UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Ph.D. Thesis

A Fast Engineering Approach to High Efficiency Power Amplifier Linearization for Avionics Applications

Author: Teng Wang

Advisor: Dr. Pere L. Gilabert

Components and Systems for Communications Research Group
Department of Signal Theory and Communications
Universitat Politècnica de Catalunya

Barcelona, October 2020

To Jesus Christ my LORD...

Behold, the heaven and the heaven of heavens is the Lord's thy God, the earth also, with all that therein is.

Deuteronomy 10:14 KJV

Abstract

This Ph.D. thesis provides a fast engineering approach to the design of digital predistortion (DPD) linearizers from several perspectives: i) enhancing the off-line training performance of open-loop DPD, ii) providing robustness and reducing the computational complexity of the parameters identification subsystem and, iii) importing machine learning techniques to favor the automatic tuning of power amplifiers (PAs) and DPD linearizers with several free-parameters to maximize power efficiency while meeting the linearity specifications.

One of the essential parts of unmanned aerial vehicles (UAV) is the avionics, being the radio control one of the earliest avionics present in the UAV. Unlike the control signal, for transferring user data (such as images, video, etc.) real-time from the drone to the ground station, large transmission rates are required. The PA is a key element in the transmitter chain to guarantee the data transmission (video, photo, etc.) over a long range from the ground station. The more linear output power, the better the coverage or alternatively, with the same coverage, better SNR allows the use of high-order modulation schemes and thus higher transmission rates are achieved. In the context of UAV wireless communications, the power consumption, size and weight of the payload is of significant importance.

Therefore, the PA design has to take into account the compromise among bandwidth, output power, linearity and power efficiency (very critical in battery-supplied devices). The PA can be designed to maximize its power efficiency or its linearity, but not both. Therefore, a way to deal with this inherent trade-off is to design high efficient amplification topologies and let the PA linearizers take care of the linearity requirements. Among the linearizers, DPD linearization is the preferred solution to both academia and industry, for its high flexibility and linearization performance. In order to save as many computational and power resources as possible, the implementation of an open-loop DPD results a very attractive solution for UAV applications.

This thesis contributes to the PA linearization, especially on off-line training for open-loop DPD, by presenting two different methods for reducing the design and operating costs of an open-loop DPD, based on the analysis of the DPD function.

The first method focuses on the input domain analysis, proposing mesh-selecting (MeS) methods to accurately select the proper samples for a computationally efficient DPD parameter estimation. Focusing in the MeS method with better performance, the memory I-Q MeS method is combined with feature extraction dimensionality reduction technique to allow a computational complexity reduction in the identification subsystem by a factor of 65, in comparison to using the classical QR-LS solver and consecutive samples selection. In addition, the memory I-Q MeS method has been proved to be of crucial interest when training artificial neural networks (ANN) for DPD purposes, by significantly reducing the ANN training time.

The second method involves the use of machine learning techniques in the DPD design procedure to enlarge the capacity of the DPD algorithm when considering a high number of free parameters to tune. On the one hand, the adaLIPO global optimization algorithm is used to find the best parameter configuration of a generalized memory polynomial behavioral model for DPD. On the other hand, a methodology to conduct a global optimization search is proposed to find the optimum values of a set of key circuit and system level parameters, that properly combined with DPD linearization and crest factor reduction techniques, can exploit at best dual-input PAs in terms of maximizing power efficiency along wide bandwidths while being compliant with the linearity specifications.

The advantages of these proposed techniques have been validated through experimental tests and the obtained results are analyzed and discussed along this thesis.

Resum

Aquesta tesi doctoral proporciona unes pautes per al disseny de linealitzadors basats en predistorsió digital (DPD) des de diverses perspectives: i) millorar el rendiment del DPD en llaç obert, ii) proporcionar robustesa i reduir la complexitat computacional del subsistema d'identificació de paràmetres i, iii) incorporació de tècniques d'aprenentatge automàtic per afavorir l'autoajustament d'amplificadors de potència (PAs) i linealitzadors DPD amb diversos graus de llibertat per poder maximitzar l'eficiència energètica i al mateix temps acomplir amb les especificacions de linealitat.

Una de les parts essencials dels vehicles aeris no tripulats (UAV) és l'aviònica, sent el radiocontrol un dels primers sistemes presents als UAV. Per transferir dades d'usuari (com ara imatges, vídeo, etc.) en temps real des del dron a l'estació terrestre, es requereixen taxes de transmissió grans. El PA és un element clau de la cadena del transmissor per poder garantir la transmissió de dades a grans distàncies de l'estació terrestre. A major potència de sortida, més cobertura o, alternativament, amb la mateixa cobertura, millor relació senyal-soroll (SNR) la qual cosa permet l'ús d'esquemes de modulació d'ordres superiors i, per tant, aconseguir velocitats de transmissió més altes. En el context de les comunicacions sense fils en UAVs, el consum de potència, la mida i el pes de la càrrega útil són de vital importància.

Per tant, el disseny del PA ha de tenir en compte el compromís entre ample de banda, potència de sortida, linealitat i eficiència energètica (molt crític en dispositius alimentats amb bateries). El PA es pot dissenyar per maximitzar la seva eficiència energètica o la seva linealitat, però no totes dues. Per tant, per afrontar aquest compromís s'utilitzen topologies amplificadores d'alta eficiència i es deixa que el linealitzador s'encarregui de garantir els nivells necessaris de linealitat. Entre els linealitzadors, la linealització DPD és la solució preferida tant per al món acadèmic com per a la indústria, per la seva alta flexibilitat i rendiment. Per tal d'estalviar tant recursos computacionals com consum de potència, la implementació d'un DPD en llaç obert resulta una solució molt atractiva per a les aplicacions UAV.

Aquesta tesi contribueix a la linealització del PA, especialment a l'entrenament fora de línia de linealitzadors DPD en llaç obert, presentant dos mètodes diferents per reduir el cost computacional i augmentar la fiabilitat dels DPDs en llaç obert.

El primer mètode se centra en l'anàlisi de l'estadística del senyal d'entrada, proposant mè todes de selecció de malla (MeS) per seleccionar les mostres més significatives per a una estimació computacionalment eficient dels paràmetres del DPD. El mètode proposat IQ MeS amb memòria es pot combinar amb tècniques de reducció del model del DPD i d'aquesta manera poder aconseguir una reducció de la complexitat computacional en el subsistema d'identificació per un factor de 65, en comparació amb l'ús de l'algoritme clàssic QR-LS i selecció de mostres d'entrenament consecutives.

El segon mètode consisteix en l'ús de tècniques d'aprenentatge automàtic pel disseny del DPD quan es considera un gran nombre de graus de llibertat (paràmetres) per sintonitzar. D'una banda, l'algorisme d'optimització global adaLIPO s'utilitza per trobar la millor configuració de paràmetres d'un model polinomial amb memòria generalitzat per a DPD. D'altra banda, es proposa una estratègia per l'optimització global d'un conjunt de paràmetres clau per al disseny a nivell de circuit i sistema, que combinats amb linealització DPD i les tècniques de reducció del factor de cresta, poden maximitzar l'eficiència de PAs d'entrada dual de gran ample de banda, alhora que compleixen les especificacions de linealitat.

Els avantatges d'aquestes tècniques proposades s'han validat mitjançant proves experimentals i els resultats obtinguts s'analitzen i es discuteixen al llarg d'aquesta tesi.

Acknowledgements

I would like to express my gratitude to all those who helped me during the writing of this thesis. The work is not achievable if without the help from any of you. My deepest gratitude goes first and foremost to Professor Pere Lluis Gilabert, my supervisor. I will never forget the patronage of him, who brought me the defining moment of my life at the Christmas 2013. Without his supports and trust, there will never be any progress of my research. He helped me on writing and polishing this dissertation with extraordinary patience and attentiveness, which is the indispensable condition to have this progress.

Second, I would like to express my heartfelt gratitude to all my colleagues in our CSC Group. To my tutor Professor Gabriel Montoro, who provided me continuous and reliably support during every experimental procedure. To my competition partner and sincere friend David Lopez, I will never forget the amazing journal in 2015. To my classmate Thi Quynh Anh Pham, it is a great memory to co-work with you on digital predistortion (DPD) research. Especially, I would like to thank Estefanía Guillena for testing the memory I-Q mesh-selecting method proposed in this thesis to train artificial neural networks (ANN) for DPD linearization. Part of the results obtained are included in Section 5.5.4 of this thesis. Her outstanding work on ANN DPD linearization can be found on her graduate thesis [Gui20]. Last but not least, Lee Tao, thanks for your efforts during the battle of internet test and go further in the future.

I would like to also thank Professor José Angel García from Universidad de Cantabria and Professor Roberto Quaglia from Cardiff University for their support, interesting discussions on power amplifier design and the donation of class-J and LMBA power amplifiers, respectively, used in our UPC test-benches to obtain the experimental results used in this thesis to validate the proposed techniques.

May my thanks go to my beloved family for their loving considerations and great confidence in me all through these years. I also owe my sincere gratitude to my friends and my fellow colleagues who gave me their help and stand with me no matter it's light or dark.

Finally, thank you Jesus Christ my Lord. I have nothing in-front of you, but you gifted me everything I have.

Teng Wang

Canton, October 2020.

Contents

Acronyms

ABS	Absolute value.
ACLR	Adjacent Channel Leakage power Ratio.
ACEPR	Adjacent Channel Error Power Ratio.
ACPR	Adjacent Channel Power Ratio.
adaLIPO	Adaptive Lipschitz Optimisation.
ADC	Analog-to-Digital Converter.
AI	Artificial Intelligence.
AICA	Artificial Intelligence Core Algorithm.
AM-AM	Amplitude Modulation to Amplitude Modulation conversion.
AM-PM	Amplitude Modulation to Phase Modulation conversion.
ANN	Artificial Neural Network.
APD	Analog Predistortion.
AWG	Arbitrary Waveform Generator.
BW	Bandwidth.
BER	Bit Error Rate.
BO	Back-off.
BPA	Balanced Power Amplifier.
BPC	Basic Predistorter Cell (or Basic Predistortion Cell).
BPSK	Binary Phase Shift Keying.
CA	Carrier Aggregation.
CALLUM	Combined Analog-Locked Loop Universal Modulator.
CCA	Canonical Correlation Analysis.
CFB	Cartesian Feedback.
CFR	Crest Factor Reduction.
CoS	Consecutive Samples.
CSP	Control Signal Power.
CW	Continuous Wave.
DAC	Digital-to-Analog Converter.

DC Direct Current.
DPD Digital Predistortion (or Digital Predistorter).
DPLS Dynamic Partial Least Squares.
DSO Digital Storage Oscilloscope.
DSP Digital Signal Processor (or Digital Signal Processing).
DUT Device Under Test.
etc. etcetera.
EE&R Envelope Elimination and Restoration.
ETSI European Telecommunications Standards Institute.
EVM Error Vector Magnitude.
FET Field-Effect Transistor.
FFT Fast Fourier Transform.
FIR Finite Impulse Response.
FOM Figure Of Merit.
FPGA Field-Programmable Gate Array.
GaN Gallium Nitrate.
GMP Generalized Memory Polynomial.
GPU Graph Process Unit.
HD Harmonic Distortion.
HDC Hybrid Digitally Controlled.
I In-phase.
IBO Input Back-off.
IEEE Institute of Electrical and Electronics Engineers.
IF Intermediate Frequency.
IFFT Inverse Fast Fourier Transform.
IIR Infinite Impulse Response.
IMD Intermodulation Distortion.
IMP Inter-modulation Products.
LASSO Least Absolute Shrinkage and Selection Operator.
LINC Linear amplification using Nonlinear Components.
LMBA Load Modulated Balanced Amplifier.
LMS Least Mean Squares.
LO Local Oscillator.
LOS Line Of Sight.
LS Least Squares.
LTE Long-Term Evolution.

LUT	Look-Up Table.
MEMS	Microelectromechanical Systems.
MeS	Mesh Selecting.
MIMO	Multiple-Input-Multiple-Output.
MISO	Multiple-Input-Single-Output.
MOSFET	Metal Oxide Semiconductor Field-Effect Transistor.
MP	Memory Polynomial.
MR	Matrix Ratio.
MSE	Mean Square Error.
NMA	Nonlinear Moving Average.
NMSE	Normalized Mean Square Error.
NN	Neural Network.
NR	New Radio.
OFDM	Orthogonal Frequency Division Multiplexing.
OMP	Orthogonal Matching Pursuit.
OP	Offset Percentage.
PA	Power Amplifier.
PAE	Power Added Efficiency.
PAPR	Peak-to-Average Power Ratio.
PCA	Principal Component Analysis.
PCB	Printed Circuit Board.
PD	Predistortion.
pdf	probability density function.
PLS	Partial Least Squares.
PHY	Physical layer.
Q	Quadrature.
QAM	Quadrature Amplitude Modulation.
QPSK	Quadrature Phase Shift Keying.
QR-RLS	QR factorization Recursive Least Squares.
RAM	Random Access Memory.
RC	Radio Control.
RDF	Reduction Factor.
RF	Radio Frequency.
rms	root mean square.
Rx.	Receiver.
SA	Simulated Annealing.

SBL	Sparse Bayesian Learning.
SC	Single Carrier.
SDR	Software-Defined Radio.
SISO	Single-Input-Single-Output.
SMA	Sub-Miniature version A.
SMD	Surface Mount Device.
SNR	Signal to Noise Ratio.
SVD	Singular Value Decomposition.
SSPA	Solid State Power Amplifier.
TDNN	Time-Delayed Neural Network.
Tx.	Transmitter.
UAV	Unmanned Aerial Vehicle.
USR	Under-Sampling Ratio.
TWTA	Traveling Wave Tube Amplifiers .
vs.	versus.
VSA	Vector Signal Analyzer.
VTB	Video Transfer Block.
W-CDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network.

Chapter 1

Introduction

1.1 Motivation

Since the time that the first airplane was first invented at the generation of Wright and Curtis until now, the airplane has gained importance as a general aviation platform. In the commercial area, airplanes usually work in two scenarios: to carry the cargo (or human, as well) to other places, and to carry the airborne electronic devices, which is also called avionics, to given places in the sky. The benefits of being carried into a high altitude boosts the performance of most electronic devices, especially in the area of aerial photography, weather forecast, typhoon observation and also emergency mobile station.

Figure 1.1: Facebook Aquila internet relay solar UAV.

With the development of aviation technology, avionics systems play an increasingly important role. For example, the most accepted concept of a fifth-generation fighter contains five elements [BRI]: stealth, high maneuverability, advanced avionics, networked data fusion from sensors and avionics, multi-role capabilities. Despite the literal word 'avionics', three of those five elements are directly determined by the performance of electronics devices on the plane.

Similarly, it shares the same trend of the unmanned aerial vehicle (UAV). At the beginning

Figure 1.2: The WV-3 Warning Star storm observe [WV2].

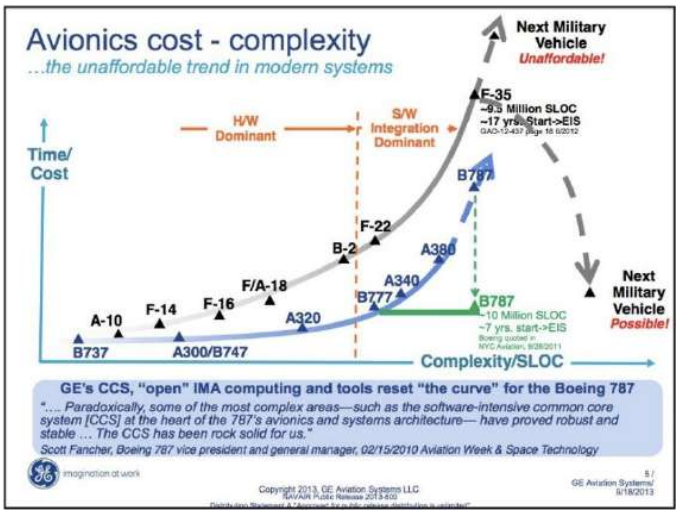

Figure 1.3: The trend of avionics cost published by General Electric [GE].

age, UAV was literally only an unmanned plane for some special usage, with unacceptable price and operation cost for commercial market. Advances in electronics have made drone control systems less expensive and more accurate, enabling drones to be adapted to increasingly complex application scenarios. With the rapid development of electronic sensors and communication systems, the market for civilian drones has begun to grow explosively in the 21^{st} century.

Technically, the vehicle itself of unmanned aerial vehicle has never been a barrier to its development in the whole history. The avionics took this job. During the early history of UAV, the limitation of the electronics was the responsible for several critical problems, e.g. the accuracy of inertial navigation system, the body size of radio devices and also the performance of the motor and battery. During most of the 20^{th} century, the UAV needed to carry a big electronic cargo in case to maintain the flight ability, which may contain inertial navigation devices, a VHF communication station, and analog controlled actuator and maybe a mechanical scanned radar.

The drone was born at a time that the technology of electronics was too well developed to lead an avionic revolution. The microelectromechanical systems (MEMS) sensor replaced most of the traditional devices with a body size that is even hard to observe; the microwave and

Figure 1.4: One of the earliest UAV of the world, designed by Charles Franklin Kettering [UAV].

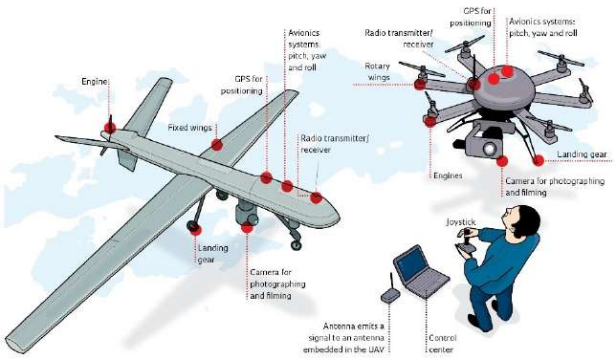

Figure 1.5: Typical avionics on UAV [RFB].

its semiconductors are commercially available; the first layer of Long-Term Evolution (LTE) standards can be easily copied and implemented by small communication companies; and also the promotion of motor and battery makes it easy to build a small power block. After 2011, the drone have been known and used by people for daily work and even leisure.

Even if we often regard UAV as a product of the aerospace industrial market, in reality, the truly technical core of the modern drone is the avionics. In comparison to the mechanical body, which works as the flight platform, the electronic sub-system of the drone provides its brain, eyes and hands.

One of the earliest avionics present in the UAV is the radio control (RC) and it is always the most irreplaceable part of an UAV. With the development of radio frequency (RF) technology, the control system also grew from the single radio command mode to more complex and flexible controlled modes. Modern control systems of UAV are usually based on a flight planning platform integrated with an auto-navigation system. The inertial and radio navigation block supports the UAV flies on a given path and velocity. The flight plan can be pre-stored into the UAV before take-off or it can be transferred to the UAV via wireless communication in real-time. The ground

Figure 1.6: A famous concept of follow-and-record of modern aerial photography drone [TDO].

station will receive lots of information from UAV during the mission procedure, e.g. current location, photographs or videos. The video information is usually the most attractive part as it can allow the operator to do more work remotely on the ground in real-time.

After 2010, a boom of commercial aerial photography drones appears on the market. The DJI Co., Ltd. is one of the most famous company in this field. The typical aerial photography drone (see Fig. 1.7 for example) use a video transfer block (VTB) to send the real time video to the control device, thus the owner can handle it to take a photo or video from above. Following this trend, the VTB became an important part of the whole UAV system, not only to support the main function of aerial photography drone, but also for industrial usage like, for example, electronic power line patrolling, as depicted in Fig. 1.8.

Figure 1.7: DJI Aerial Photography Drone [DJI].

Unlike control signal, transferring videos from the drone to the ground station requires big data rates. This becomes the main challenge for the whole RF system. The simplest VTB is an analog modulator - demodulator with RF front-end. It offers a data rate of 300-400 Kbps thus the video quality is lower than 320p, which cannot meet the requirement of most video-demanding applications. The original 1024p quality video has a data rate of 500 Mbps, which is even hard for CAT5 cable to transfer. The modern video compression technology, like H.264, offers a data compress ratio of 80, but it takes time for calculation. The VTB for commercial usage transfer the H.264 video with 200 ms calculation delay on a data rate around 7 Mbps. As

Figure 1.8: The drone operates for Electric power line patrol [WET].

far as the RF front-end is concerned, a wide-band power amplifier (PA) is also required at the final stage of VTB. In addition, the PA should be linear because there is mostly no linearization technology embedded on the VTB.

The transfer range of VTB is also an attractive performance, especially for industrial usage. Due to the requirement of signal bandwidth, the VTB usually works on the L and S band (1 − 6 GHz). Thus the transfer loss of atmosphere is more significant than at lower frequencies. A high PA is required when the aerial photography drone is designed to work over a long range from the ground station. Combining requirements of bandwidth and linearity, a class A linear PA often appears on VTBs, which is featured with low efficiency and massive heat emission.

As the final stage of the data flow between the drone and ground station, the PA design has to take into account the compromise among bandwidth, in-band and out-of-band linearity (i.e., evaluated in terms of error vector magnitude (EVM) or adjacent channel power ratio (ACPR)) and power efficiency. The classic straight forward design topology is to use high linearity PA without focusing on its power efficiency. Although comparing with the propulsion power of drones, the heat emitting of PAs will not impact the flight endurance of large size UAVs, but it becomes significant because it decreases UAVs' take-off weight. The high emitting heat of high power amplifier also affects other electric systems since the small size UAVs usually do not have enough space for complex cooling system.

The last generation of industrial video transfer takes advantage of 3rd Generation Partnership Project (3GPP) Long-Term Evolution (LTE) techniques. The data rate increased significantly by using advanced coding and modulation methods. The Orthogonal Frequency Division Multiplexing (OFDM) is one of the most efficient way to increase the physical layer performance, e.g. high spectral efficiency, robustness against interference and facilitates single frequency networks. However, the use of OFDM-based signal aggravates the inherent linearity and efficiency trade-off in PAs, due to the high peak-to-average power ratios (PAPR) that present that kind of waveforms. The linearity challenges become even worse when facing the next generation com-

munication techniques, e.g., 5G New Radio (NR), with wider bandwidths at higher frequencies but with similar high PAPRs. Ensuring that the whole system operates with certain linearity levels (required to guarantee a certain quality of service), has become one of the most challenging objectives in system design, since it cannot longer be solved by simply sacrificing the PA's power efficiency for linearity.

New design amplification topologies tend to involve specialized linearization methods to improve the linearity of the power amplifier. High efficient amplification architectures based on dynamic supply or dynamic load modulation techniques have been proposed to be implemented in the modern communication systems. Among other linearization methods such as feedback or feed-forward, digital predistortion (DPD) has been widely recognized as an effective and flexible approach to cope with the increasing demands of linearity in PAs.

The principle of DPD linearization is straightforward, consists in preceding the PA with a subsystem (nonlinear function in a digital signal processor) that counteracts to the nonlinear characteristic of the PA. Therefore, a lot of research has been devoted to find mathematical models capable to characterize and compensate for the PA nonlinear behavior and its dynamics [Kat16]. Fig. 1.9 shows the block diagram of an adaptive DPD. Operating the identification system in real-time increases the system complexity by involving a feedback observation path to acquire the PA output characteristics in real-time [Woo17]. From a high level view, a communication system with adaptive DPD means that it has one more receiver than normal systems. The feedback observation path provides real-time awareness of the PA non-linear behavior, which allows the DPD algorithm to continuously adapt to any time-variant scenario. Thus, adaptive DPD shows high robustness and high reliability. The DPD adaptation process, however, can be designed to follow different strategies depending on the timing constraints imposed by the changes in the PA behaviour. Therefore, real-time adaptation will be necessary when the PA behaviour changes very fast, while non real-time adaption can be considered when the PA changes in a more relaxed time scales, for example, due to temperature or aging.

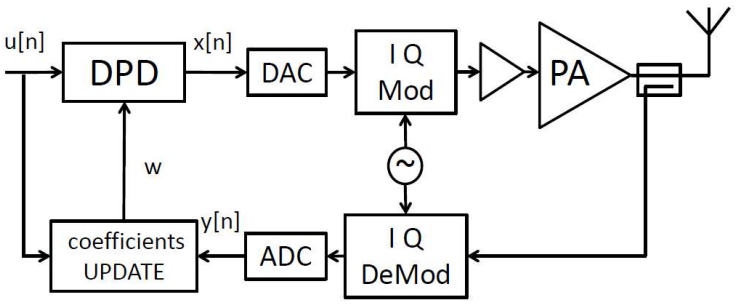

Figure 1.9: Block of typical adaptive digital predistortion.

Focusing again in the linearity requirements of the UAV video transmitting communication

system, in order to save as many computational and power resources as possible, the implementation of an open-loop DPD results a very attractive solution. With an open-loop DPD we can save the additional circuitry required in the feedback or observation path (i.e., downconverters, ADC) and the computational complexity associated with the parameter's estimation. However, the price to pay is the loss of robustness and flexibility to adapt to a time-variant scenario. The considered operating scenario is pretty time-invariant contemplating line of sight (LOS) UAV transmissions. In this context, open-loop DPD can be considered as a suitable solution and the linearization performance will rely on the accuracy of the off-line (factory) training.

Therefore, one of the main purposes of this dissertation, is to provide a series of methodologies for off-line DPD training oriented at improving the performance and robustness of open-loop DPD and reducing the computational complexity of the identification/adaptation subsystem. In addition, taking into account the vast amount of literature dedicated to propose single-input single-output (SISO) behavioral models [Sch09], or multiple-input single-output (MISO) behavioral models [Gil19], in this dissertation, a machine learning based method is proposed and tested aiming to simplify the DPD design procedure in a scenario with several free-parameters to tune.

1.2 Outline of the Thesis

In the general context of designing open-loop DPD linearizers for high efficient but non-linear power amplifiers for UAV communications, in this thesis we propose two different techniques for reducing the design and operating costs in practical DPD designs based on the analysis of the DPD function. The first method focuses on the input domain analysis, aiming to provide an efficient way to reduce the computational of the DPD adaptation function and enhance the robustness of the estimation by proposing a mesh-selection technique to select the most relevant training samples for the DPD parameter extraction. The second method involves the use of machine learning techniques in the DPD design procedure, to enlarge the capacity of the DPD algorithm and to maximize the power efficiency in dual-input PAs when several free-parameters need to be properly configured.

The main body of this Thesis is schematically depicted in Fig.1.10, and it is organized as follows.

First, Chapter 1 presents the motivation of the work, the outline of the dissertation and a list of research publications related to the work developed by the PhD candidate.

Chapter 2 presents an overview of high efficient PA topologies and summarizes the linearity versus efficiency trade-off in PAs.

Chapter 3 includes the state-of-the art in DPD and presents some design challenges regarding

the open-loop DPD. The main objectives pursued in this thesis will be detailed in this chapter.

Chapter 4 discusses the effects of applying under-sampling techniques in the DPD adaptation subsystem and presents a selecting method, named mesh-selecting, to find a small subset of the original test signal for DPD coefficients extraction while keeping the same accuracy level. Experimental results will show the advantages in computational complexity reduction of the proposed memoryless mesh-selecting technique when combined with the principal component analysis (PCA) dimensionality reduction technique.

Chapter 5 presents a series of improved mesh-selecting methods to provide enhanced selecting performance. The reduction capabilities vs. accuracy degradation of the new mesh-selecting methods are evaluated and compared. Moreover, focusing on the method with better selecting performance (the memory I-Q mesh-selecting method), experimental results will show the computational complexity reduction achieved when this mesh-selecting method is combined with feature extraction dimensionality reduction techniques.

Unlike the two previous Chapters focused on the testing signal for DPD coefficients adaptation, in Chapter 6 and Chapter 7 we propose a machine learning based DPD tuning method to assist the DPD design procedure. A global optimization algorithm is first applied to find the optimal configuration of parameters of a Generalized Memory Polynomial (GMP) PA behavioral model to linearize a Class-J PA. Then, in Chapter 7 the optimization method is extended to operate in a scenario where the free-parameters to tune correspond to a dual-input load modulated balanced amplifier (LMBA). Experimental results, will show the capability of the proposed optimization method to find the best configuration of the LMBA from the power efficiency point of view while the linearity levels are guaranteed by means of DPD linerization.

