to the operation region of the transistor. For the desired linear region, the conduction resistance can be seen as a linear resistance with its approximated value given by equation 3.39.

$$R_{on} = \frac{1}{g_{ds}} = \frac{L}{W\mu C_{ox}(V_{GS} - V_{TH})}. \tag{3.39}$$

where g_{ds} refers to the drain-to-source conductance, μ to the mobility in the channel, V_{GS} the gate-to-source voltage, V_{TH} the threshold voltage and V_{DS} the drain-to-source voltage. Usually, when MOSFETs operate as a switch (or when they are used for digital signals) the length L of the transistor is set to the minimum value allowed by the technology process (typically to guarantee speed, as $v \propto 1/L^2$). This means that, only the width W is tuned. From the formula of the parasitic capacitances, a smaller W results in a smaller parasitic capacitance. At the same time, from the formula of the on-resistance, a smaller W implies a bigger resistance. This suggests that, in order to have a smaller on-resistance the parasitic capacitances will be bigger, leading to power losses and to a lower efficiency.

Bearing this in mind, one can now identify the main parasitics connected at the output of the class-D PA when the transistors work in the linear region. Considering that this topology is essentially an inverter connected to a filter, let us simply analyze the case of an inverter. Figure 3.22 illustrates the main parasitic capacitances of this block when the PMOS (on the left) and the NMOS (on the right) are selected. Let us assume for now that the transistors are never conducting at the same time. As shown in picture, only the gate-to-drain capacitances C_{GD1} and C_{GD2} and the overlap capacitances C_{DB2} and C_{DB1} are connected to the output of the stage. For the reasons previously mentioned, only capacitances C_{GD12} are considered in the model.

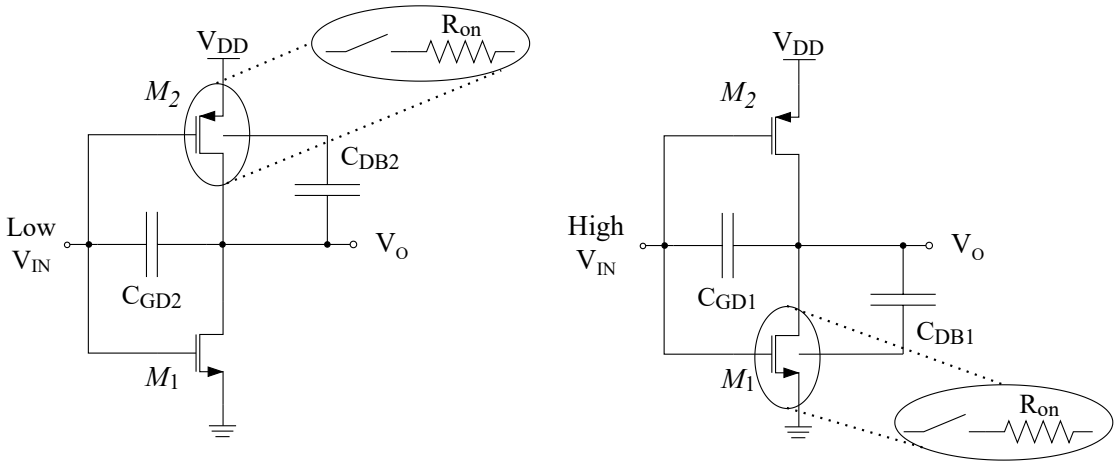


Figure 3.22: Main parasitic capacitances connected to the output node of a simple inverter stage: on the left, when PMOS transistor is conducting; on the right, when NMOS transistor is conducting.

In order to facilitate the analysis, it is helpful to move the parasitic capacitance C_{GD1} and C_{GD2} between the gate and the drain terminals and place it between the output node

and ground. This can be done by applying Miller's effect. Thus, the approximated output capacitance of a simple class-D PA is obtained in 3.40.

$$C_o = 2C_{GD1} = 2C_{GD2} = C_{ox}WL + 2C_oW \propto C_{ox}WL \tag{3.40}$$

3.4.3 Cascoded Class-D Losses

The non-idealities of a cascoded structure can be analysed by looking at its non-ideal model presented in Fig. 3.23.

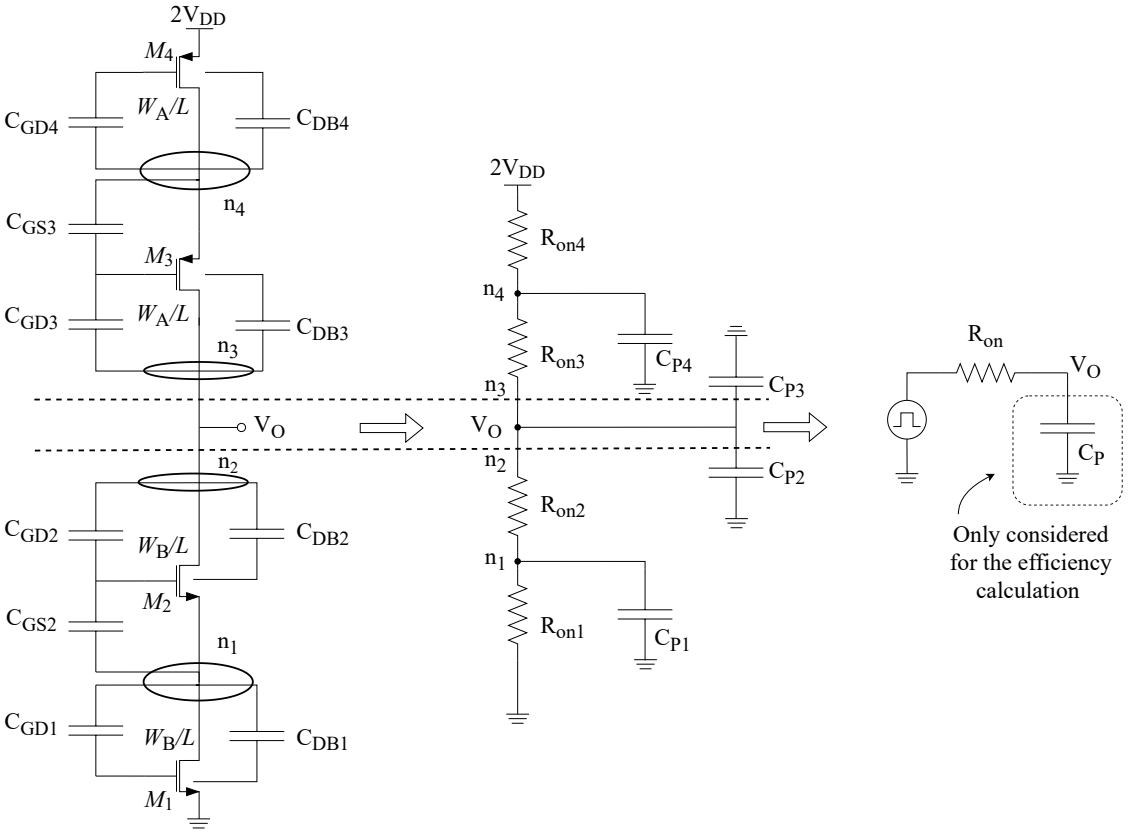


Figure 3.23: Simplification of the cascoded class-D PA non-idealities (the output filter is omitted).

In such model, the parasitic capacitances connected at the internal nodes n4 (C_{GD4}, C_{DB4} and C_{GS3}), n3 (C_{GD3} and C_{DB3}), n2 (C_{GD2} and C_{DB2}) and n1 (C_{GD1}, C_{DB1} and C_{GS2}) are identified. Since either the two PMOS or the two NMOS transistors are conducting at the time, only the parasitic capacitances connected at the nodes n4 and n3 or the nodes n2 and n1 will be charged and discharged every clock cycle, respectively. For this reason, let us simply analyse the case where the two PMOS are conducting. To better estimate the impact of these capacitances in the efficiency, it is desirable to group together all the capacitances connected to the nodes n4 (let us call it C_{P4}) and n3 (let us call it C_{P3}), so their effect can be the sum in a single parasitic capacitance that it is charged (C_P).

To simplify the analysis and similarly to what was done in the previous section, let us ignore the effect of the parasitic capacitances C_{DB4} and C_{DB3}. By doing this, only the capacitances C_{GD4} and C_{GS3} and the capacitance C_{GD3} are connected to the nodes n4 and n3, respectively. To simplify the analysis and similarly to what was done in the previous section, let us ignore the effect of the parasitic capacitances C_{DB4} and C_{DB3}. By doing this, only the capacitances C_{GD4} and C_{GS3} and the capacitance C_{GD3} are connected to the nodes n4 and n3, respectively. In order to have a single capacitance C_{P4} placed between the node n4 and ground and a single capacitance C_{P3} placed between the node n3 and ground, the miller's effect can be applied to these capacitances. Capacitances C_{P2} and C_{P1} can be obtained by applying the same reasoning to the nodes n2 and n1 when the two NMOS are conducting. In this way, the new capacitances C_{P4} and C_{P3} are obtained in 3.41 and 3.42.

$$C_{P4} = C_{P2} = C_{GD4,2} \parallel C_{GS3,1} = 2C_{ox}WL + 4C_oW \propto 2C_{ox}WL \tag{3.41}$$

$$C_{P3} = C_{P1} = C_{GD3,1} = C_{ox}WL + 2C_oW \propto C_{ox}WL \tag{3.42}$$

The total parasitic capacitance can then be obtained, by summing the capacitances that are charged as shown in Eq.3.43. Note that this expression is only valid to calculate the total power losses due to the capacitances.

$$C_P = C_{P4,2} + C_{P3,1} \propto 3C_{ox}WL \tag{3.43}$$

On the other hand, the total on-resistance can now be estimated as the sum of the on-resistance of the two PMOS or the two NMOS, as described in the equation 3.44,

$$R_{on} = R_{on4} + R_{on3} = R_{on2} + R_{on1} = \frac{2L}{W\mu C_{ox}(V_{GS} - V_{TH})}. \tag{3.44}$$

It is important to note that the on-resistance of the two PMOS and the two NMOS transistors must be the same. This means that, the width W_A and the width W_B must be choosen in order to guarantee the same resistive value. In pratical implementations, W_A is usually two to three times bigger than W_B, depending on the technology and node used.

3.4.4 Modelling

Let us now understand better what is the impact of replacing the ideal switch with a cascoded class-D PA in the model of the SCPA studied in section 3.3.4.

As mentioned in the previous section, a cascoded class-D PA can be simply seen as switch in series with an on-resistance and the parallel of the parasitic capacitances and the series of the filter and the load. In the SCPA model, the resistive losses of the cascoded structure were already considered in order to properly estimate the output power and the

efficiency. However, the losses due to the intrinsic parasitic capacitances of the cascoded structure were not taken into account. Note that, these losses do not affect the output power directly. They affect the power consumption only. For this reason, it is important to estimate their impact on the efficiency.

In this way, the power needed to charge and discharge the parasitic capacitances can be estimated by using expression 3.45,

$$P_{CP} = \frac{1}{2} n C_P (V_{DD})^2 f, \tag{3.45}$$

where C_P can be estimated using Eq.3.43. As a result, the efficiency of the SCPA will now be affected by the power needed to charge and discharge the array, as shown in 3.46.

$$PAE_{model,2} = \frac{P_o}{P_o + P_{Ron} + P_{RL2} + P_{RL3} + P_{SC} + P_{CP}}. \tag{3.46}$$

Let us now evaluate the impact of considering these losses in the PAE model. Figure 3.24 presents the curves of the ideal SCPA, the PAE_{model1} and the PAE_{model2}. Looking into the pictures is possible to see that the efficiency dropped considerably when adding the parasitic capacitances of the cascoded class D.

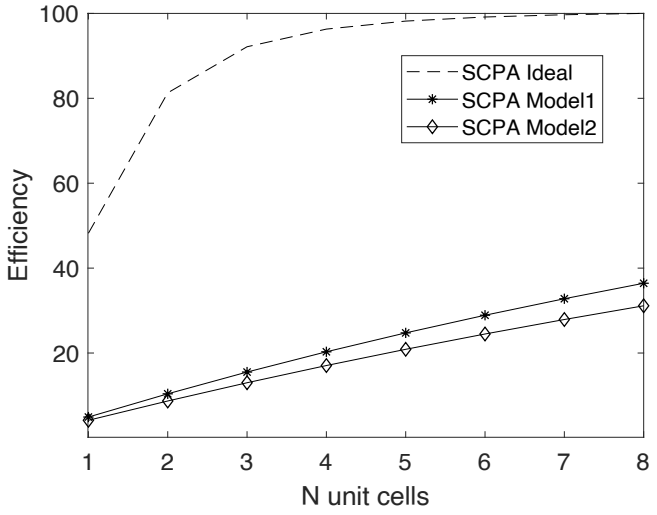


Figure 3.24: Impact of adding the losses of the parasitic capacitances of the cascoded class-D in the PAE.

3.4.5 Simulation Results

Now that the cascoded class-D and its intrinsic parasitic losses were already presented, let us analyse the simulation results of the SCPA architecture when a real switch and a real matching network are implemented. In order to have a better understanding of how the circuit behaves when operating at RF frequencies, two cases are presented: the first,

where the switch uses "Normal" Transistors and, the second, where it uses RF Transistors. Before looking into the simulations, let us understand how to size the block.

As mentioned in section 3.4, each cascoded class-D unit cell is designed for a 17.6Ω on-resistance (corresponding to an 8.8Ω on-resistance per transistor) in order to ensure a total resistance of 2.2Ω in full configuration (all 8 PA cells on). Table 3.4 presents the optimization of each cell for both the RF and normal standard transistors that guarantee the desired resistance. Since the cascoded class-D receives two clock phases (one for the PMOS and the other for the NMOS), the time guard between the clocks is set to 15ps and the rising and falling times are set to 10ps.

Table 3.4: Components' dimensions and respective on-resistance for both normal and RF transistors (P_12_RF and N_BPW_12_RF transistors).

Normal Transistors	Size [μ]	Number of Fingers	Ron [Ω]
$M_{1,2}$	48.7/0.120	10	8.79
$M_{3,4}$	209.2/0.120	10	8.80
RF Transistors	**Size [μ]**	**Number of Fingers**	**Ron [Ω]**
$M_{1,2}$	6.92/0.120	7	8.81
$M_{3,4}$	9.58/0.120	21	8.80

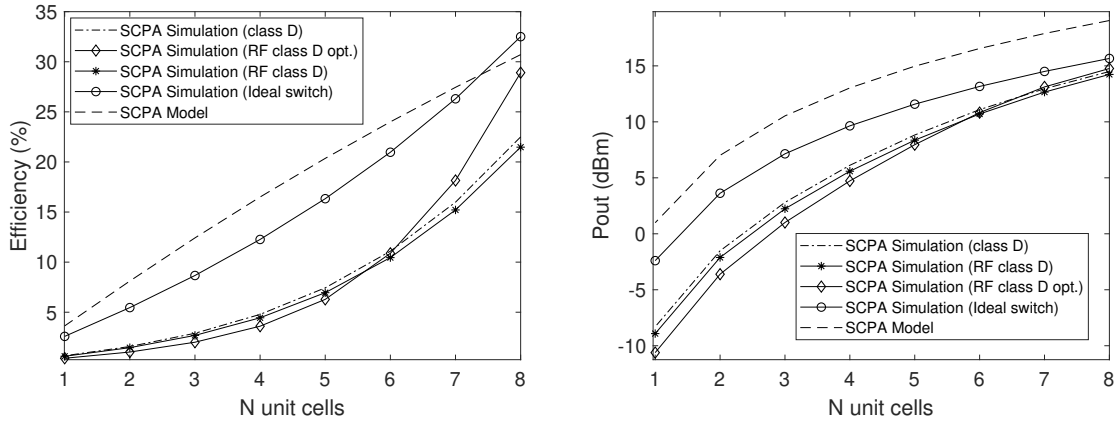
When RF transistors are used, an additional optimization in the matching network is performed in order to improve the results. Considering that the matching network was initially designed to convert the load resistance into the optimal resistance, the intrinsic parasitic capacitances of the cascoded class D lead to inevitable mismatches that impact the performance of the PA. Table 3.5 presents the dimensions of the components for both cases, where the initial sizes refer to the components values of the matching network when the cascoded topology uses normal transistors and RF transistors and the final sizes correspond to the optimized values of the matching network when the cascoded uses RF transistors.

Table 3.5: Components' dimensions of the matching network before and after the optimization.

Inductors	Initial Size [nH]	Final Size [nH]
L0	0.58	0.58
L1	1.0	1.50
L2	1.58	2.08
L3	3.69	3.69
Capacitors	**Initial Size [pF]**	**Final Size [pF]**
C0	7.40	10.0
C1	16.35	10.9
C2	5.40	5.0

Figure 3.25 presents the simulated efficiency and output power of the SCPA, per number of active PAs, when the cascoded class-D PA is added using normal transistors, RF transistors and RF transistors with the optimized matching network. The three curves are compared to the curve of the SCPA using an ideal switch and to the modelled curve.

As expected, both the output power and the efficiency drop when adding the real class D PA. This difference increases when RF transistors are used instead of normal transistors. Since RF transistors are better modelled for RF simulations, these losses are more close to what would happen in layout simulations. After the optimization, it is possible to observe that both output power and efficiency are increased when 6, 7 or 8 unit PA cells are used.



(a) Efficiency

(b) Output Power

Figure 3.25: Efficiency and Output power of the SCPA when: using the model; using an ideal switch; using a cascoded class-D with normal transistors; using a cascoded class-D with RF transistors; using a cascoded class-D with RF transistors and optimized matching network.

Figure 3.26 presents the simulated THD of the SCPA when the cascoded class-D PA is added using normal transistors, RF transistors and RF transistors with the optimized matching network. Similarly to what is observed for the efficiency and output power simulations, the THD increases when using RF transistors. In the optimized case, the SCPA show better results when using 6 up to 8 active PA cells. In general, the values of THD are acceptable for all the configurations of the PA, being the highest value 2.68% when 3 active PA cells are used, in the optimized case.

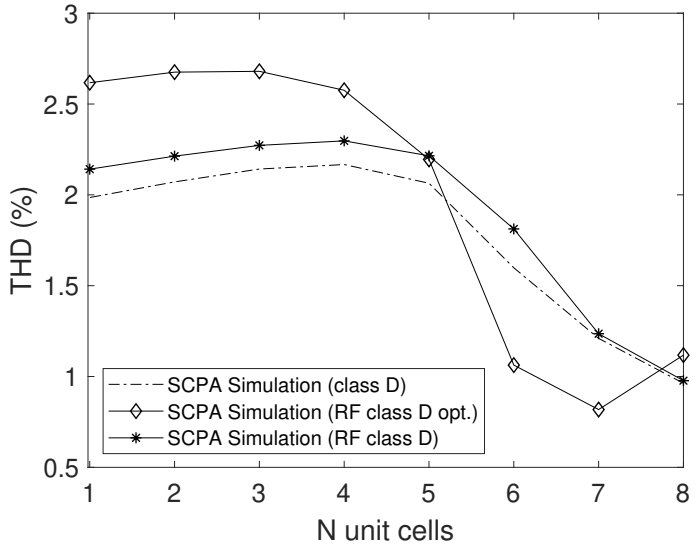
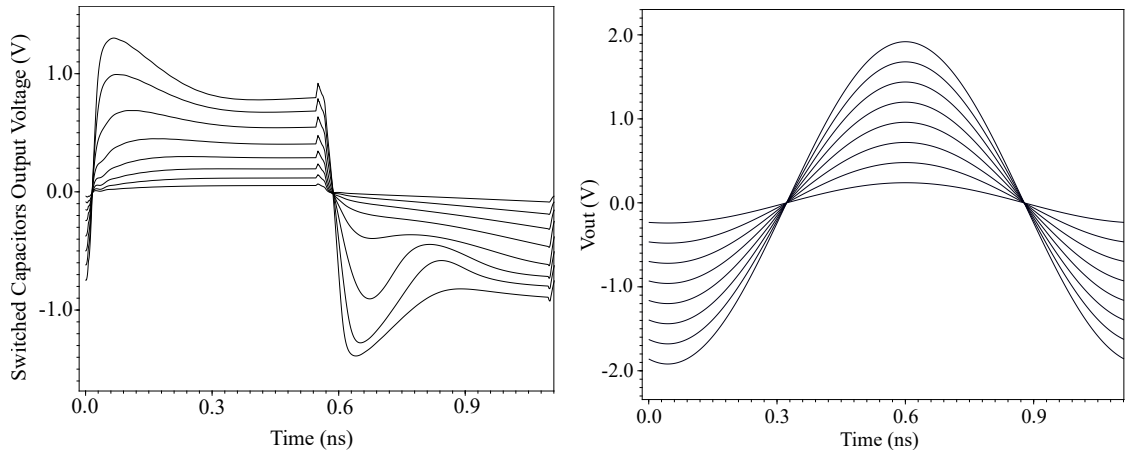


Figure 3.26: Total Harmonic Distortion of the SCPA when adding: a cascoded class-D with normal transistors; a cascoded class-D with RF transistors and a cascoded class-D with RF transistors and optimized matching network.

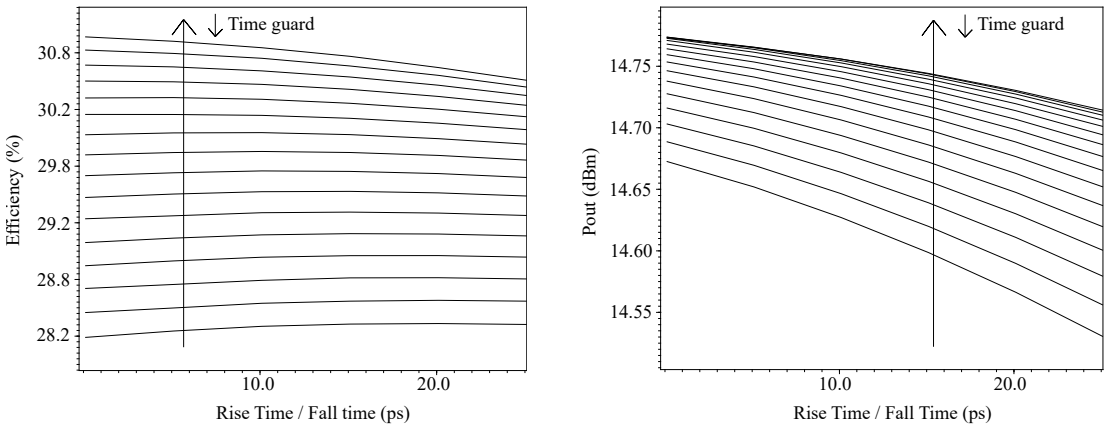
The simulated voltage at the top plate of the capacitors and at the load per number of active PAs are presented in Fig. 3.27.



(a) Voltage at the top plate of the capacitors (b) Output Voltage

Figure 3.27: Voltage at the top plate of the capacitors and Output voltage per number of active cells.

Although the time guard, rising time and falling time parameters of the phase clock were chosen in order to guarantee satisfactory results, it is important to emphasize that a small variation in these parameters will harm PA's performance. Figures 3.28 and 3.29 present the impact of the different clock parameters in the efficiency, output power and total harmonic distortion.



(a) Efficiency (b) Output Power

Figure 3.28: Impact of the time guard and rising time, the falling time between clocks in the efficiency and output power of the SCPA architecture (all PA cells on).

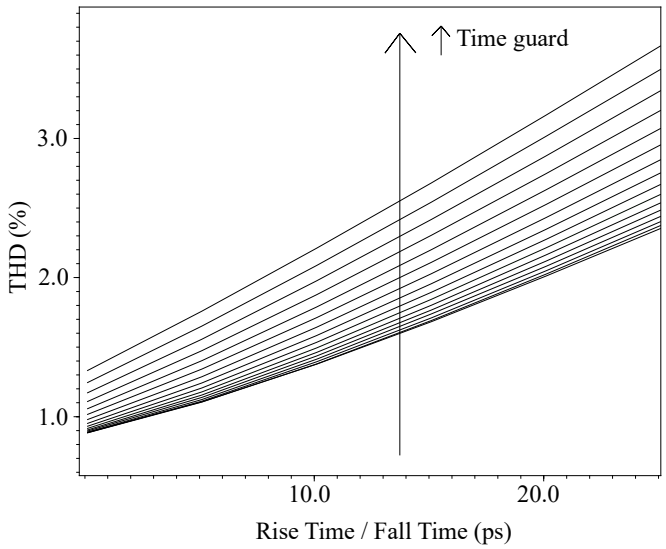


Figure 3.29: Impact of the time guard, rising time and falling time between clocks in the THD (all PA cells on).

Looking into the figures, it is possible to conclude that increasing the time guard leads to a lower efficiency, a lower output power and a higher THD. These results suggest that it is better to have a small interval in time where the two PMOS and two NMOS are conducting at the same time (thus, consuming more), than having two non-overlap clocks. Furthermore, it is also possible to observe that small rising and falling times are crucial to guarantee a lower THD and a higher output power.

The output spectrum of the architecture in full configuration when adding the optimized RF cascoded class-D PA is presented in Figure 3.30. It achieves a maximum amplitude of 14.75dBm at the carrier frequency, an HD2 of -41.58dBc, an HD3 of -42.56dBc and a THD of 1.12%.

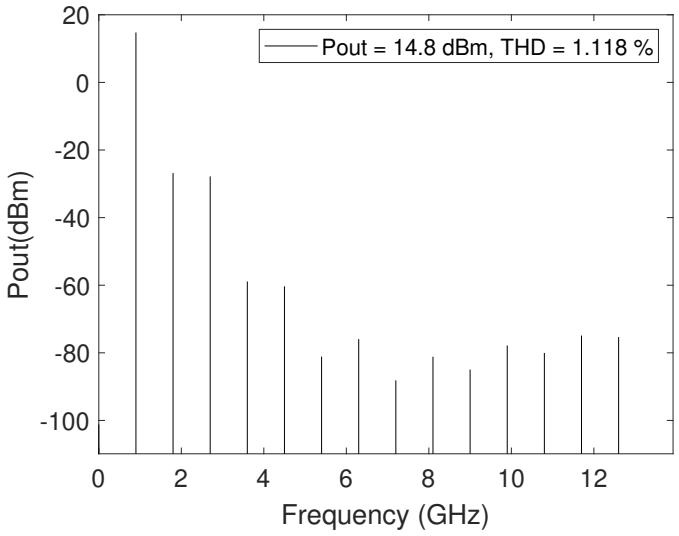


Figure 3.30: Output spectrum of the SCPA architecture when adding a RF cascoded class-D PA with matching network optimization (all unit PA cells on).

Table 3.6 presents the main results of the SCPA for the 3 cases analysed when all PA cells are on. For the architecture, the solution that uses the RF transistors with the optimized matching network is employed.

Table 3.6: Simulated results of the SCPA when using the cascoded class D PA made of: Normal Transistors, RF Transistors, RF Transistors optimized, when all cells are on

	Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	Efficiency [%]
Normal Transistors	14.50	0.96	1.19	23.73	22.50
RF Transistors	14.22	0.98	1.15	22.99	21.48
RF Transistors (opt.)	14.76	1.12	1.22	24.47	28.91

3.5 Switch Driver

A direct consequence of sizing up the transistors to obtain the desired optimum resistance is the increase of the input capacitance of the Cascoded Class-D Power Amplifier. This means that an additional circuit is usually needed at the input of the PA in order to provide enough current to charge its large input load and guarantee a minimum delay response. The drive of the input capacitance is usually done by adding a chain of N inverters, as represented in figure 3.31, where C_L corresponds to the load capacitance seen by the last stage of the inverter chain and C_{g1} corresponds to the input capacitance of the first stage.

The idea behind the chain of inverters is to have an optimum number of stages sized in such a way that every stage is strong enough to drive the previous stage and the last one

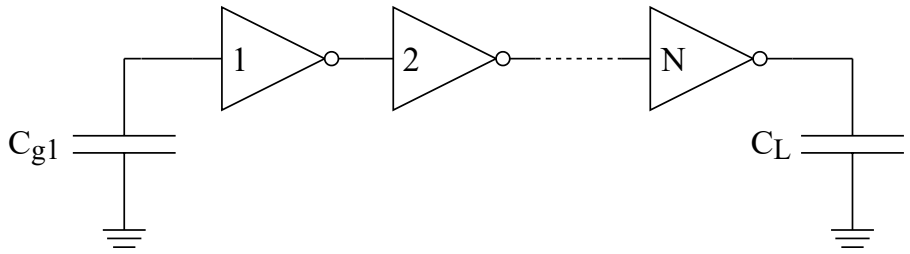


Figure 3.31: Chain of N Inverters.

can drive the main PA, assuring that the delay through the inverter chain is minimized. Let us now analyse in more detail how the optimum number can be obtained and what is the best way of sizing these stages.

Reference [35] shows that each stage has to be sized up by the same factor f with respect to the previous stage, in order to guarantee the same effective fanout f_i [1] and hence the same delay between stages. The size factor is given by the following equation:

$$f = \sqrt[N]{\frac{C_L}{C_{g1}}} = \sqrt[N]{F}, \tag{3.47}$$

where F corresponds to the overall effective fanout of the circuit (given by C_L/C_{g1}) and N corresponds to the number of stages. Considering this, the minimum delay through the chain can be obtained as a function of the size factor, as expressed by equation 3.48:

$$t_p = N t_{p0}\left(1 + \frac{f}{\gamma}\right) = N t_{p0}\left(1 + \frac{\sqrt[N]{F}}{\gamma}\right), \tag{3.48}$$

where t_{p0} corresponds to the intrinsic delay of each inverter of the chain and γ corresponds to the proportionality factor between the input gate capacitance and the intrinsic output capacitance of each stage.

By analysing equation 3.48, one can see that the optimum number of stages that minimize the delay through the chain of inverters can be estimated by differentiating the minimum delay expression by the number of stages and equal the result to zero, as shown in expression 3.49.

$$\gamma + \sqrt[N]{F} + \frac{\sqrt[N]{F} ln F}{N} = 0 \tag{3.49}$$

By solving the above equation, one can conclude that the optimum number of stages is obtained when $N = ln(F)$ and the effective fanout of each stage is given by $f = 2.71829 = e$ [35]. Considering this result, the chosen number of stages for the project was 3, as well as the size factor between each stage, being the last stage the bigger stage of the chain.

[1] The effective fanout corresponds to the ratio between the external load capacitance and the input capacitance of a stage.

Figure 3.32 illustrates the architecture of the driver stage connected to the PMOS transistor (at the top) and the NMOS transistor (at the bottom) of the Cascoded Class-D PA.

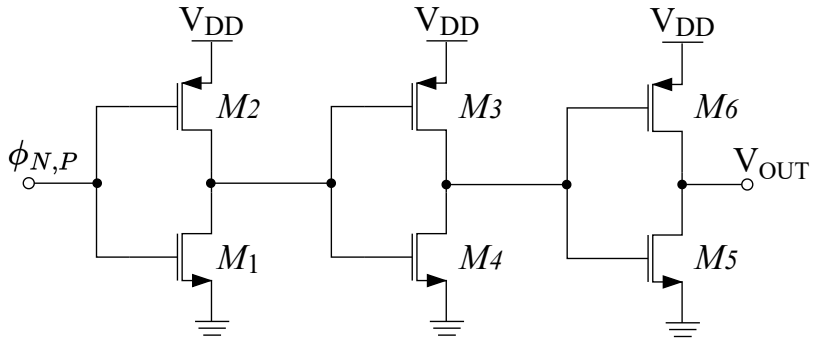


Figure 3.32: Driver stage architecture: Cascaded inverters.

3.5.1 Simulation Results

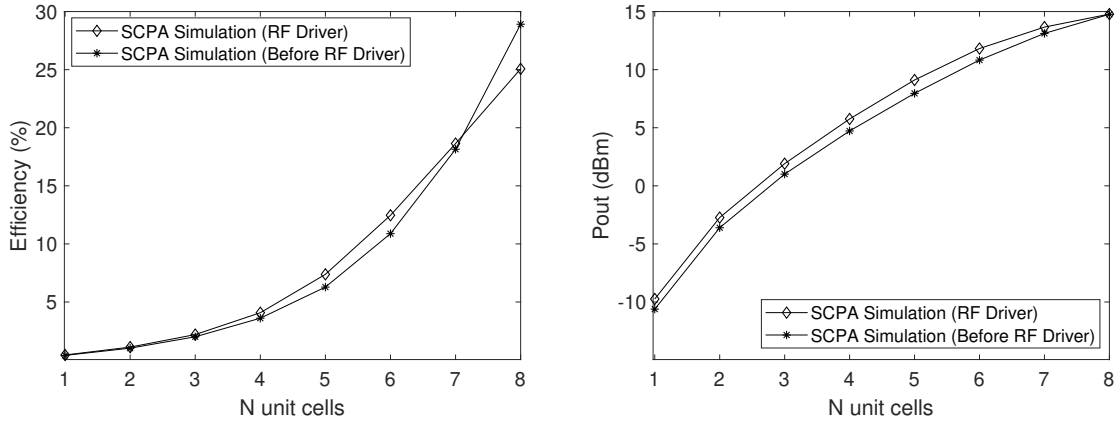
Let us now analyse the results of the SCPA when the driver stage is added to the architecture. Following the study done in the previous section, the 3 driver stages should be sized with a factor of 3 between them. However, due to the sizing limitations imposed by the technology (meaning, the minimum size for the specified number of fingers was reached), this rule cannot be applied. For this reason, a different ratio between inverters is used: the last inverter is 3 times smaller than the cascoded topology, the second inverter is 2 times smaller than the last one, and the first inverter has the same size as the second inverter. Table 3.7 presents the components dimensions of the driver. Note: Although a smaller number of fingers in the PMOS would not allow reaching the ratio of 3 between the inverters, it could potentially lead to better results. This solution will be further analysed in the section 3.8. For now, let us look into the results when using these dimensions.

Table 3.7: Components' dimensions of the driver stages.

Transistor	Size [μm]	Number of Fingers
M_1	1.15/0.12	7
M_2	1.6/0.12	21
M_3	1.15/0.12	7
M_4	1.6/0.12	21
M_5	2.3/0.12	7
M_6	3.2/0.12	21

Figure 3.33 presents the simulated curves of the efficiency and output power per number of active PA cells, before and after adding the RF Driver.

Looking first at the simulated curves of the efficiency, it is possible to see that the efficiency of the SCPA is better when the RF driver is added, in the cases where 7 or less active PAs are selected. On the other hand, in full configuration the efficiency is relatively lower, dropping from 28.91% (when no RF Driver is used) to 25.07% (when the RF Driver is used). Regarding the simulated curves of the output power, one can observe that the output power is slightly higher when the RF driver is introduced. Despite this, in full configuration the produced power is very similar between the two cases, being 14.76dBm (when no RF Driver is used) and 14.78dBm (when the RF Driver is used). The small improvements in the efficiency and output power are expected since the driver is indeed driving the gate capacitances of the cascoded RF class D PA, as analysed in the previous section.



(a) Efficiency (b) Output Power

Figure 3.33: Efficiency and Output power of the SCPA when the RF driver is implemented (per number of active PA cells).

Let us now understand what happens to the THD of the circuit. Figure 3.34 shows the simulated curves of the THD before and after adding the RF Driver in the SCPA architecture. Looking at the figure, it is possible to see that the total harmonic distortion increases significantly when the RF Driver is added to the architecture for almost all the configurations. In fact, only when the maximum number of active PAs is used the THD is smaller, dropping from 1.118% (when no RF Driver is used) and 1.057 % (when the RF Driver is used). In the remaining cases, the THD is higher, being above 3% when less than 7 unit PA cells are selected.