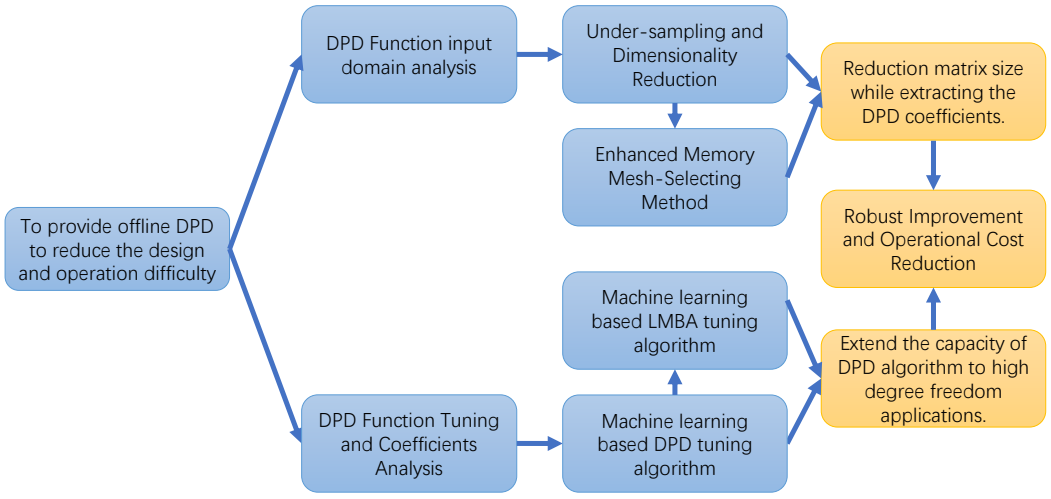

Figure 1.10: Overview of main chapters of this dissertation.

Finally, Chapter 8 gives the conclusion on the dissertation and discusses possible future research lines in the field of DPD input signal analysis and artificial intelligence applications.

1.3 Research Contributions

This Ph.D. dissertation contributes to the PA linearization, specially on off-line training for open-loop digital predistortion. As described in previous subsection, mesh-selecting methods are proposed to accurately select the proper samples for DPD coefficients extracting calculation and machine learning assist methods are presented to support DPD in scenarios with large free-parameters to be configured. The presented solutions reduce the computational and tuning complexity of the open-loop DPD linearization system and, at the same time, increase its robustness. Experimental tests have been carried out to validate the performance of these techniques.

The research reported in this thesis has generated publications in international conferences and journal papers. The publications are listed in the following.

- **T. Wang**, P. L. Gilabert and G. Montoro, "Under-sampling effects and computational cost reduction in RF power amplifier behavioral modeling," *in Proc. 10th European Microwave Integrated Circuits Conference* (EuMIC), Paris, France, Sept. 2015, pp. 57-60.

- G. Montoro, **T. Wang**, M. N. Ruiz, J. A. García and P. L. Gilabert "Reducción de la frecuencia de muestreo en los conversores ADC y DAC usados en predistorsionadores digitales,"*in Proc. XXX Simposium Nacional de la Unión Científica Internacional de Radio* (URSI), Pamplona, Spain, Sept. 2015, pp. 1-4.

- D. López-Bueno, **T. Wang**, P. L. Gilabert and G. Montoro, "Amping Up, Saving Power: Digital Predistortion Linearization Strategies for Power Amplifiers Under Wideband 4G/5G Burst-Like Waveform Operation," *in IEEE Microwave Magazine*, vol. 17, no. 1, pp. 79-87, Jan. 2016.

- P. L. Gilabert, G. Montoro, **T. Wang**, M. N. Ruiz and J. A. García, "Comparison of model order reduction techniques for digital predistortion of power amplifiers," *in Proc. 46th European Microwave Conference* (EuMC), London, UK, Oct. 2016, pp. 182-185.

- Q. A. Pham, D. López-Bueno, **T. Wang**, G. Montoro, and P. L. Gilabert, "Multidimensional LUT-based digital predistorter for concurrent dual-band envelope tracking power amplifier linearization," *in Proc. IEEE Topical Conf. on RF/Microw. Power Amplifiers for Radio and Wireless Applications.* (PAWR), Anaheim, USA, Jan. 2018, pp. 47-50.

- Q. A. Pham, D. López-Bueno, **T. Wang**, G. Montoro, and P. L. Gilabert, "Partial least squares identification of multi look-up table digital predistorters for concurrent dual-band envelope tracking power amplifiers," *in IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 12, pp. 5143-5150, Dec. 2018.

- **T. Wang**, P. L. Gilabert, R. Quaglia, G. Montoro, "Machine-Learning Assisted Optimisation of Free-Parameters of a Dual-Input Power Amplifier for Wideband Applications" *Submitted for publication in IEEE Transactions on Microwave Theory and Techniques.* First submission rejected. Pending to be re-submitted in Nov. 2020.

- **T. Wang**, P. L. Gilabert, "Mesh Selecting for Computational Efficient Power Amplifier Behavioral Modeling and Digital Predistortion Linearization" *IEEE Microwave and Wireless Components Letters.* Accepted for publication in Oct. 2020 (to be published in late 2020 or early 2021).

1.4 Awards

- David López and **Teng Wang**. First prize in the *Power Amplifier (PA) Linearization Through Digital Pre-distortion (DPD) Linearization international competition.* The event took place in the frame of the Student Design Competitions organized by the IEEE Microwave Theory and Techniques Society (MTT-S) during the International Microwave Symposium (IMS) held in Phoenix, AZ on 17-22 May 2015.

Figure 1.11: First prize winners of the 2015 Power Amplifier (PA) Linearization Through Digital Predistortion (DPD) Linearization international competition.

Chapter 2

High Efficiency Power Amplifier Architectures and Digital Predistortion Linearization

2.1 High Efficiency Power Amplifier Architectures

The power amplifier (PA) is one of the most critical components present in the transmitter of both modern wireless and wired communications. In macro base-stations the PA is responsible for most of the direct current (DC) consumption of the transmitter. In handsets and small cells, the processing part has also a significant impact in the DC consumption but, in any case and independently on the size and power of the base-station, the PA is one of the main sources of signal distortion in the whole communications system. The simple way to avoid introducing nonlinear distortion is to operate the PA far from saturation, i.e., with significant power back-off levels. In order to prevent the peaks of the signal operate in the compressed region, the back-off level of operation has to be similar to the peak-to-average power ratio (PAPR) of the transmitted signal. This means that, when considering linear but inefficient Class-A or -AB PAs, for amplitude and phase modulated signals with high PAPR, linear amplification can be achieved only at the price of a significant power efficiency degradation, since the PA power efficiency rapidly decays when backing-off from compression.

System-level linearization approaches for PA, such as digital pre-distortion (DPD), can extend the linear range of the PA. Further more, by properly combining linearization with the crest factor reduction (CFR) techniques, we can drive the PA further into compression while meeting the linearity requirements [Lop14]. Therefore, DPD and CFR techniques are usually applied to overcome or at least mitigate the trade-off between efficiency and linearity of the whole power amplification sub-system. However, the power efficiency obtained by applying linearization techniques to PA operated as a controlled current source (e.g., Class A, B, AB) is

limited. To avoid wasting excessive power resources when processing high PAPR signals, the current source mode PA can be forced to follow its envelope (dynamic supply solutions), or a high-efficient switch mode amplification stage can be introduced. Among all the techniques targeting dynamic bias or load modulation, envelope tracking (ET) PA [Wan15b, Pop17, Wat18], Doherty PA [Pen16b, Dar16], load modulated balanced amplifiers (LMBA) [Qua18] and linear amplification with nonlinear components (LINC) or outphasing PA [Bar16, Pop18] are the most widely researched in literature. In either case, these efficient topologies require linearization techniques to ensure the degree of linearity specified in communication standards.

2.1.1 Class-E PA

PAs are classified into different classes according to the conduction angle of the drain current. Thus, the power transistor operation class is determined by the fraction of the RF cycle over which the power transistor conducts. Theoretically, 100% (conduction angle 2π) corresponds to class A, 50% (conduction angle π) to class B, between 50 % and 100 % to class AB, and finally less than 50 % to classes C, D, E and F. The performance trade-off among the different operation classes includes efficiency, linearity, power gain, signal bandwidth and output power [Raa02].

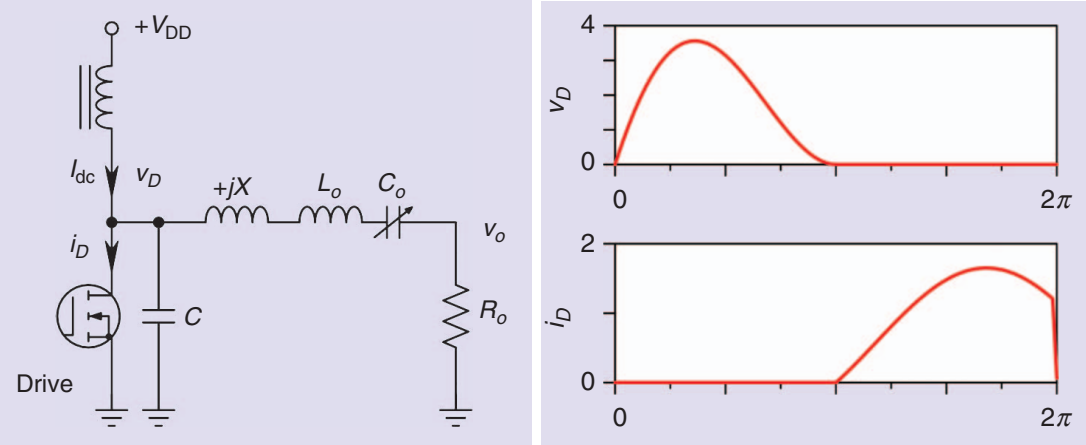

Figure 2.1: Class-E PA simplified circuit and waveforms [Med18].

Since Nathan Sokal's invention of the class-E PA, several designs have been reported in literature ranging from kHz and up to mm-wave [Pop18]. To achieve high efficiency, the transistor is driven to act as a switch, thus avoiding that high drain voltage and drain current appear at same time to reduce power loss. A typical configuration of class-E PA is showed in Fig. 2.1. As described in [Med18], the circuit in Fig. 2.1 has an RF choke for DC feed, drain-shunt capacitance, and a series-tuned output. The series-tuned output circuit passes the fundamental-frequency current but blocks harmonic currents. In contrast, a class-B or class-C amplifier has a parallel-tuned output circuit. In an ideal Class-E amplifier, the drain voltage drops to zero and has zero slope at the instant the transistor turns on as shown in Fig. 2.1. Because drain current

flows only when the drain voltage is zero, the drain voltage is nonzero only when there is no drain current. Consequently, if there is no discharge loss, the efficiency of an ideal class-E PA is 100%.

As it will be shown in the following subsections, an ideal class-E amplifier can be used in the design of high efficient amplifying topologies, such as, for example, the outphasing PA [Rui16] [Bel09], [Veg17].

2.1.2 Doherty PA

The Doherty PA, named after W.H. Doherty from Bell Labs [Doh36], is one of the most popular solutions for high efficiency amplification in cellular base stations because thanks to its particular configuration is able to accommodate signals with high PAPR whilst still retaining a good level of power efficiency. The Doherty configuration uses two amplifiers biased differently in order to carry out different functions. The carrier amplifier normally operates in a linear class of operation and provides gain at any power level. While, the peaking amplifier is only functioning when the carrier amplifier is getting towards its limits. Therefore, the peaking amplifier complements the carrier amplifier by providing the extra power capability that the carrier amplifier on its own cannot provide. As shown in Fig. 2.4, two separate power amplifier are assembled by quarter wave length transmission-line ($\lambda/4$ T-line). The most significant aspect in the operation of the Doherty amplifier is to activate the peaking amplifier only when it is required.

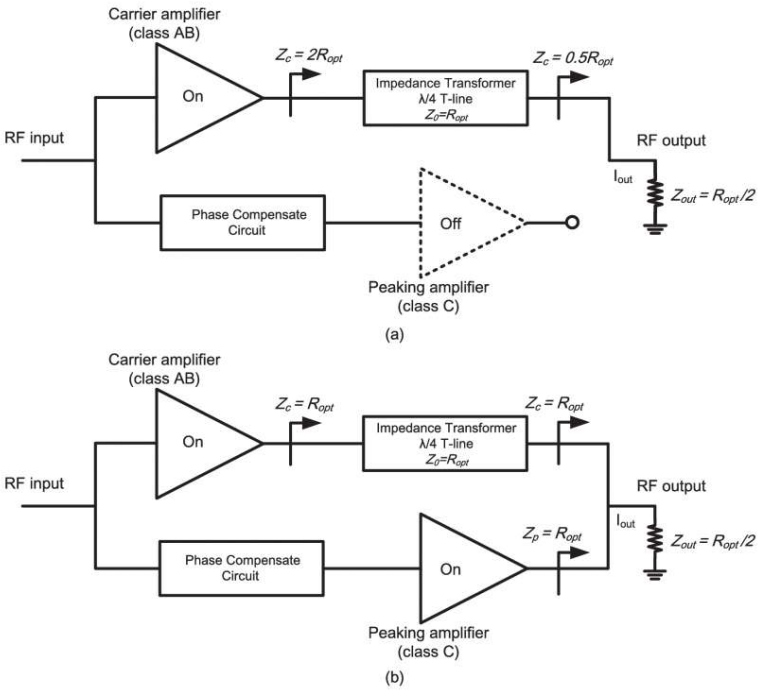

Figure 2.2: Classical topology of a Doherty PA [Zha17].

The output load of the Doherty PA is modulated by the current ratio between the carrier and the peaking amplifiers. Only the carrier PA is turned on during low-power operation, while the two PAs are turned on at high power. Therefore, following the dynamic load modulation concept, it achieves high efficiency at backed-off output power level as well as at the peak power. The carrier amplifier usually operates in Class-AB, while the peaking amplifier works in Class-C to boost the whole output power when the input is higher than 6 dB back-off [Zha17, Pen16b].

The Doherty PA suffers from nonlinear distortion, mainly due the gain compression that appear just before the peaking amplifier starts conducting. This phenomenon of double gain compression can be compensated using DPD linearization [Dar16].

2.1.3 Outphasing Power Amplifier

The outphasing modulation topology was first conceived by H. Chireix in 1935 [Chi35], and later revised and re-introduced by D. Cox in 1974 [Cox74] under the name of linear amplification with nonlinear components (LINC). Basically, the outphasing architecture is designed to amplify an amplitude-modulated signal by combining the amplification of two phase-modulated representations of the original input signal.

As show in Fig. 2.3, in the LINC transmitter the input signal is converted into two constant envelope signals. These two signals are amplified independently by two high-efficient switched-mode PAs in parallel branches. Because both signals have constant envelope, the two branch PAs can operate close to saturation. At the PA outputs, both signals are added in a two-to-one combiner obtaining a linear amplified replica of the input signal.

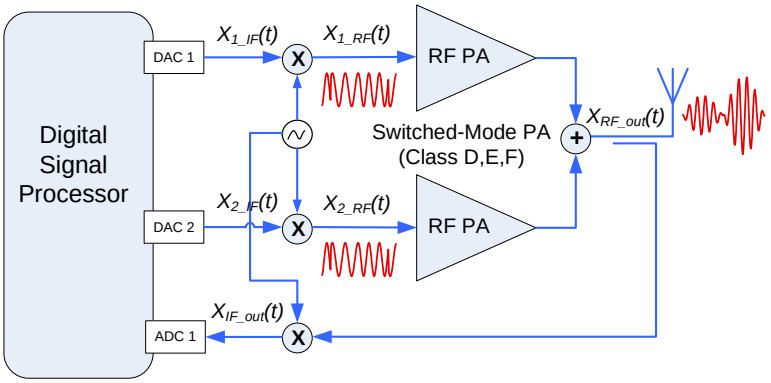

Figure 2.3: Simplified block diagram of a LINC amplifier [Gil11].

There are two popular approaches to the outphasing amplifier, the LINC and the Chireix outphasing. The main difference is the type of employed power combiner: isolated or non-isolated. It also affects to the system operation. The LINC combines the amplified constant envelope signals with an isolated combiner, which has the benefits of obtaining a linear amplified replica of the original signal at the output, but at the price of dissipating power at the combiner. Instead,

when considering the outphasing topology but with a Chireix (non-isolated) combiner, the power efficiency can be remarkable high but at the price of having severe nonlinear distortion. [El-12] [Rui16].

The linearization of a Chireix outphasing PAs in pure outphasing mode can be very challenging when the signal bandwidth is in the order of tenths of MHz. Hybrid outphasing operation modes have been proposed, which when properly combined with a DPD linearization can help to mitigate the linearity versus efficiency trade-off [Gil20].

2.1.4 Load Modulation Balanced Amplifier

The load modulated balanced amplifier (LMBA) was first introduced in [She16] by Shepphard et al., and it was based on a balanced PA where a control signal power injected at the isolated port of the output 90° coupler modulated the load at each balanced device.

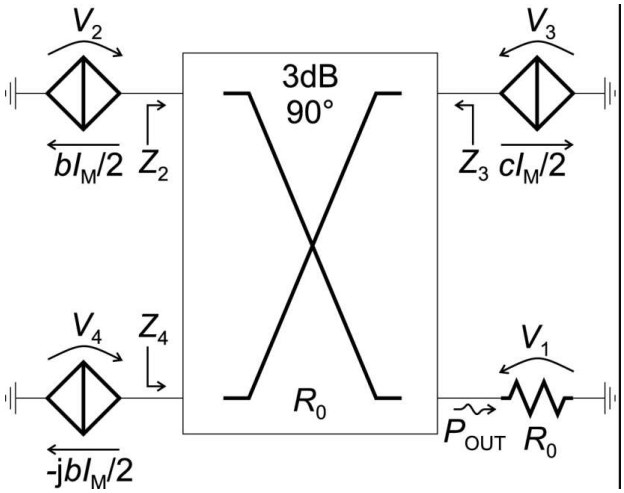

Figure 2.4: Classical topology of Load Modulated Balanced PA [Qua18].

As described in [Qua18], the key feature of the LMBA is that the control signal power can be fully recovered at the output of the LMBA. and thus it is always contributing positively to the total ouptput power no matter what load modulation it is imposing. In the Doherty PA instead, the auxiliary output phase determines the load modulation on the main, but its power is only fully recovered when the phase is aligned with the main. Therefore, unlike in the Doherty PA, in the LMBA the control signal power recovery is independent to the load modulation. Another key property of the LMBA is its wider bandwidth of operation, since the load modulation can be applied for the whole frequency band of the 90° coupler, that is normally larger than the bandwidth of a Doherty combiner.

The additional degrees of freedom offered by a dual-input LMBA can be used to optimize the performance on the same or larger bandwidth, or to improve other performance metrics such

as linearity and average efficiency. In Chapter 7 of this thesis, we will present an auto-tuning approach to take advantage of the possibilities given by the extra degrees of freedom in dual-input LMBAs to enhance the power efficiency figures along wide signal bandwidths and high PAPR values, while meeting the linearity requirements through DPD linearization.

2.1.5 Envelope Tracking

PA dynamic supply modulation techniques can achieve better power efficiency figures than the ones that can be achieved with fixed supply, even when using linearization techniques to extend the linear output power dynamic range.

One popular dynamic supply solution is envelope tracking (ET) [Wan15b], where the RF PA device is operating as a current source amplifying an amplitude and phase modulated signal while it is dynamically supplied, as shown in Fig. 2.5. The supply modulator plays a key role to determine the efficiency of the whole ET transmitter because it is defined as the product between the RF PA efficiency and the supply modulator power efficiency. One of the main limitations of ET is that the supply modulator has to efficiently supply the power required by the PA transistor at the speed (i.e. slew-rate) required by the signal's envelope. In ET the supply voltage of the PA is adjusted to save energy and thus the envelope tracking speed can be selected to trade-off the bandwidth/slew-rate requirements of the supply modulator (or envelope amplifier) and the overall ET power efficiency.

Some shaping strategies can be applied to the signal's envelope to reduce its bandwidth or slew rate [Mon10a], while still improving the overall power efficiency in comparison to fixed supply. However, the use of slew-rate or reduced bandwidth envelopes generates a particular nonlinear distortion. The slow envelope dependent nonlinear behaviour can be compensated by means of DPD linearization [Mon11, Gil12a].

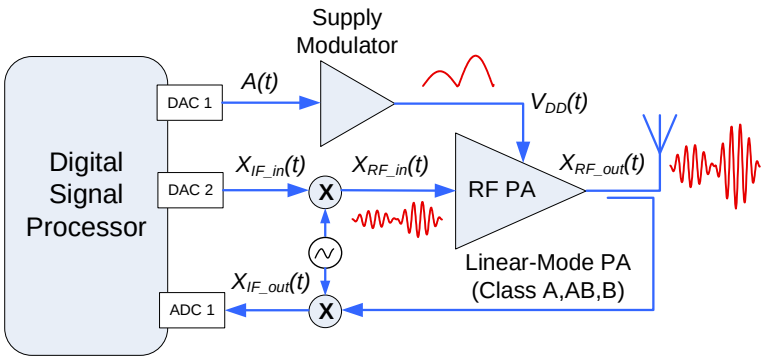

Figure 2.5: Block diagram of an envelope tracking PA [Gil11].

2.2 Predistortion Linearization Method

The aforementioned high power efficient amplification architectures based on dynamic load or dynamic supply modulations are oriented at maximizing the PA's power efficiency, however, in order to guarantee the linearity levels specified in the communications standards (e.g., in LTE we have to meet at least -45 dBc of adjacent channel power ration (ACPR)), linearization techniques, such as predistortion, have to be included. In the following subsections we will introduce predistortion linearization in both its analog and digital form.

2.2.1 Analog Predistortion

To compensate for the nonlinear distortion introduced by the inherent nonlinear behavior of the PA, an intuitive idea consists of compensating this non-ideal behavior by adding a nonlinear circuit that generates the inverse behavior working in series with the PA. The additional nonlinear circuit will first distort the input signal and then the PA, by amplifying the previously distorted signal will generate a linear amplified version at the output. The basic idea behind predistortion (independently if it is analog or digital) is depicted in Fig. 2.6.

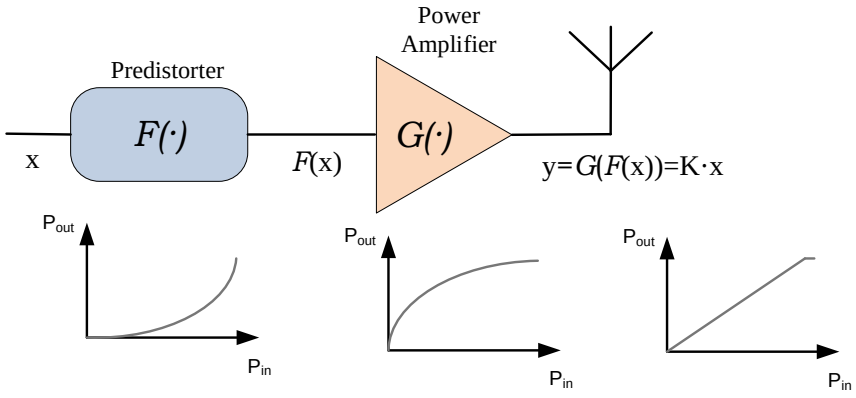

Figure 2.6: The concept of predistortion.

Analog predistortion (APD) uses an analog circuit to predistort the input signal. One of the advantages of ADP is that the bandwidth expansion is taking place in the analog domain, which relaxes the clock rate requirements in the digital-to-analog converter (DAC) to generate the analog transmitted signal. For the design of analog predistorters, the key point is to tune a nonlinear circuit that has to show the exact inverse nonlinear characteristic as the targeted PA. Several solutions have been proposed to generate the inverse characteristic using analog circuitry. In the following, some examples are discussed.

The company Scintera (acquired by Maxim Integrated), for example, commercialized an APD solution based in Volterra series polynomials implemented with analog multipliers. Alternatively,

some solutions considered the use of the nonlinear characteristics of diodes [Yam96, Yu99].

As explained in [Den18] , composite right-left-handed transmission lines (CRLHTLs) were used for ADP purposes since they can vary its nonlinear transmission characteristics with the input power or bias voltage (see Fig. 2.7). The authors reported 8.1 dB enhancement in the inter-modulation distortion at -6 dB output power back-off when linearizing a C-band 40 W solid-state PA.

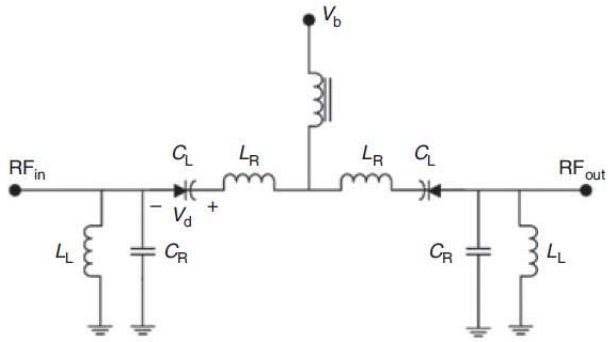

Figure 2.7: The non-linear CRLHTL circuit predistorter [Den18].

Another commonly used solution to generate the nonlinear characteristic using analog circuitry is the use of transistors. In [Qi 16] a transistor-based APD was presented. The proposed APD (see Fig. 2.8) showed that considering a two-tone test, more than 10 dB IMD3 improvement was reported. In addition, the output 1 dB gain compression point (P1dB) was increased from 35 dBm to 38.5 dBm.

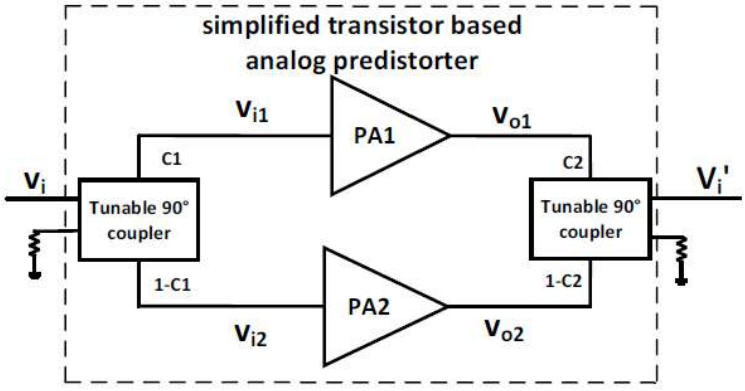

Figure 2.8: The transistor based Analog Predistorter Schematics [Qi 16].

Another straightforward benefit of APD is that can be designed to be implemented in the same analog RF circuitry as the PA [Gao18, Gum18], although this solution provides less flexibility in comparison to adaptive digital predistortion.

2.2.2 Digital Predistortion

With the development of signal processing hardware (e.g. system-on-chip (SoC) and field programmable gate arrays (FPGAs)) with high speed processing and memory capabilities, digital predistortion linearization has become the dominant linearization technique.

The principles of DPD are the same as in APD and summarized in Fig. 2.6. However, unlike in APD solutions, the predistortion of the transmitted signal is carried out at baseband in a digital signal processor. By applying DPD to the complex digital base-band signal, the predistorter signal to be transmitted suffers a bandwidth expansion (e.g., 3× to 5× the original signal bandwidth) that has to be taken into account to determine the DAC clock rate. In addition, as depicted in the block diagram of Fig. 2.9, a feedback or observation path is necessary to monitor the amplified signal and estimate the coefficients of the DPD. Again, the ADC in the observation path has to be able to capture the output signal taking into account the spectrum regrowth that appears due to the PA nonlinear behavior.

The digital predistorter is usually described by behavioral models or black-box models, that are mathematical descriptors of the inverse non-ideal behavior of power amplifiers. Unlike physical models, where it is necessary to know the electronic elements that comprise the PA, their constitutive relations and the theoretical rules describing their interactions, the extraction of PA or DPD behavioral models relies only on a set of input-output observations. Consequently, as it will be discussed in the following Chapters the performance of the DPD depends on the adopted model structure and the parameter extraction procedure.

In literature it is possible to find a vast amount of published PA and DPD behavioral models to address not only SISO systems [Sch09] but also MISO systems, for example when having to characterize concurrent multi-band transmissions or dynamic supply modulation strategies for the PA [Gil19]. In Chapter 3 of this thesis the principles of digital baseband predistortion will be described and analyzed more in depth.

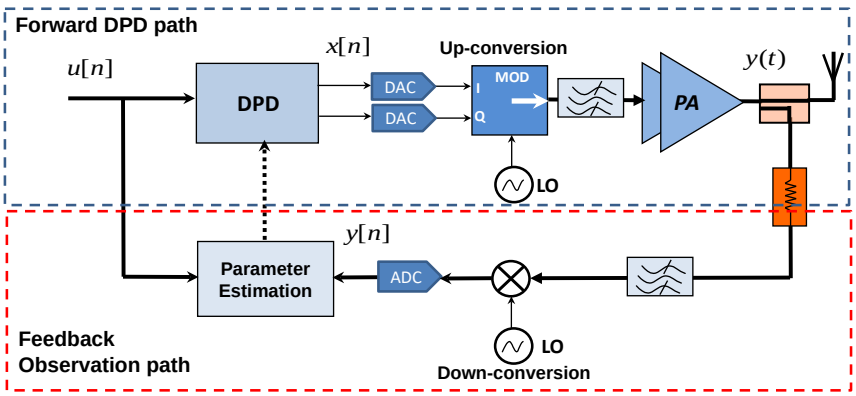

Figure 2.9: Simplified block-diagram of an adaptive DPD.