For this reason, unlike the output power and efficiency metrics, the THD exhibits very poor results. However, since other important blocks need to be added to the architecture, it is still early to take final conclusions regarding this block. Although the THD is high, let us keep this block unchanged and understand later the impact of it on the whole architecture.

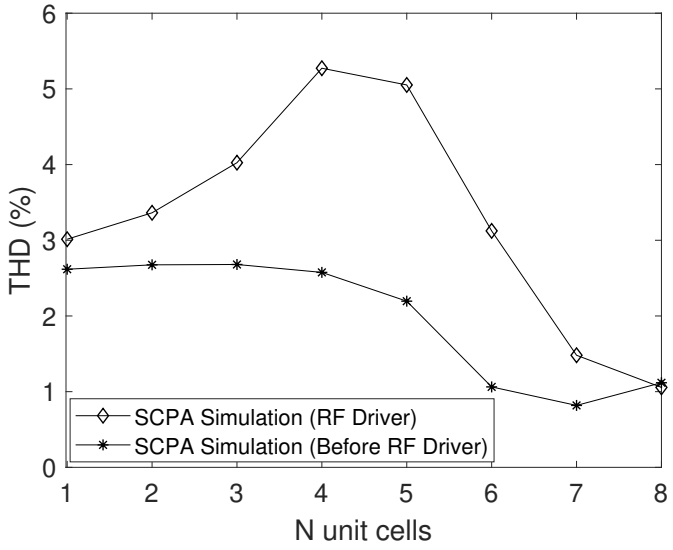


Figure 3.34: Total Harmonic Distortion of the SCPA when the RF driver is implemented (per number of active PA cells).

The main signals of the RF driver stages can be found in Figure 3.35. In the figure, one can identify four main signals: the input and output voltages of the NMOS RF driver (at the bottom of the figure), where the dash signal corresponds to the input signal; and the input and output voltages of the PMOS RF Driver (at the top of the figure), where the dash signal corresponds to the input signal.

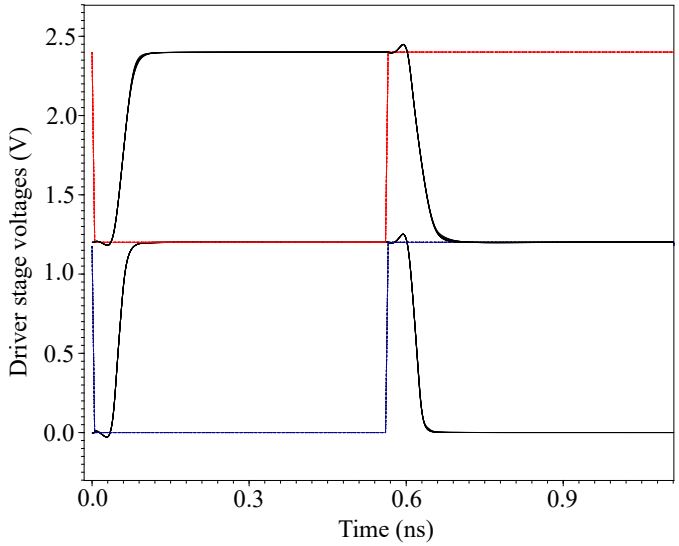


Figure 3.35: Input and output voltages of the driver stages. At the top, the output and input voltages that drive the PMOS of the cascoded topology. The NMOS counterparts are at the bottom.

From the figure, one can confirm the proper behaviour of the driver stages. Since 3

inverters are used in cascade configuration, the output voltage corresponds to the inverse of the input signal. It is also possible to observe that the circuit is slower when driving the PMOS than the NMOS. This is related to the fact the PMOS has bigger parasitic capacitances than the NMOS.

The simulated voltage at the load, for all the SCPA configurations, is presented in Figure 3.36. It is possible to see that the output voltage is perfectly filtered for all the cases. However, unlike the previous cases, it is possible to see some misalignment between the curves when a lower number of unit PA cells is used. In full configuration, the output voltage reaches a maximum of 1.23 Vrms.

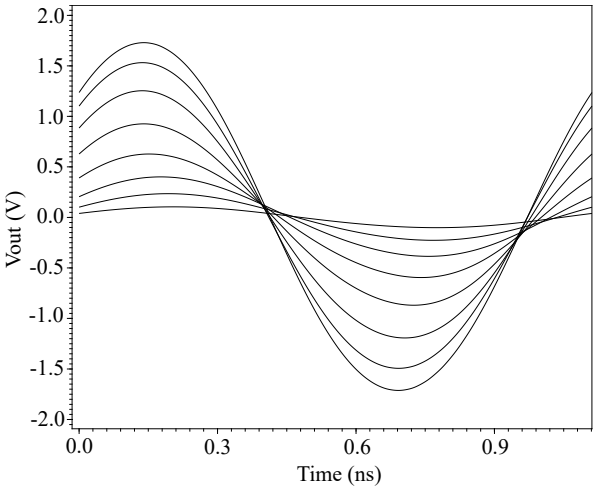


Figure 3.36: SCPA output voltage when adding the RF driver, per number of active PAs.

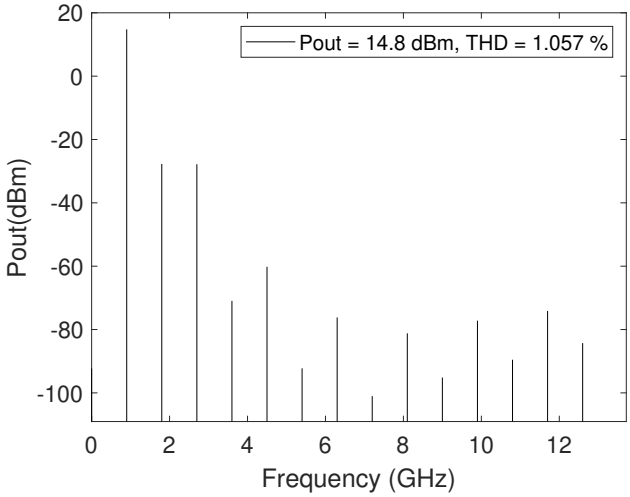
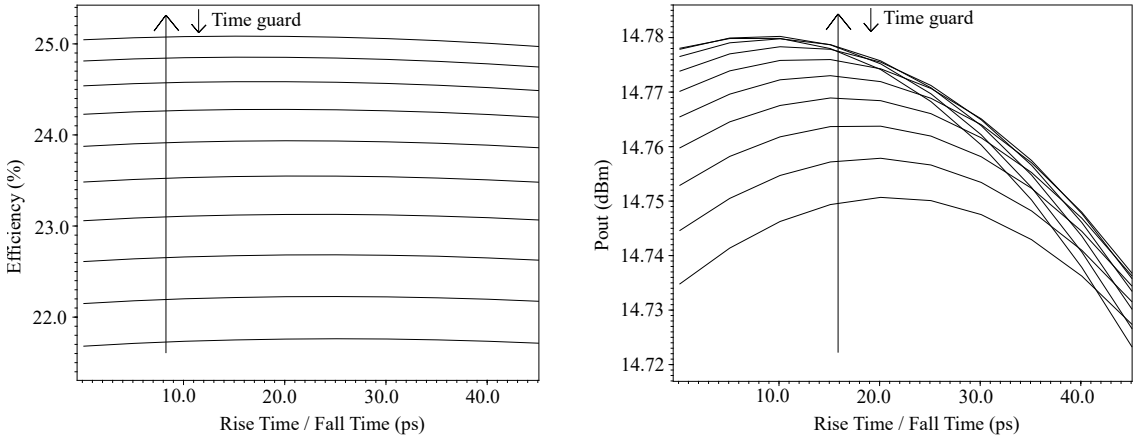


Figure 3.37: Output spectrum of the SCPA architecture when adding a RF driver (all unit PA cells on.)

The output spectrum of the architecture when used in full configuration is depicted in Fig.3.37. As already discussed, the output power achieves the maximum value of 14.78dBm (previously 14.75dBm) at the carrier frequency. The HD2 and HD3 correspond to -42.48dBc (previously -42.58dBc) and -42.58dBc (previously -42.56dBc), respectively, and the THD is given by 1.06% (previously 1.12%).

Similarly to what was done in other blocks within the architecture, let us understand better the impact of the phase clock parameters in the efficiency, output power and THD of the SCPA. Figure 3.38 presents the simulated curves of the efficiency and output power, whereas figure 3.39 presents the simulated curves of the THD. All the abovementioned metrics were analysed for different rising/falling and time guard times.



(a) Efficiency (b) Output Power

Figure 3.38: Impact of the time guard, rising time and falling time between clocks in the efficiency and output power of the SCPA architecture (all PA cells on).

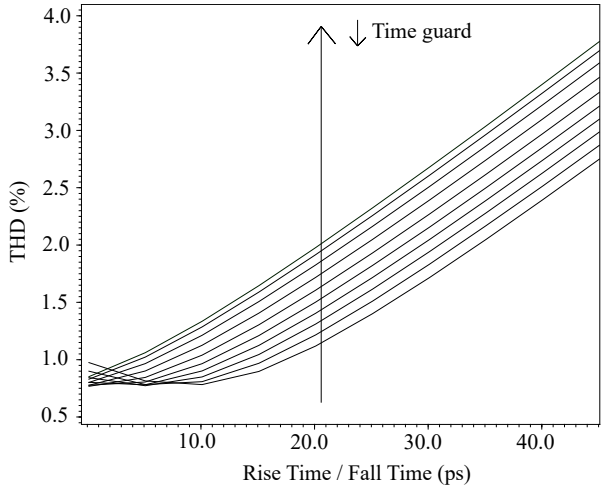


Figure 3.39: Impact of the time guard, rising time and falling time between clocks in the THD of the SCPA architecture (all PA cells on).

Looking into the efficiency curves, it is possible to see that the lower the time guard and the rising/falling times, the better the efficiency. Regarding the output power and the THD distortion, better results are achieved when a lower time guard is applied, along with a rising/falling around 10ps. This means that rising and falling times of 10 ps are preferable to obtain better results. All the remaining combinations will lead to worse results.

Finally, table 3.8 presents the main results of the architecture when the RF Driver is designed in full configuration. In full configuration, the architecture produces an output power of 14.78dBm with an efficiency of 25% and a THD of 1.06%.

Table 3.8: Simulated results of the SCPA when adding the RF driver, when all PA unit cells are on.

	Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	Efficiency [%]
RF Transistors	14.78	1.06	1.23	24.52	25.07

3.6 Level Shifter

The proposed architecture for the level shifter circuit consists of two branches of cascoded transistors that are differentially switched [40]. Such topology is depicted in figure 3.40, where the first branch is composed by transistors M_{1A}, M_{2A}, M_{3A}, M_{4A} and the second branch is composed by transistors M_{1B}, M_{2B}, M_{3B}, M_{4B}. In this scheme, two capacitors C_{UP1} and C_{UP2} are used to couple the gates of the transistors M_{1A} and M_{1B} with the gates of the transistors M_{4A} and M_{4B}, respectively, and an inverter is connected between the gates of the transistors M_{4B} and M_{4A}.

The working principle is the following: when the input voltage is high, the transistor M_{1A} switches on and the source voltage of transistor M_{2A} is pulled down to ground. At the same time, the inverter used at the input of M_{1B} forces the transistor to be in the "off" state and the source voltage of the transistor M_{2B} is charged until V_{DD}. Since the gate voltage of transistors M_{2A} and M_{2B} are connected to V_{DD}, then both transistors are switched to the "on" and to the "off" state, respectively. As a result, the source voltage of transistors M_{3A} and M_{3B} are pulled down to ground and pulled up to a voltage of $2V_{DD}$, accordingly. Note that, both M_{4B} and the M_{4A} drain voltages will never drop below V_{DD} as both gates of transistors M_{3A} and M_{3B} of are set to V_{DD}. Hence, considering that M_{4A} and M_{4B} are cross coupled, the drain voltage at M_{4A} switches on the transistor M_{4B} that will charge the output until $2V_{DD}$ and, ultimately, switch off the transistor M_{4A}.

In this way, when the input voltage is set to high voltage, the output voltage V_{OUT} will be $2V_{DD}$. The same reasoning can be applied when the input voltage is set to the low voltage, in which case the output voltage V_{OUT} will be V_{DD}.

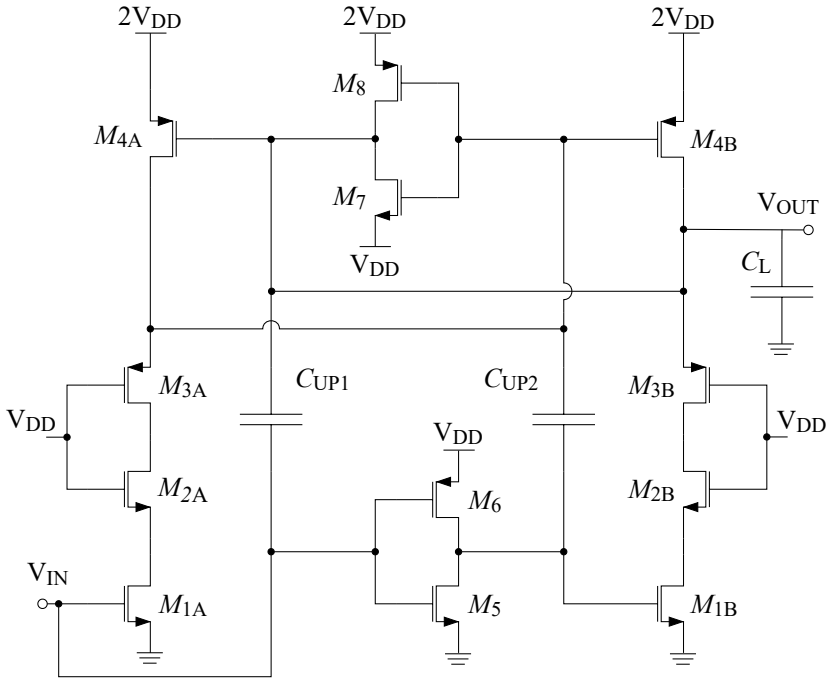
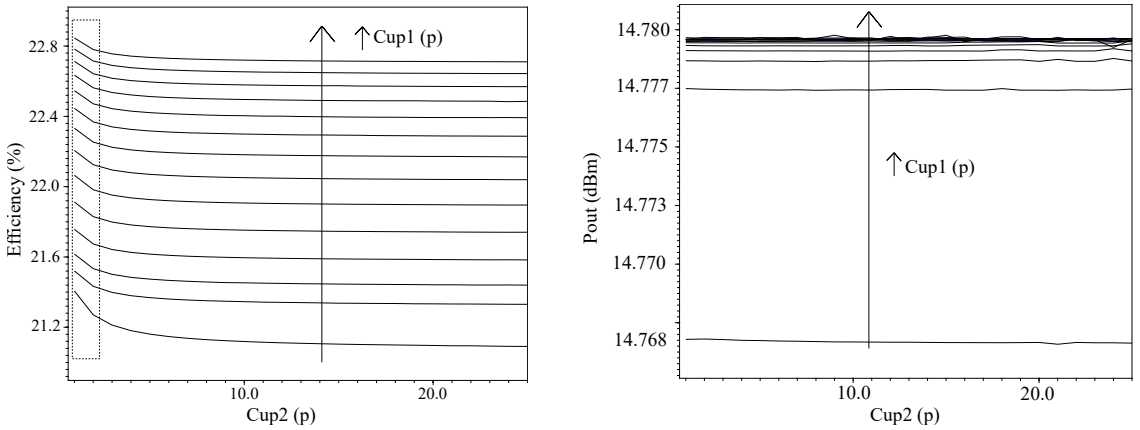


Figure 3.40: Level shifter architecture. Note: The PMOS and NMOS transistors with undefined bulk have their bulks connected to $2V_{DD}$ (M_{4A}, M_{4B}, M_{3A}, M_{3B} and M_8), V_{DD} (M_6 and M_7) and ground (M_{1A}, M_{1B}, M_{2A}, M_{2B} and M_5).

This particular topology has two main advantages. First, it uses coupling capacitances, C_{UP1} and C_{UP2}, that guarantee that the switching transitions in the cross couple pair are faster. Second, it uses an inverter between the gate of the transistors M_{4A} and M_{4B}, that not only helps charge and discharge the output capacitance C_L (that corresponds to the gate capacitance of the driver) but also improves the function of C_{UP1}.

3.6.1 Simulation Results

Now that the level shifter was carefully studied, let us analyse which component values are the best fit for the architecture. In order to guarantee the maximum speed, and similarly to what has been done in other blocks, the length of the transistors is kept with the minimum value allowed by the technology. The width, on the other hand, is sized with the same value as the first inverter of the driver. Besides the transistors, also the capacitance values of C_{up1} and C_{up2} need to be decided. Figures 3.41 and 3.42 present the simulated curves of efficiency, output power and THD when C_{up1} are swept between 1pF and 14pF and C_{up2} are swept between 1pF and 25pF. Note that, this analysis is performed when the SCPA architecture is operating in full configuration, meaning all the 8 unit PA cells are on.



(a) Efficiency (b) Output Power

Figure 3.41: Impact of Cup1 and Cup2 capacitances in the efficiency and output power of the SCPA architecture (all PA cells on).

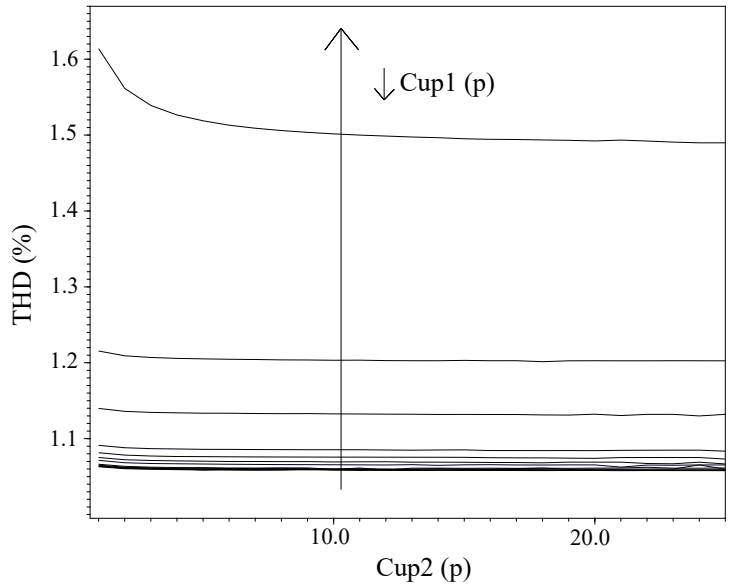


Figure 3.42: Impact of Cup1 and Cup2 capacitances in the total harmonic distortion of the SCPA architecture (all PA cells on).

Looking into figure 3.41, one can see that better results of efficiency and output power are achieved when C_{up1} assumes higher capacitance values and C_{up2} assumes lower capacitance values. The same behaviour can be observed when looking into the THD curves presented in figure 3.42. Since the results do not improve much for capacitance values of C_{up1} higher than 10pF, the adopted value at this stage is 10pF. The C_{up2}, on the other hand, is set to 1pF. It is important to bear in mind that these results may vary when other blocks are added to the architecture. For this reason, these parameters will be reviewed when all the blocks are integrated within the SCPA architecture. The components' dimensions of the level shifter are presented in table 3.9.

Table 3.9: Components' dimensions of the level shifter.

Transistor	Size [μm]	Number of Fingers
$M_{1A,1B}$	1.15/0.12	7
$M_{2A,2B}$	1.15/0.12	7
$M_{3A,3B}$	1.6/0.12	21
$M_{4A,4B}$	1.6/0.12	21
$M_{5,7}$	1.15/0.12	7
$M_{6,8}$	1.6/0.12	21

Capacitor	Size [pF]
C_{UP1}	10
C_{UP2}	1

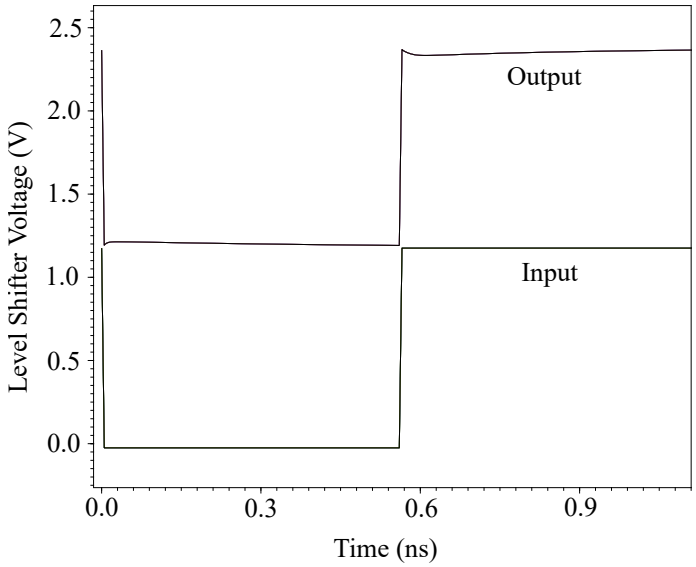
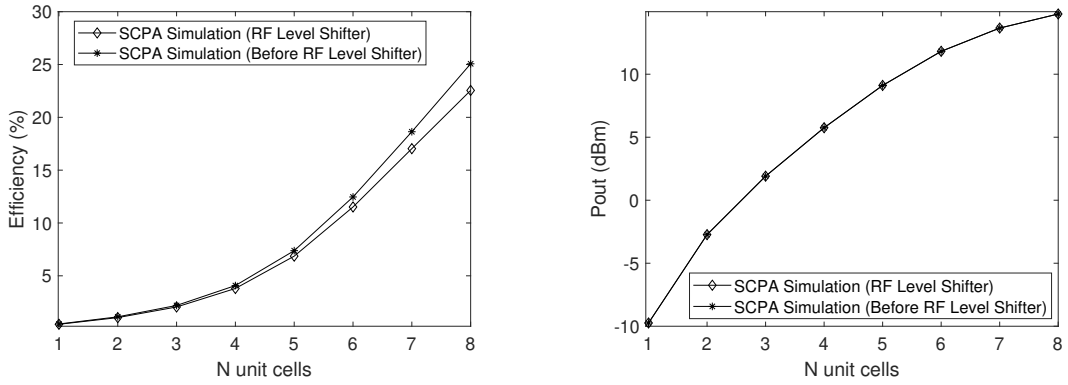


Figure 3.43: Output and input waveforms of the level shifter in a unit PA cell.

Now that the values of capacitances C_{up1} and C_{up2} were decided, let us take a look into the input and output voltages of the level shifter. Figure 3.43 presents the two abovementioned curves of the level shifter for one active PA cell branch. From the figure, one

can clearly see that the level shifter is operating properly: the output voltage corresponds to the input phase clock shifted 1.2V.

Considering the results obtained so far, let us now evaluate the SCPA performance, across all the number of PA cells, when the RF level shifter is implemented. In order to understand better the impact of adding the RF level shifter in the architecture, the results will be compared to the ones obtained before adding the RF level shifter. Figure 3.44 presents the simulated curves of the efficiency and output power per number of active PA cells, before and after adding the RF level shifter.



(a) Efficiency (b) Output Power

Figure 3.44: Efficiency and Output Power of the SCPA when adding the RF Driver, per number of active PA cells.

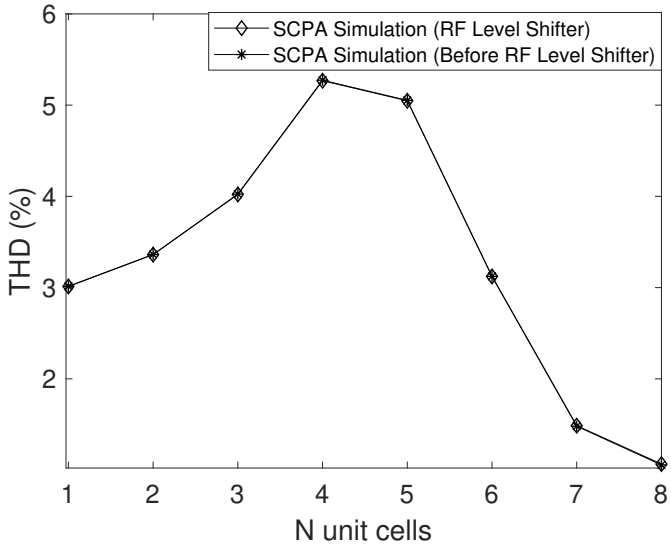


Figure 3.45: THD of the SCPA when adding the RF Driver, per number of active cells.

From the simulated curves, it is possible to conclude that adding the RF level shifter does not impact the value of the output power. In fact, this result is expected since the level shifter it is only used to obtain the correct phase clock amplitudes to be used by the PMOS transistor within the cascoded structure. However, adding this additional block slightly impacts the efficiency of the architecture, as shown in the figure. This result is also expected since adding the level shifter implies a higher number of components that consume power, but the produced output power remains the same. Naturally, a higher consumption for the same output power means lower efficiency. In full configuration, the efficiency drops from 25.07% (before adding the RF level shifter) to 22.55% (when adding the RF level shifter).

Figure 3.45 presents the simulated curves of the THD before and after adding the RF driver, per number of active PA cells. Looking into the figure, one can see that the THD does not change after adding the level shifter block. Although the THD assumes high values when 6 or less active PAs are selected, no improvements will be done at this stage.

The output voltage at the load per number of active cells is presented in figure 3.46. From the figure, one can observe that the output voltage is the same as the one obtained when the RF driver was added in the architecture, where the output voltage is perfectly filtered for all the cases, but some dis-alignment is happening when less number of PAs are selected.

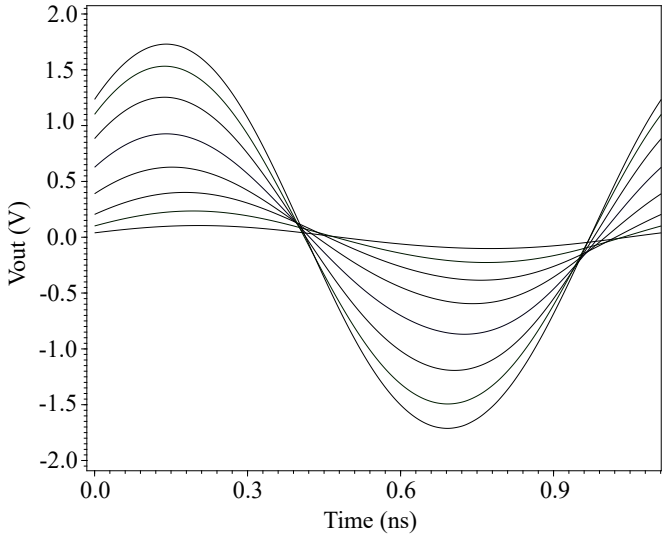


Figure 3.46: Output voltage of the SCPA architecture when adding the level shifter per number of PA cells.

The simulated output spectrum of the architecture when used in full configuration is depicted in figure 3.47. The output power achieves a maximum value of 14.78dBm (same as before). The HD2 and HD3 correspond to -42.32dBc (previously, -42.48dBc) and -42.59dBc (previously, -42.58dBc), respectively, and the THD is around 1.07% (previously, 1.06%).

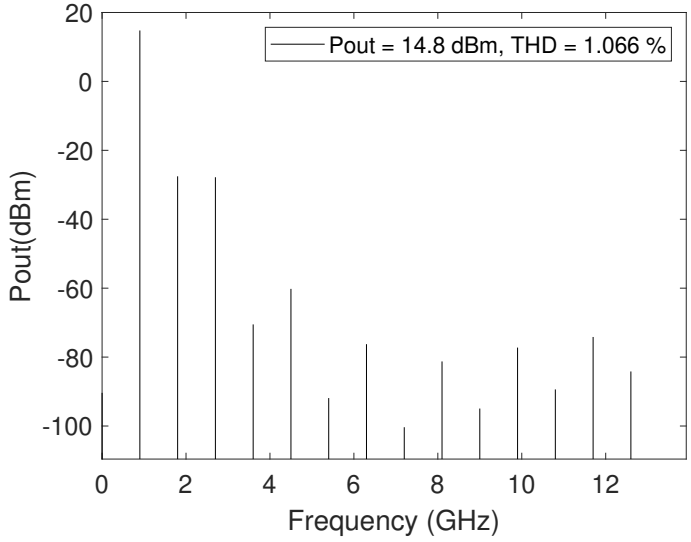


Figure 3.47: Output spectrum of the SCPA architecture when adding an RF level shifter (all unit PA cells on).

Table 3.10 presents the main results of the architecture when the RF shifter is designed. For full configuration, the architecture produces an output power of 14.78dBm with an efficiency of 22.55% and a THD of 1.07%.

Table 3.10: Simulated results of the SCPA when adding the RF level shifter, when all PA unit cells are on.

	Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	Efficiency [%]
RF Transistors	14.78	1.07	1.23	24.52	22.55

3.7 Selection Logic

The number of unit PA cells selected to produce the desired output power is determined by using a selection logic.

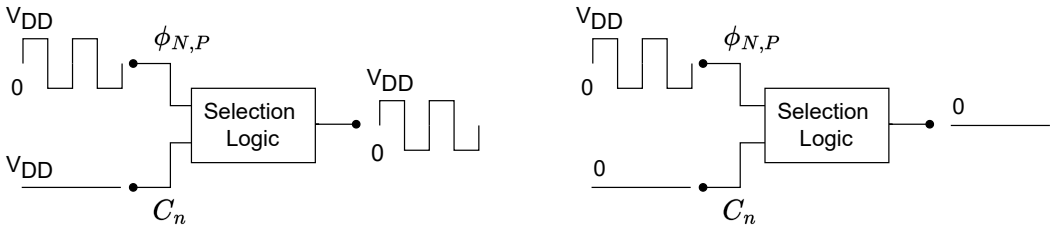


Figure 3.48: Selection logic used to obtain the waveform in each unit PA cell.

Figure 3.48, illustrates the logic used in each unit PA cell branch, where the selection logic block is used to either drive a square waveform or produce a low output voltage, depending on the applied input code. The result output waveform will then be used to drive the unit PA cell branch ("on" cell) or connected to ground ("off" cell). In this work, two implementations of this selection logic are studied: the AND gate and the transmission gate.

Figure 3.49 shows the implemented architecture of the AND gate, which can be seen as a NAND gate stage followed by an inverter stage. Figure 3.50 shows the implemented architecture of the transmission gate.

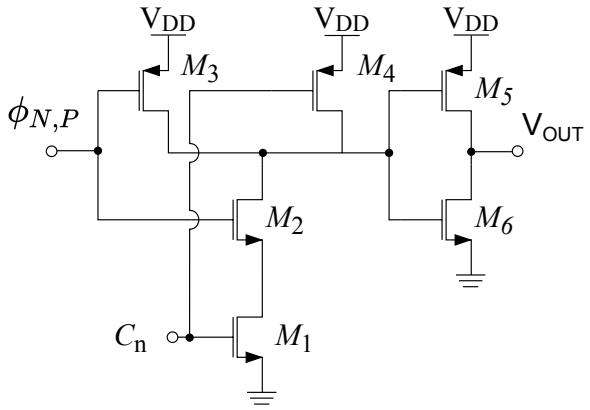


Figure 3.49: AND gate architecture. Note: The PMOS and NMOS transistors with undefined bulk have their bulks connected to V_{DD} and ground.

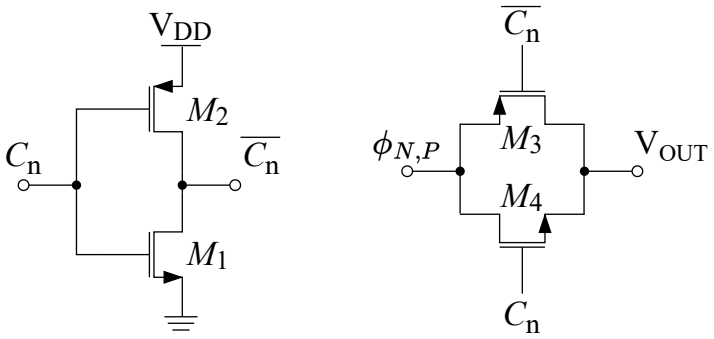


Figure 3.50: Transmission gate architecture. Note: The PMOS and NMOS transistors with undefined bulk have their bulks connected to V_{DD} and ground.

3.7.1 Simulation Results

Before adding the components' dimensions of the selection logic blocks, some additional adjustments were performed in the previous blocks. The main changes are related to the number of fingers used and the dimensions of the cascoded class-D transistors. For simplicity and considering that a final optimization will be performed after choosing the

selection logic that will integrate the architecture, these adjustments are not going to be presented here. Bearing this in mind, the components' dimensions of the AND block and the transmission gate block are presented in table 3.11 and table 3.12.

Table 3.11: Components' dimensions of the AND gate.

Transistor	Size [μm]	Number of Fingers
M_1	5/0.12	7
M_2	5/0.12	7
M_3	7/0.12	15
M_4	7/0.12	15
M_5	7/0.12	15
M_6	5/0.12	7

Table 3.12: Components' dimensions of the transmission gate.

Transistor	Size [μm]	Number of Fingers
M_1	5/0.12	8
M_2	7/0.120	30
M_3	7/0.120	8
M_4	5/0.120	30

In order to understand which selection block is more appropriate for the architecture, let us first take a look at the main waveforms of the blocks. Figure 3.51 presents the output voltages of both the RF Transmission gate and the AND gate. From inspection, it is possible to observe that the transmission gate block is slightly faster than the AND gate. Besides, since this block uses less components than the AND gate, using the transmission gate may be a better option in order to consume less power.

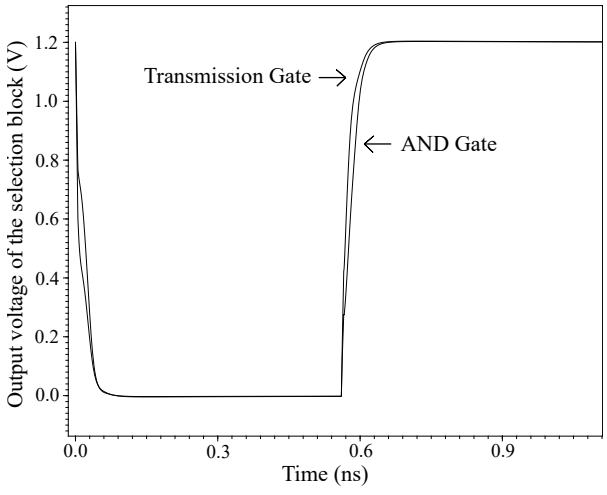
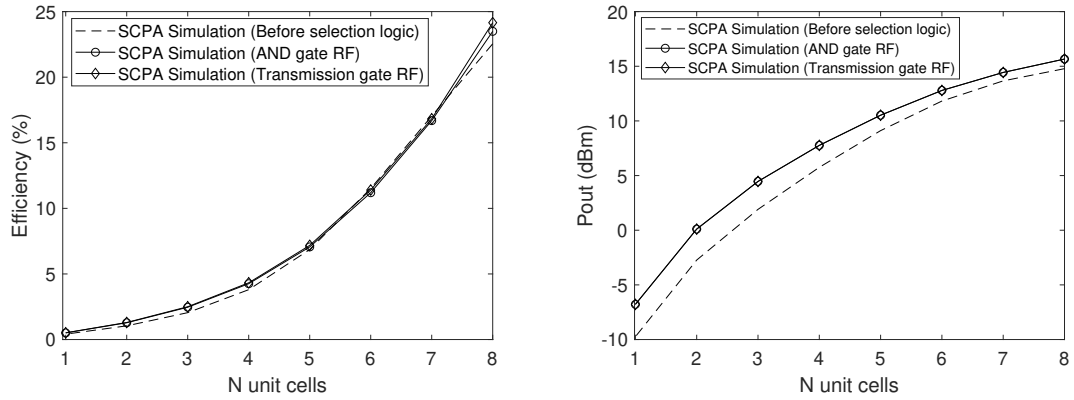


Figure 3.51: Output voltage of the transmission and AND gates in a single PA cell.