2.2.3 Hybrid Predistortion

With the increasing signal bandwidth requirements in 5G New Radio (NR), specially at mm-wave bands (i.e., several hundreds of MHz), new hybrid analog/digital predistortion solutions have been proposed.

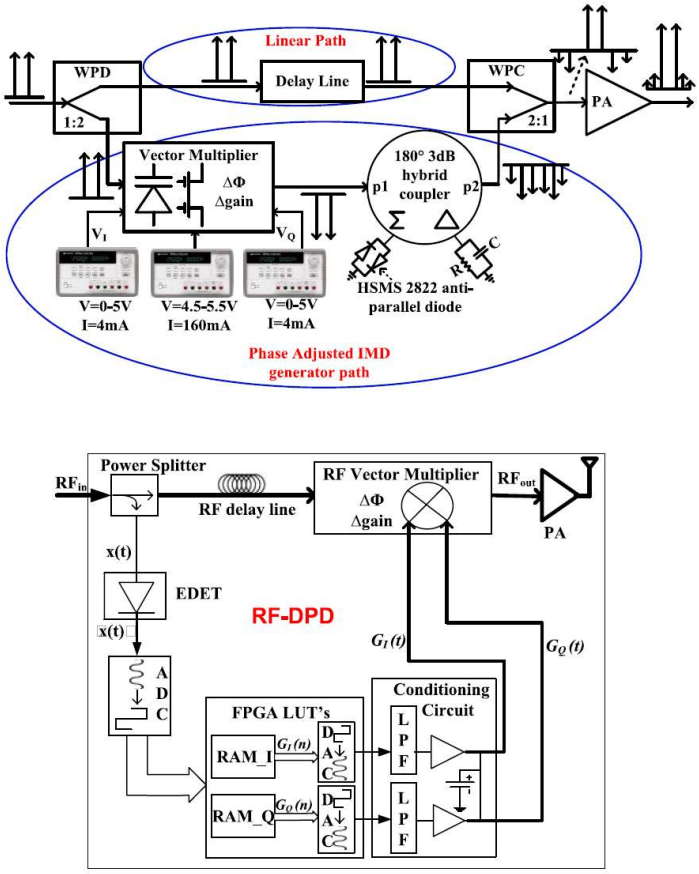

Figure 2.10: The block diagram of the hybrid digitally controlled analog predistorter [Gum18].

Usually, the most powerful predistortion algorithms are achieved by employing digital signal processing able to predistort complex baseband I-Q signals. With DPD we can compensate for both static and dynamic nonlinearities (i.e., memory effects). However, one of the main drawbacks of DPD compared to APD, is that requires higher baseband sampling and processing speed with respect to the nominal processing performed in the modem and highly demanding matrix operations, which increases the radio equipment power consumption. In order to relax the baseband sampling requirements (i.e., DACs and ADCs) of DPD, hybrid solutions combining APD with some sort of baseband processing have been proposed.

In [Gum18], for example, the authors present an hybrid digitally controlled (HDC) analog

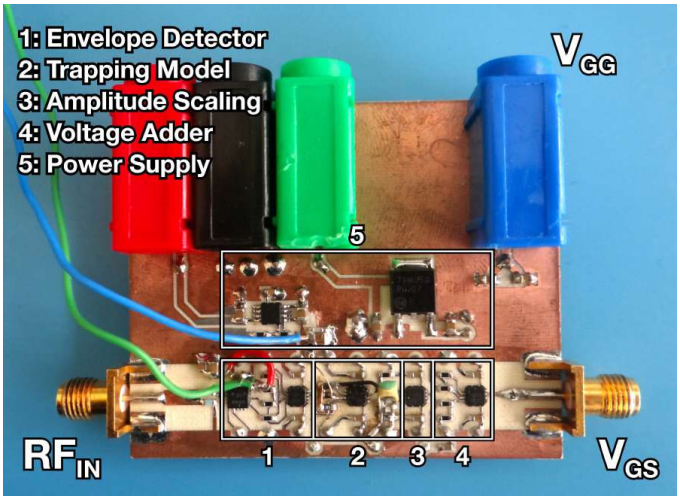

Figure 2.11: Photograph of the analog compensation circuit [Tom19].

predistorter (APD) in which linearization is performed by RF components in the analog domain. The hybrid predistorter eliminates the constraint on the system bandwidth of the conventional DPD. The requirements of data converters and reconstruction filters are relaxed in the APD architecture, whereas HDC-APD eliminates the need of analog components by compensating the delay digitally. The APD architecture alone is able to deliver ACPR of -46.4 dBc with an ACPR improvement of 13.4 dB at 1 dB back-off. Furthermore, with digital intervention, the HDC-APD further provides a significant improvement in the linearization performance and delivers an ACPR of -53.5 dBc with an ACPR improvement of 20.5 dB.

A different approach is considered in [Tom19], where a hybrid analog/digital linearization scheme for gallium nitride high-electron-mobility transistor (GaN HEMT) PAs is proposed. The authors combine the conventional DPD based on a generalized memory polynomial (GMP) behavioral model, with an analog feed-forward circuit specifically design for the compensation of long-term memory effects due to the self-biasing caused by electron trapping in GaN HEMTs. Experimental results demonstrate that by combining the analog predistortion with digital, it achieves a level of inter-modulation distortion 6.8 dB better than the GMP DPD only.

Chapter 3

Principles and Challenges in Open-Loop DPD

3.1 Overview on Digital Predistortion Linearization

Digital predistortion (DPD) is one of the most extended linearization techniques used to mitigate the well-known power amplifier (PA) trade-off between linearity and power efficiency. One of the major benefits of digital predistortion is that it separates linearization and PA design into two independent parts, which allows researchers to focus in what they are good at. It also boosts the complexity of both parts. For more researchers come up with innovative, complex and efficient power amplifiers, people in the field of DPD will also offer new coming promoted algorithms to respond to the challenges posed by the new PAs.

DPD can be adaptive (closed-loop) or non-adaptive (open-loop). To be able to compensate for eventual changes of the PA conditions (temperature, aging, etc), the DPD incorporates a feedback path to allow adaptation of the DPD coefficients. Sometimes, the adaptation may not be carried out in real-time but in a slower time-scale. Including a closed-loop adaptation subsystem, inevitably increases the cost, usually on the bill-of-materials (BOM), and energy consumption of the digital processing part. However, the consumption related to the adaptation part can be mitigated by properly defining an adaptation policy in which, for example, some subsystems can stay in idle mode when they are not used. In any case, in both open-loop and closed-loop DPD the identification of the coefficients has to be properly carried out to ensure the robustness of the DPD. That is, once properly estimated, the DPD coefficients should be the optimum ones in a global (not local) sense for a given transmitted signal with certain statistical features.

In general, a memoryless DPD can be seen as a single input single output (SISO) system, where the output of the DPD function depends only on the nonlinear transformation of the

actual input signal,

$$x = f_{DPD}(u) \tag{3.1}$$

When targeting also the compensation of the PA memory effects, the DPD function can be seen as a multiple input single output (MISO) system, where the output of the DPD depends on several input signals that are delayed versions on the actual input signal.

$$x = f_{DPD}(u_0, u_1, \cdots, u_k, \cdots, u_K) \tag{3.2}$$

where $u_k = u[n - \tau_k]$ with $k = 0, \cdots, K$ and where τ_k ($\tau_0 = 0$) are the most significant sparse delays characterizing the memory of the PA. We now rewrite (3.2) as

$$x = f_{DPD}(\boldsymbol{\nu}) \tag{3.3}$$

where $\boldsymbol{\nu} = (u_0, u_1, \cdots, u_k, \cdots, u_K)$ is the $(K+1) \times 1$ vector of (multiple) input signals to the nonlinear DPD function f_{DPD}.

3.1.1 DPD behavioral models

The DPD model is the basic part of the whole DPD algorithm. In general, it can be divided into two major families: memoryless behavioral models and dynamic (with memory effects) models. In memoryless model, the PA behavior is regarded as simple nonlinear mapping between the input and output signal, then the reversed mapping will be exactly the pre-distorting function to generate the PA input signal. The Saleh Model [Sal81] was initially applied to characterize traveling wave tube (TWT) amplifiers. The two general functions to approximate the amplitude-amplitude (AM-AM) and the amplitude-phase (AM-PM) inverse envelope characteristics of the PA are:

$$f_A(|u[n]|) = \frac{\alpha_A |u[n]|}{1 + \beta_A |u[n]|^2} \tag{3.4}$$

$$f_\phi(|u[n]|) = \frac{\alpha_\phi |u[n]|^2}{1 + \beta_\phi |u[n]|^2} \tag{3.5}$$

where α_A, β_A, α_ϕ, β_ϕ are constant parameters chosen to approximate the inverse PA characteristics.

Together with Saleh's models, memoryless polynomial approximations are used to estimate the PA inverse nonlinear static behavior,

$$x[n] = \sum_{p=0}^{P} \alpha_p u[n] |u[n]|^p \tag{3.6}$$

where P the order of the polynomial function and α_p complex coefficients.

However, to accurately compensate for the unwanted PA distortion we have to consider that RF power amplifiers are not only a function of the input signal amplitude at the same instant, but also related with the recent history of the input signal, especially when the input signal has a large bandwidth. This phenomenon is usually called the memory effects of PA. The memory effects model is designed to also consider the output signal of PA is the mapping of several input signals in a short time period, which end up with a MISO function.

It is possible to find in literature an enormous amount of publications on PA or DPD behavioral modeling to address not only SISO systems [Sch09] but also MISO systems, for example when having to characterize and compensate the nonlinear distortion in concurrent multi-band transmissions or in dynamic supply modulation strategies for the PA [Gil19]. Some of the most commonly used polynomial-based behavioral models can be regarding as a simplified approximation of the general Volterra series. Volterra series are aimed at describing time-invariant nonlinear systems with fading memory. The discrete-time low-pass equivalent Volterra series formulation considering complex signals is described in the following. The DPD output $x[n]$ taking into account the full Volterra series is defined as

$$x[n] = \sum_{p=1}^{P} \sum_{q_1=0}^{Q-1} \sum_{q_2=q_1}^{Q-1} \cdots \sum_{q_p=q_{p-1}}^{Q-1} \cdots \sum_{q_{2p-1}=q_{2p-2}}^{Q-1} h_{2p-1}(q_1, q_2 \cdots, q_{2p-1}) \prod_{i=1}^{p} u[n-q_i] \prod_{j=p+1}^{2p-1} u^*[n-q_j] \tag{3.7}$$

The series is composed by $2P-1$ kernels of increasing dimensional order. The main drawback of using the full Volterra series is that the number of parameters grows exponentially when considering higher order kernels and typical communication signals do not present enough richness to fully excite these kernels, which ultimately may lead to an ill-conditioned problem.

One of the most commonly used behavioral models found in literature is the memory polynomial (MP) [Kim01], mainly for its simplicity to characterize the PA nonlinear behaviour and its memory effects. An extension of the MP is the generalized memory polynomial (GMP) behavioral model, proposed in [D. 06]. Unlike the MP, the GMP has bi-dimensional kernels (considering cross-term products between the complex signal and the lagging and leading envelope terms) which increases the accuracy of the modeling at the price of increasing the number of parameters. The input-output relationship in the DPD taking into account a GMP model is defined as:

$$
\begin{aligned}
x[n] = & \sum_{i=0}^{N_a-1} \sum_{p=0}^{P_a-1} \alpha_{pi} \cdot u[n-\tau_i^a] \left| u[n-\tau_i^a] \right|^p \\
& \sum_{j=1}^{M_b} \sum_{i=0}^{N_b-1} \sum_{p=1}^{P_b-1} \beta_{pij} \cdot u[n-\tau_i^b] \left| u[n-\tau_i^b-\tau_j^b] \right|^p \\
& \sum_{j=1}^{M_c} \sum_{i=0}^{N_c-1} \sum_{p=1}^{P_c-1} \gamma_{pij} \cdot u[n-\tau_i^c] \left| u[n-\tau_i^c+\tau_j^c] \right|^p
\end{aligned}
\tag{3.8}
$$

where P_a, P_b, P_c are the nonlinearity orders of the polynomials, N_a, N_b, N_c, M_b, M_c are the lengths of memories. α_{pi}, β_{pij} and γ_{pij} are the complex coefficients describing the model, and τ^a, τ^b and τ^c (with $\tau \in \mathbb{Z}$ and $\tau_0 = 0$) are the most significant non-consecutive delays of the input signal $u[n]$ that better contribute to characterize memory effects. The total number of coefficients of GMP model is $O = P_a M_a + P_b M_b N_b + P_c M_c N_c$. Note that the first branch (out of three) of the GMP behavioral model in (3.8) corresponds to the definition of the MP behavioral model.

3.1.2 Closed-Loop Adaptive DPD

The block diagram of a closed-loop adaptive DPD is shown in Fig. 3.1. In the forward path, the input-output relationship at the DPD block can be described as

$$x[n] = u[n] - d[n] \tag{3.9}$$

where $u[n]$ is the input signal, $x[n]$ is the predistorted signal and $d[n]$ is the additional distortion term, that can be described by any of the proposed DPD behavioral models published in literature.

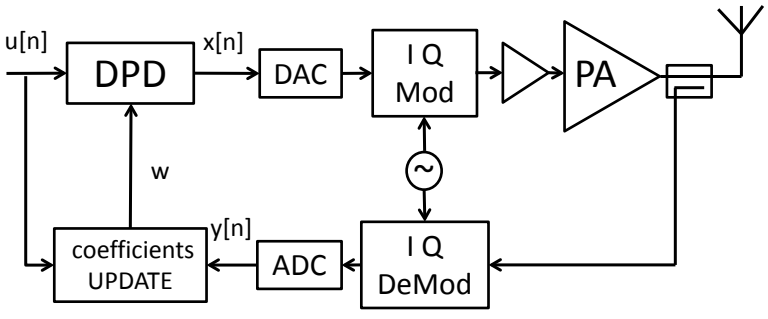

Figure 3.1: Block diagram of a closed-loop DPD.

Following a matrix notation, the $L \times 1$ vector $\mathbf{d} = \big(d[0], \cdots, d[n], \cdots, d[L-1]\big)^T$, with L being the number of samples, is defined as

$$\mathbf{d} = \mathbf{U}\mathbf{w} \tag{3.10}$$

where the DPD coefficients, $\mathbf{w}$, are iteratively extracted following a gradient descent approach as,

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \mu\Delta\mathbf{w} \tag{3.11}$$

with μ being a weighting factor and where the gradient $\Delta\mathbf{w}$ is calculated using the LS algorithm.

$$\Delta\mathbf{w} = \left(\mathbf{U^H U}\right)^{-1}\mathbf{U^H e} \tag{3.12}$$

where $\mathbf{U}$ is the $L \times O$ data matrix,

$$\mathbf{U} = \begin{pmatrix} \psi_1[0] & \cdots & \psi_q[0] & \cdots & \psi_O[0] \\ \vdots & & \vdots & & \vdots \\ \psi_1[n] & \cdots & \psi_q[n] & \cdots & \psi_O[n] \\ \vdots & & \vdots & & \vdots \\ \psi_1[L-1] & \cdots & \psi_q[L-1] & \cdots & \psi_O[L-1] \end{pmatrix}$$

with $\psi_q[n]$ $(q = 1, \cdots, O)$ being each of the specific basis functions at time n. For example, taking into account the GMP behavioral model in (3.8), $O = N_aP_a + M_bN_b(P_b-1) + M_cN_c(P_c-1)$ represents the order (i.e., number of columns) of the DPD function and $\mathbf{w} = \left(w_1, \cdots w_q, \cdots w_O\right)^T$ is the $O \times 1$ vector of coefficients. The mapping between the GMP specific coefficients $(\alpha_{pi}, \beta_{pij}, \gamma_{pij})$ in w_q and the general purpose DPD coefficients w_q in (3.10) is shown in the following:

$$w_q = \begin{cases} \alpha_{pi} & \text{with } q = \sum_{i=0}^{N_a-1}\sum_{p=0}^{P_a-1} iP_a + (p+1) \\ & \text{if } q \le N_aP_a \\ \beta_{pij} & \text{with } q = N_aP_a + \sum_{j=1}^{M_b}\sum_{i=0}^{N_b-1}\sum_{p=1}^{P_b-1}(j-1)(P_b-1)N_b + i(P_b-1) + p \\ & \text{if } N_aP_a < q \le N_aP_A + M_bN_b(P_b-1) \\ \gamma_{pij} & \text{with } q = N_aP_a + M_bN_b(P_b-1) + \\ & \sum_{j=1}^{M_c}\sum_{i=0}^{N_c-1}\sum_{p=1}^{P_c-1}(j-1)(P_c-1)N_c + i(P_c-1) + p \\ & \text{if } N_aP_A + M_bN_b(P_b-1) < q \le N_aP_a + M_bN_b(P_b-1) + M_cN_c(P_c-1) \end{cases}$$

Finally, the $L \times 1$ error vector is defined as

$$\mathbf{e} = \frac{\mathbf{y}}{G_0} - \mathbf{u} \tag{3.13}$$

with $\mathbf{e} = \big(e[0], \cdots, e[n], \cdots, e[L-1]\big)^T$. Similarly, $\mathbf{u}$ is the $L \times 1$ input signal vector, $\mathbf{y}$ is the $L \times 1$ PA's output signal vector and G_0 is the targeted linear gain.

For adaptive DPD, when the adaptation process is finished and the coefficients values have converged, the DPD will perform well (with low ACLR and NMSE value) for the same input signal. But when considering different test signals with the same bandwidth, same mean power and similar PAPR, the DPD linearization has different performance with each set of signal.

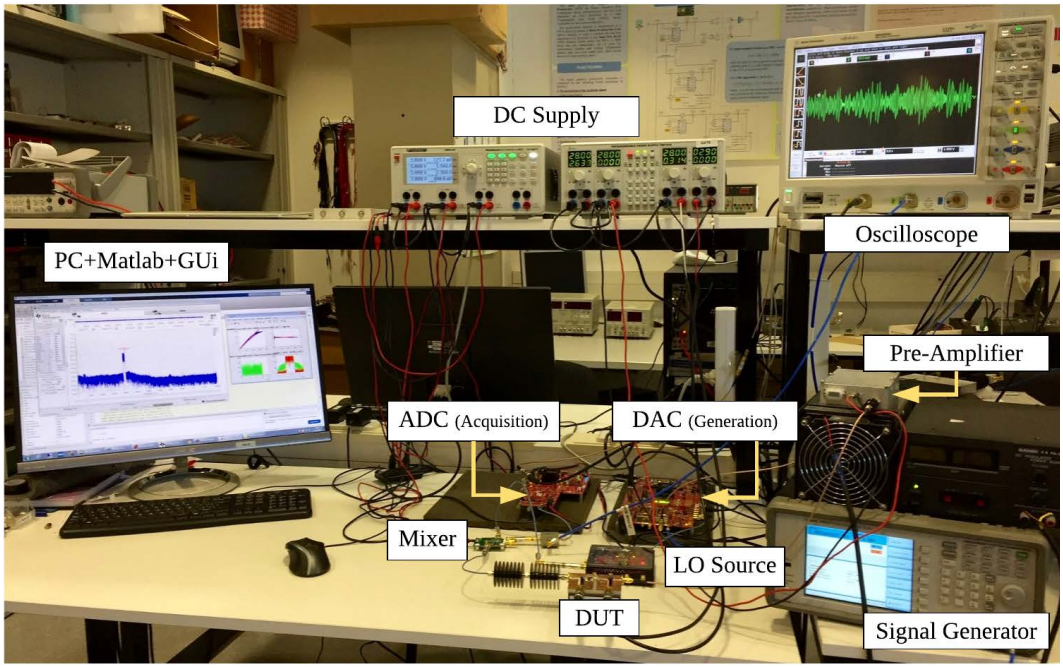

Figure 3.2: Typical test-bench for DPD linearization of an RF PA using commercial boards from Texas Instruments for signal generation and capturing [Flo17].

A test or design bench for DPD is usually assembled in a way to have the higher reliability. The typical closed-loop DPD test bench consists in a powerful computer as the computational core, with Matlab as the software platform, the device-under-test (DUT) and instrumentation or commercial boards to generate and capture the RF signal, as shown in Fig. 3.2. The PA under test will be connected to an arbitrary waveform generator (AWG) or commercial boards (with DACs and up-converters) to generate the RF signal and, to close the loop, to a digital storage oscilloscope (DSO) or commercial boards (ADCs) to digitize the RF signal with high accuracy. An FPGA is often used to implement the DPD function to operate in real-time. The powerful parallel computing capabilities of FPGAs make them a suitable platform to run the DPD function. In this thesis however, the DPD function will be also carried out in Matlab running in a PC. The code that will be written in Matlab includes the signal alignment, normalization and

coefficients extraction. The DPD model and parameters need to be precisely adjusted to match the specific nonlinear behavior of the PA under test and the specific characteristics of the input signal. This is usually the most important part.

3.1.3 Look-Up Table Implementation of DPD Models

In a mathematical sense, functions have three expressions: parsing, tables, and images. Given a precision, known function parsing equations can be converted into tables, which is called the look-up table (LUT).

The programmable logic (PL) unit, for example, of a system-on-chip (SoC) FPGA device, can be used as the platform when targeting the DPD forward path into a digital signal processing platform implementation. For example, by following a LUT approach as in [Gil08], [Mol17], or by considering a polynomial approach using the Horner's rule as in [Mra12], or by combining both complex multipliers/adders and memory as in [Cao17]. As a result, those approaches provide simplified design methods (i.e., including the minimum and most relevant basis functions) for the DPD function in the forward path to save as many hardware logic resources and memory as possible.

For DPD function, the look-up table between input and predistorted signal can be converted according to a given set of coefficients. Particularly, the authors in [Ma13] discussed the advantages of LUT-based over polynomial/parsing based implementation of a DPD model. First, the computational complexity of LUTs is much lower in terms of using fewer multipliers, while reducing the number of multipliers means reducing the hardware resources requirements for the DPD implementation. Secondly, also thanks to the lower number of multipliers, LUTs are more numerically stable than polynomials, especially when the order of the polynomials is high. For example, a 32-bit processor cannot handle polynomials greater than sixth order according to practical experience. Thus, with higher order polynomials LUT takes further advantage as the numerical instability becomes apparent. Third, the change of numerical scale will deteriorate the predistorter's robust performance when using polynomial multipliers, in the case that the signal peak power changes significantly over time.

In hardware implementation, FPGA is the most attractive solution for implementing the DPD function, thanks to its fast prototyping, high-speed processing, high density integration, flexible implementation and parallel operation mechanisms [Gua10]. The DPD function can be implemented in FPGA in form of either the polynomial-based or the LUT-based methods. The polynomial-based implementation requires several complex multiplications and additions [Kwa13]. function.

Due to the significant advantages of the LUT solution that discussed above, most research has focused on the LUT direction. Several papers have shown the advantages of LUT-based

implementations [Gil08, Gil11].

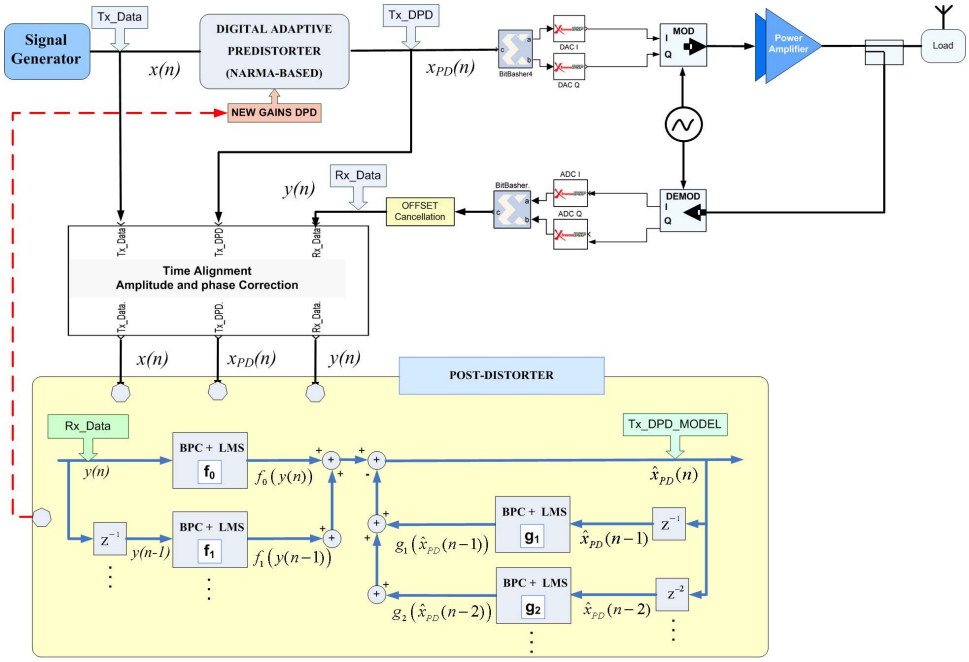

Figure 3.3: Block diagram of adaptive DPD on FPGA implementation [Gil11]

3.1.4 Sampling Rate Requirements in the Observation Path

The behavior of PAs usually changes with temperature and aging. Moreover, behavioral models are very dependent on the signal characteristics, i.e., power levels, signal bandwidth, PAPR, modulation type. It is therefore not possible to create a unique DPD function to predict and compensate for all PA unwanted effects in the whole life time of a communication system. A better idea is to monitor the changing of PA behavior and then update the DPD function accordingly. The adaptive DPD is designed based on this idea, to make the DPD function adaptive to the PA changes.

In DPD, the inverse behavior of PA is extracted with input-output data observations of the PA. Thus, in adaptive DPD implementations an observation path is included to acquire the PA output signal. This signal is then properly attenuated and time-aligned to compare it to the original transmitted signal and thus evaluate the linearity of the whole system. The feedback or observation path usually includes a analog down converter to move the RF output of the PA into base-band or IF, as shown in Fig. 3.4. In case of down-converting it to IF, after the ADC, further digital down-conversion to base-band is required. The most expensive component is the ADC and the price is proportional to the sampling rate. Intuitively, to reduce the required sampling rate at the observation path has significant economic benefits.

It is well known that due to the nonlinear behavior of the PA, the PA output signal suffers

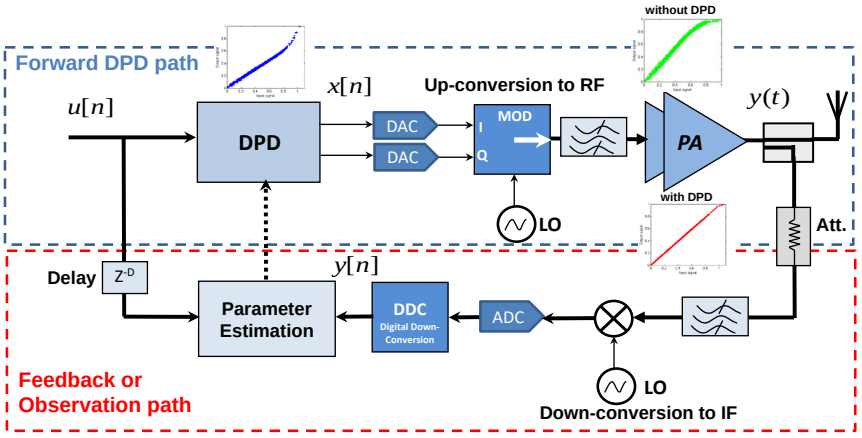

Figure 3.4: Block diagram of adaptive DPD.

spectral regrowth (3x to 5x the bandwidth of the original signal, depending on the severity of the nonlinear distortion). Therefore, in order to identify the PA nonlinear behavior and train the DPD, it is common to sample the output signal at twice the maximum output frequency to meet the Shannon-Nyquist criteria and capture the out-of-band distortion without aliasing. First Zhu in [Zhu92] for nonlinear memoryless systems and later Frank in [Fra96] extending it for nonlinear systems with memory, formulated a general sampling theorem showing that it is sufficient to sample at twice the maximum input frequency to reconstruct undersampled signals [Zhu92].

Recent research has demonstrated that several techniques can be used to reduce the sampling rate requirement of DPD. For example, constraining the bandwidth of the DPD model and linearize the PA within a limited bandwidth [Yu12]. In this way, the feedback signal can be filtered before sampling without affecting system performance. A different approach consists in filtering also the feedback signal before sampling, but the DPD produces full-band signals in [Gua12a, Liu14, Ma14, Liu15, Bra15]. In this case, the model extraction algorithms need to apply spectral extrapolation to the loss function to achieve full-band distortion cancellation. Another method is to remove the anti-aliasing filter before sampling [Wan15a, Wan16, Koe06, Wan17, Che17]. The aliased feedback signal is then used together with specially processed input signals to extract model coefficients. The under-sampling method do not effect the accuracy of the sampled signal in time domain, which means the original undersampled signal is able to be used for DPD update estimation (see Fig. 3.5).