Figure 3.52 presents the simulated curves of efficiency and output power for both the AND and transmission gates. The two curves are compared to the efficiency and output power simulated curves obtained before adding the selection logic.



(a) Efficiency (b) Output Power

Figure 3.52: Efficiency and Output Power of the SCPA when adding the RF AND gate and the RF transmission gate (per number of active PA cells).

Looking into the simulated curves, one can see that adding the selection logic block improves both the efficiency and output power of the SCPA. These improvements are more clear when looking into the output power values. Comparing the curves obtained using the AND gate and the transmission gate, it is possible to see that the results are very similar: both the output power and the efficiency barely change from one case to another.

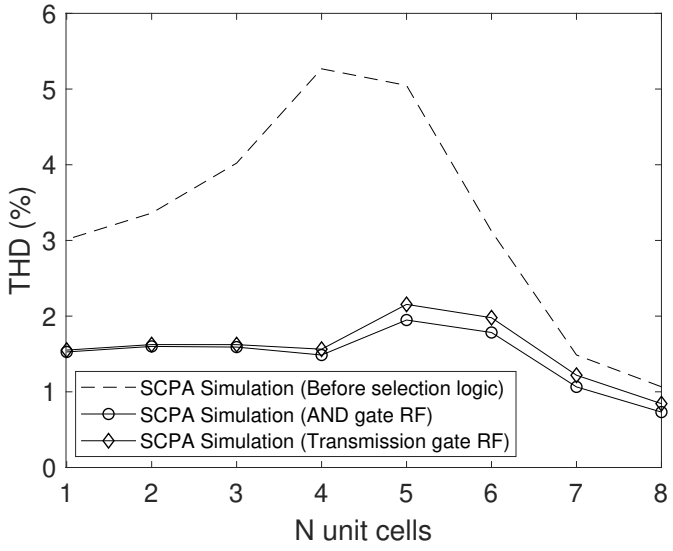


Figure 3.53: Total Harmonic Distortion when adding: the RF AND gate; the RF transmission gate.

Figure 3.53 shows the THD simulated curves of the SCPA for the same abovementioned cases. Results show a big improvement in the THD when the selection logic is introduced, dropping from a maximum of 5.27% when 4 unit PA cells are selected (and no selection gate is implemented) to a maximum of 2.16% when 5 unit PA cells are selected, for the transmission gate implementation, and a maximum of 1.95% when 5 unit PA cells are selected, for the AND gate implementation. Although the THD increases when using the transmission gate, the faster response and the reduced area of this circuit make it a better option for the architecture. For this reason, the transmission gate is the adopted block for the architecture.

Let us now analyse the output voltage when the final block of the SCPA is added in the architecture. Figure 3.54 presents the simulated voltage at the load, per different number of active PA cells.

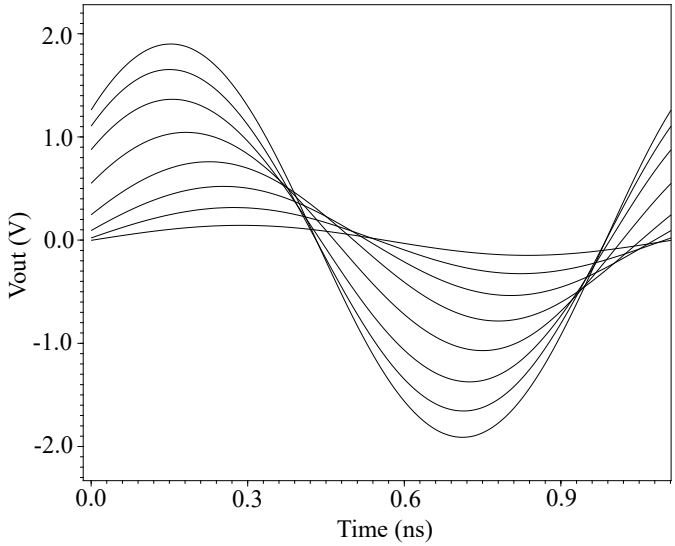


Figure 3.54: Output voltage of the SCPA architecture when adding the RF transmission gate, per number of PA cells.

From the figure, one can see that the output voltage is perfectly filtered for all the number of selected PAs. However, adding the transmission gate means a slower response of the architecture that it becomes even more slower when a lower number of unit PA cells is selected.

The output spectrum of the final architecture when used in full configuration is depicted in figure 3.55. In the final configuration the output power achieves a maximum value of 15.7dBm. The HD2 and HD3 correspond to -51.55dBc and -43.39dBc, respectively, and the THD is around 0.84%. It is important to note that, unlike what one would expect, the second harmonic is actually lower than the third harmonic. In fact, when adding the selection logic, the second harmonic dropped more than the third harmonic, giving the idea of an increase in the third harmonic. However, both the two harmonics are lower when the block is added.

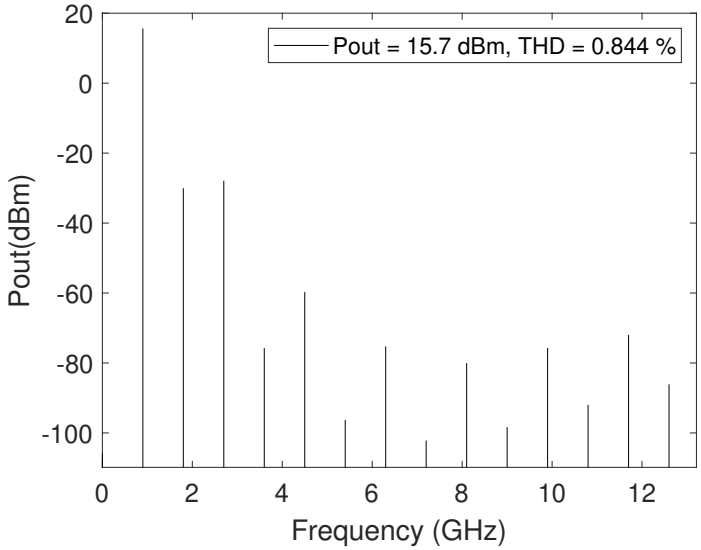


Figure 3.55: Output spectrum of the SCPA architecture when adding an RF transmission gate (all unit PA cells on.)

Finally, table 3.13 presents the main values of the SCPA for both the AND gate and the transmission gate. The adopted architecture uses the transmission gate. In this final configuration, the SCPA produces an output power of 15.66dBm with a PAE of 24.15% and a THD of 0.84%.

Table 3.13: SCPA architecture simulation results when adding selection logic (all PA cells on).

	Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	PAE [%]
AND gate	15.66	0.73	1.36	27.13	23.50
Transmission gate	15.66	0.84	1.36	27.13	24.16

3.8 Final Optimization

In order to boost the SCPA performance, a final optimization is performed. This optimization process is done in two steps: the first step, where the components dimensions of the different blocks are reviewed and, the second step, where the impact of varying the number of inverters of the driver block is analysed.

3.8.1 Final Sizing

In the following tables 3.14, 3.15, 3.16, 3.17 and 3.18, the final dimensions of the matching network, cascoded class-D PA, driver stage, level shifter and transmission gate, respectively.

Note: All the components are real, namely, the inductors (L_C20K_RF), the capacitors ($MIMCAPS\_RF$) and the transistors (P_12_RF and $N\_BPW\_12\_RF$).

A. Matching network and Filter

Table 3.14: Final components' dimensions of the matching network.

Inductors	Final Size [nH]
L0	0.57
L1	1.50
L2	2.07
L3	3.69
Capacitors	**Final Size [pF]**
C0	10.05
C1	10.75
C2	5.0

B. Cascoded Class-D PA - Per unit

Table 3.15: Final components' dimensions of the cascoded RF class D.

Transistor	Final Size [μm]	Number of Fingers
$M_{1,2}$	15.6/0.12	7
$M_{3,4}$	15/0.12	15

C. Driver Stage - Per unit

Table 3.16: Final components' dimensions of the RF driver stage.

Transistor	Size [μm]	Number of Fingers
M_1	1.15/0.12	6
M_2	1.6/0.12	8
M_3	1.15/0.12	6
M_4	1.6/0.12	8
M_5	2.3/0.12	6
M_6	3.2/0.12	8

D. Transmission gate - Per unit

Table 3.17: Final components' dimensions of the RF transmission gate.

Transistor	Size [μm]	Number of Fingers
M_1	5/0.12	8
M_2	7/0.12	30
M_3	7/0.12	8
M_4	5/0.12	30

E. Level shifter - Per unit

Table 3.18: Final components' dimensions of the RF level shifter.

Transistor	Initial Size [μm]	Number of Fingers
$M_{1A,1B}$	1.15/0.12	7
$M_{2A,2B}$	1.15/0.12	7
$M_{3A,3B}$	1.6/0.12	10
$M_{4A,4B}$	1.6/0.12	10
$M_{5,7}$	1.15/0.12	7
$M_{6,8}$	1.6/0.12	10

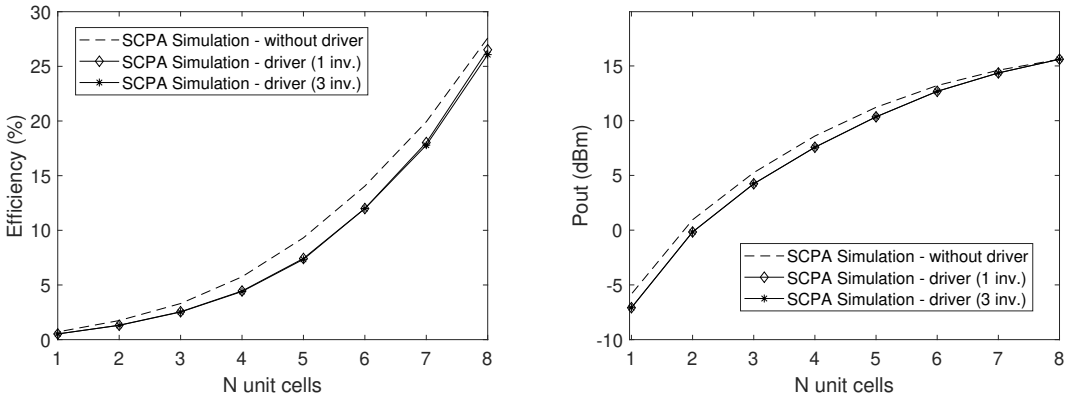
Capacitor	Final Size [pF]
C_{UP1}	17
C_{UP2}	0.9

3.8.2 Driver stage impact

Another way of improving the SCPA performance is by analysing more closely the impact of the driver stage within the architecture. As discussed in section 3.5, the optimum number of inverters and the sizing factor between each inverter is of 3. This means that, in order to optimally drive a big capacitance (that for this project corresponds to the gate capacitance of the cascoded class-D PA), an optimum number of 3 inverters should be implemented and each inverter should be 3 times bigger than the previous one (being the one closer to the driven capacitance the bigger). However, since for this project the dimensions of the cascoded class-D transistors are not typically big, the sizing factor between the inverters cannot be met. This suggests that a lower number of inverters stages should be used. For this reason, three additional simulations are performed and compared: the first, where the driver is removed from the architecture; the second, where the driver presents only 1 inverter; and the third, where the 3 inverters are used.

Figure 3.56 presents the simulated curves of the efficiency and output power for all three abovementioned cases. Looking into the results, one can easily see that both the efficiency and the output power curves are very similar when a driver with one inverter and a driver with three inverters are used. When the driver is removed from the architecture, both the measured output power and efficiency slightly improve when compared to the remaining cases.

Figure 3.57 shows the results of the simulated THD. From the figure, one can see that although removing the driver from the architecture increases the efficiency and output power, the THD also increases. On the other hand, results show that using a driver with one inverter or three inverters leads to very similar values of THD, which are significantly better when compared to the results obtained when no driver is used.



(a) Efficiency (b) Output Power

Figure 3.56: Efficiency and Output Power of the SCPA after final optimization when: no driver is used; a driver with one inverter is used; a driver with three inverters is used.

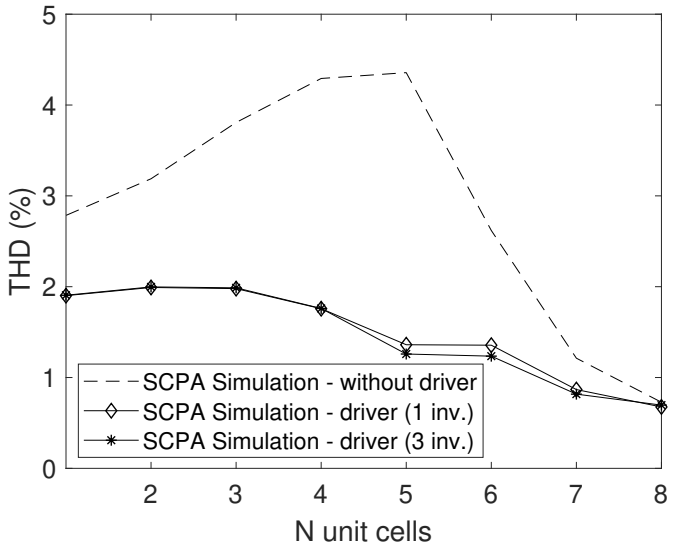


Figure 3.57: THD of the SCPA after final optimization when: no driver is used; the driver with one inverter is used; the driver with three inverters is used (per number of active PA cells).

Considering these results, the final implementation is modified in order to include a driver with only 1 inverter, instead of the 3 inverters initially designed. Figure 3.58 presents the simulated voltage at the load for the optimized configuration, per different number of active PA cells.

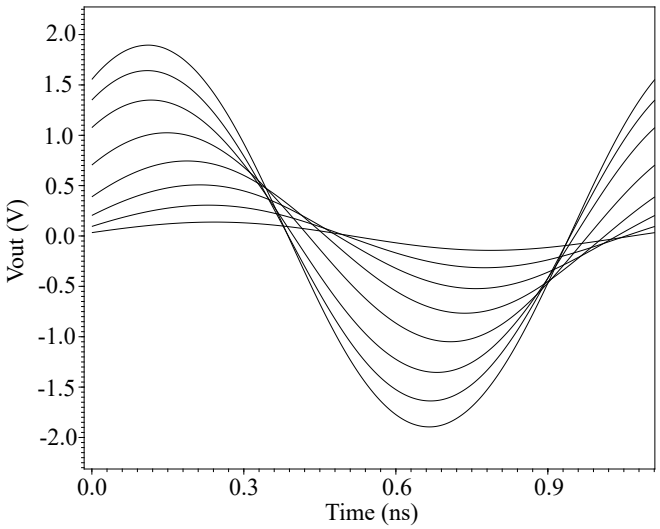


Figure 3.58: Output voltage after the final optimization, per number of PA cells.

The output spectrum of the optimized architecture when used in full configuration is depicted in figure 3.59. In this configuration, the output power achieves a maximum value of 15.61dBm (previously 15.66dBm). The HD2 and HD3 correspond to -70.23dBc (previously -51.55dBc) and -43.41dBc (previously -43.39dBc), respectively, and the THD is around 0.68% (previously 0.84%).

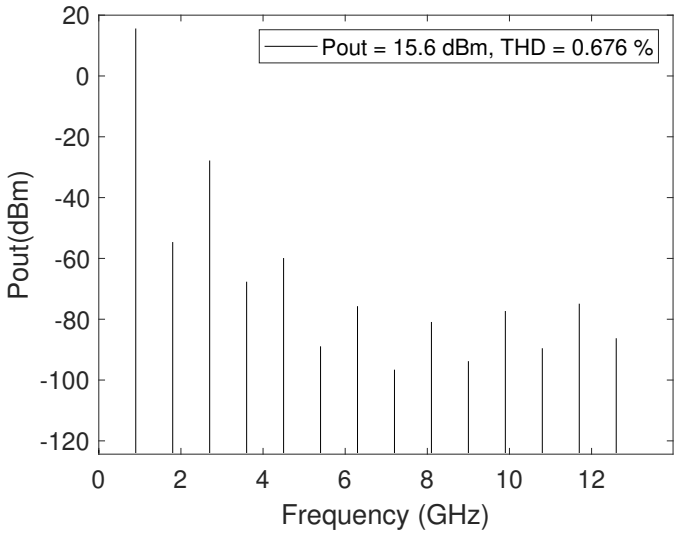


Figure 3.59: Output spectrum of the SCPA architecture after the final optimization (all unit PA cells on.).

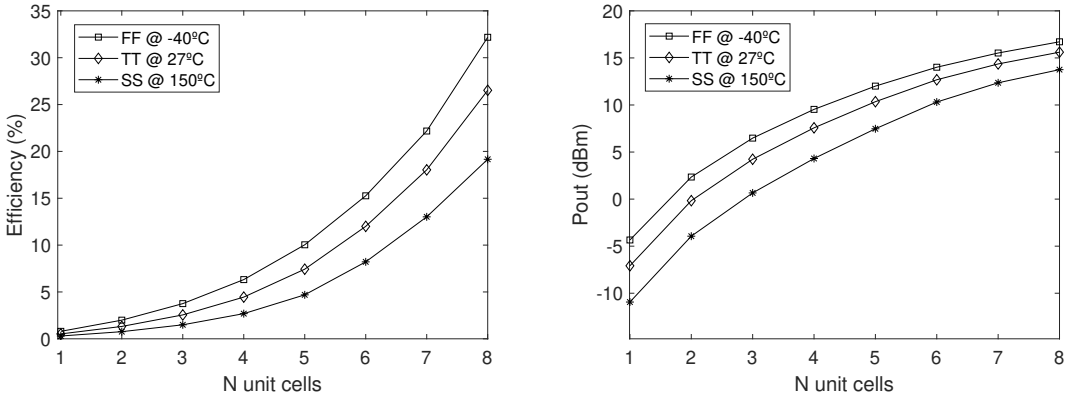
Lastly, table 3.19 presents the main values of the SCPA for the optimized architecture. In this final configuration, the SCPA produces an output power of 15.61dBm with a PAE of 26.52% and a THD of 0.68%.

Table 3.19: Final SCPA architecture simulation results (all PA cells on).

Output Power [dBm]	THD [%]	Voltage [Vrms]	Current [Irms]	PAE [%]
15.61	0.68	1.35	26.98	26.52

3.8.3 Performance Across PVT Corners

All the simulations presented so far were performed under particular PVT conditions (TT at 27ºC and 1.2V/2.4V supply voltage). However, in practical cases, variations in the process, temperature and voltage must be taken into account, as they will impact the overall performance of the projected circuit. For this reason, let us analyse the SCPA performance for different processes (FF, SS, TT) when the temperature varies from -40ºC to 150ºC. Figure 3.60 presents the simulated results of output power and efficiency for the most relevant PVT corners. (Note: Variations in the supply voltage are not considered.)



(a) Efficiency (b) Output Power

Figure 3.60: Efficiency and Output Power simulated values across the most relevant PVT corners (for different numbers of active PA cells).

Looking at the figure, one can easily see that the worst results are obtained for a SS process at 150ºC. In this corner, the efficiency drops to 19.2% and the output power reaches its maximum value of 13.75dBm (when the PA is used in a full configuration). Note that, this particular corner (along with FF at -40ºC) corresponds to a very extreme case, that hardly would happen in reality. The best results are obtained for a FF corner at -40ºC, conditions at which the efficiency reaches 32.18% and the output power 16.71dBm. These results suggest that the PA operates better for lower temperatures and faster corners,

as expected. In fact, since one of the main constraints of the PA is the power dissipation, higher temperatures will lead to bigger losses. Besides, a faster process means having faster transistors and hence a faster charge and discharge of the output nodes.

Figure 3.61 presents the simulated results of the THD for the same PVT corners previously shown. Unlike the previous cases, the FF corner at -40ºC seems to have a higher range of THD values and the SS at 150ºC seems to have a lower range. The bigger the number of unit PA cells used, the better the results when using FF process at -40ºC. However, the lower the number of unit PA cells used, the worse the results are using the same corner. The opposite happens in the slow corner at 150ºC.

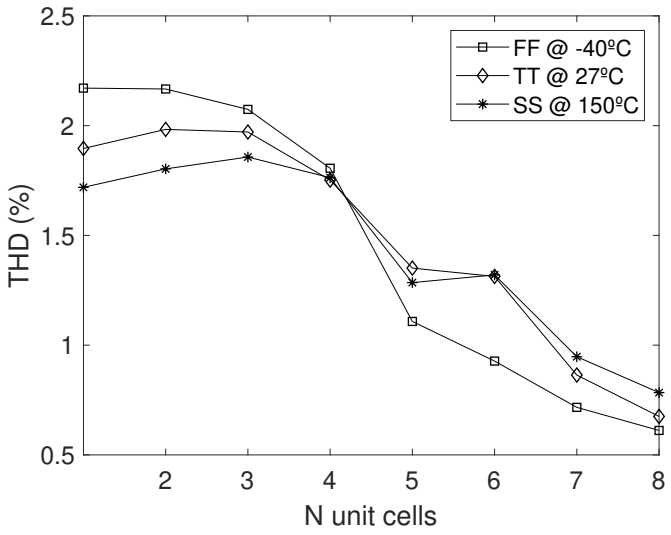


Figure 3.61: THD simulated values across the most relevant PVT corners (for different numbers of active PA cells)

Table 3.20: Final SCPA architecture simulation results across some PVT corners (all PA cells on).

	Output Power [dBm]	Voltage [Vrms]	Current [Irms]	PAE [%]	THD [%]
TT -40°C	16.52	1.50	29.96	31.41	0.62
TT 27°C	15.61	1.35	26.99	26.52	0.68
TT 150°C	14.16	1.14	22.83	20.08	0.77
FF -40°C	16.71	1.53	30.63	32.18	0.61
FF 27°C	15.83	1.38	27.66	27.27	0.67
FF 150°C	14.39	1.17	23.45	20.70	0.76
SS -40°C	16.19	1.44	28.83	30.32	0.64
SS 27°C	15.24	1.29	25.85	25.47	0.69
SS 150°C	13.75	1.09	21.79	19.16	0.78

Table 3.20 presents the main results of the three corners when operating at different temperatures, when all the PA cells are on.

The above-mentioned results provide a clear idea of how the architecture behaves when operating in extreme conditions (meaning, all the NMOS and PMOS are identically slow, fast or typical). Although it can happen to fall within one of those cases, it is unlikely that it would happen. In reality, it may occur that after the fabrication process some transistors will be slower and others faster.

A simple way of evaluating the impact of different processes and mismatches is by performing a Monte Carlo simulation. In this work, a Monte-Carlo corner is analysed. This simulation uses as a starting point all PMOS and NMOS in typical conditions and, every run, introduces both process variations and mismatches across different transistors. Additionally, this model also introduces process variations and mismatches in the capacitors. In this way, a Monte Carlo simulation is performed for 30 iterations (referred to in the documentation of the technology as the iteration number necessary to prove the proper behaviour of the circuit, meaning that over 80% of all possible component values operate correctly), a sigma of 3 (parameter that dictates the standard deviation of the mean value, covering 99.7% of the possible cases) and at 27ºC.

Figures 3.62, 3.63 and 3.64 present the Monte Carlo simulation results for the PAE, output power and THD, respectively.

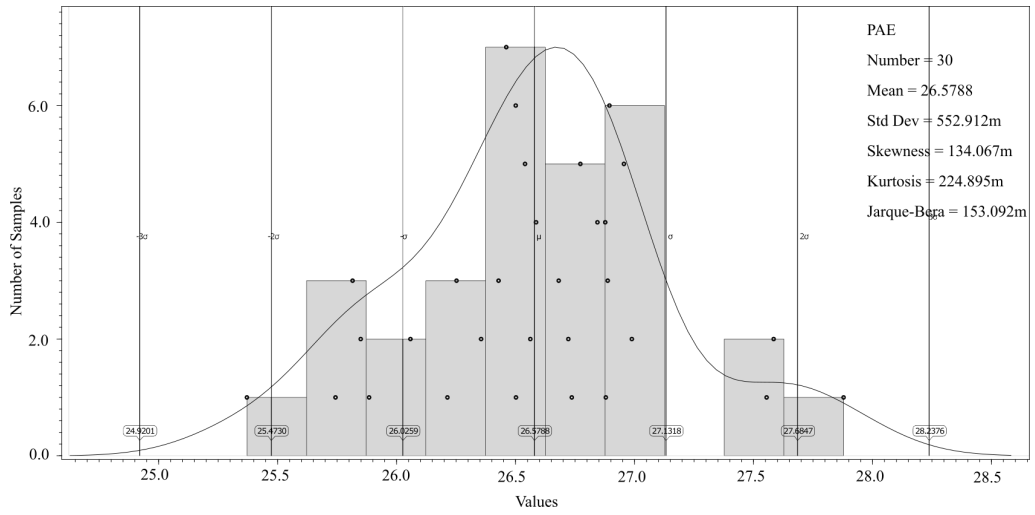


Figure 3.62: Monte Carlo Simulation for the Efficiency (all PA cells on).

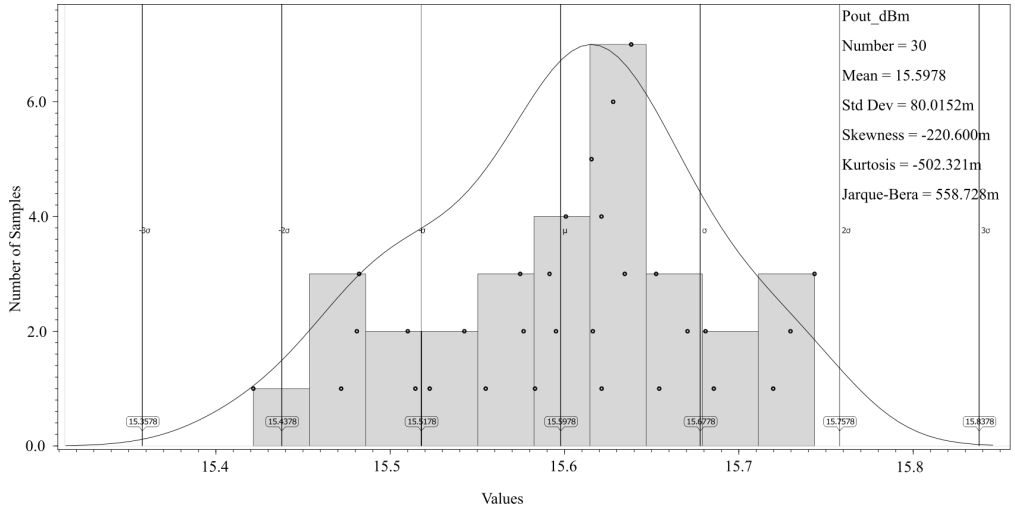


Figure 3.63: Monte Carlo Simulation for the Output Power (all PA cells on).

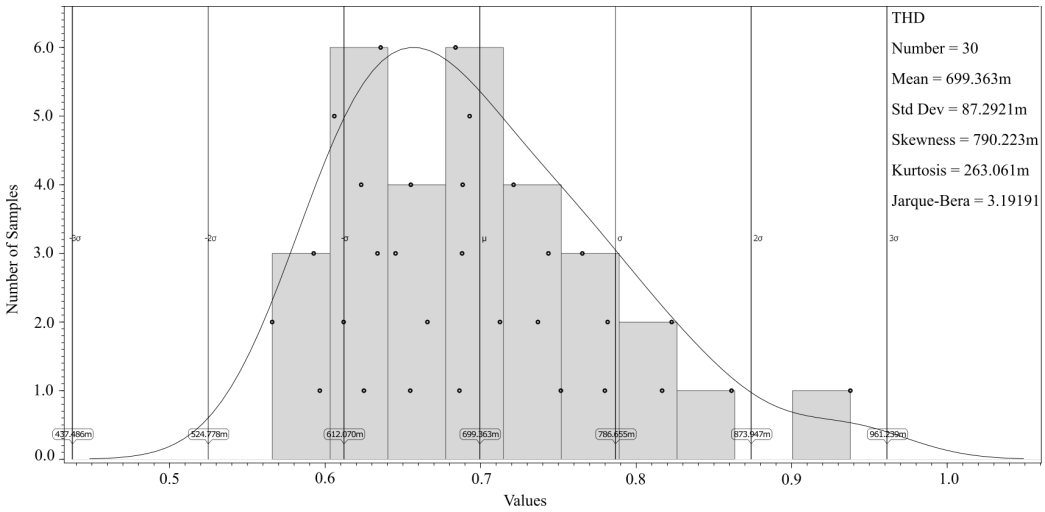


Figure 3.64: Monte Carlo Simulation for the THD (all PA cells on).

Table 3.21 presents the main values extracted from the aforementioned simulations, where μ corresponds to the mean value and σ refers to the standard deviation.

Table 3.21: Monte Carlo Simulation Results (all PA cells on).

	Pout [dBm]	PAE [%]	THD [%]
μ	15.6	26.58	0.7
σ	0.08	0.55	0.09

3.9 Modulation Stage

In order to understand how the SCPA behaves when integrated into the transmission architecture, an additional modulation stage is implemented.

Although for NB-IoT applications the modulation employed is typically QPSK, in this project the 16-QAM modulation is presented. The main reason behind this choice is related to the fact that in QPSK modulation only one amplitude is used, which means that the SCPA would be always operating with a fixed number of unit PA cells and only the phase clock would change. On the other hand, in the 16 QAM modulation, 3 different amplitudes are used combined with 12 different phases, allowing to study better the behaviour of the PA. The adopted modulation stage is presented in Figure 3.65.

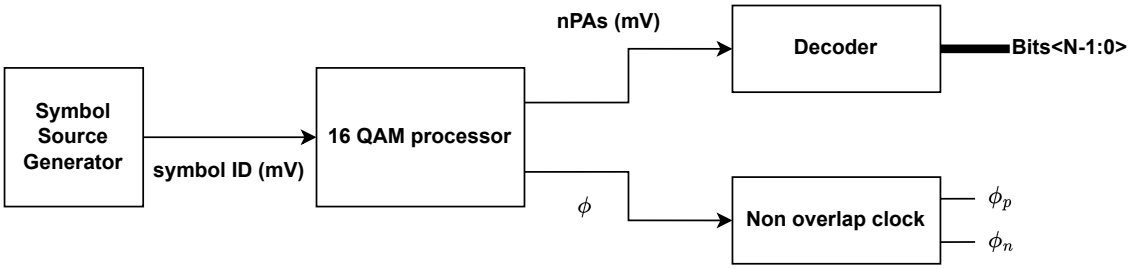


Figure 3.65: Proposed modulation stage of the SCPA.

Broadly speaking, the idea of the modulation stage is to have a symbol source generator (that in this project is done using an analogue component) that generates a random symbol from the 16-QAM constellation. This symbol is then sent to the 16QAM processor, that will generate a number (nPAs) that represent the PAs that have to be selected and a phase clock (Φ). From here, the number of PAs signal is consumed by the decoder block that, according to the number received (that can go from 1 up to 8), will select the respective unit PA cells. The phase clock, on the other hand, will be consumed by a non-overlap clock block in order to generate the phase to be used by the PMOS and by the NMOS of the cascoded topology. However, since in this project it was verified that using only one clock to feed both the PMOS and the NMOS is preferable, the non-overlap clock block is not used. Notwithstanding, the reasoning behind this block will be discussed. All the blocks in this section were designed using VerilogA.

3.9.1 16 QAM Processor

Let us now analyse how the 16QAM processor is implemented. As already discussed, this block is responsible for translating the symbol ID into a phase clock and a number of PAs. The adopted modulation scheme is presented in table 3.22, where either 8, 6 or 3 unit PA cells are selected. Note that the choice of PAs to be selected was arbitrary. Eventually, using 8, 7 or 6 unit PA cells would allow us to achieve better results.

Table 3.22: 16 QAM Processor.

Symbol ID	Generated phase	Number of PAs selected (mV)
1	2.36	8
2	2.82	6
3	-2.36	8
4	-2.82	6
5	1.89	6
6	2.36	3
7	-1.89	6
8	-2.36	3
9	0.79	8
10	0.32	6
11	-0.79	8
12	-0.32	6
13	-1.25	6
14	0.79	3
15	-1.25	6
16	-0.79	3

Depending on the generated phase (Φ) and on the number of PAs selected (A) a carrier waveform is obtained, according to the equation presented in 3.50:

$$carrier = A\dot{c}os(2\pi f_c t + \Phi), \tag{3.50}$$

where the f_c corresponds to the carrier frequency, that for this project is given by 0.9G. This equation allows to obtain the phase clock by using the following reasoning: when the amplitude of the carrier signal is higher than 0, the clock is set to V_{DD}, when the amplitude of the carrier signal is below than 0, the clock is set to ground. In this way, the information of both the frequency and the phase are automatically transferred to the phase clock. The number of selected PAs is sent directly to the following block.

3.9.2 Non Overlapping Clocks

The non-overlapping clocks can be obtained by applying the reasoning illustrated in figure 3.66. As mentioned before, this block receives as an input the phase clock of the symbol to be transmitted. Let us represent this signal in the figure as the initial phase. Let us now obtain a delayed version of this signal, where the delay corresponds to the desired time guard between the two clocks. These two signals are now going to be used to obtain both the NMOS and the PMOS phases.

From inspection, it is possible to see that the PMOS phase can be set to V_{DD} when the amplitude of the initial phase is higher than half of V_{DD} (marked in the figure as Vtrans) and set to 0 when the delayed phase is below Vtrans. In the same manner, the NMOS phase can be set to V_{DD} when the amplitude of the delayed phase is higher than half of V_{DD} and set to 0 when the initial phase is below Vtrans.

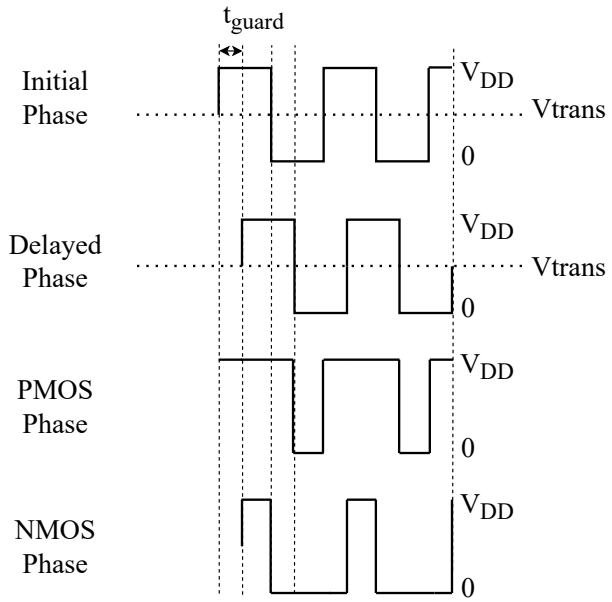


Figure 3.66: Illustration of how the PMOS and NMOS phases are generated.

By doing this, two non-overlapping clocks can be obtained. Note that, as already mentioned, a single-phase clock is used in this project.

3.9.3 Decoder

The truth table behind the decoder is very straightforward and it is presented in table 3.23. As already discussed, albeit the table presents all the possible cases for the 16 QAM modulation, only 3 cases are used (3, 6 and 8 unit PAs). In such configuration, each output of the decoder is connected directly to a unit PA cell.