In [Wan16, Wan15a] researchers from Ireland and Spain published the idea to re-define the sampling problem for the extraction of the DPD coefficients. For example, in [Che17] the authors showed the possibility of using up to 80 times lower Nyquist frequency to retrieve the non-linear characteristics of the PA and fulfill the requirements for DPD parameters extraction (see Fig. 3.6). Applying under-sampling methods to DPD parameter extraction facilitates the use of ADCs with significantly lower sampling frequency, which offers cost advantages. In [Wan16], Z. Wang

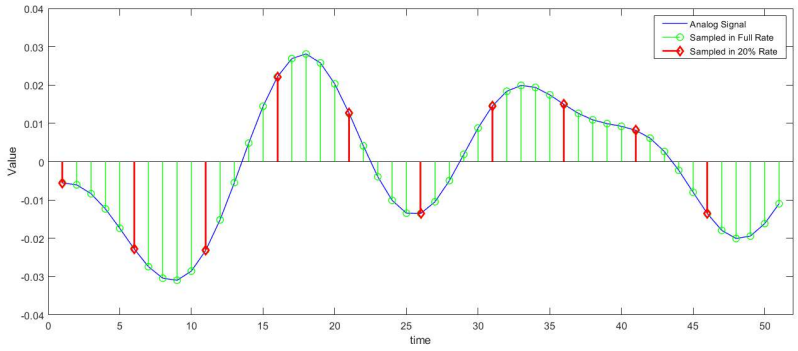

Figure 3.5: The undersampled signal in time domain

et al. showed that if the acquired signal from the PA output has the same normalized magnitude distribution as the original input signal, this is sufficient to update the DPD function. In other words, it has nothing to do with the sampling rate. In [Wan15a], T. Wang et al. proposed a mesh-method to accurately choose (consecutive or non-consecutive) samples from the PA output signal to reduce the computational cost of the parameters extraction. Therefore, the selection results in decimated training signals that present aliasing and yet are valid to obtain effective parameter identification. The mesh-selection method will be further discussed in Chapters 4 and 5 of this thesis.

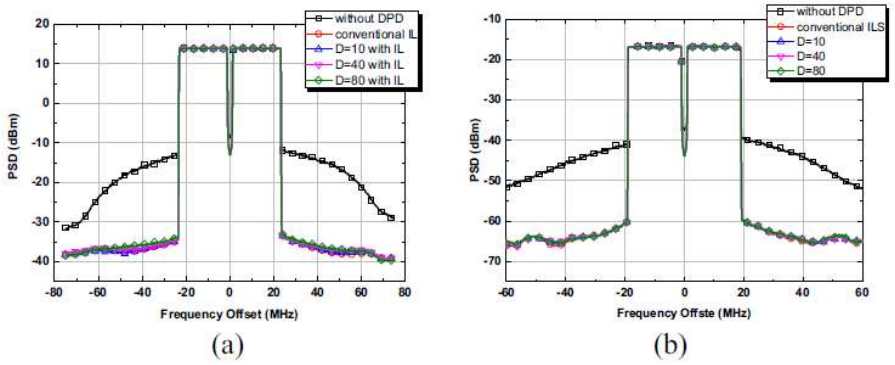

Figure 3.6: Performance of under-sampling method on (a) Class AB PA (B) Doherty PA [Che17].

3.1.5 Dimensionality Reduction in DPD Linearizers

Dimensionality reduction techniques aim at analyzing a given number of features, such as dimensions, variables, basis functions or components, under consideration in a given set of data. Based on the result of such analysis, outstanding features are able to be highlighted while the redundant or irrelevant ones will be eliminated. Thus, the model's performance is improved or kept at the same level at least [Pha19a]. These techniques can be sorted into:

- *Feature selection*, selecting the most relevant variables from a random set of original variables.
- *Feature extraction*, creating a reduced set of new variables that are linear or nonlinear combinations of the original variables.

Feature selection techniques are used in the field of DPD linearization to reduce the number of parameters of the DPD function. The objective of feature selection is to enforce the sparsity constraint on the vector of parameters by minimizing the number of active components (i.e., ℓ_0-norm) subject to a constraint on the ℓ_2-norm squared by the identification error. In literature several approaches have been proposed targeting both a well-conditioned identification and model order reduction, such as: least absolute shrinkage and selection operator (LASSO), used for example by Wisell et al. in [Wis08] and consisting in a ℓ_1-norm regularization; the Ridge regression, used for example by Guan et al. in [Gua12b] and consisting in a ℓ_2-norm regularization; the sparse Bayesian learning (SBL) algorithm, used by Peng et al. in [Pen16a]; or the orthogonal matching pursuit (OMP), a greedy algorithm for sparse approximation used in [Rei15] by Reina et al. to select the most relevant basis functions of the DPD function.

A different approach to reduce the parameters of the DPD system, are the feature extraction techniques. Unlike feature selection techniques, feature extraction techniques do not reduce the number of coefficients of the DPD function in the forward path. With techniques such as the principal component analysis (PCA) [Gil13a] or the partial least squares (PLS) [Pha18a], it is possible to create a new reduced set of orthogonal components, which are linear combinations of the original basis functions, and then apply dimensionality reduction in the DPD adaptation system. These techniques can be seen as an alternative to common solutions used to solve the least squares regression problem, such as the QR factorization combined with recursive least squares (QR-RLS) [Mur06]. Feature extraction techniques ensure a proper well-conditioned estimation and a reduction in the number of parameters in the identification process. Alternatively, both feature selection and future extraction techniques can be properly combined as in [Pha18b] by:

- Doing an a priori off-line search (e.g., OMP, LASSO) to reduce the number of basis functions of the DPD function in the forward path.
- Using PCA or PLS techniques for the parameter extraction in the adaptation path.

3.2 Challenges of Open-Loop DPD

Even though, as stated above, DPD linearization is able to boost PA power efficiency while guaranteeing linearity values (even when considering high bandwidth signals) most of the avionics

companies refuse to import this technology into their product development chain. This phenomenon is especially obvious in commercial VTB companies, even the top level video transfer providers, like the AMIMO company of Israel and DJI of China. This means the DPD method still presents some drawbacks for being considered in large scale implementation in the field of avionics.

3.2.1 High Technical Requirements

From the industry perspective, to build a team for designing a transmitter with DPD (even an open-loop one where the parameters are pre-calculated off-line in the factory), it is necessary to have engineers that master Matlab-controlled RF system design to assemble a test-bench for validation and engineers that master real-time implementations with system-on-chip (SoC) FPGAs. The first ones have to be familiar with academic papers to prototype the DPD algorithms and the RF system, while the second ones have to have good FPGA coding skills.

In addition, the instrumentation required to assemble a laboratory test-bench to carry out DPD research can be significantly expensive. For signal generation, arbitrary waveform generators are required. While, high resolution oscilloscopes, high-speed digitizers or even spectrum analyzers are necessary to capture the PA output data. Yet most commercial VTB companies only need to deal with the RF front-end after the transceiver chip.

Last but not least, a radio frequency (RF) transmitter with DPD linearization incorporated means that is going to have significantly higher complexity than a transmitter without linearization. Not only due to the additional observation path including a down-converter and an ADC, which can be deminished in open-loop approach, but also due to the additional computation required in the FPGA and the higher DAC clock requirements (due to the bandwidth expansion after DPD) in the forward path. The cost will grow up with the new system requirements and taking into account the extra investment on its development team.

3.2.2 Non-Robustness

Even with a trained team and sufficient tests, the open-loop DPD is still a high risk approach due to its non-adaptive nature. The PA behavior may change with the fluctuation of temperature, aging and other factors. Thus, most non-adaptive DPD cannot compromise the avionics scenario, where the size and cooling power is limited in the sky. When it goes to the adaptive DPD, the DPD function is updated all the time to track possible variations in the PA characteristics. But if the input signal goes out of the range of the test area of the DPD method, unlike the behavior of analog chips, the digital part has a possibility to generate unwanted results (i.e., bad points). These bad points will appear as extra peaks of the RF signal, extending the PAPR or peak to unexpected power levels, which may lead irreversible damage of the PA. Moreover, if the DPD is

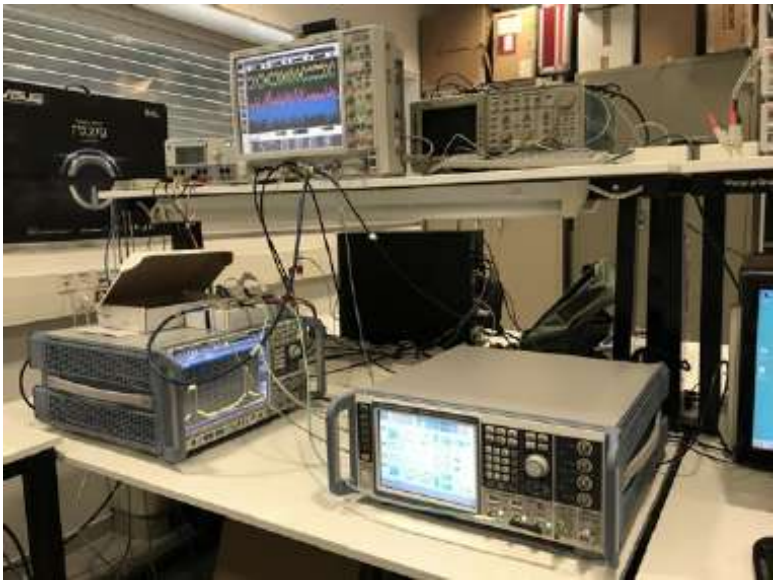

Figure 3.7: DPD linearization test-bench implemented for the IMS2017 student design competition.

not carefully designed, when the PA enters in hard saturation, it may make the DPD algorithm to diverge.

Recent Studies of the DPD Robustness

DPD behavioral models become more and more complex in order to achieve better performance and to compensate for unwanted distortion in more complex amplification scenarios. In addition, some researches has dedicated their efforts to improve the DPD robustness.

For example, the authors in [Isl17] focus on the numerical problem of the DPD procedure. They identify the cause of divergence in dynamic rational function, i.e. the function of the DPD behavioral model, and propose a new computational process called constrained coefficient identification procedure to avoid the divergence. The experimental results on a GaN Doherty PA with 40 MHz and 80 MHz bandwidth signals proved the validation of the constrained identification procedure.

It is not very difficult for researchers and engineers to have an accurate DPD behavioral modelling performance in the laboratory by considering the PA behavior to be time-invariant. However, the evolution of the PA non-linear behavior will directly deteriorate the performance of DPD in practical cases. In adaptive implementation, the DPD method is able to capture eventual changes of the PA behavior through a feedback path and then update the DPD function coefficients to compensate for it. But if the PA variant behavior approximates the adaptation rate of the DPD algorithm on a time scale, it causes a relatively obvious robustness problem.

To enhance DPD robustness in front of thermal variations, a direct solution is to take the

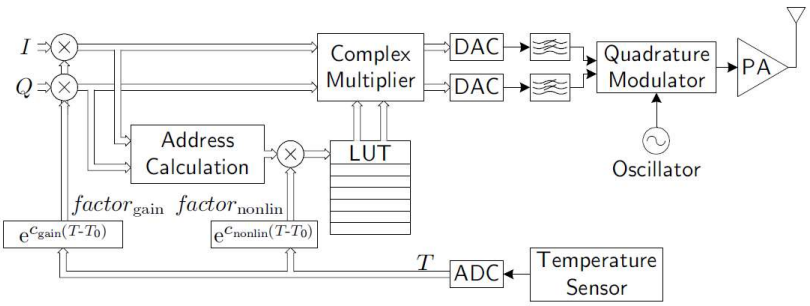

Figure 3.8: Scaling of the AM-AM characteristic to equalize the small signal RF gain feature. [Wol10]

temperature as an independent variable to the DPD behavioral model. In [Yu17], for example, the authors propose a canonical piecewise linear (CPWL) functions based DPD modeling to compensate thermal memory effects. In [Wol10], instead, the authors provide a simple way to linearize a power amplifier with only static characteristics (AM-AM and AM-PM curve) over the entire temperature range. The author illustrated the variance of PA gain feature which significantly effects the DPD performance (see in Fig. 3.8). The system will first gather all characteristic curves at different temperatures and map them by means of scaling factors into a combined resultant curve. The inverse static curve is mapped into a LUT and properly combined with the temperature scaling factors are able to implement a DPD that takes into account temperature variations, as shown in Fig. 3.9. As a result, the linearization performance is shown to be constant over the entire temperature range from -30℃ to 90℃.

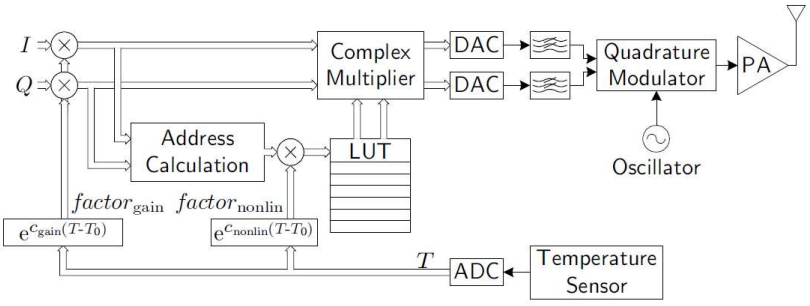

Figure 3.9: Temperature independent predistortion system using only one LUT for all temperatures [Wol10]

Robustness Analysis of the DPD Function

In PA behavioral modeling or DPD linearization, the mathematical model (or fitting function) consists of a linear combination of nonlinear basis functions composed not only of the actual sample but also past samples of the input (in order to take into account memory effects). The

coefficients of this fitting function are extracted from the original input data matrix (containing the basis functions of the PA behavioral model or the DPD linearizer) through some adaptive algorithms, either sample-by-sample or block-by-block stochastic gradient algorithms. The widely used LS algorithm for linear regression (despite the fact that the DPD fitting function is nonlinear, it is however linear in parameters) can be seen as a simplified alternative to the Wiener filter theory. As explained in [Hay91], the Wiener filter is derived from a set of averages with the result that its coefficients are optimum (in the probabilistic sense) and are obtained for all realizations of the operational environment (assumed to be wide-sense stationary). Instead, the LS method is considered a deterministic approach, involving the use of time averages, with the result (optimal filter coefficients) strongly depend on the number of samples used in the computation.

Therefore, a priory, the LS estimator may need a large number of data samples (or equations) to obtain the best approximation of the coefficients of the fitting function. As reported in [Gua12b], using short sequences of data samples leads to two consequences: a) the ill-conditioning problem due to the rank deficiency of LS matrices; and b) the statistical mismatch problem, because the short data sequence often cannot fully represent (accuracy problem) the statistical property of the PA output signal. Some efforts have been made to solve the ill-conditioning problem as well as to reduce the model extraction errors when using a small number of data samples to characterize PA behavioral models [Gua12b]. Alternatively, reducing the order of the DPD model has beneficial effects in both the computational complexity and in the conditioning of the data matrices. The idea of using dominant eigen-values/eigen-vectors for reducing the model order was presented in [Bra08, Gil13b]. However, to the best author's knowledge, few works have addressed the problem of determining statistical features that allow avoiding the accuracy problem in DPD linearization.

Because the fitting function is based on the data matrix, the domain of this function is also determined by the data matrix. The data matrix is generated from several sets of training signal, thus the range and distribution of the training signal determines the domain of the fitting function. The relation between the training signal and the fitting function shows how to analyse and evaluate its domain.

3.2.3 Design Complexity

As an edge technology, DPD now is still a hot topic in academia research. Although there are always several new publications per year and lots of them are shining, it is still difficult for engineers from industry to find a proper way to develop the DPD method taking into account the academic research. In other words, the product development of DPD implementation is kind of semi-research procedure, technicians need to search and test lots of methods and behavioral models proposed in all the research papers to figure out the proper fit for their own product. So

the project timeline becomes much more unpredictable at the same time.

All the drawbacks mentioned above does not mean that DPD linearization has no usage or market in the avionics sector, but indicates that there is still a long way to go to make it more popular and easy to use in the commercial industrial avionics. Furthermore, the main point of this journey is not purely academic DPD research, but the link between research and industrial implementation.

3.3 Design, Test and Implementation Strategies to Facilitate Open-loop Industrial Development

This thesis aims to contribute to facilitate an engineering approach for industrial DPD implementation in the avionics field, addressing some of the limitations discussed in the preceding subsection. This research focuses on two main parts, first is to reduce the computational consumption of the DPD algorithm and the update procedures, second is to provide an artificial intelligence reinforced DPD tuning method.

3.3.1 Test Signal Evaluation and Computational Reduction

Before being implemented in a FPGA platform for real-time linearization of the PA, the DPD is first prototyped and tested in a test-bed. As explained before, this test-bed is usually based on instrumentation (or commercial boards) connected to a PC running Matlab (or Python) to carry out the digital signal processing. In any case, during the test some signal will be first used as the training signal to train the DPD function. Then, several validation signals will be sent to the PA via the DPD linearizer to validate the robustness and stability of the DPD.

In order to train a DPD to be able to operate in open-loop, the test signal is of crucial importance, since its characteristics, i.e., mean power, bandwidth, type of modulation, PAPR will determine the coefficients of the specific DPD. Since the DPD model is signal dependent (since the PA behaviour is dependent on the type of excitation), the off-line DPD training should take into account all possible configurations at which the open-loop DPD will operate. Therefore, the training signal should be enough representative (from a statistical point of view) to be able to capture the all the features of the signal that will be later used in the real scenario, otherwise it will lead to a non-robust DPD identification.

One of the contributions of this thesis is to propose a new method based on mesh-selecting to generate a test signal that is able to capture the required statistical information with the minimum number of samples, as it will be detailed in chapters 4 and 5. Considering a LS extraction of the DPD coefficients, the proposed mesh-selecting methods will contribute to decrease the number of rows of the data matrix and thus reduce the computational load.

3.3.2 Artificial Intelligence Reinforced DPD Tuning

In a first approach to the use of artificial intelligence (AI) to help with the DPD design, the weblab-based DPD test mode will be widely used in this research, which contains a different way to organize and link the experimental set up, the digital processing computer (or the server in the cloud platform) and the physical researchers all together.

A time demanding part of DPD design is to choose a proper DPD behavioral model and then adjust its parameters. And it usually includes a combining of different methods together to compensate wider bandwidth and higher PAPR. This procedure not only requires some knowledge of DPD algorithm and experience adjusting the DPD linearizer, but also dealing with some non-logical phenomena, such as the tuning of CFR and internal signal gain. Thus, actually the trial and error is one of the popular approach in real DPD design/implementation, which results time consuming and sometimes leads to unpredictable results. Besides, for obvious reasons, usually it is not possible to try all the parameter configurations by hand in the laboratory and test them properly with large amount of signals to have statistically representative results. The final result is often a trade-off between performance and time limitation.

An artificial intelligence structure will be established as an auto-design platform. The global optimization methods (e.g., the simulated annealing, genetic algorithms, adaLIPO or ant colony algorithms) will be first implemented to match PAs with DPD algorithms and tuning the parameters automatically. The algorithm will automatically test different DPD models with different parameters combined with a large amount of representative test signals. The chosen DPD behavioral model and its adjusted parameters will appear as the solution when the artificial intelligence algorithm finishes all its test and analysis.

The state of the art in AI strongly rely on the computing power. An effective AI approach, such as YOLO v3 [YOL] in the field of object detection, requires large amount of storage space and parallel computing power, which ends up with random access memory (RAM) space and a big number of graphics process unit (GPU). Thus, a cloud-based platform is better for implementing the AI software rather than to buy more and more computers in the microwave laboratory. The prerequisite is to link the RF platform containing the device-under-test (DUT) to the cloud-based AI software. The concept of a webLab provides an internet data interface to exchange I-Q data between a RF test-bench containing the DUT and the digital processing machine where the base-band processing algorithms will run. In this thesis, most of experimental results were obtained in webLabs with different PAs: Class-AB PAs, class-J PAs or even load modulated balanced amplifiers. The details of the webLabs used to obtain the experimental results are given in section 5.4 of Chapter 5

The long-term objective (or the prospect vision) is that multiple remote test-beds will be allowing access to the artificial intelligence core algorithm (AICA). Researchers from all over the

world will be able to upload their new idea of DPD model or linearization method. The AICA will test the new methods/models with all available PAs (test-beds) automatically and then will show the researcher a wide variety of results to evaluate the potential linearization performance. PAs designers will receive a suitable DPD solution to linearize their designs that will be chosen automatically by AICA selected from the most relevant academic publications.

In this dissertation, the author is focusing on the basic approach of the vision. The global optimization algorithm is implemented to optimize power amplifier behavior modeling, DPD linearization and high complex PA tuning. The analysis and experimental results will be shown in Chapter 6 and Chapter 7. The result of global optimization algorithm illustrates the advantage of applying artificial intelligence in the digital communication front-end and also envisages the benefits of the AICA vision.

Chapter 4

Under-sampling Effects and Complexity Reduction Techniques for the Parameters Identification Subsystem

4.1 Introduction

In adaptive digital predistortion (DPD) linearization, one of the key elements is the observation or feedback path. It has to be linear to allow a correct identification of the DPD coefficients and, in order to be competitive, it has to be as cheap as possible.

On the one hand, several efforts [Zhu92, Fra96, Bra11] have been dedicated to overcome the standard approach where, the power amplifier output signal is down-sampled and digitized at a sampling rate that is several multiples of the Nyquist rate to capture the out-of-band distortion (bandwidth expansion 5x according to the rule of thumb) without aliasing. Since the price of an ADC increases with the sampling rate and resolution, it is possible to reduce costs by using under-sampling ADCs in the observational path. For example, considering a 10 MHz band-pass bandwidth LTE signal and a bandwidth expansion after the PA of 5 times, we may need a minimum sampling rate of 50 MS/s. However, as it will be shown later in the paper it is possible to identify the coefficients with significantly lower sampling rates reducing the cost of the ADC several factors (e.g. using a 4.25 MS/s instead of a 105 MS/s ADC with the same resolution can reduce costs at least by a factor of 5).

On the other hand, the adaptation algorithm has to be as computationally efficient as possible. As discussed in Chapter 3, in PA behavioral modeling or DPD linearization, the mathematical model consists of a set of basis waveforms from nonlinear transformations of the input (or several inputs) signals. To estimate the coefficients describing the PA behavioral model or its

inverse DPD function, we have to address a quadratic minimization problem, which typically is conducted by a least squares (LS) o recursive least squares (RLS) linear regression, since these models are composed by nonlinear basis functions but linear in parameters.

The number of parameters required by these behavioral models grows dramatically when considering memory effects. This has a negative impact in the LS-based model extraction because it increases the computational complexity and compromises the conditioning of the LS estimation. Because the data samples used for the LS regression are not independent from each other, the LS estimator needs a large number of data samples (or equations) to obtain the best approximation of the model coefficients. As reported in [Gua12c], using short sequences of data samples leads to two consequences: a) the ill-conditioning problem due to the rank deficiency of LS matrices; and b) the statistical mismatch problem, because the short data sequence often cannot fully represent (accuracy problem) the statistical property of the PA output signal.

Some efforts have been made to solve the ill-conditioning problem as well as to reduce the model extraction errors when using a small number of data samples to characterize PA behavioral models [Gua12c]. Alternatively, reducing the order of the DPD model has beneficial effects in both the computational complexity and in the conditioning of the data matrices. The idea of using dominant eigenvalues/eigenvectors for reducing the model order was presented in [Gil13b].

In this Chapter we present a first attempt of reducing the computational complexity of the least squares (LS) identification of the parameters describing PA behavioral models. To reduce the dimensions of the input data matrix, two strategies are proposed:

i. Model order reduction based on the principal component analysis (PCA) theory.

ii. Apply a mesh-selecting method to reduce the number of required equations.

In this context, the effect of using under-sampling ADCs for the LS parameter extraction aiming at reducing the costs of PA identification is also discussed. Finally, the trade-off between the cost/complexity reduction and quality (or identification accuracy) loss is evaluated. The proposed strategies can also be considered for low-computational cost digital predistortion implementations.

As it will be described in the following, in this Chapter we combine an AM-AM mesh-memoryless selecting method to reduce the number of rows (i.e., equations) with the PCA feature selection technique, oriented at reducing the number of columns (i.e., basis functions) of the data matrix. Chapter 5 will focus in the reduction of the number of required samples to extract the PA or DPD behavioral models, proposing improved mesh-selecting techniques. We will combine these mesh-selecting techniques with dimensionality reduction techniques to reduce the computational complexity of the adaptation subsystem. The topic of dimensionality reduction techniques, however, falls out of the main scope of this thesis. Instead, we will used

the research developed by a member of our research group, Dr. Thi Quynh Anh Pham, in her PhD Thesis [Pha19c].

4.2 Complexity Reduction Techniques for the Parameters Identification Subsystem

4.2.1 The Application of Under-Sampling ADC

The Shannon-Nyquist theorem ensures perfect reconstruction of a band-limited signal from samples taken at rate of $2f_{max}$. According to [Fra96], for modeling a Volterra-based nonlinear system excited by a band-limited signal, it is sufficient to sample the continuous-time input and output signals at twice the maximum input frequency rather than double the output frequency. As a consequence, we may only need to sample the extended bandwidth (due to the spectral regrowth caused by the PA nonlinear behavior) of the output signal at twice its maximum baseband input frequency.

At this point, we will define the under-sampling ratio (USR) as the ratio between the Nyquist-rate ADC sampling frequency required for capturing the extended PA output signal ($f_{s_{Extended\ BW}}$) and the actual ADC sampling frequency ($f_{s_{ADC}}$) employed to capture the PA output signal:

$$USR = \frac{f_{s_{Extended\ BW}}}{f_{s_{ADC}}} \tag{4.1}$$

As it will be shown later in the experimental results section, since the PA output signal will be only used for the parameter extraction (we do not need to reconstruct it in the analog domain), the aliasing will not be an issue. Therefore, generally speaking, there is no USR limit in order to be capable of extracting the PA behavioral model (or DPD) parameters. The only constraint for a good identification comes from the minimum amount of equations (or samples) required for the LS extraction. The number of linear independent equations relies on the ability to capture the statistics of the output signal. Therefore, the accuracy of the LS estimation will not rely on the sampling frequency, but on the fact that the number of captured samples are enough statistically rich. As a consequence, if we will use very high USR, the only problem is that it will take more time to gather a representative set of equations to perform an accurate LS identification.

4.2.2 Dimensionality Reduction Using Principal Component Analysis

Principal component analysis (PCA) is is a statistical learning technique suitable for converting a basis of observed and eventually correlated data into a basis of uncorrelated data, named principal components. The principal components are linear combinations of the original basis functions oriented at capturing the maximum variance of the data contained in the data matrix.

Consequently, it is possible to apply dimensionality reduction by discarding those components (eigenvectors) with smaller eigenvalue.

An example on how to apply the PCA theory to reduce the model order of a DPD was published by the authors in [Gil13a]. With this technique, we can perform a change of basis where the number of required coefficients decreases by a certain reduction factor (RDF), i.e., number of coefficients of the original basis divided by the number of coefficients of the new basis. The reduced order LxR data matrix $\hat{\boldsymbol{X}}$ is defined as

$$\hat{\boldsymbol{X}} = \boldsymbol{X}\boldsymbol{P} \tag{4.2}$$

where $\boldsymbol{X}$ is the $L \times O$ original data matrix and $\boldsymbol{P}$ is the $O \times R$ transformation matrix obtained by finding the eigenvalues and eigenvectors of the covariance matrix $\boldsymbol{X}^{\boldsymbol{H}}\boldsymbol{X}$.

With PCA, thanks to the orthogonality property of the resulting transformed matrix, the DPD coefficients extraction can be carried out with simple dot products (avoiding the Moore-Penrose matrix inversion in the LS solution). In addition, by using the adaptive PCA technique [LB18] it is possible to apply dimensionality reduction in the coefficients estimation by selecting only the minimum necessary number of principal components required to meet the target linearity levels, specified in terms of adjacent channel power ratio (ACPR) and normalized mean square error (NMSE).

4.2.3 AM-AM Memoryless Mesh-Selecting Method

For PA behavioral modeling and DPD linearization, it is essential to collect enough statistically representative input and output data samples. Normally, this data collection is carried out by taking all the samples in a given time period and taking into account a certain sampling rate. If the number of collected samples is not sufficiently high, it may not be enough to properly represent the PA nonlinear behavior and its dynamics (accuracy problem). Besides, the number of collected samples is directly related to the computational complexity of the LS extraction. In order to reduce the input data matrix dimensions, we propose a method that consists in creating a mesh to properly select the most significant input-output data according to the statistical distribution of the PA AM-AM characteristic.

The principle of this mesh selecting method consists in dividing the AM-AM characteristic into several segments and calculating the number of samples in every segment, obtaining thus a two-dimensional histogram of the PA AM-AM characteristic, as shown in Fig. 4.1. Then, in the mesh-selecting process, the data is collected according to the probability of each segment of the mesh. Therefore, for a given or fixed amount of data (or number of equations) to extract the PA coefficients, instead of using sets of consecutive data samples (as it would be the case shown in Fig. 4.2), by using the mesh-selecting method the data is gathered with the objective of fitting

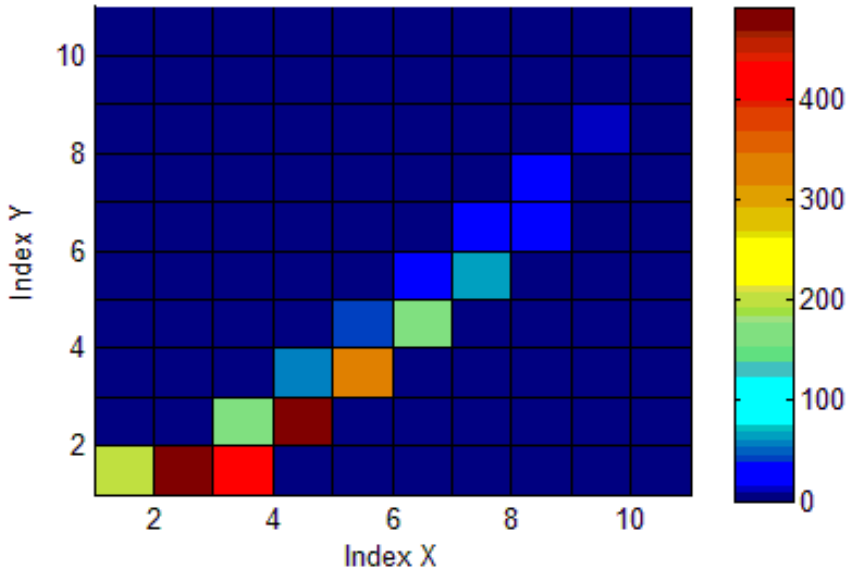

Figure 4.1: 2-D histogram of the PA AM-AM characteristic taking into account 25, 600 samples.

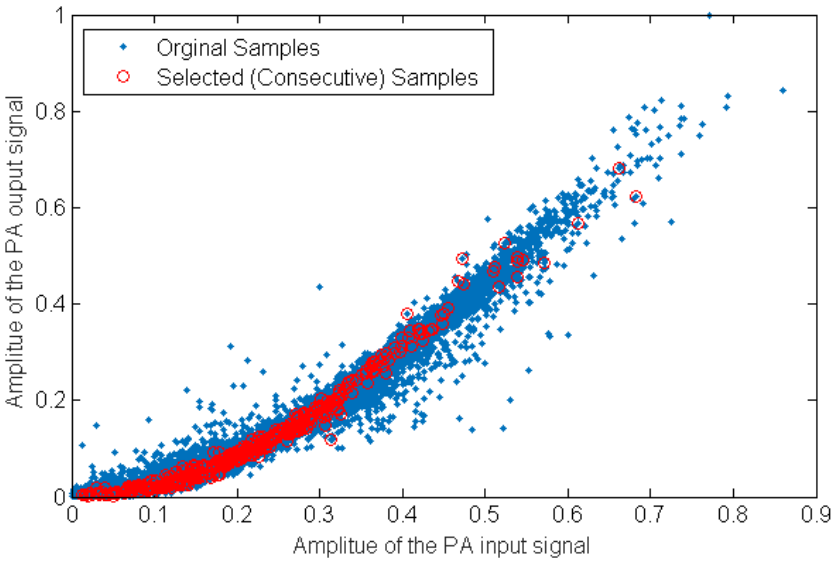

Figure 4.2: Collection of a fixed number of consecutive data samples for the LS coefficients extraction.

the original PA AM-AM characteristic histogram (see Fig. 4.3). As depicted in Fig. 4.3, with the same number of collected samples as in Fig. 4.2, thanks to the mesh selecting method it is possible to better represent the PA AM-AM characteristic histogram.

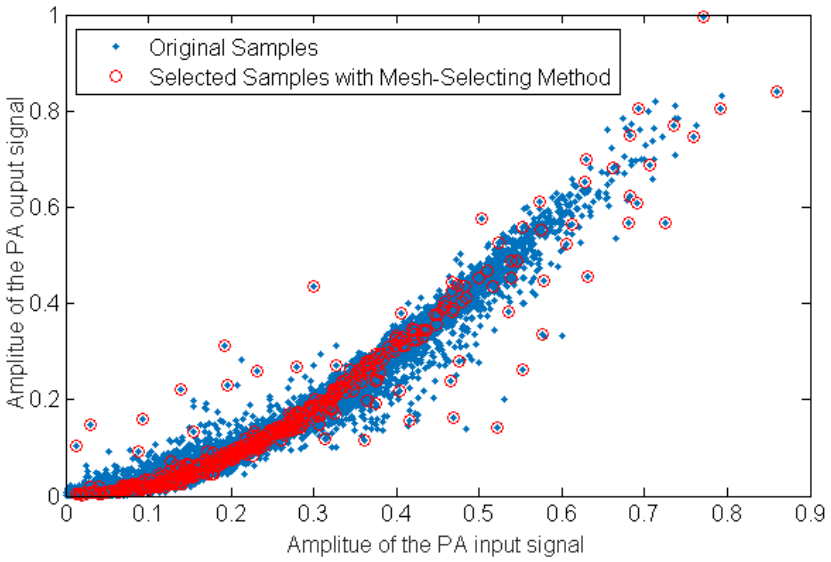

Figure 4.3: Collection of a fixed number of data samples for the LS coefficients extraction using the mesh-selecting method.

4.3 2-Dimensional MISO Model for Envelope Tracking PA behavioural modeling and DPD linearization

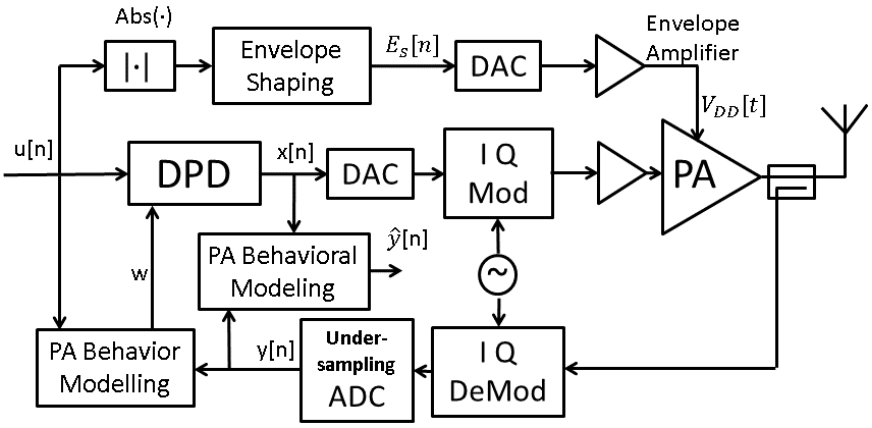

Figure 4.4: Block diagram of an envelope tracking PA system with DPD.

The block diagram of the envelope tracking (ET) system with DPD that we have used to capture input-output data records to run the experimental tests is depicted in Fig. 4.4 and a photo of the test-bench is shown in Fig. 4.5. This system is intended to be used for DPD purposes, however, for the sake of simplicity, in this Chapter we will be showing results for PA behavioral modeling, while in Chapter 5, results using mesh-selecting techniques for DPD linearization will be shown.

The well-known and widely used metrics to evaluate the accuracy of the behavioral models

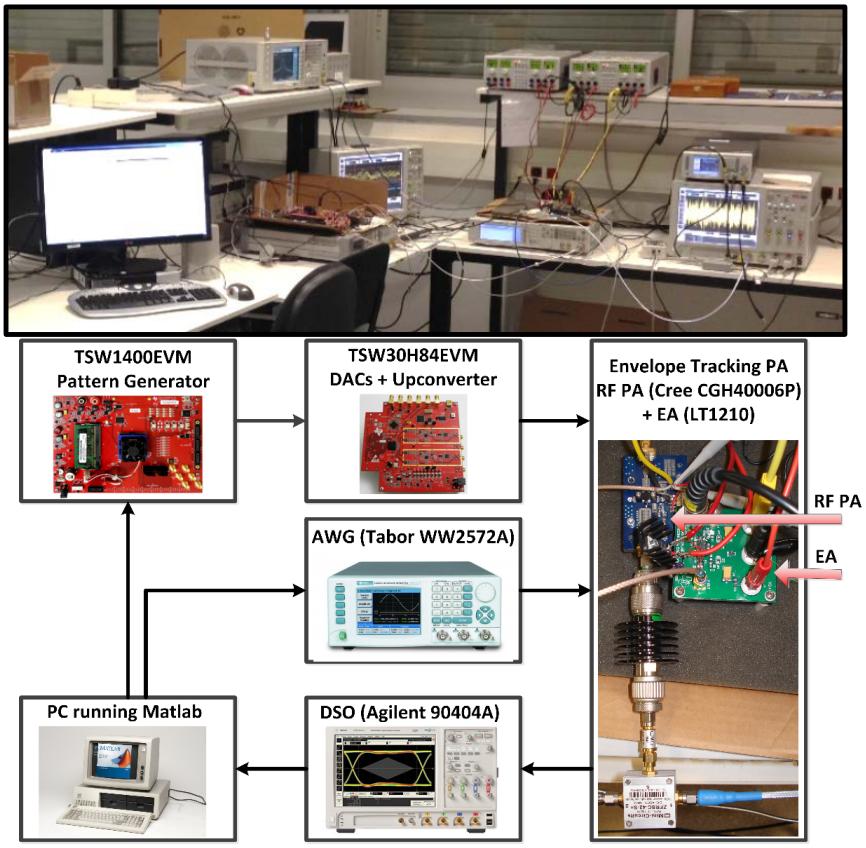

Figure 4.5: Picture of the experimental ET PA test-bench [Gil15].

are the normalized mean square error (NMSE) and the adjacent channel error power ratio (ACEPR). Finally, the test signal is a LTE signal with 10 MHz and an extended bandwidth after the PA amplification of around 50 MHz. To properly characterize the nonlinear distortion of the PA with dynamic supply we will consider the behavioral model described in [Gil12b], with the following input-output relationship.

$$\hat{y}[n] = \sum_{j=0}^{M-1} \sum_{q=0}^{Q-1} \sum_{i=0}^{N-1} \sum_{p=0}^{P-1} \alpha_{piqj} \ \phi_{piqj}[n] \tag{4.3}$$

Note that two-dimensional (2-D) basis functions are used to characterize the ET PA behavioral modeling:

$$\phi_{piqj}[n] = (E_s[n-\tau_j])^q \ x[n-\tau_i] \ |x[n-\tau_i]|^p \tag{4.4}$$

with $E_s[n]$ being the envelope of the input signal after the shaping function [Mon10b], $x[n]$ being the input signal and τ being the most significant sparse delays. Following a matrix notation, the $L \times 1$ estimated output vector results

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{w} \tag{4.5}$$

with $\boldsymbol{w} = \left(\alpha_{0000}, \cdots, \alpha_{(M-1)(Q-1)(N-1)(P-1)}\right)^T$ being the $O \times 1$ vector of coefficients of the model and where $O = P \cdot N \cdot Q \cdot M$. The input data matrix

$$\boldsymbol{X} = \begin{pmatrix} \phi_{0000}[0] & \cdots & \phi_{(M-1)(Q-1)(N-1)(P-1)}[0] \\ \vdots & & \vdots \\ \phi_{0000}[n] & \cdots & \phi_{(M-1)(Q-1)(N-1)(P-1)}[n] \\ \vdots & & \vdots \\ \phi_{0000}[L-1] & \cdots & \phi_{(M-1)(Q-1)(N-1)(P-1)}[L-1] \end{pmatrix}$$

has dimensions $L \times O$, where L is the number of data samples or equations. Finally, the least squares solution is defined as,

$$\boldsymbol{w} = (\boldsymbol{X}^{\boldsymbol{H}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{H}}\boldsymbol{y} \tag{4.6}$$

with $\boldsymbol{y}$ being the $L \times 1$ output data vector captured with the under-sampling ADC, and where the error vector is $\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{y}}$.

4.4 Experimental Results

Fig. 4.6 shows the NMSE and ACEPR for different reduction factors (RDFs) when applying model order reduction based on the PCA theory. The original number of coefficients was $O = 144$. Without PCA reduction ($RDF = 1$), the LS identification may be ill-conditioned when considering a small number of equations, while for low RDF values (e.g. $RDF = 3 \rightarrow R = 48$ coefficients) the ACEPR and NMSE are kept almost the same. Unfortunately, for high RDF values the identification accuracy will be degraded.

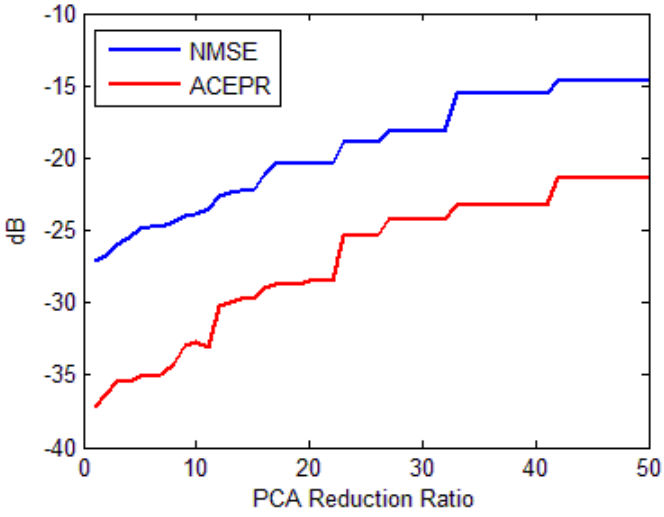

Figure 4.6: NMSE and ACEPR for different RDFs.

To evaluate the data matrix dimensions taking into account the complexity of the model (in terms of number of coefficients), we define the data matrix ratio as the ratio between the number of equations (rows) and the number of coefficients (columns).

Fig. 4.10 and Fig. 4.8 show the NMSE and ACEPR for different USRs taking into account different number of equations. The PA behavioral model after PCA model order reduction has 48 coefficients. As observed, at least 1000 equations (matrix ratio of around 20) are required in order to obtain good NMSE and ACEPR values. Moreover, the variability shown when considering a fixed number of equations is not due to the USR, or for not being compliant with Nyquist. Instead, this variability is due to the fact that for some specific set of captured data it is not possible to properly represent the statistical property of the PA output signal.

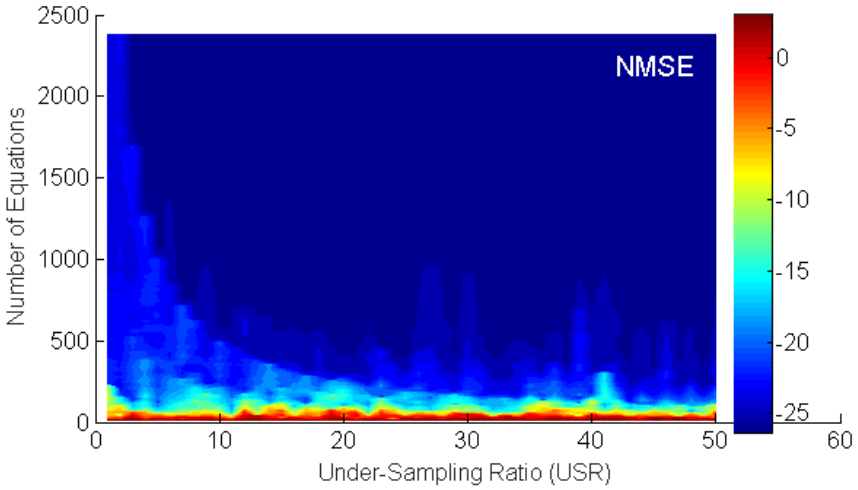

Figure 4.7: NMSE for different USRs and data matrix sizes.

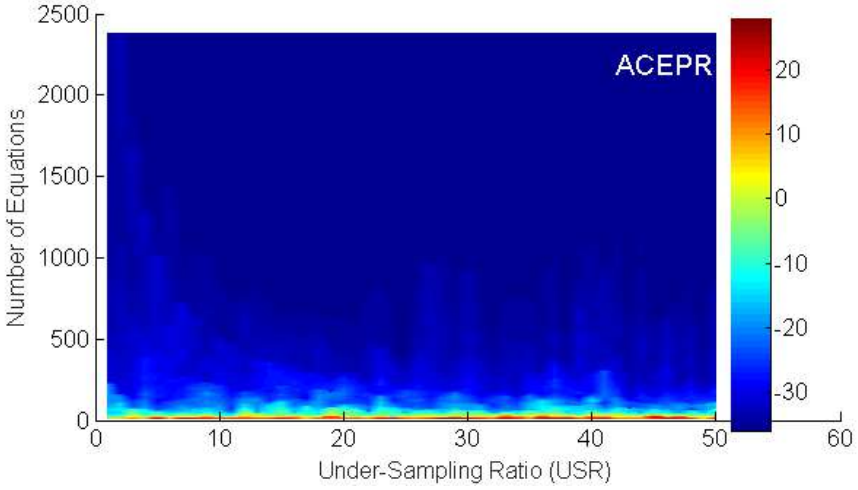

Figure 4.8: ACEPR for different USRs and data matrix sizes.

Fig. 4.6 shows the NMSE and ACEPR for different PCA reduction factors and number of equations. It can be observed that for low values of PCA it is necessary to use a lot of equations and yet we can end up with an ill-conditioned LS identification. On the other hand, as shown before, the NMSE/ACEPR is degraded with high RDFs no matter how many equations we use.

Finally, with a proper RDF it is possible to keep the NMSE/ACEPR accuracy and significantly reduce the number of required equations necessary to extract the behavioral model coefficients.

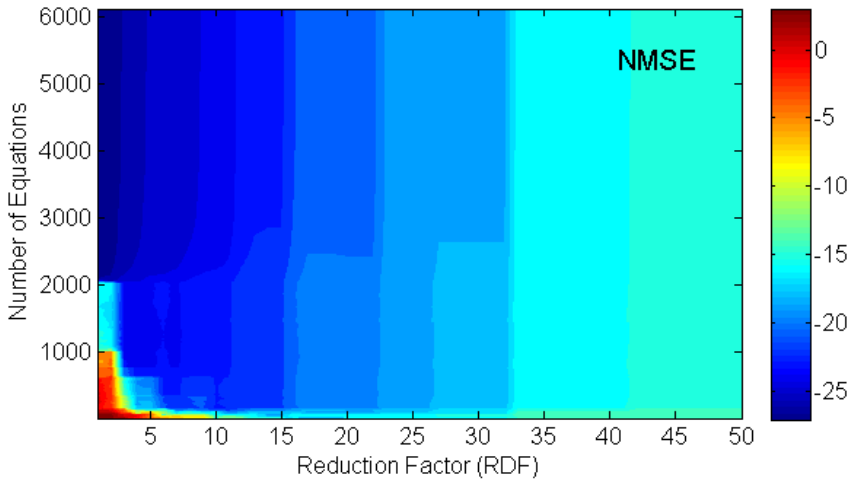

Figure 4.9: NMSE for different RDFs and data matrix sizes.

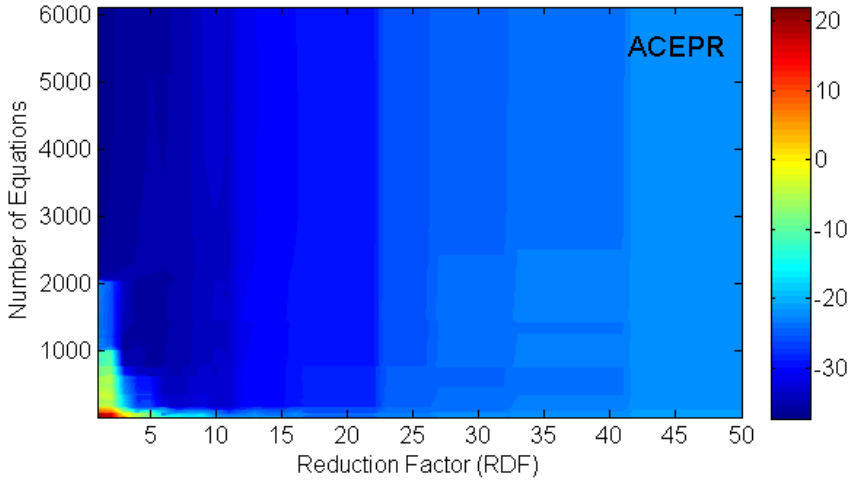

Figure 4.10: ACEPR for different RDFs and data matrix sizes.

The performance of the proposed AM-AM memoryless mesh-selecting method is depicted in Fig. 4.11 and Fig. 4.12. The NMSE and ACEPR values are plotted for different data matrix ratios and considering different mesh densities. The advantage of using the mesh-selection method in comparison to using the consecutive original data, is that we can slow-down the drop of accuracy (measured in terms of NMSE and ACEPR) when decreasing the matrix ratio. In other words, thanks to the proposed AM-AM memoryless mesh-selecting method we can reduce the number of required equations to extract the parameters of the PA behavioral model and thus reduce the computational complexity of the LS identification. As expected, the higher the mesh density the better the NMSE/ACEPR for a given matrix ratio. However, the improvement achieved

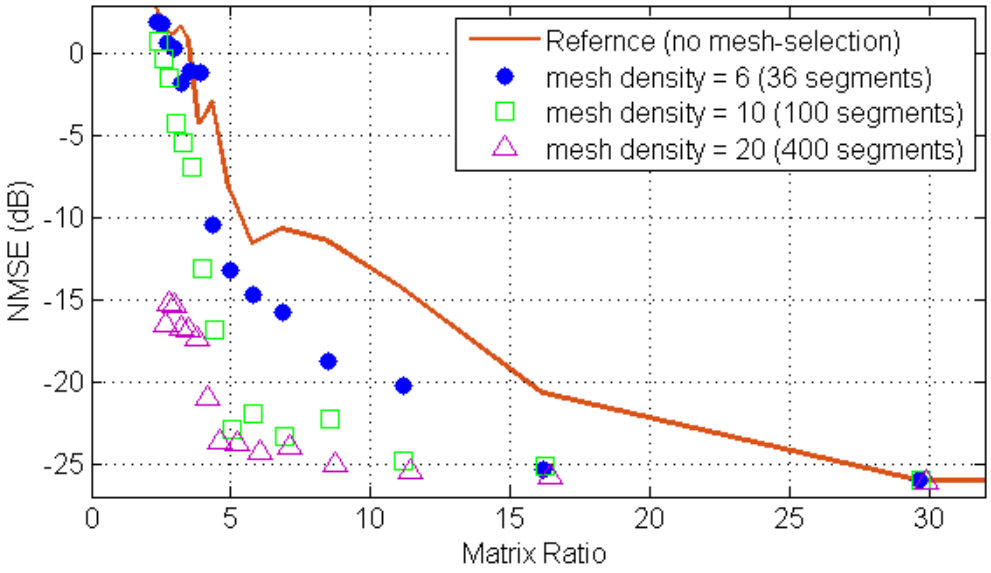

Figure 4.11: NMSE for different data matrix sizes considering the AM-AM memoryless mesh-selecting method with different mesh densities

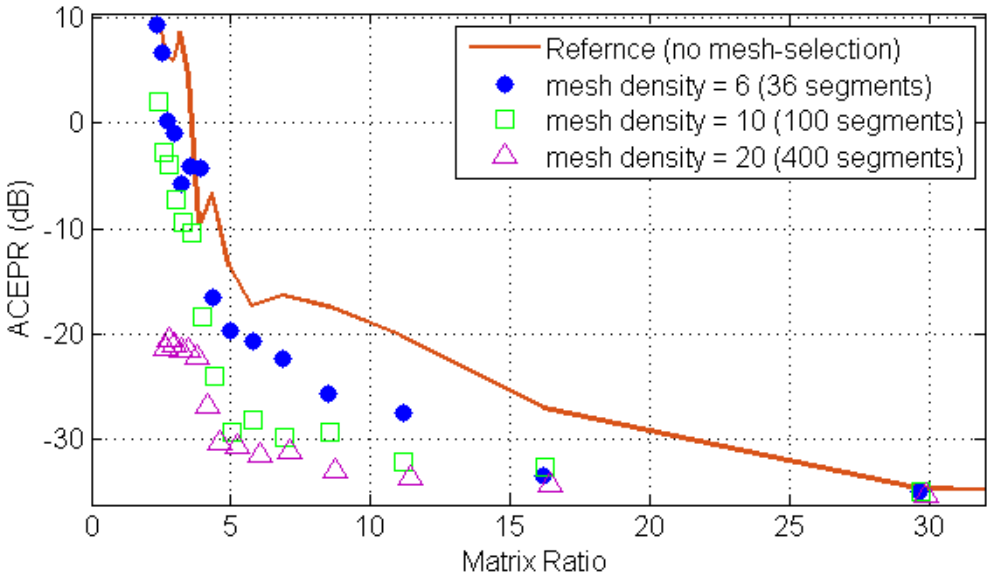

Figure 4.12: ACEPR for different data matrix sizes considering the AM-AM memoryless mesh-selecting method with different mesh densities

when considering high density mesh values is not significant enough to compensate for the extra complexity introduced.

4.5 Conclusion

In this Chapter we have discussed how there is not apparent limit for the maximum USR that can be applied to properly identify the coefficients of a PA behavioral model. Instead, the key factor relies on both the basis functions (and thus number of parameters) used to characterize the PA nonlinear behavior; and the minimum amount of equations (or samples) required for the LS extraction. Consequently, it should be possible to significantly relax the specifications (and costs) of the ADC.

By using model order reduction based on the PCA theory combined with a proper selection of the equations through the AM-AM memoryless mesh-selecting method, we can significantly reduce the data matrix dimensions as well as to improve the matrix conditioning. However, the price for reducing the computational complexity is some accuracy loss. Because of this trade-off, the amount of quality loss that we can allow will determine the bottleneck for computational cost reduction. For example, considering our 2-D PA behavioral model, an original data matrix of dimensions 3500×144 can be reduced up to 384×48 with less than 3 dB of accuracy loss in both NMSE and ACEPR.

In next Chapter we will present and compare alternative mesh-selecting methods and we will take the I-Q with memory mesh-selecting method as a reference to properly combine it with another feature extraction dimensionality reduction technique in the context of DPD linearization.

Chapter 5

Enhanced Mesh-Selecting Methods

5.1 Introduction

The parameters identification of power amplifier (PA) behavioral models or digital predistortion (DPD) linearizers rely on measured input-output data observations and the specific mathematical models chosen to characterize the PA nonlinear behavior and its dynamics. These models are usually described as linear combinations of nonlinear basis functions, defining a linear regression problem where the data matrix is tall, i.e., significantly more data samples than basis functions. Therefore, the problem of behavioral modeling or DPD linearization has no exact solution since it is over-determined (i.e. more equations than unknowns).

While the method of least squares (LS) performs well to extract the coefficients of over-determined systems when considering big data sets, it may face the risk of over-fitting when the behavioral model contains more parameters than the model really needs. In addition, when the number of training data samples is not statistically rich enough, the LS solution may suffer from uncertainty leading to an inaccurate parameter extraction. Reducing the computational complexity in the identification subsystem and avoid the poor predictive performance derived from over-fitting or the lack of significant training data is one of the objectives pursued in the field of PA behavioral modeling and DPD linearization.

Reducing the tall data matrix in the dimension of the columns (i.e., basis functions) has been widely addressed in literature to both reduce the computational complexity and avoid over-fitting problems (e.g., [Pha19b, LB18, Pha19c, RT15]). Reducing the number of rows (i.e., equations) has been indirectly addressed by several papers targeting undersampling strategies to reduce the sampling rate of the A/D converters in the observation path. For example, techniques based on filtering the feedback signal before sampling and then applying some additional processing to be able to achieve full-band distortion cancellation in the forward path [Liu15, Bra15]; or directly linearizing the PA within a limited bandwidth [Yu12]; or even including techniques

based on poly-phase filters to reduce the DPD clock rate in the forward path [Li20]. In this dissertation however, we assume that the behavioral model coefficients can be extracted with aliased signals [Wan15a,Wan17] obtained by applying undersampling techniques with no filtering in the observation path.

Assuming, therefore, that the PA or DPD behavioral model coefficients can be extracted with aliased signals obtained by applying undersampling techniques, some sample selection methods were presented in [Wan15a, Kra20]. The genetically optimised histogram (GOH) in [Kra20], calculates an optimized histogram of signal magnitudes taking into account both PA input-output signal characteristics. The GOH method shows very good sample reduction performance but at the price of a high computational cost required for the optimisation. In the previous Chapter we have presented a mesh-selecting (MeS) method [Wan15a] where the selection of samples is carried out taking into account the PA AM-AM static characteristic.