Table 3.23: 16 QAM Decoder.

Number of selected PAs (mV)	B0	B1	B2	B3	B4	B5	B6	B7
1	V_{DD}	0	0	0	0	0	0	0
2	V_{DD}	V_{DD}	0	0	0	0	0	0
3	V_{DD}	V_{DD}	V_{DD}	0	0	0	0	0
4	V_{DD}	V_{DD}	V_{DD}	V_{DD}	0	0	0	0
5	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	0	0	0
6	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	0	0
7	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	0
8	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}	V_{DD}

3.9.4 Simulation Results

Figure 3.67 presents the simulation of the overall system when 4 random symbols from the 16-QAM modulation are transmitted. A complete simulation of all the symbols was also performed and it is presented in section C.1 of Appendix B. For now, let us consider this simple case.

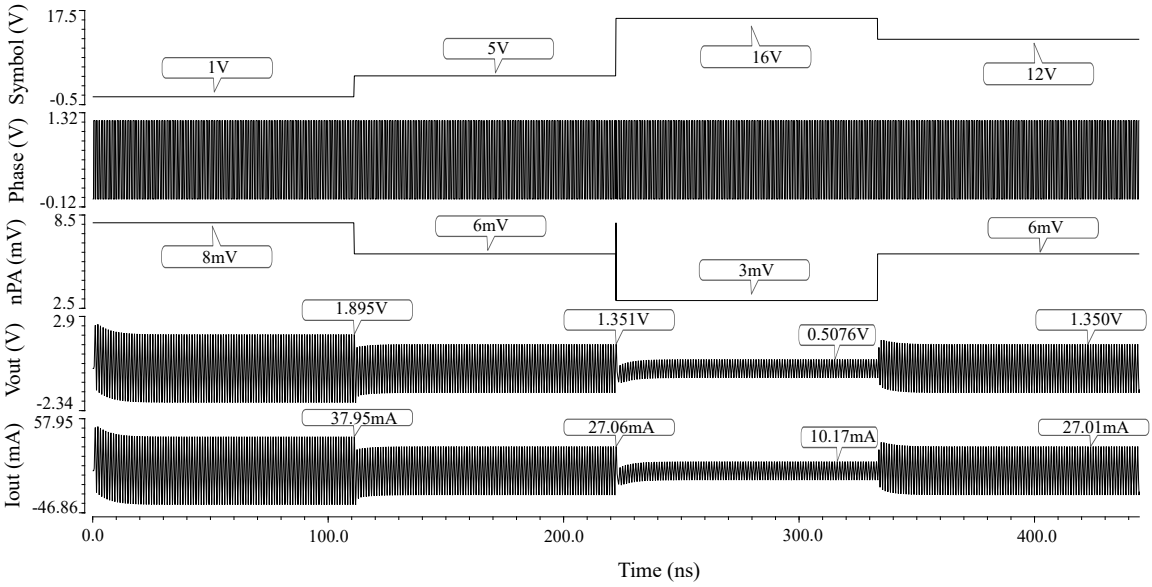


Figure 3.67: Example of the modulation scheme for 4 symbols.

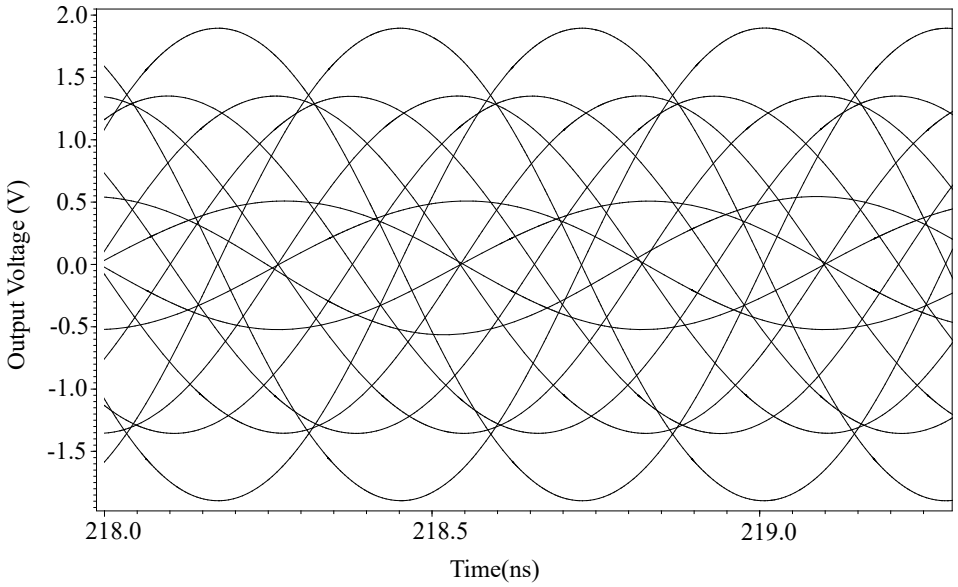


Figure 3.68: Eye diagram of the output voltage for the different 16 symbols.

Looking at figure 3.67, one can see the 4 main signals in the architecture. The first signal, Symbol, identifies the symbol that is being transmitted. In this case, the symbols being transmitted are 1, 5, 16 and 12. The second signal, Phase, corresponds to the clock phase of each of the signals being transmitted. Although it is not clear in the figure, the clock phase is changing when every time a new symbol is transmitted. Recalling the values presented in the table 3.22, one can confirm that all the symbols present a different phase, meaning the clock is changing between the 4 symbols. The third signal, nPA, corresponds to the number of unit PA cells that are selected for each transmitted symbol. In the simulation, they correspond to 8, 6, 3 and 8, respectively. Finally, the signals Vout and Iout, correspond to the simulated output voltage and output current obtained for the different symbols. These values are very similar to the final values obtained in section 3.8. Note that, the second symbol (symbol 5) and the last symbol (symbol 12) present the same amplitude value, but different phases. The same behaviour could be obtained by looking into the current signal.

In order to have a better understanding of the different phases, let us use the eye diagram tool to plot the different symbols being transmitted. Using the symbol time duration, it is possible to superimpose the different symbols, each one at the top of the other, giving a clear perception of the different phases and amplitudes. Figure 3.68 presents the output voltages for all 16 symbols.

4

Conclusions and Future Work

4.1 Conclusions

In Chapter 2, the concept behind the main transmitter architectures was analysed, along with the role of the main blocks involved in the process of the signal being transmitted. Furthermore, the metrics used to evaluate the non-idealities of the Power Amplifier and to properly measure its efficiency and linearity were carefully studied. In addition, the different classes of PAs and the common efficiency and linearization architectures used to overcome their limitations were presented.

In Chapter 3, the theoretical analysis and implementation methodology of each block within the SCPA is presented. The SCPA is first implemented with ideal components in order to validate the model. The real blocks are then added, one by one, and the impact of each block in the architecture is carefully analysed. A final modulation stage is designed for a 16-QAM modulation in order to analyse the performance of the architecture when different symbols are transmitted.

In the ideal configuration with modelled losses, the SCPA is able to produce the desired output power of 19.05dBm with the expected PAE of 36.43%, in full configuration. When replacing the ideal components of the matching network with the real components, the output power and efficiency drop considerably. An additional optimization is performed to improve the results, allowing to reach a maximum output power of 15.66dBm, a maximum PAE of 32.50% and a maximum THD of 0.45%.

After sizing the matching network, the cascoded class-D PA switch was introduced in the architecture. In order to have a better understanding of the impact of using normal transistors and RF transistors in the switch cell, the two abovementioned cases were compared. In the first case, the simulated maximum output power and PAE are of 14.50dBm and 23.73% and, in the second case, of 14.22dBm and 21.48%. Since the architecture is designed with better-modelled RF transistors, the matching network was readjusted for this case in order to boost the results, leading to a maximum output of 14.76dBm, PAE of 28.91% and THD of 1.18% in full configuration. It was also observed that using the same clock in both the NMOS and PMOS branches is preferable to using a non-overlap

clock, since increasing the time guard leads to worst results. The same is observed when looking into the rising and falling time of the phase clock: the lower the time, the better the results.

Having added the cascoded class-D PA, the next block being implemented was the switch driver, consisting of 3 inverters connected in a cascade configuration. The simulations show that adding this block leads to an output power of 14.78dBm, a PAE of 25.07% and a THD of 1.057%, in full configuration. Unlike the previous cases, introducing this block increases significantly the total harmonic distortion when the SCPA is used in other configurations, achieving a maximum value higher than 5%.

The next block being designed was the level shifter, used to shift the phase clock of the PMOS by 1.2V. The circuit was analysed for different combinations of C_{up1} and C_{up2}. As expected, adding this block impacted the efficiency of the system, which dropped to 22.50%. The remaining metrics showed similar results.

Finally, the selection logic was introduced in the architecture and a final optimization was performed. At this point the study of the optimum number of inverters was conducted, showing that using 1 inverter is sufficient to obtain good results. This solution not only decreases the area of the whole system, but also reduces the THD. The complete architecture can produce 15.61dBm of output power with a PAE of 26.52% and a THD of 0.68%, when in full configuration. In all configurations, the THD is always below than 2.5%. Considering these results, an additional study was conducted to evaluate the performance of the SCPA across PVT corners and an additional Monte Carlo analysis was carried out. Results show the SCPA is robust when operating in different conditions.

The final overall system is implemented after adding the modulation stage. This modulation stage allows to give a context on how the PA is inserted in the transmission architecture. A final simulation was performed where the PA had to transmit the 16 symbols of the constellation.

Table 4.1 presents a comparison of this work with some published works. Note that, all the works presented in this table were chosen in order to be used by NB-IoT applications, where the maximum output power cannot exceed 23 dBm.

Table 4.1: Comparison of published RF PAs and this work.

Reference	Process [nm]	Frequency [GHz]	Supply Voltage [V]	Output Power [dBm]	PAE [%]
[51]	16	0.9	0.85/0.3	11.5	51
[7]	40	0.750-0.930	1	8	45
[42]	130	0.8	1.2	19	18
[54]	180	0.750-0.930	1.7	13.6	25.25
[14]	180	0.9	3.3/2.8	23	44
[19]	600	0.9	3	20	30
This work	130	0.9	1.2/2.4	15.61	26.52

4.2 Future Work

Following the work presented in this document, future improvements can be further investigated. Firstly, a global optimization of the matching network should be performed in order to guarantee a better impedance matching between the SCPA architecture and the load. This would allow not only to achieve better efficiency and higher output power, but also to reduce the total harmonic distortion of the topology. The optimization could be extended to the remaining blocks of the architecture, in order to improve their area while maintaining efficiency. Genetic algorithms or similar optimization techniques could be a helpful and powerful tool in the design of the PA. The optimization could be performed aiming for a higher output power that could be comparable with more projects.

Secondly, a differential implementation of the SCPA should be considered in order to achieve better linearity (typically by reducing the second harmonic of the output spectrum) and higher output power.

Another interesting improvement that could be analysed is the implementation of the architecture in a more advanced technology node. Since the SCPA employs a switching capacitor technique, the CMOS scaling can help achieve better results by reducing the parasitic losses. This would allow to achieve faster transitions, reduce power dissipation and increase efficiency.

Moreover, a further examination of the transmitted constellation points could be performed in order to understand the impact of the previous symbols in the current transmission. By transforming the output signals into symbols and by comparing them with the initial symbols, one could analyse the non-idealities added by topology, namely the impact on the linearity and the memory effect.

Finally, the layout of the different blocks that integrate the SCPA should be performed in order to evaluate more accurately the PAs performance.

Bibliography

[1] "A 30.5 dBm 48% PAE CMOS Class-E PA With Integrated Balun for RF Applications". In: (Aug. 2008). ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.925605 (cit. on p. 19).

[2] "A 65 nm CMOS 30 dBm Class-E RF Power Amplifier With 60% PAE and 40% PAE at 16 dB Back-Off". In: (May 2009). ISSN: 0018-9200. DOI: 10.1109/JSSC.2009.2020680 (cit. on p. 19).

[3] "A Fully-Integrated High-Power Linear CMOS Power Amplifier With a Parallel-Series Combining Transformer". In: (Mar. 2012). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2011.2180977 (cit. on p. 22).

[4] "A highly linear 25dBm outphasing power amplifier in 32nm CMOS for WLAN application". In: ISSN: 1930-8833. Sept. 2010. DOI: 10.1109/ESSCIRC.2010.5619705 (cit. on p. 23).

[5] "A Multiphase Switched Capacitor Power Amplifier". In: (2017). ISSN: 1558-173X. DOI: 10.1109/JSSC.2016.2626277 (cit. on p. 28).

[6] "An Efficient Mixed-Signal 2.4-GHz Polar Power Amplifier in 65-nm CMOS Technology". In: (Aug. 2011). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2011.2155790 (cit. on p. 18).

[7] A. Ba et al. "26.3 A 1.3nJ/b IEEE 802.11ah fully digital polar transmitter for IoE applications". In: *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. 2016, pp. 440–441. DOI: 10.1109/ISSCC.2016.7418096 (cit. on p. 96).

[8] R. N. Braithwaite. "Using a cascade of digital and analog predistortion to linearize a dual-band RF transmitter". In: *2017 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*. 2017, pp. 77–80. DOI: 10.1109/PAWR.2017.7875578 (cit. on p. 23).

[9] J. Crols and M. Steyaert. *CMOS Wireless Transceiver Design*. 2003 (cit. on p. 6).

[10] Z. Deng et al. "9.5 A dual-band digital-WiFi 802.11a/b/g/n transmitter SoC with digital I/Q combining and diamond profile mapping for compact die area and improved efficiency in 40nm CMOS". In: Jan. 2016. ISBN: 9781467394673. DOI: 10.1109/ISSCC.2016.7417962 (cit. on p. 27).

[11] Y. Fei et al. "Class F and Inverse Class F Dual Modes Dual Bands Power Amplifier". In: *2020 IEEE MTT-S International Wireless Symposium (IWS)*. 2020, pp. 1–3. DOI: 10.1109/IWS49314.2020.9360107 (cit. on p. 20).

[12] J. Fritzin et al. "Phase Predistortion of a Class-D Outphasing RF Amplifier in 90 nm CMOS". In: (). ISSN: 1558-3791. DOI: 10.1109/TCSII.2011.2164149 (cit. on p. 18).

[13] J. Fuhrmann et al. "A 15-bit 28nm CMOS fully-integrated 1.6W digital power amplifier for LTE IoT". In: Sept. 2017. ISBN: 9781509050253. DOI: 10.1109/ESSCIRC.2017.8094560 (cit. on p. 27).

[14] T. Gao et al. "A novel CMOS transmitter front-end for mobile RFID reader". In: *2009 IEEE Radio Frequency Integrated Circuits Symposium*. 2009, pp. 105–108. DOI: 10.1109/RFIC.2009.5135500 (cit. on p. 96).

[15] A. Ghahremani, A.-J. Annema, and B. Nauta. "A 20dBm outphasing class E PA with high efficiency at power back-off in 65nm CMOS technology". In: Honolulu, HI, USA, June 2017. ISBN: 9781509046263. DOI: 10.1109/RFIC.2017.7969087 (cit. on p. 23).

[16] A. R. Ghorbani and M. B. Ghaznavi-Ghoushchi. "A 35.6dB, 43.3% PAE class E differential power amplifier in 2.4GHz with cross coupling neutralization for IoT applications". In: May 2016. ISBN: 9781467387897. DOI: 10.1109/IranianCEE.2016.7585571 (cit. on p. 19).

[17] M. Gilasgar, A. Barlabé, and L. Pradell. "A 2.4 GHz CMOS Class-F Power Amplifier With Reconfigurable Load-Impedance Matching". In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 66.1 (2019), pp. 31–42. DOI: 10.1109/TCSI.2018.2851608 (cit. on p. 20).

[18] P. A. Godoy et al. "A 2.4-GHz, 27-dBm Asymmetric Multilevel Outphasing Power Amplifier in 65-nm CMOS". In: (Oct. 2012). ISSN: 1558-173X. DOI: 10.1109/JSSC.2012.2202810. URL: http://ieeexplore.ieee.org/document/6241449/ (cit. on p. 23).

[19] R. Gupta, B. Ballweber, and D. Allstot. "Design and optimization of CMOS RF power amplifiers". In: *IEEE Journal of Solid-State Circuits* 36.2 (2001), pp. 166–175. DOI: 10.1109/4.902757 (cit. on p. 96).

[20] T.-P. Hung et al. "CMOS Outphasing Class-D Amplifier With Chireix Combiner". In: (Aug. 2007). ISSN: 1531-1309. DOI: 10.1109/LMWC.2007.901800 (cit. on p. 23).

[21] H. Jin, D. Kim, and B. Kim. "Efficient Digital Quadrature Transmitter Based on IQ Cell Sharing". In: (May 2017). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2017.2655058 (cit. on p. 28).

[22] S. Jin et al. "A Highly Efficient CMOS Envelope Tracking Power Amplifier Using All Bias Node Controls". In: (Aug. 2015). ISSN: 1531-1309, 1558-1764. DOI: 10.1109/LMWC.2015.2440652 (cit. on p. 26).

[23] Jinsung Choi et al. "Optimized Envelope Tracking Operation of Doherty Power Amplifier for High Efficiency Over an Extended Dynamic Range". In: (June 2009). ISSN: 0018-9480, 1557-9670. DOI: 10.1109/TMTT.2009.2020674 (cit. on p. 26).

[24] D. Kang et al. "Envelope-Tracking CMOS Power Amplifier Module for LTE Applications". In: (Oct. 2013). DOI: 10.1109/TMTT.2013.2280186 (cit. on p. 26).

[25] A. Kavousian et al. "A Digitally Modulated Polar CMOS Power Amplifier With a 20-MHz Channel Bandwidth". In: (Oct. 2008). ISSN: 0018-9200. DOI: 10.1109/JSSC.2008.2004338 (cit. on pp. 26, 34).