In this Chapter we first define and compare alternative mesh-selecting methods to the one proposed in Chapter 4, by including memory and re-defining the signals involved to generate the mesh. Later, we focus in the in-phase (I) and quadrature (Q) with memory MeS method that significantly improves the reduction performance of our previous work in [Wan15a] when considering wideband signals. In addition, unlike in [Kra20], no optimization process is required and the selection of samples only relies on the transmitted data characteristics, being therefore agnostic to the PA used. Consequently, it is significantly less computationally complex than the GOH method in [Kra20], but requires generating a bigger mesh than in [Wan15a], since memory is included in the mesh-selection to ensure the tracking of very fast amplitude variations. Experimental results will show how the proposed I-Q with memory MeS method can significantly reduce the overall computational load in the parameters extraction, it is independent on the solver used and it can be properly combined with dimensionality reduction techniques.

5.2 General Description of PA Behavioral Modeling and DPD Linearization with Mesh-Selecting

In this section we review the mathematical description of the PA behavioral modeling and DPD linearization in section 3.1.2 taking into account a general mesh-selecting method used to select the most relevant samples to estimate the PA or DPD model coefficients.

Following the notation in Fig. 5.1, the estimated PA behavioral model output $\hat{y}[n]$ (for $n = 0, 1, \cdots, N-1$), can be defined following a matrix notation as

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{w_{pa}} \tag{5.1}$$

where $\boldsymbol{w_{pa}} = \left(w_1^{pa}, \cdots, w_i^{pa}, \cdots, w_M^{pa}\right)^T$ is the $M \times 1$ vector of parameters and $\boldsymbol{X}$ is the $N \times M$

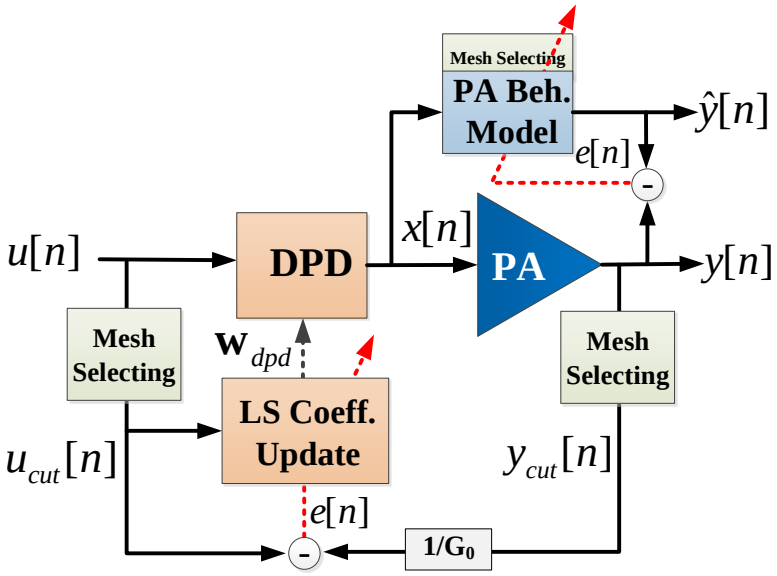

Figure 5.1: PA behavioral modeling and DPD linearization with mesh-selecting.

data matrix (with $N \gg M$) containing the basis functions. The data matrix is defined as

$$\boldsymbol{X} = \Big(\boldsymbol{\varphi}_x[0], \boldsymbol{\varphi}_x[1], \cdots, \boldsymbol{\varphi}_x[n], \cdots, \boldsymbol{\varphi}_x[N-1]\Big)^T \tag{5.2}$$

where $\boldsymbol{\varphi}_x[n] = \Big(\phi_1^x[n], \cdots, \phi_i^x[n], \cdots, \phi_M^x[n]\Big)^T$ is the $M \times 1$ vector of basis functions $\phi_i^x[n]$ (with $i = 1, \cdots M$) at time n (with $n = 0, \cdots, N-1$). This general equation can be particularized for any behavioral model.

After applying the mesh-selecting method, the least squares (LS) solution to extract the coefficients $\boldsymbol{w_{pa}}$ is given by

$$\boldsymbol{w_{pa}} = (\boldsymbol{X_{cut}}^H \boldsymbol{X_{cut}})^{-1} \boldsymbol{X_{cut}}^H \boldsymbol{y_{cut}} \tag{5.3}$$

where the $T \times M$ matrix $\boldsymbol{X_{cut}}$ (with $T \ll N$) is the data matrix $\boldsymbol{X}$ containing the basis functions after a proper selection of equations (rows), for example:

$$\begin{aligned}\boldsymbol{X_{cut}} = \Big(&\boldsymbol{\varphi}_x[0], \boldsymbol{\varphi}_x[1], \varphi_x[2], \cdots, \cancel{\boldsymbol{\varphi}_x[n-1]}, \cancel{\boldsymbol{\varphi}_x[n]}, \\ &\boldsymbol{\varphi}_x[n+1] \cdots, \cancel{\boldsymbol{\varphi}_x[N-3]}, \cancel{\boldsymbol{\varphi}_x[N-2]}, \cancel{\boldsymbol{\varphi}_x[N-1]}\Big)^T\end{aligned} \tag{5.4}$$

and with $\boldsymbol{y_{cut}}$ being the corresponding output data vector after selecting the most relevant equations,

$$\begin{aligned}\boldsymbol{y_{cut}} = \Big(&y[0], y[1], y[2], \cdots, \cancel{y[n-1]}, \cancel{y[n]}, \\ &y[n+1] \cdots, \cancel{y[N-3]}, \cancel{y[N-2]}, \cancel{y[N-1]}\Big)^T\end{aligned} \tag{5.5}$$

Similarly, as shown in Fig. 5.1, the input-output relationship in the DPD forward path is described as

$$\boldsymbol{x} = \boldsymbol{u} - \boldsymbol{U}\boldsymbol{w}_{\boldsymbol{dpd}} \tag{5.6}$$

where $\boldsymbol{x} = (x[0], \cdots, x[n], \cdots, x[N-1])^T$ and $\boldsymbol{u} = (u[0], \cdots, u[n], \cdots, u[N-1])^T$ are the pre-distorted and input $N \times 1$ signal vectors, respectively. Moreover, $\boldsymbol{w}_{\boldsymbol{dpd}}$ is the $M \times 1$ vector of coefficients and the $N \times M$ data matrix $\boldsymbol{U}$, similarly to $\boldsymbol{X}$, contains the basis functions of the DPD, $\phi_i^u[n]$ (with $i = 1, \cdots M$), at time n. Targeting an adaptive DPD direct learning approach, the LS solution to extract the DPD coefficients after mesh-selecting is defined as

$$\boldsymbol{w}_{\boldsymbol{dpd}}^{(\boldsymbol{j+1})} = \boldsymbol{w}_{\boldsymbol{dpd}}^{(\boldsymbol{j})} + \mu(\boldsymbol{U}_{\boldsymbol{cut}}{}^H \boldsymbol{U}_{\boldsymbol{cut}})^{-1} \boldsymbol{U}_{\boldsymbol{cut}}{}^H \boldsymbol{e}_{\boldsymbol{cut}} \tag{5.7}$$

where the $T \times M$ matrix $\boldsymbol{U}_{\boldsymbol{cut}}$ (with $T \ll N$) is the data matrix $\boldsymbol{U}$ after mesh-selecting,

$$\begin{aligned} \boldsymbol{U}_{\boldsymbol{cut}} = \Big(&\boldsymbol{\varphi}_u[0], \boldsymbol{\varphi}_u[1], \varphi_u[2], \cdots, \cancel{\boldsymbol{\varphi}_u[n-1]}, \cancel{\boldsymbol{\varphi}_u[n]}, \\ &\boldsymbol{\varphi}_u[n+1] \cdots, \cancel{\boldsymbol{\varphi}_u[N-3]}, \cancel{\boldsymbol{\varphi}_u[N-2]}, \cancel{\boldsymbol{\varphi}_u[N-1]}\Big)^T \end{aligned} \tag{5.8}$$

and with the residual error after mesh-selecting, $\boldsymbol{e}_{\boldsymbol{cut}} = \boldsymbol{y}_{\boldsymbol{cut}}/G_0 - \boldsymbol{u}_{\boldsymbol{cut}}$, being

$$\begin{aligned} \boldsymbol{e}_{\boldsymbol{cut}} = \Big(&e[0], e[1], e[2], \cdots, \cancel{e[n-1]}, \cancel{e[n]}, \\ &e[n+1] \cdots, \cancel{e[N-3]}, \cancel{e[N-2]}, \cancel{e[N-1]}\Big)^T \end{aligned} \tag{5.9}$$

5.3 Improvements to the Mesh-Selecting Method

The basic mesh selecting method presented in Chapter 4 consisted in obtaining a two-dimensional histogram of the PA AM-AM static characteristic. Therefore, instead of selecting sets of consecutive data samples we proposed selecting samples with the objective of fitting the original PA AM-AM characteristic histogram. The proposed method worked well for a specific PA operated with a narrow-band signal and combining it with the principal component analysis (PCA) reduction technique. However, as it will be shown in the following, when considering wide-band signals, the reduction performance of this specific memoryless mesh-selecting method is limited. Therefore, with the new mesh-selecting approaches that will be presented in this Chapter, we can outperform the reduction capabilities achieved by the previous AM-AM mesh-selecting method.

In modern telecommunication systems using digital modulations (e.g., in 5G NR or former LTE, etc.) the RF signal is modulated from complex amplitude and phase modulated signals, i.e., from base-band signals defined by their I-Q components. Therefore, one of the improvements of the new mesh-selecting method will consist in considering the numerical value of each I-Q component separately to build the mesh.

In addition, as discussed in Chapter 3, current PA and DPD behavioral models include memory terms to characterize the PA memory effects. Consequently, the distribution of the input

signal should contain not only the information of the present sample but also of past samples. Thus, another improvement that can included in the mesh-selecting algorithm is to extend the coverage taking into account historical samples, which is called mesh-selecting method with memory.

These two new approaches, i.e., considering the I-Q components and including past samples, will be combined in different mesh-selecting strategies, tested and compared in the following sections of this Chapter.

5.3.1 Mesh-Selecting Method with Memory

As previously discussed in Chapter 3 the general input to a given DPD function can be defined as $\boldsymbol{\nu} = \big(u_0, u_1, \cdots, u_k, \cdots, u_K\big)$, where $u_k = u[n - \tau_k]$ with $k = 0, \cdots, K$ and where τ_k (with $\tau_0 = 0$) are the most significant sparse delays characterizing the memory of the PA. Therefore, $\boldsymbol{\nu}$ is a $(K+1) \times 1$ vector of (multiple) input signals to the nonlinear DPD function f_{DPD}, with K defining the memory depth of this given f_{DPD}. When $K = 0$ the system is considered to be memoryless.

The result of the mesh-selecting is not related to the choice of the specific DPD function f_{DPD}, i.e., the mesh-selecting is an accurate description of the statistic feature of the f_{DPD} input domain independently of the function itself. Let's take K = 3 as an example. For a given

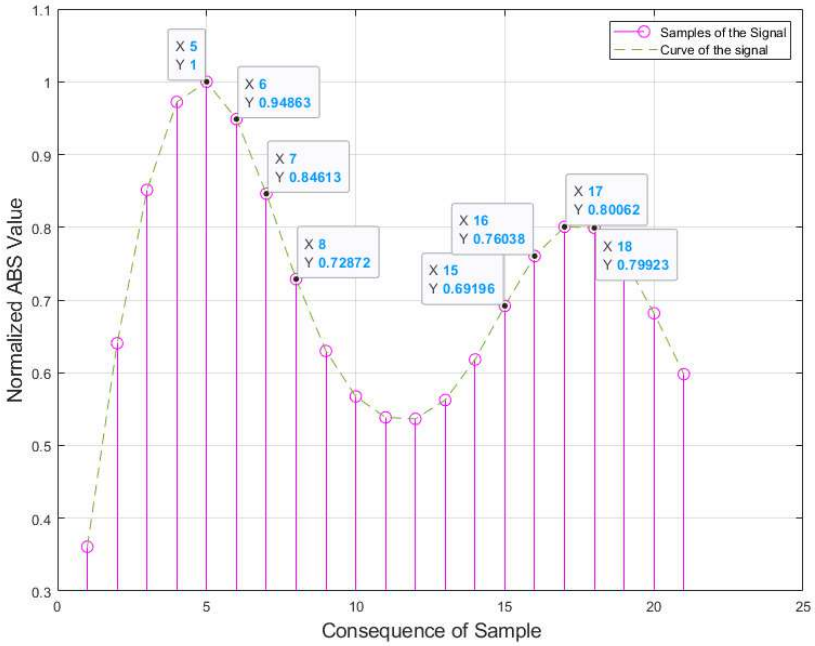

Figure 5.2: Part of a signal sequence

$\boldsymbol{\nu} = \big(u_6, u_7, u_8, u_9\big)$, the present sample is u_6 while u_7, u_8, u_9 are past samples (i.e., the historical ones). Let's assume the use of the simple mesh-selecting method described in Chapter 4 with a mesh density of 8 for values ranging from 0.3 to 1.1, as shown in Fig. 5.2. The sample u_6 is allocated in mesh segment 7, just like sample u_4. The index of u_4 and u_6 in the mesh will be 7, while the indexes of u_7, u_8 and u_9 are $6, 5$ and 4, respectively. From the perspective of $\boldsymbol{\nu} = \big(u_6, u_7, u_8, u_9\big)$, the index of $\boldsymbol{\nu}$ is the combination of indexes of $\big(u_6, u_7, u_8, u_9\big)$, i.e. $(7, 6, 5, 4)$. Since the mesh density is assumed to be 8 in this example, the index vector can be changed into an octal number O(7654) for simplicity.

In general, given a mesh-selecting method with memory, with a mesh density of D and a memory depth of K, an index vector of $(K+1) \times 1$ is used to describe the allocation of the equivalent $\boldsymbol{\nu} = \big(u_0, u_1, \cdots, u_k, \cdots, u_K\big)$. The index vector can be translated into a K-digit N-decimal number to defined the mesh index number. Further more, the distribution of this index number also represents the distribution of the original input vector $\boldsymbol{\nu} = \big(u_0, u_1, \cdots, u_k, \cdots, u_K\big)$.

5.3.2 I-Q Mesh-Selecting Method

Unlike the memoryless AM-AM mesh-selecting method in Chapter 4, the I-Q with memory mesh-selecting method proposed in this Chapter relies only on the transmitted signal characteristics. Consequently the mesh values do not depend on the PA output (it is independent on the system). Instead, the mesh is created relying only in the distribution of the complex input data and its memory which improves the reduction capabilities when considering wide-band signals.

The base tool of the proposed mesh-selecting method for evaluating the transmitted signal is inspired in the traditional histogram calculation (i.e., uniformly dividing the magnitude of the input signal values into different segments and counting the proportion of samples in each segment). However, the memoryless AM-AM mesh-selecting method, the I-Q with memory mesh-selecting method consists in evaluating each complex value of the multiple input vector to the PA behavioral model or DPD function (i.e., $f_{DPD}(u_0, u_1, \cdots, u_K)$) and build a multi-dimensional histogram (mesh) by addressing each independent value into its corresponding mesh bin. The number of segments or mesh bins is

$$q = 2^{2(K+1)n_{bits}} \tag{5.10}$$

with K being the memory depth (i.e., number of memory taps) and n_{bits} the number of bits defining the histogram resolution.

Fig. 5.3 outlines the principles of the I-Q mesh-selecting method but, in order to simplify the mesh representation, depicting only a memoryless approach. Therefore, the I-Q mesh-selecting method can be described as follows:

i. The multi-dimensional histogram is created considering both I-Q components. It is a bi-

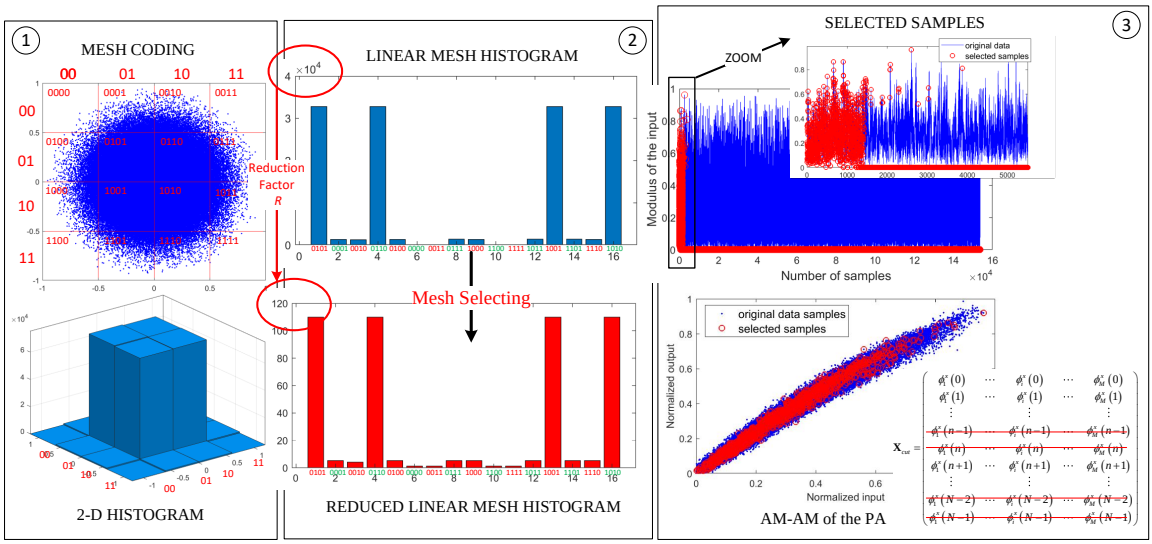

Figure 5.3: Principles of the mesh-selecting algorithm for a complex-valued input signal and considering a memoryless system.

dimensional histogram when considering a memoryless approach (as in the simplified example in Fig. 5.3), it will be a tetra-dimensional histogram if it includes one memory tap (u_1), etc.

ii. Each I-Q bi-dimensional histogram is converted into a uniform linear mesh histogram (i.e., taking into account all the bi-dimensional histograms for u_0, u_1, u_2, etc.). Then, the number of samples per bin are reduced by a factor R, by keeping the first T_i samples from the original N_i samples at mesh bin i. The minimum number of samples per mesh bin is 1. Thus, the number of samples per mesh bin is reduced proportionally respecting the original multi-dimensional distribution, i.e., $\frac{N_i}{T_i} \approx \frac{N_j}{T_j} \approx R$.

iii. Finally, with the remaining samples it is possible to build the corresponding data matrix ($\boldsymbol{X_{cut}}$ in (5.3) or $\boldsymbol{U_{cut}}$ in (5.7)) and select the corresponding output samples ($\boldsymbol{y_{cut}}$ in Fig. 5.1) to extract the coefficients of the PA or DPD model.

Regarding the way in which the method selects the samples inside each specific bin, again, for simplicity, is done sequentially. However, we could consider any other approach as long as we select the required number of samples that corresponds to this specific bin. The advantage of doing it this way is that, if we consider that we process the samples in order of appearance, the faster we have the minimum number of samples per bin, the faster we can proceed to the adaptation.

As discussed in Chapter 4, reducing the number of samples to carry-out the coefficients extraction impacts the computational complexity independently on the solver used. For example, the computational complexity of the QR decomposition using Givens rotation is $\mathcal{O}(NM^2)$

[Wes96]. While the computational complexity of using a feature extraction dimensionality reduction approach such as the principal component analysis combined with dynamic partial least squares (PCA-DPLS) for the coefficients identification is approximately $\mathcal{O}(NML + M^2 + ML)$. As proved in [Pha19b], if the number of estimated transformed coefficients is significantly smaller than the number of coefficients in the forward path (i.e., $L \ll M$), the PCA-DPLS approach will introduce significantly less computational complexity than the QR-LS. By reducing the number of samples, i.e., $T \ll N$, the computational complexity of the identification subsystem will decrease proportionally. In section 5.5.3 of this Chapter, we will show experimental results that will evidence the computational complexity reduction that can be achieved properly combining both dimensionality reduction and mesh-selecting approaches.

5.3.3 Description of Different Mesh-selecting Methods

As previously discussed, by introducing memory to the mesh-selecting approach and considering the I-Q components instead of the absolute value of the signals, it is possible to define several alternative mesh-selecting approaches to the one first proposed in Chapter 4. In this subsection we will introduce several mesh-selecting methods combining the aforementioned approaches. A simple-select method will be included to facilitate the comparison of the proposed mesh-selecting methods. A comparison of the performance obtained with each of these methods will be provided in the following subsections.

To allow a fair comparison among the different mesh-selecting methods, a mesh density of 8 is set as default. For every method, the histogram will be cut to 10 %. Accurately, the percentage of selected samples cannot be exactly the same among different methods due to the effect when transferring a fractional number into a integer. For example, if one segment of the histogram for method A has 101 samples and another for method B has 99. By keeping 10 % of the samples this will result in 10.1 and 9.9 respectively. Practically the ground value will be use as 10 and 9. This phenomenon induces a slightly difference of the number of selected samples, in small quantity level.

Method 1: AM-AM or ABS Memoryless Mesh-selecting

The first method is the basic AM-AM Memoryless mesh-selecting method which only considers the absolute value of the present input sample and the corresponding PA output value, as presented in Chapter 4. Neither memory nor complex I-Q components are taken into consideration. In the following, this method will be considered as the reference method to later try to further enhance the performance of the mesh-selecting methods. In addition, to unify the nomenclature, this method will be also referred in the results section as the memoryless absolute value (ABS) method. The selecting result is shown in Fig. 5.4. The selected sub-set contains 10 % percent of

the original signal samples.

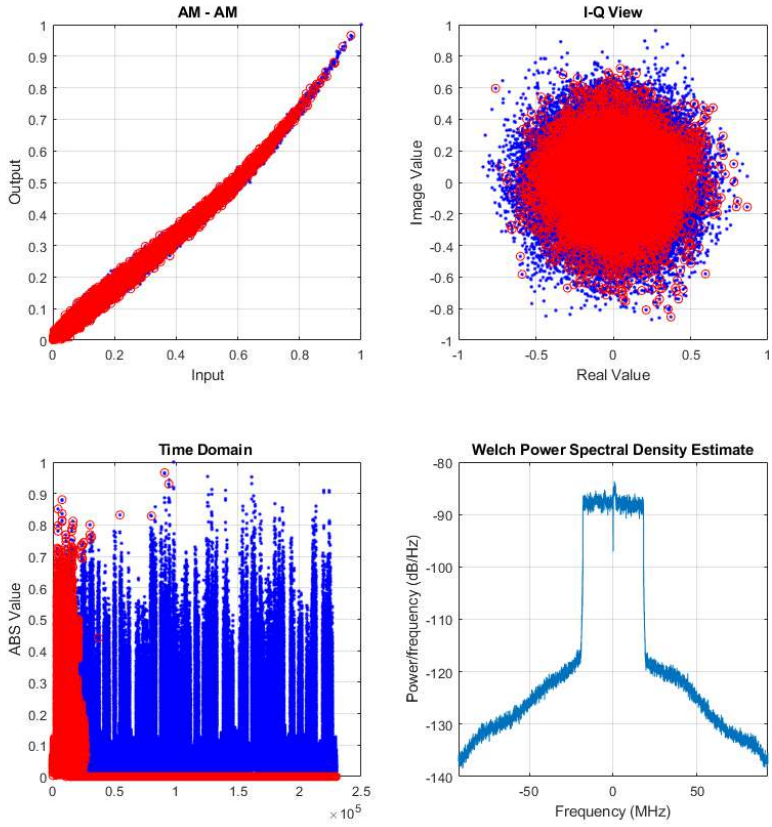

Figure 5.4: Selecting result of Memoryless ABS method

Method 2: Memoryless IQ Mesh-selecting

For the complex signal, which is widely used in modern telecommunication systems, the memoryless I-Q mesh-selecting considers the real and imaginary part of the PA input signal independently. A two-dimensional mesh is used to map each sample of the PA input signal according to the value of real part and imaginary part, which provides additional accuracy than the previous one. The I-Q memoryless method is schematically described in Fig. 5.3. Similarly to the previous method, since it is a memoryless method, the historical samples are not taken into consideration.

The selecting result is shown in Fig. 5.5. The selected sub-set contains 10 % of the original signal samples. In comparison to the memoryless ABS mesh-selecting, we can see from the time domain figure that there are more samples in the high amplitude area that have been selected, which predicts that by considering the separation of I-Q parts, the mesh-selecting method provides enhanced accuracy.

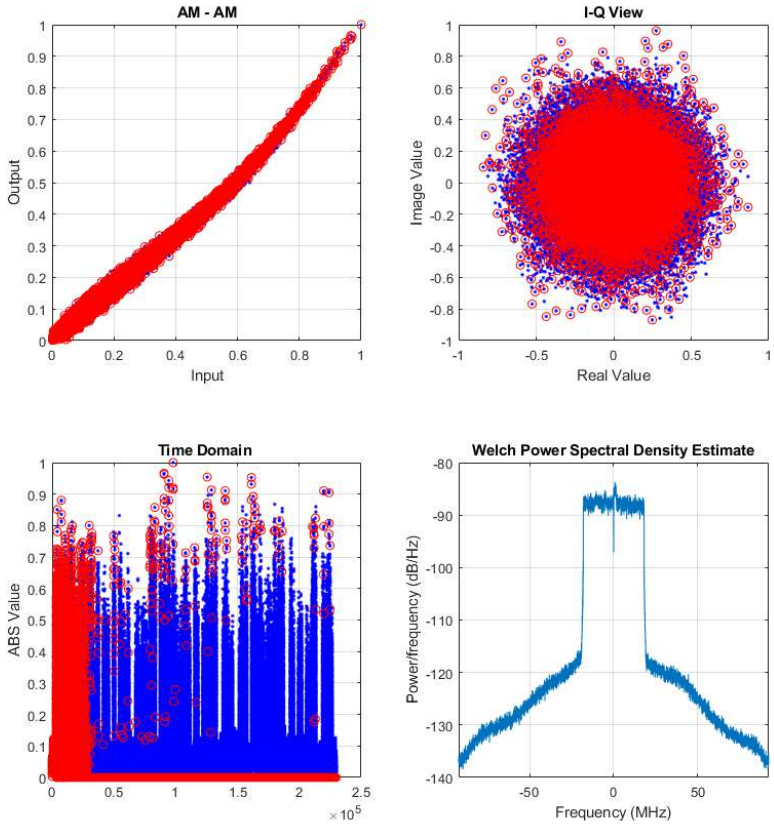

Figure 5.5: Selecting result of Memoryless IQ method

Method 3: Memory ABS Mesh-selecting

When considering DPD linearization, most of the DPD behavioral models include memory terms to be able to compensate for the memory effects in PAs. As we have discussed previously, the other approach to enhance the accuracy of mesh-selecting is to consider not only the present sample but also historical samples. The histogram of the memory mesh-selecting method will be multi-dimensional and will be determined by the memory length of the chosen DPD function. For example, if a DPD function contains a memory depth of 5 taps, the histogram will be 6-dimensional, since it includes also the present sample value.

It is important to note that, for Memory ABS MeS, the PA output value is not taken into consideration. Thus it is not a simple extension from the previous AM-AM MeS method by adding memory samples.

The selecting result is shown in Fig. 5.6. The selected sub-set contains 10 % of the original signal samples.

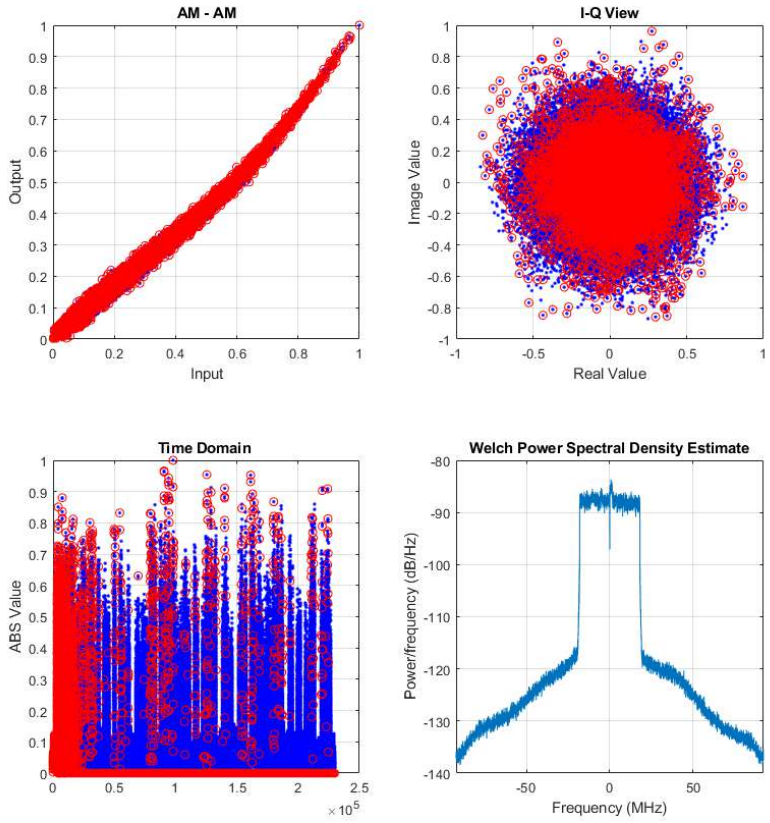

Figure 5.6: Selecting result of Memory ABS method

Method 4: Memory I-Q Mesh-selecting

This method has been already introduced in section 5.3.2, despite the fact that Fig. 5.3 was only schematically describing the memoryless version for simplicity.

The number of dimensions of the mesh will be doubled in comparison to the previous method, i.e., the Memory ABS Mesh-selecting (method 3). It is important to note that in a real-case implementation of the memory I-Q mesh-selecting, the consumption of the RAM storage space will increase exponentially. The total RAM requirement of a mesh-selecting method with mesh density of 10 and memory depth of 5 could be thousands of Gigabytes (on MatLab platform), which will never be acceptable for implementation even on a cloud platform. Thus, the mesh density and memory depth have to be selected accurately to achieve the properly performance under the trade-off with affordable RAM storage consumption.

The selecting result is shown in Fig. 5.7. The selected sub-set contains 10 % of the original signal. It is obviously that the selection of samples in the time domain is the most homogeneous in comparison to the previously described methods.

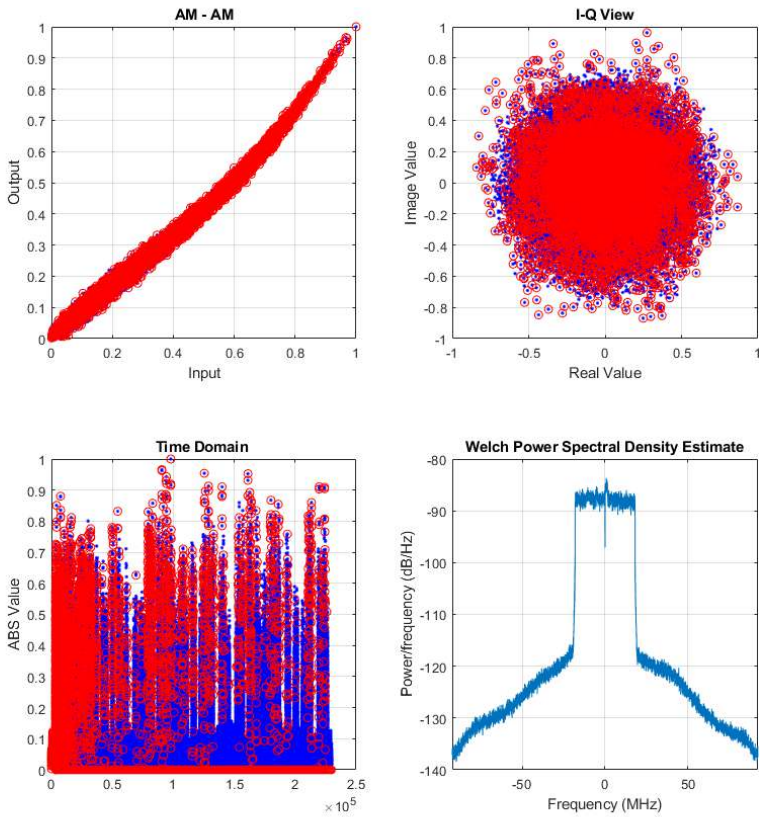

Figure 5.7: Selecting result of Memory I-Q method

Method 5 and 6: Memory-Y Approaching Mesh-selecting

For cross comparison, the memory-Y series mesh-selecting are designed to take also into consideration the PA output signal. The histogram is extended to involve the value of the output signal, taking into account both its absolute value (method 5) and its I-Q components (method 6).

The selecting result is shown in Fig. 5.8 and Fig. 5.9. The selected sub-set contains 10 % of the original signal samples. Note that by the display of the Matlab figure, the coverage area of a dot or a circle is much larger than the correspondingly accurate math value on the coordinate system. Besides, the overlap of dots and circles will neither induce darker color nor other displayable hints. Thus, it appears like that the number of selected samples is much larger if samples are distributed on a large coverage of the figure, in comparison to a set of narrow distributed samples, even if they have the same number. This phenomenon can be observed by comparing the selecting result of Fig 5.8 and Fig 5.9.

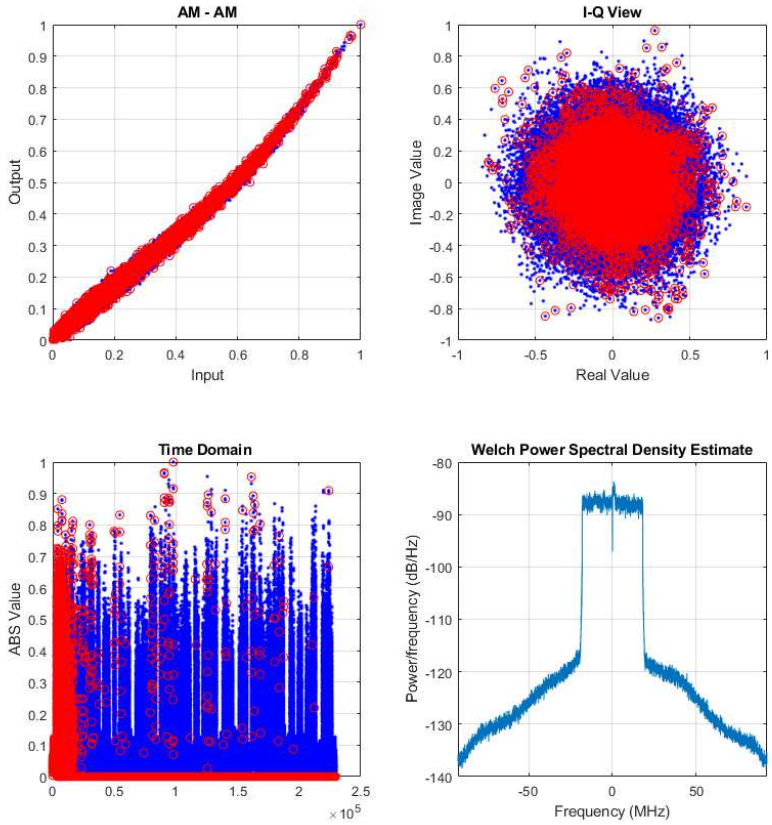

Figure 5.8: Selecting result of Memory-Y ABS method

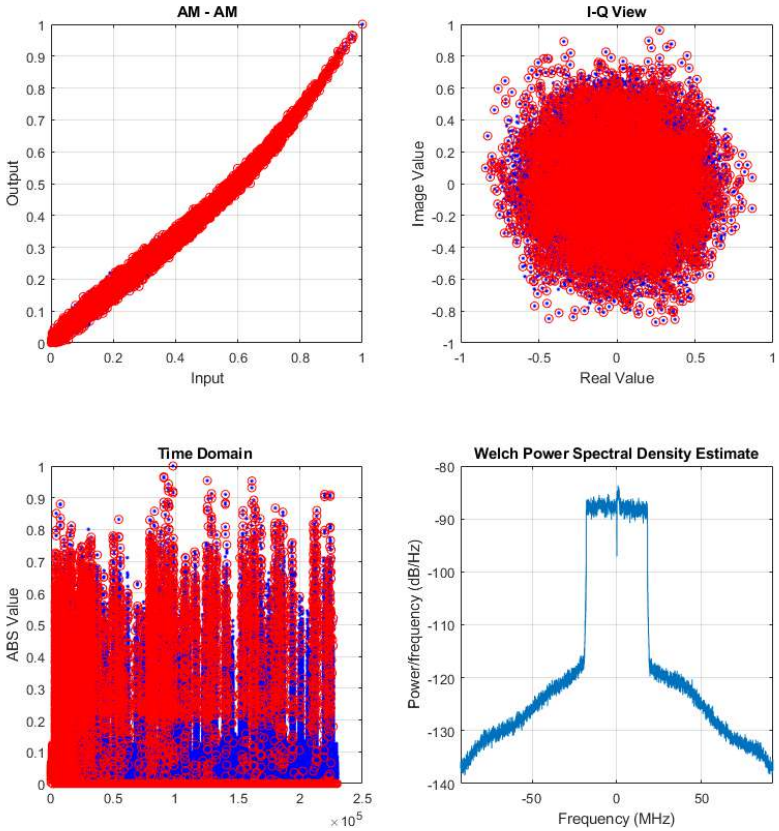

Figure 5.9: Selecting result of Memory-Y I-Q method

Method 7: Simple-select

The mesh-selecting method, as presented above, is designed to select a sub-set of the original transmitted signal not only covering the range of sample values of the input signal, but also keeping the distribution of the original histogram. For cross-comparing purposes, a simple-select method is designed to only consider the range of input sample values but not taking into account the distribution of the original histogram. Consequently, with the simple-select method, the number of samples in each segment of the multi-dimensional histogram is irrelevant, because this method will use a uniform given number to replace it and then do the selection. This method is designed to validate the importance of keeping the distribution of the multi-dimensional histogram of the mesh-selecting.

The selecting result is shown in Fig. 5.10. The selected sub-set contains 10 % of the original signal samples.

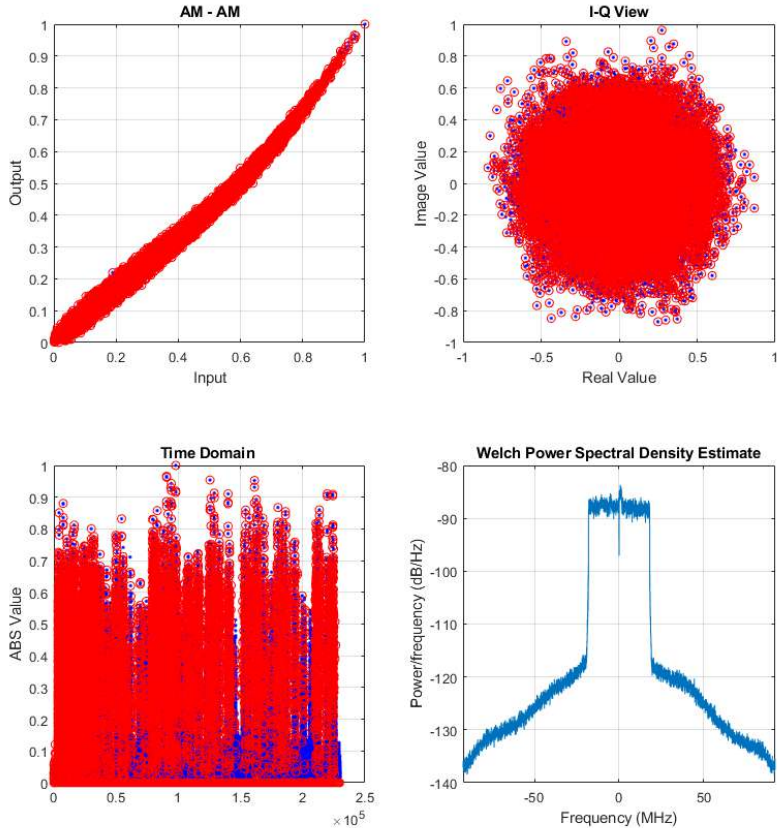

Figure 5.10: Selecting result of Simple select method

5.4 Experimental test-benches

In order to validate the proposed techniques in this thesis, experimental results were obtained using 3 different test-beds implemented for DPD linearization, that were accessible remotely. This means that the digital processing computing could run in a cloud and then, the access to the specific remote Weblab (containing the DUT and instrumentation) is done though the internet, exchanging PA input and output I-Q signals.

Two of the remote Weblabs used in this thesis are located in the research laboratory of the CSC research group of the UPC, the other one is the remote Weblab located at Chalmers University [Cha].

5.4.1 UPC Weblab

The permanent UPC Weblab based on a transceiver board (Zynq-AD936 platform from Analog devices) is depicted in Fig. 5.11. The AD9361 transceiver evaluation board is used as ADC and DAC while connecting with Zynq-7000 board.

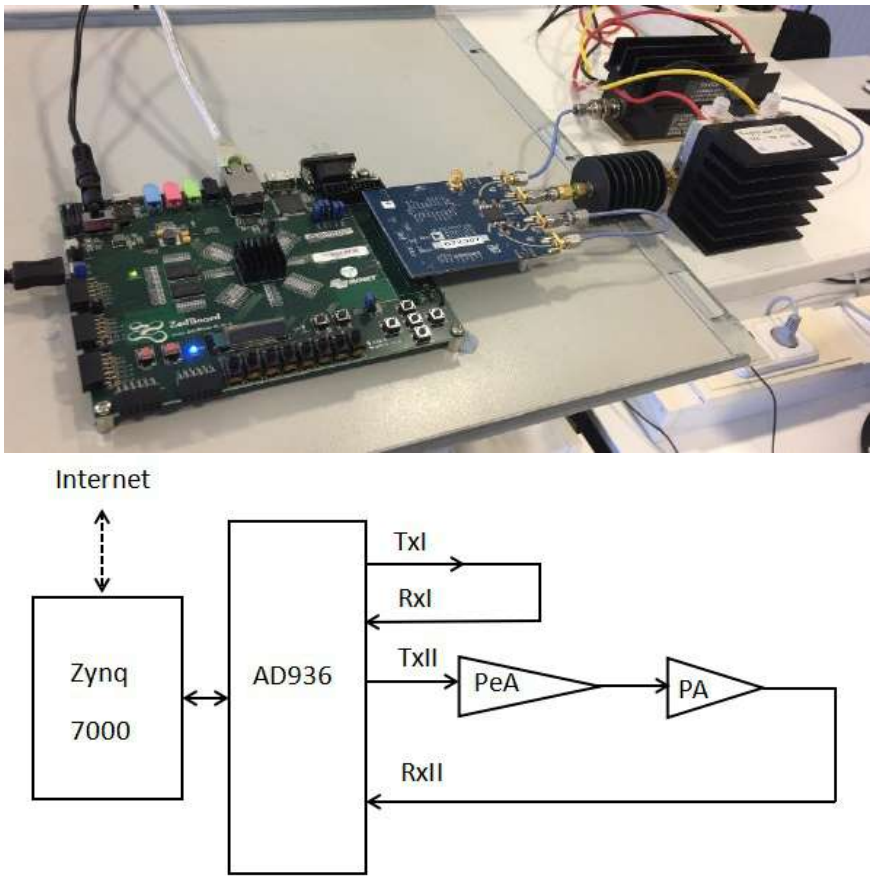

Figure 5.11: The Zynq-AD936 platform assembled with pre-amplifier and PA and it data flow

The FPGA - Linux platform is used to replace the expensive loop for signal generation and acquisition based on instrumentation (e.g., an arbitrary waveform generator (AWG) and a digital storage oscilloscope (DSO)). An internet interface is set to allow users to send their test signals and receive the PA output. This part is now implemented on Zynq 7000 FPGA-Linux board to establish the internet connection. The board also works as a driver to the AD936 direct converting transceiver to access the DUTs, composed by a driver PA from Mini-circuits and a class-J PA operating at center frequency of 900 MHz. The high efficient Class-J PA was designed by José Angel García's research group from the University of Cantabria. The AD936 transceiver allows RF signal generation from 70 MHz to 6.0 GHz with up to 56 MHz bandwidth [ANA] which allows around 20 MHz DPD operation bandwidth.

5.4.2 Chalmers University Weblab

Chalmers University organized the International Microwave Symposium (IMS) Digital Pre-Distortion Student Design Competition in 2014 and 2015. For the competition they assembled a remote Weblab to allow PhD students all around the world to test their DPD algorithms, as depicted in Fig. 5.12. Since then, they made this Weblab permanent and accessible to all through Matlab and an internet connection.

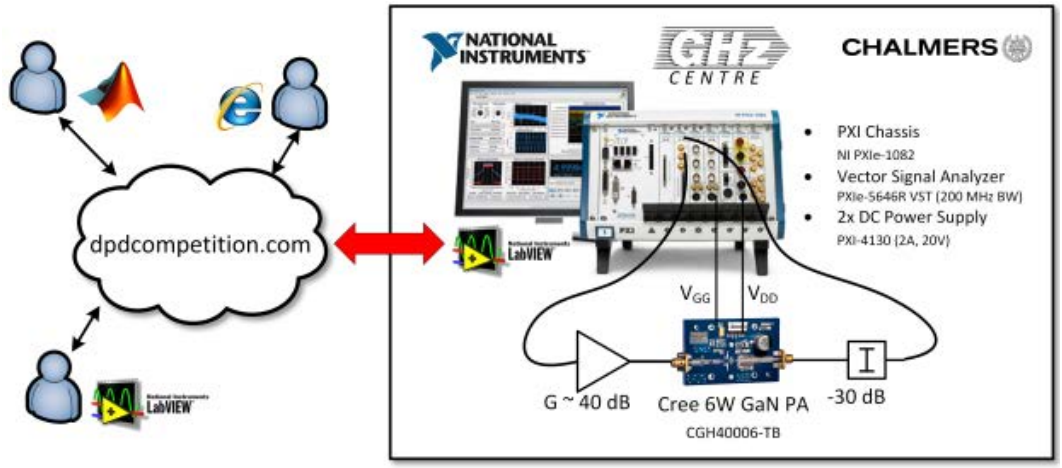

Figure 5.12: The Test bench of Chalmers WebLab [Cha]

The measurement setup is based on a PXI Chassis (PXIe-1082) with embedded host PC from National Instruments. The chassis is equipped with a Vector Signal Transceiver (PXIe-5646R VST) with 200 MHz instantaneous bandwidth. The signal generated (center frequency 2.14 GHz) from the VST transmitter is fed to a linear driver amplifier before the GaN PA DUT (Cree CGH40006-TB, testboard for the transistor CGH40006P). A 30 dB RF attenuator is placed at the DUT output from which the output signal is connected to the VST receiver. A PC embedded in the PXI chassi is used to control the instruments and to down- and upload

data files at request from the users. The DUT is supplied by a power source module (PXI-4130) which also measures the current consumption of the power amplifier [Cha].

5.4.3 LMBA test-bench at UPC

This instrument-based test-bench consisted in a Matlab-controlled setup interfacing the Keysight M8190A AWG for signal generation and the Keysight 90404A DSO for direct RF analog to digital conversion, as shown in Fig. 5.13.

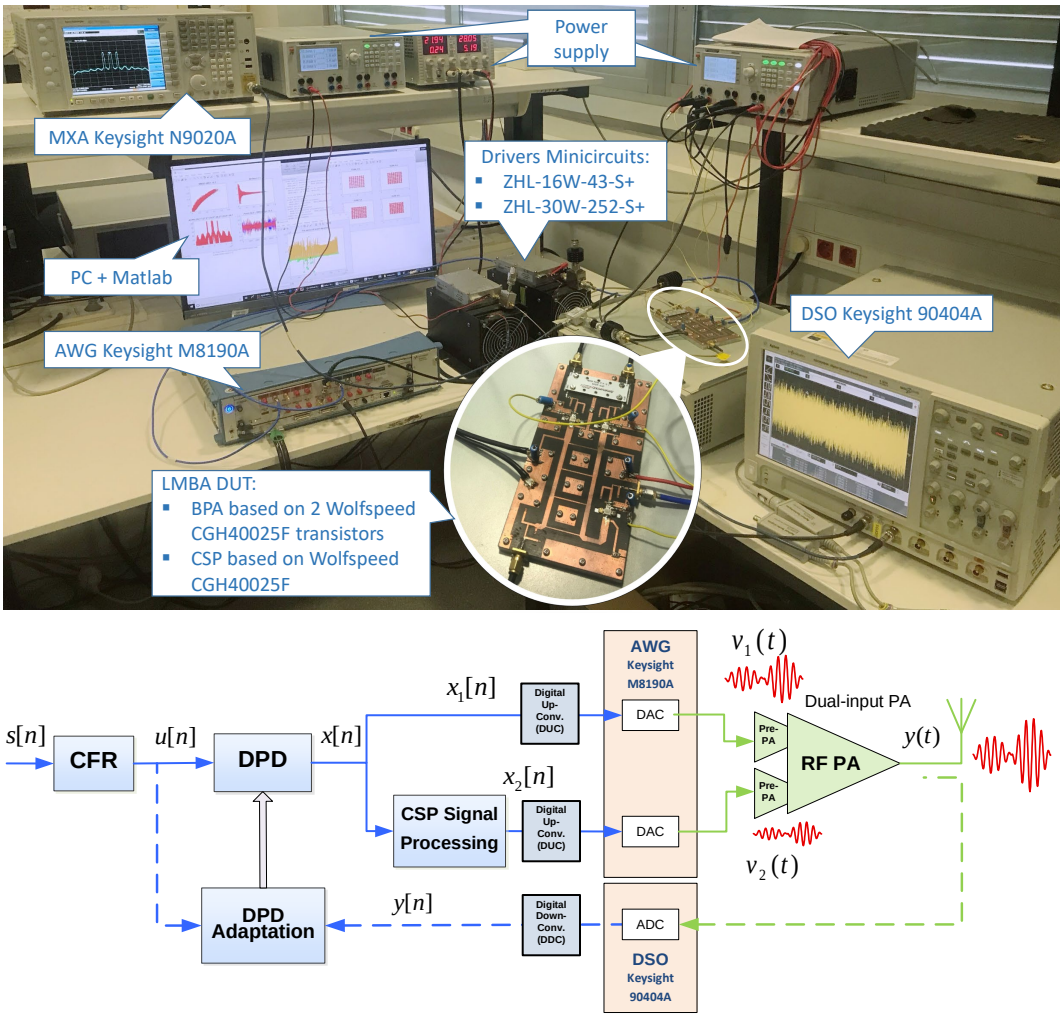

Figure 5.13: Picture of the test setup employed for LMBA experimental validation.

The baseband processing in Matlab is carried out with a baseband clock of 614.4 MSa/s in order to be able to process predistorted signals of hundreds of MHz of bandwidth. The DUT is a dual-input load-modulated balanced amplifier (LMBA) designed by Roberto Quaglia et al. from Cardiff University in [Qua18]. Therefore, the main and auxiliary signals required by the dual-input LMBA are generated, predistorted, time-aligned and phase-shifted in Matlab to later

digitally up-convert them to the 2 GHz RF frequency and digitally-to-analog convert them (i.e., both main and auxiliary RF signals) using the two RF channels of the Keysight M8190A AWG with a clock rate of 7.9872 GHz and 14 bits. Both RF signals are pre-amplified with two high gain linear amplifiers from Minicircuits to feed the LMBA. The PA output signal is RF sampled with a Keysigh 90404A DSO at 20 GSa/s with 8-bit resolution, digital down-converted and re-sampled for DPD processing. A Keysigh N9020A MXA signal analyzer was used to characterize the spectrum and ACPR at the output of the PA.

5.5 Experimental results

In order to evaluate and compare the mesh-selecting techniques proposed in this Chapter, experimental results were obtained using the 3 aforementioned test-benches implemented for DPD linearization, that were accessible remotely.

First, a whole coverage experiment is made using the UPC WebLab to cross compare the performance of the different mesh-selecting approaches presented before. Then, some preliminary tests, conducted in the Chalmers WebLab from 2014 to 2015, are shown. Theses results focus on the comparison of the performance between the ABS mesh-selecting approach and the I-Q mesh-selecting approach for different mesh-densities. Then, more in deep tests were made on the UPC LMBA test-bench. The tests were aimed at showing the benefits of the memory I-Q mesh-selecting when properly combined with dimensionality reduction techniques when taking into account both PA behavioral modelling and also DPD linearization.

Similarly to the methodology used in Chapter 4 to illustrate the performance of the AM-AM (or ABS) memoryless mesh-selecting, the performance of the memory mesh-selecting method will be evaluated in terms of NMSE and ACLR degradation when considering different data matrix ratios (MRs). In this chapter we will also focus on the tuning parameters of the mesh-selecting method itself, such as for example, the influence of the number of mesh-segments (also called as mesh density).

5.5.1 Evaluation of the Proposed Mesh-Selecting Methods

A full scale test comparing the different mesh-selecting methods was run in the UPC WebLab. The test-bench is depicted in Fig. 5.11 and described in section 5.4.1 of this Chapter.

Taking into account the bandwidth expansion of the predistorted signal and given the bandwidth limitations of the Zynq-AD936 platform (up to 56 MHz bandwidth with a clock rate of 61.14 MSa/s), the test signal used was a LTE 10 MHz waveform with 131,072 samples. The DPD was based in a GMP behavioral model with a total of 80 coefficients.

Experimental results of MemoryLess ABS Mesh-selecting

As shown in Fig. 5.14 and Fig. 5.15, by using the MemoryLess ABS mesh-selecting method we can only achieve a MR of 400 with less than 5 dB of performance loss. We can also observe how the performance is highly dependent on the mesh density, which will reduce the mathematical robustness.

Experimental results of MemoryLess I-Q Mesh-selecting

As shown in Fig. 5.16 and Fig. 5.17 with the MemoryLess I-Q Mesh-selecting it is only possible to achieve a MR of 400 with less than 5 dB performance loss, which is not significantly different from the previous MemoryLess ABS MeS. The number of mesh density has been limited to the same level of the MemoryLess ABS mesh-selecting method for a fair comparison. From the figures we can see that the performance is highly dependent on the mesh density, but the fluctuation of performance is smaller than with the ABS method.

Experimental results of Memory ABS Mesh-selecting

As shown in Fig. 5.18 and Fig. 5.19, Memory ABS Mesh-selecting it possible to achieve a MR of 200 with less than 3 dB performance loss in ACPR. The performance is even better in NMSE. With high mesh density tuning, it is possible to go beyond a MR of 100 almost without accuracy loss in NMSE and only 2 dB in ACPR. The effect of mesh density is demising with the numerical increase. This property provides a mathematical stability to the designers to choose the mesh density in real-case DPD implementation.

Experimental results of Memory I-Q Mesh-selecting

As shown in Fig. 5.20 and Fig. 5.21 with the Memory I-Q mesh-selecting method it is possible to achieve smaller MRs than 100 with less than 2 dB performance loss in ACPR. The performance is even better in NMSE. In high mesh density tuning, it is able to achieve less than a MR of 50 almost without accuracy loss in NMSE and only 1 dB in ACPR. The effect of mesh density is becoming demising with the numerical increase.