[26] E. Kaymaksut and P. Reynaert. "Transformer-Based Uneven Doherty Power Amplifier in 90 nm CMOS for WLAN Applications". In: (July 2012). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2012.2191334 (cit. on p. 22).

[27] M. K. Kazimierczuk. *RF power amplifiers*. Wiley, 2008. ISBN: 9780470779460 (cit. on p. 19).

[28] C.-y. Liu et al. "A 2.4 GHz CMOS Doherty Power Amplifier". In: June 2006. ISBN: 9780780395411. DOI: 10.1109/MWSYM.2006.249834 (cit. on p. 22).

[29] *NB-IoT 3GPP release*. https://haltian.com/resource/nb-iot-3gpp-release-14-what-are-the-new-features/ (cit. on p. 2).

[30] K. Oishi et al. "A 1.95 GHz Fully Integrated Envelope Elimination and Restoration CMOS Power Amplifier Using Timing Alignment Technique for WCDMA and LTE". In: (Dec. 2014). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2014.2358554 (cit. on p. 25).

[31] L. B. Oliveira. *Eletronic IV Slides: Analysis and project elements for passive and active RF-Microwave circuits*. 2018 (cit. on pp. 39, 40).

[32] L. B. Oliveira et al. *Analysis and Design of Quadrature Oscillators*. Springer Netherlands, 2008. ISBN: 9781402085154. DOI: 10.1007/978-1-4020-8516-1 (cit. on p. 6).

[33] *PA Survey*. https://ideas.ethz.ch/research/surveys/pa-survey.html (cit. on p. 29).

[34] A. Passamani et al. "A linear model of efficiency for Switched-Capacitor RF Power-Amplifiers". In: *2014 10th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME)*. 2014, pp. 1–4. DOI: 10.1109/PRIME.2014.6872748 (cit. on p. 41).

[35] J. M. Rabaey, A. Chandrakasan, and B. Nikolic. *Digital Integrated Circuits: A design perspective*. 2003 (cit. on pp. 52, 53, 63).

[36] B. Razavi. "RF transmitter architectures and circuits". In: San Diego, CA, USA, 1999. ISBN: 9780780354432. DOI: 10.1109/CICC.1999.777273 (cit. on p. 6).

[37] J. REN and X. YANG. "A Narrow-Band Dithering Technique in Feedback Channel Quantization for Improving Effect of Digital Predistorter". In: *2021 IEEE International Conference on Consumer Electronics (ICCE)*. 2021, pp. 1–5. DOI: 10.1109/ICCE50685.2021.9427724 (cit. on p. 23).

[38] P. Reynaert and M. Steyaert. "A 1.75-GHz polar modulated CMOS RF power amplifier for GSM-EDGE". In: (Dec. 2005). ISSN: 0018-9200. DOI: 10.1109/JSSC.2005.857425 (cit. on p. 25).

[39] H. S. Ruiz and R. Berenguer. *Linear CMOS RF power amplifiers: a complete design workflow*. Springer, 2014. ISBN: 9781461486572 (cit. on pp. 8–12, 17–19, 21–25, 30).

[40] B. Serneels, M. Steyaert, and W. Dehaene. "A High speed, Low Voltage to High Voltage Level Shifter in Standard 1.2V 0.13μm". In: *2006 13th IEEE International Conference on Electronics, Circuits and Systems*. 2006, pp. 668–671. DOI: 10.1109/ICECS.2006.379877 (cit. on p. 69).

[41] R. Shrestha. "RF Power Amplifier Techniques for Spectral Efficiency and Software-Defined Radio". PhD thesis. University of Twente, 2010. DOI: 10.3990/1.9789036530767 (cit. on p. 21).

[42] N. Singhal et al. "A 19 dBm 0.13µm CMOS parallel class-E switching PA with minimal efficiency degradation under 6 dB back-off". In: *2011 IEEE Radio Frequency Integrated Circuits Symposium*. 2011, pp. 1–4. DOI: 10.1109/RFIC.2011.5940714 (cit. on p. 96).

[43] Z. Song et al. "A Low-Power NB-IoT Transceiver With Digital-Polar Transmitter in 180-nm CMOS". In: (Sept. 2017). ISSN: 1549-8328, 1558-0806. DOI: 10.1109/TCSI.2017.2707412 (cit. on p. 27).

[44] V. Vorapipat, C. S. Levy, and P. M. Asbeck. "A Class-G Voltage-Mode Doherty Power Amplifier". In: (Dec. 2017). ISSN: 1558-173X. DOI: 10.1109/JSSC.2017.2748283 (cit. on p. 28).

[45] J. S. Walling. "The Switched-Capacitor Power Amplifier: A Key Enabler for Future Communications Systems". In: Sept. 2019. ISBN: 9781728115504. DOI: 10.1109/ESSCIRC.2019.8902890 (cit. on pp. 25, 27).

[46] J. S. Walling et al. "A Class-E PA With Pulse-Width and Pulse-Position Modulation in 65 nm CMOS". In: 44 (June 2009). ISSN: 0018-9200. DOI: 10.1109/JSSC.2009.2020205 (cit. on p. 27).

[47] J. S. Walling et al. "A Switched-Capacitor RF Power Amplifier". In: (). ISSN: 1558-173X. DOI: 10.1109/JSSC.2011.2163469 (cit. on pp. 26, 28, 35).

[48] F. Wang et al. "A Monolithic High-Efficiency 2.4-GHz 20-dBm SiGe BiCMOS Envelope-Tracking OFDM Power Amplifier". In: (June 2007), pp. 1271–1281. ISSN: 0018-9200. DOI: 10.1109/JSSC.2007.897170 (cit. on p. 26).

[49] F. Wang et al. "An Improved Power-Added Efficiency 19-dBm Hybrid Envelope Elimination and Restoration Power Amplifier for 802.11g WLAN Applications". In: (Dec. 2006). ISSN: 0018-9480. DOI: 10.1109/TMTT.2006.885575 (cit. on p. 26).

[50] H. Wang and K. Sengupta. *RF and mm-wave power generation in silicon*. 2016. ISBN: 9780124095229 (cit. on pp. 10, 26, 27).

[51] K. Xu et al. "A 0.85mm2 51%-Efficient 11-dBm Compact DCO-DPA in 16-nm FinFET for Sub-Gigahertz IoT TX Using HD2 Self-Suppression and Pulling Mitigation". In: *IEEE Journal of Solid-State Circuits* 54.7 (2019), pp. 2028–2037. DOI: 10.1109/JSSC.2019.2906803 (cit. on p. 96).

[52] R. Yogi. "Design of a power amplifier for NB-IoT". In: (2018). URL: http://lup.lub.lu.se/student-papers/record/8957461 (cit. on p. 17).

[53] S.-M. Yoo et al. "A Class-G Switched-Capacitor RF Power Amplifier". In: (May 2013). ISSN: 0018-9200, 1558-173X. DOI: 10.1109/JSSC.2013.2252754 (cit. on p. 28).

[54] X. Yu et al. "A Sub-GHz low-power transceiver with PAPR-tolerant power amplifier for 802.11ah applications". In: *2015 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. 2015, pp. 231–234. DOI: 10.1109/RFIC.2015.7337747 (cit. on p. 96).

[55] W. Yuan et al. "A quadrature switched capacitor power amplifier in 65nm CMOS". In: May 2015. DOI: 10.1109/RFIC.2015.7337723 (cit. on p. 28).

[56] Li-Yuan Yang, Hsin-Shu Chen, and Y.-J. Chen. "A 2.4 GHz Fully Integrated Cascode-Cascade CMOS Doherty Power Amplifier". In: (Mar. 2008). ISSN: 1558-1764. DOI: 10.1109/LMWC.2008.916812 (cit. on p. 22).

A

Schematic

A.1 SCPA Schematic

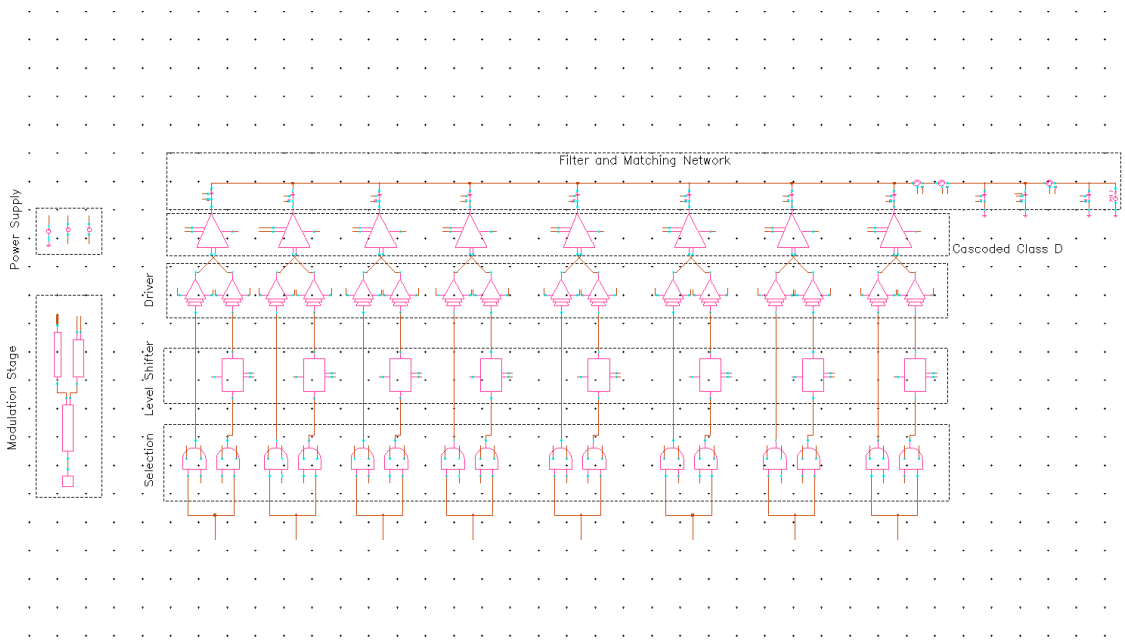


Figure A.1: Schematic of the top level SCPA and Modulation Stage. Note: The driver blocks present 2 views (using 3 inverters and using 1 inverter), along with the selection logic block (using the AND gate and using the transmission gate).

B

VerilogA

B.1 Modulation Stage VerilogA

B.1.1 16 QAM Processor

```
// VerilogA for Tese, Processor, veriloga

`include "constants.vams"
`include "disciplines.vams"

module processor16QAM_cell(symbol_id, phi, nPAs);

input symbol_id;
electrical symbol_id;
output phi;
output nPAs;
electrical phi, nPAs;
parameter real carrier_frequency = 0.9G;
parameter real avcc = 1.2;
parameter real avss = 0;
real phase = 0;
real envelope = 1;
real carrier_wave;

analog begin
        case (floor(V(symbol_id)))
                1: begin phase=2.3562; V(nPAs)<+0.008; end
                2: begin phase=2.8198; V(nPAs)<+0.006; end
                3: begin phase=-2.3562; V(nPAs)<+0.008; end
                4: begin phase=-2.8198; V(nPAs)<+0.006; end
                5: begin phase=1.8925; V(nPAs)<+0.006; end
                6: begin phase=2.3562; V(nPAs)<+0.003; end
                7: begin phase=-1.8925; V(nPAs)<+0.006; end
                8: begin phase=-2.3562; V(nPAs)<+0.003; end
                9: begin phase=0.7854; V(nPAs)<+0.008; end
                10: begin phase=0.3218; V(nPAs)<+0.006; end
                11: begin phase=-0.7854; V(nPAs)<+0.008; end
                12: begin phase=-0.3218; V(nPAs)<+0.006; end
                13: begin phase=1.2490; V(nPAs)<+0.006; end
```

```
                14: begin phase=0.7854; V(nPAs)<+0.003; end
                15: begin phase=-1.2490; V(nPAs)<+0.006; end
                16: begin phase=-0.7854; V(nPAs)<+0.003; end
            default: begin phase=999; V(nPAs)<+0.999; end
        endcase

        carrier_wave = cos(2*`M_PI*carrier_frequency*$abstime + phase);
        if (carrier_wave > 0)
                V(phi) <+ avcc;
        else
                V(phi) <+ avss;
        $bound_step(0.8/(carrier_frequency*40));
end
endmodule
```

B.1.2 Decoder

```
// VerilogA for Tese, Decoder, veriloga

`include "constants.vams"
`include "disciplines.vams"

module decoder_cell(nPAs, Bit);

input nPAs;
electrical  nPAs;
output [7:0] Bit;
electrical [7:0] Bit;

parameter real avcc = 1.2;

    analog begin
        case(floor(V(nPAs)*1000))
                1: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ 0; V(Bit[2]) <+ 0;
                        V(Bit[3]) <+ 0; V(Bit[4]) <+ 0; V(Bit[5]) <+ 0;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                2: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ 0;
                        V(Bit[3]) <+ 0; V(Bit[4]) <+ 0; V(Bit[5]) <+ 0;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                3: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
                        V(Bit[3]) <+ 0; V(Bit[4]) <+ 0; V(Bit[5]) <+ 0;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                4: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
                        V(Bit[3]) <+ avcc; V(Bit[4]) <+ 0; V(Bit[5]) <+ 0;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                5: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
                        V(Bit[3]) <+ avcc; V(Bit[4]) <+ avcc; V(Bit[5]) <+ 0;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                6: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
                        V(Bit[3]) <+ avcc; V(Bit[4]) <+ avcc; V(Bit[5]) <+ avcc;
                        V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
                7: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
```

```
                        V(Bit[3]) <+ avcc; V(Bit[4]) <+ avcc; V(Bit[5]) <+ avcc;
                        V(Bit[6]) <+ avcc; V(Bit[7]) <+ 0; end
                8: begin V(Bit[0]) <+ avcc; V(Bit[1]) <+ avcc; V(Bit[2]) <+ avcc;
                        V(Bit[3]) <+ avcc; V(Bit[4]) <+ avcc; V(Bit[5]) <+ avcc;
                        V(Bit[6]) <+ avcc; V(Bit[7]) <+ avcc; end
        default: begin V(Bit[0]) <+ 0; V(Bit[1]) <+ 0; V(Bit[2]) <+ 0;
                       V(Bit[3]) <+ 0; V(Bit[4]) <+ 0; V(Bit[5]) <+ 0;
                       V(Bit[6]) <+ 0; V(Bit[7]) <+ 0; end
        endcase
    end

endmodule
```

B.1.3 Non Overlapping Clocks

```
// VerilogA for Tese, NonOverlapPhase, veriloga

`include "constants.vams"
`include "disciplines.vams"

(*ignore_hidden_state*) module non_overlap_phase_cell(phi, phi_n, phi_p);

electrical phi, phi_n, phi_p;
input phi;
output phi_n;
output phi_p;

parameter real Vhigh = 1.2;
parameter real Vlow = 0;
parameter real Vtrans = 0.6;
parameter time_guard = 100f; //5p

real outN = Vlow;
real outP = Vlow;
real delayed_phi = 0;

analog begin
        delayed_phi = transition( V(phi), time_guard );

        @( cross( V(phi) -Vtrans, +1)) begin
            outP = Vhigh;
        end
        @( cross( V(phi) -Vtrans, -1)) begin
                outN = Vlow;
        end
        @( cross( delayed_phi -Vtrans, +1)) begin
                outN = Vhigh;
        end
        @( cross( delayed_phi -Vtrans, -1)) begin
                outP = Vlow;
        end
```

```
    V(phi_n) <+ transition( outN, 0, 2.5p ); //2.7778p
    V(phi_p) <+ transition( outP, 0, 2.5p );
end
endmodule
```

C

System Simulation

C.1 Final Simulation Of The Overall System

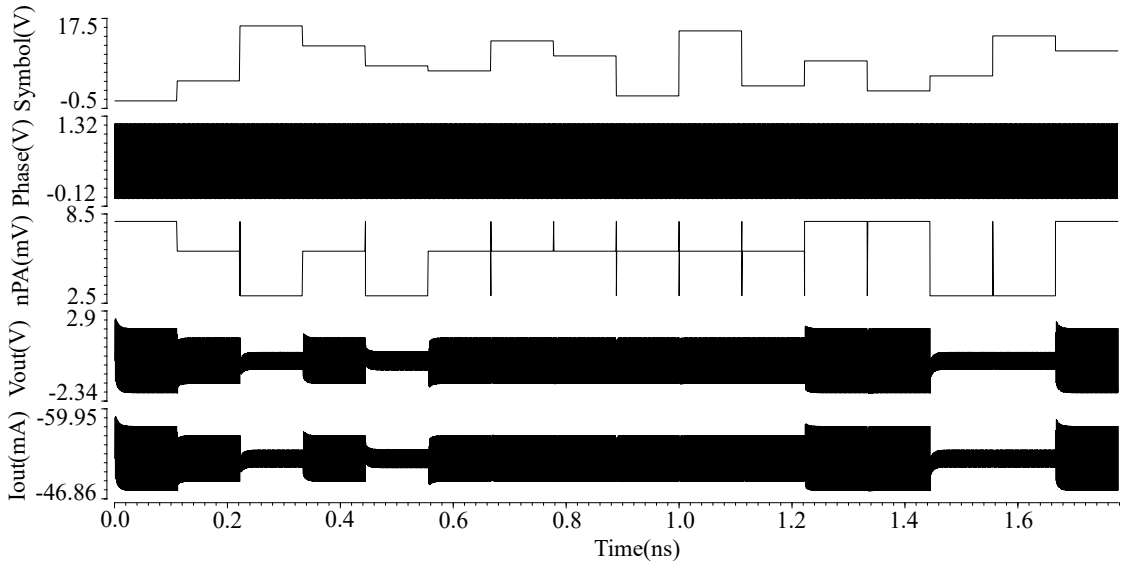


Figure C.1: Example of the transmission of the 16 QAM symbols.