Experimental results of Memory-Y ABS Mesh-selecting

As shown in Fig. 5.22 and Fig. 5.23, the Memory-Y ABS mesh-selecting method presents similar reduction capabilities than the Memory I-Q mesh-selecting method. Thus, it can achieve smaller MRs than 100 with less than 2 dB performance loss in ACPR. The performance is even better in NMSE. In high mesh density tuning, it is able to achieve MRs less than 50 almost without

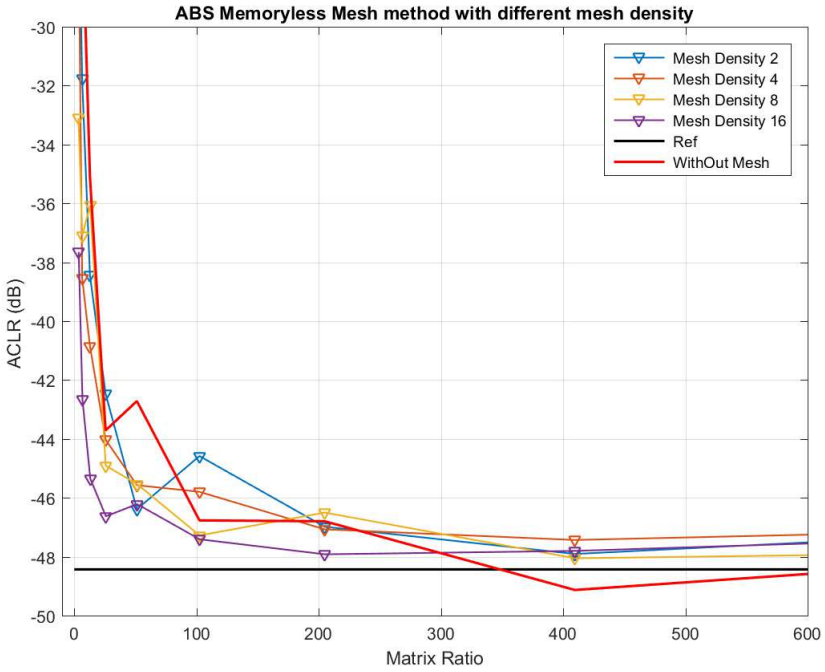

Figure 5.14: Performance of Memoryless ABS Mesh-selecting Method: ACPR values vs. Matrix Ratio

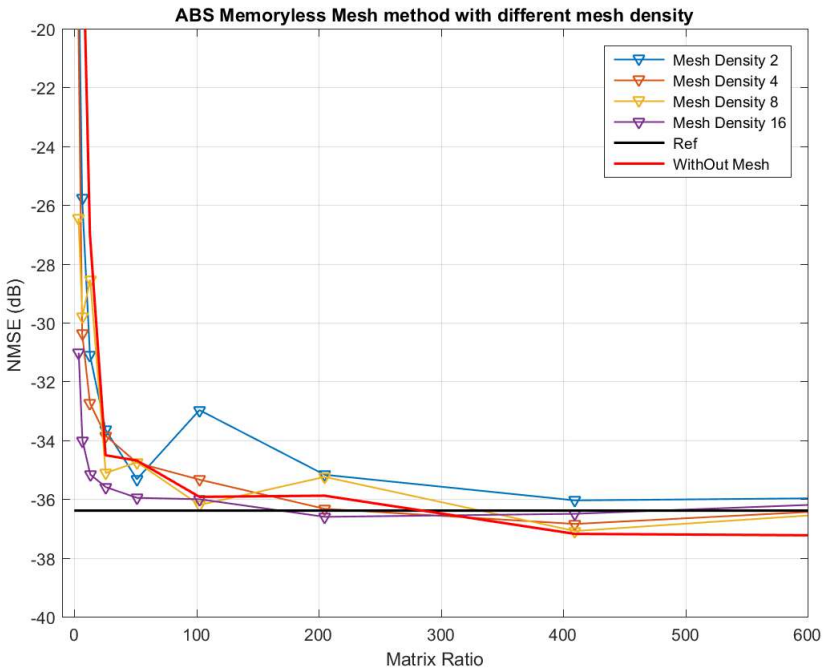

Figure 5.15: Performance of Memoryless ABS Mesh-selecting Method: NMSE values vs. Matrix Ratio

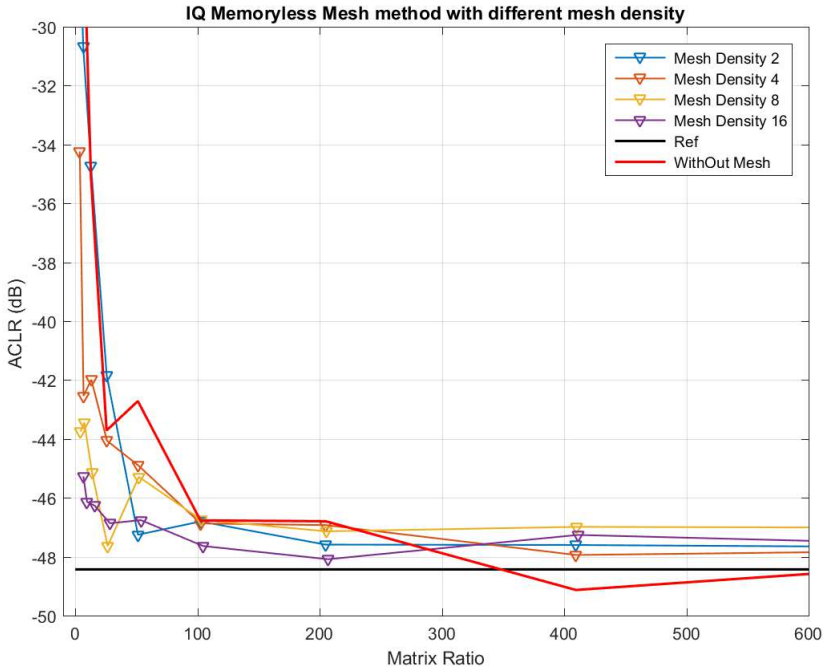

Figure 5.16: Performance of Memoryless I-Q Mesh-selecting Method: ACPR values vs. Matrix Ratio

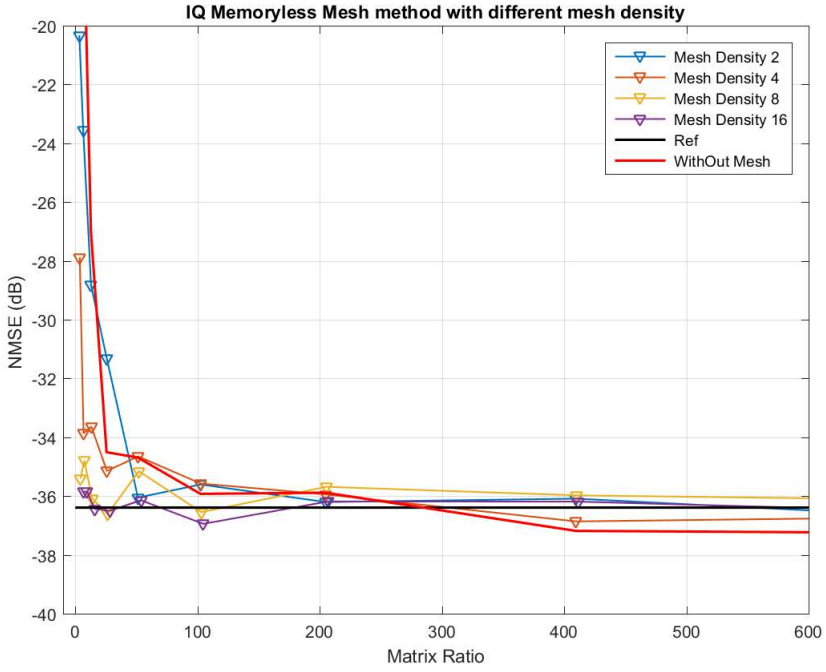

Figure 5.17: Performance of Memoryless I-Q Mesh-selecting Method: NMSE values vs. Matrix Ratio

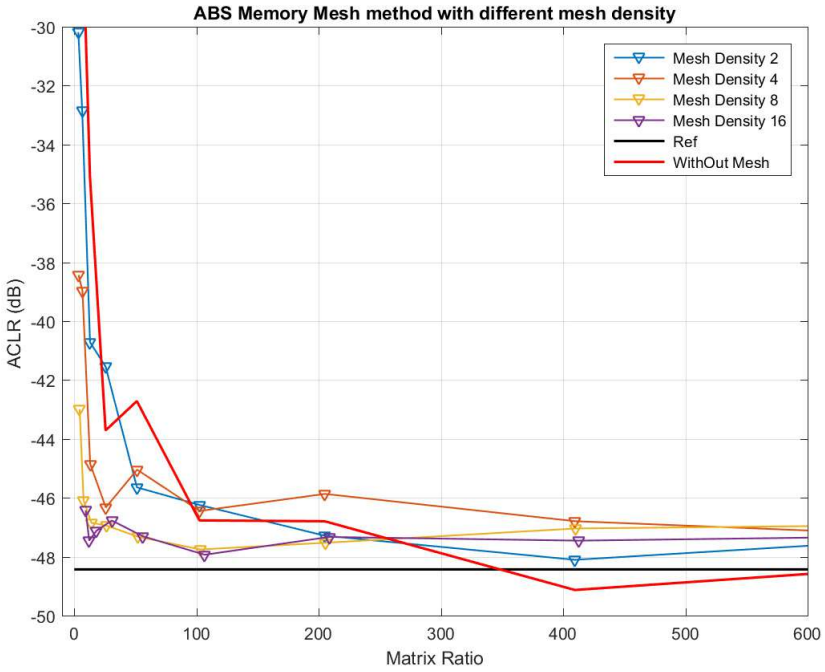

Figure 5.18: Performance of Memory ABS Mesh-selecting Method: ACPR values vs. Matrix Ratio

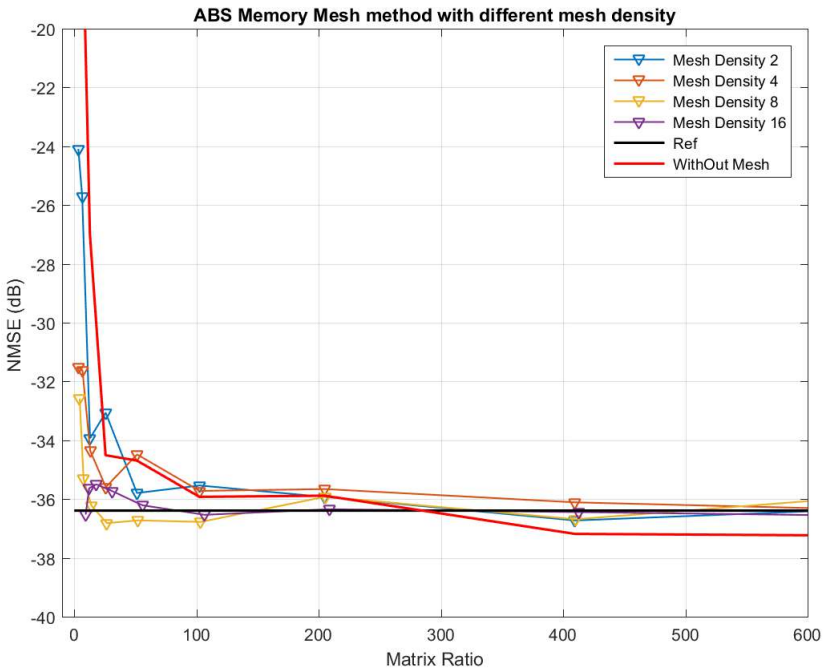

Figure 5.19: Performance of Memory ABS Mesh-selecting Method: NMSE values vs. Matrix Ratio

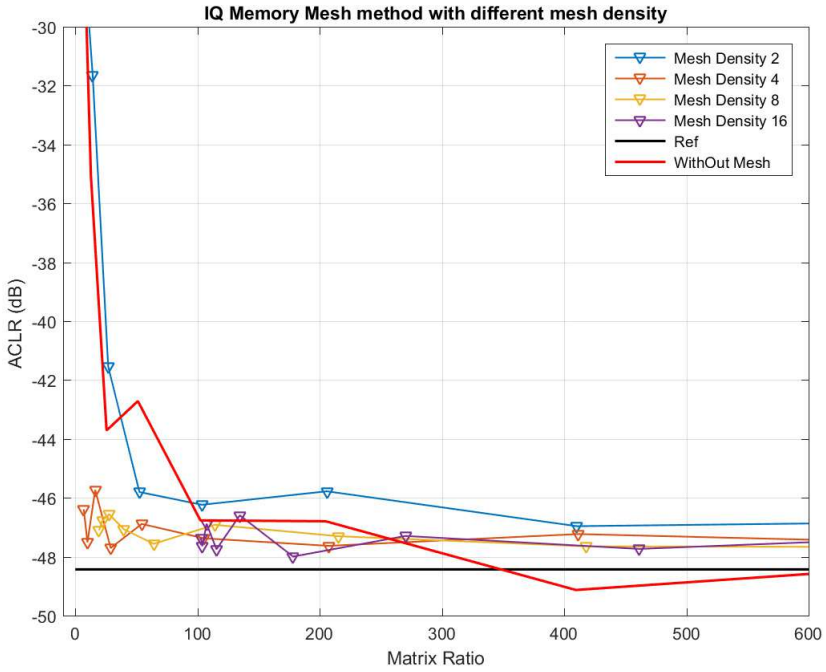

Figure 5.20: Performance of Memory I-Q Mesh-selecting Method: ACPR values vs. Matrix Ratio

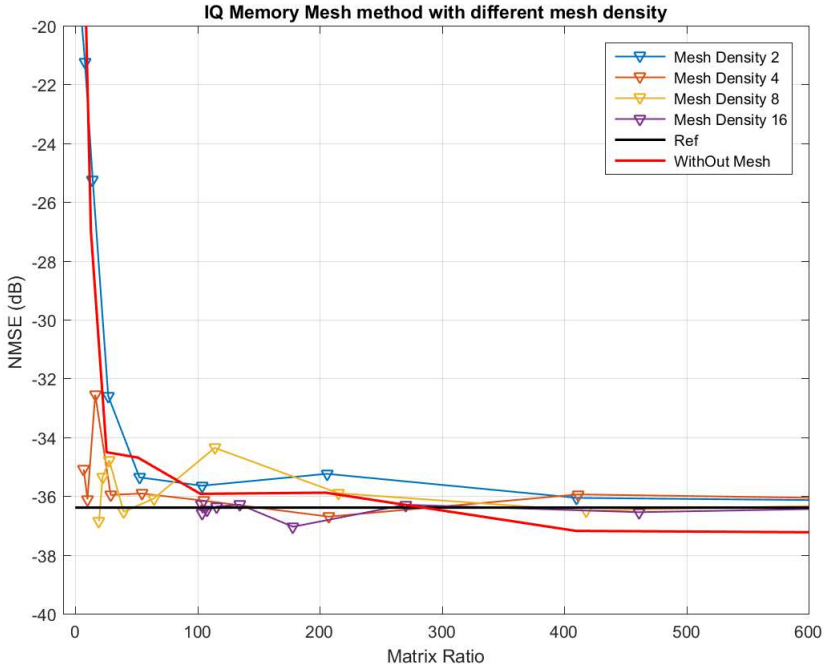

Figure 5.21: Performance of Memory I-Q Mesh-selecting Method: NMSE values vs. Matrix Ratio

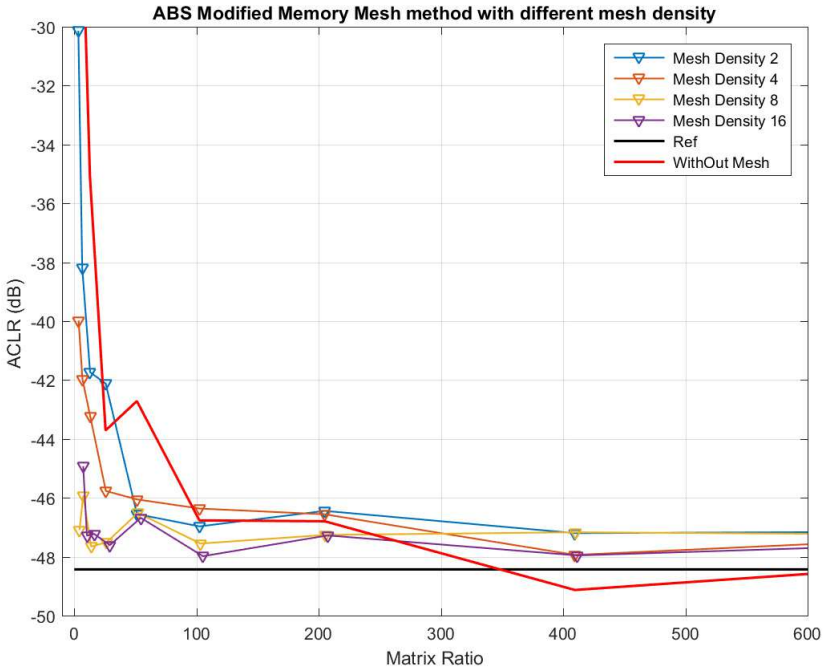

Figure 5.22: Performance of Memory-Y ABS Mesh-selecting Method: ACPR values vs. Matrix Ratio

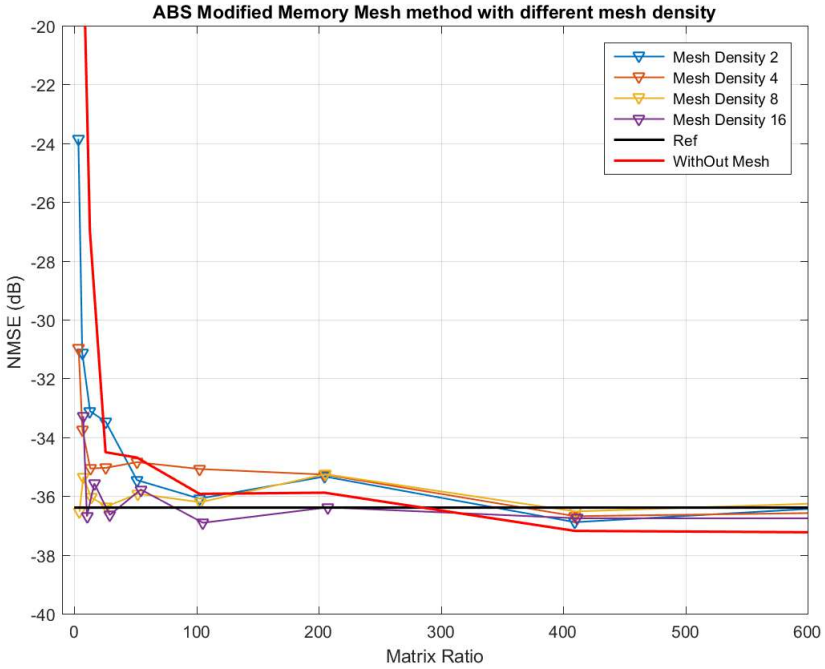

Figure 5.23: Performance of Memory-Y ABS Mesh-selecting Method: NMSE values vs. Matrix Ratio

accuracy loss in NMSE and only 1 dB in ACPR. The effect of mesh density is becoming demising with the numerical increase.

Experimental results of Memory-Y I-Q Mesh-selecting

As shown in Fig. 5.24 and Fig. 5.25, by using the Memory-Y I-Q Mesh-selecting method it is possible to achieve MRs smaller than 100 with less than 2 dB performance loss in ACPR. The performance is even better in NMSE. With high and even moderate mesh density values, it is able to achieve less MRs than 50 almost without accuracy loss in NMSE and only 1 dB in ACPR. It means the selection range of the mesh density has become wider by adding the I-Q components into consideration along with the PA output value.

Experimental results of Simple-select Mesh-selecting

The performance of the control experiment is illustrated in Fig. 5.26 and Fig. 5.27. It is obvious that when ignoring the multi-histogram distribution of the transmitted signal, it is hard to achieve an acceptable DPD performance. With moderate mesh density values we are able to achieve a performance loss of 3 dB, but it is not predictable with neither the increasing of the mesh density nor by reducing the MR. This performance is not a properly guidance to DPD engineers for a real case implementation.

Conclusions after the Evaluation of the Different Mesh-Selection Methods

As shown above, the performance of mesh-selecting increases with respect to the original approach in Chapter 4 when including memory and consider I-Q signals. However, when increasing the mesh-density and memory depth the demands for storage space grows exponentially. It is therefore not possible to build a mesh with high mesh-density, long memory depth while considering both I-Q components. In practical applications, trade-offs must be made among these factors. As a summary of the experimental results shown above, some trends can be pointed out as a design guide:

- The simple select method, which demands almost the highest storage requirements when applied in I-Q mode in same memory depth, provides nearly the lowest selecting performance. This phenomenon illustrates the importance of keeping the distribution of the multi-dimensional histogram of mesh-selecting. It is one essential property of the mesh-selecting methods.
- The benefits of increasing mesh density fades away as the number rise. To choose a middle level mesh density is a recommendable trade-off between complexity and performance.

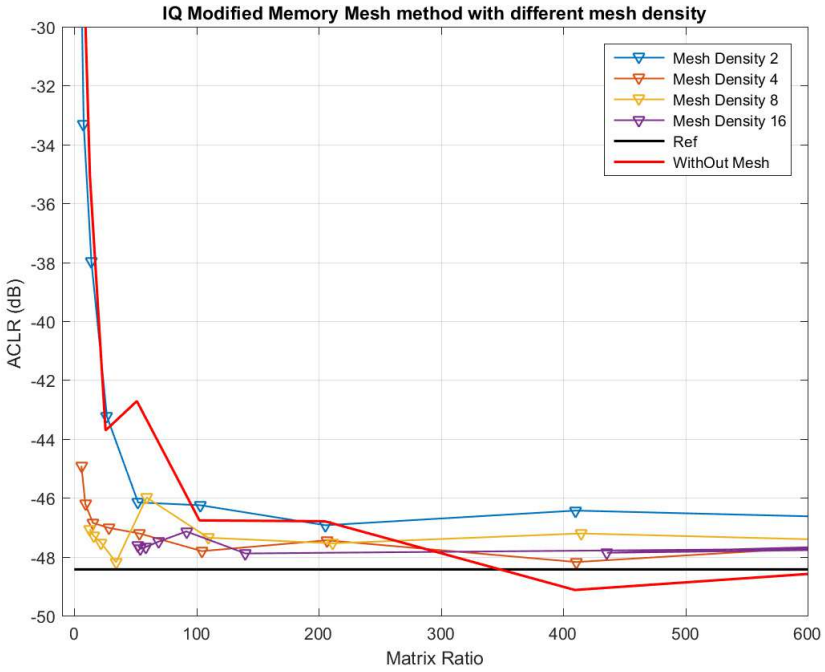

Figure 5.24: Performance of Memory-Y I-Q Mesh-selecting Method: ACPR values vs. Matrix Ratio

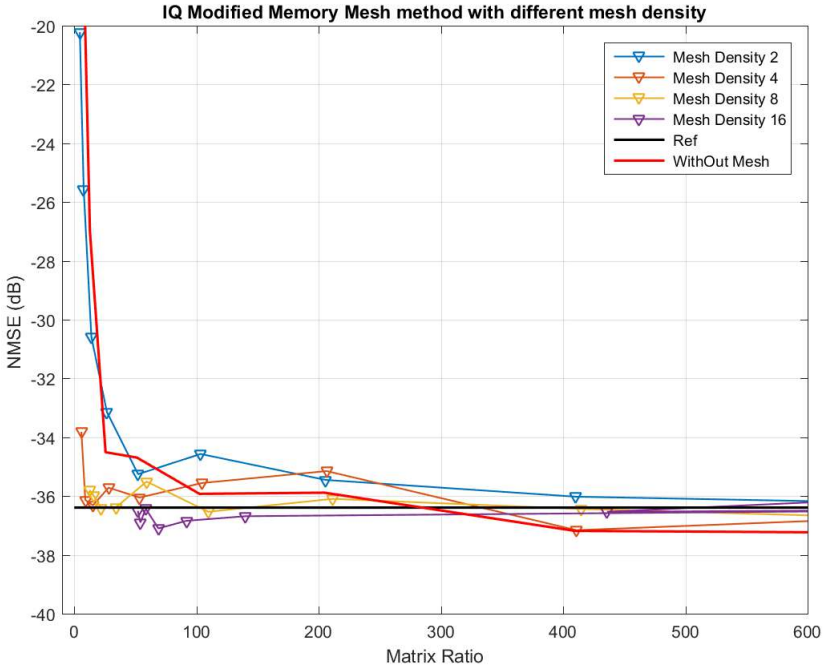

Figure 5.25: Performance of Memory-Y I-Q Mesh-selecting Method: NMSE values vs. Matrix Ratio

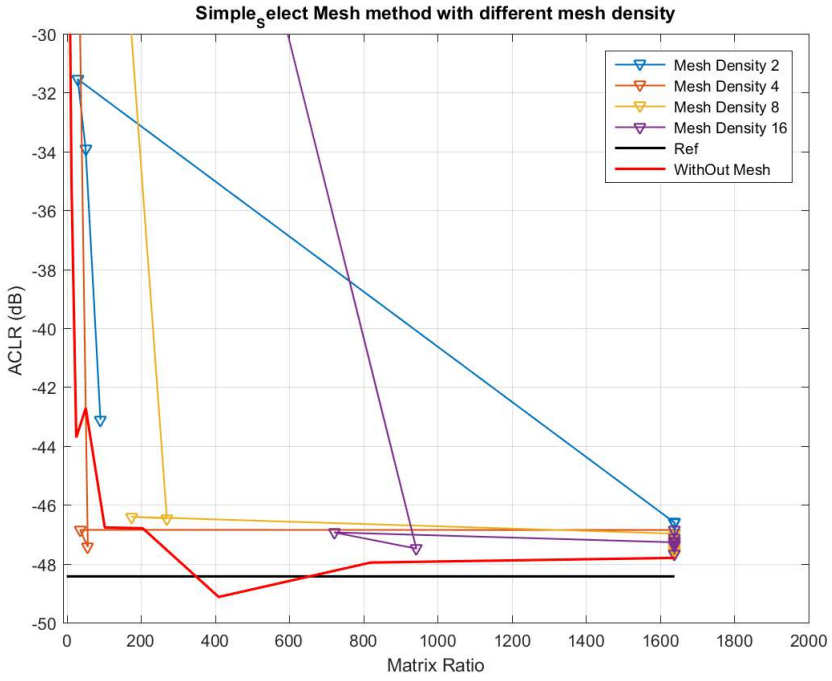

Figure 5.26: Performance of Simple-Select Mesh-selecting Method: ACPR values vs. Matrix Ratio

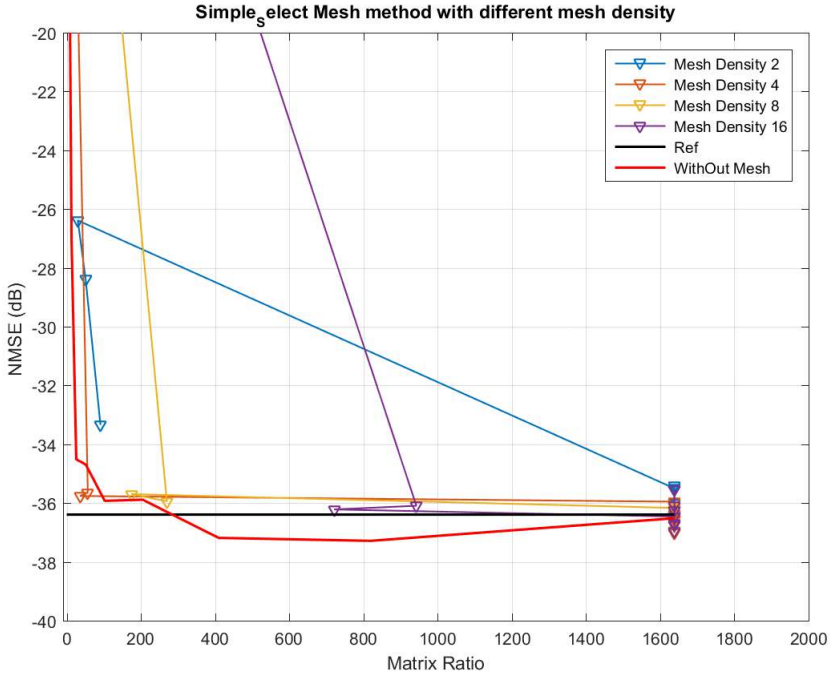

Figure 5.27: Performance of Simple-select ABS Mesh-selecting Method: NMSE values vs. Matrix Ratio

It is also notable that the mathematical stability of the mesh density is very good at middle-high range.

- Including the PA output signal in the mesh-selecting process (i.e., Memory-Y ABS MeS and Memory-Y I-Q MeS) in some cases can slightly enhance the samples reduction capabilities when estimating the DPD coefficients, but considering the exponentially growth of storage consumption, it is a low priority option when applying mesh-selecting.

Above all, the memory mesh-selecting is the base for all mesh-selecting methods during the practical DPD coefficients estimation, by properly choosing the memory length and mesh density to control storage level. Deploying an I-Q approach instead of ABS is recommendable if enough storage resources are available.

5.5.2 Advantages of Memory I-Q vs. ABS memoryless Mesh-Selecting

As concluded in the previous subsection, the memory enhancement of MeS is the most efficient way to provide better performance on the DPD implementation. On the contrary, considering the PA output samples to build the mesh is not a suitable option when targeting the MeS implementation, since it unnecessarily increases the required memory resources. Considering the I-Q components instead of the absolute value, seems to be a trade-off solution, and it is suggested to be taken into consideration when enough storage resources are available. A further test is operated on the Chalmers Weblab with more complex PA input signal. The test will also help us to see the detail performance of I-Q enhancement for MeS, both on technical indicators and robustness.

Some experimental results were obtained using Chalmers Weblab to detail compare the performance between the ABS mesh-selecting approach and the I-Q mesh-selecting approach . The Chalmers Weblab is depicted in Fig. 5.12 and described in section 5.4.2 of this Chapter. A LTE 20 MHz bandwidth burst-mode signal is used to feed the power amplifier. The spectrum and time domain waveform is shown in Fig. 5.28. The original signal contains 230k samples.

Again, the DPD used to linearize the PA was based in a GMP behavioral model (described in (3.8)). The number of coefficients increased to 296 in order to compensate the nonlinear distortion introduced by the DUT in the Chalmers Weblab when excited with the burst-mode signal.

Fig. 5.29 and Fig. 5.30 show how by using Memory ABS mesh-selecting it is possible to achieve a MR of 100 with less than 2 dB performance loss. When considering high mesh densities, it is possible to achieve MRs smaller than 50 almost without accuracy loss in ACLR and only 1 dB in NMSE. Another interesting performance is that the effect of mesh density demised

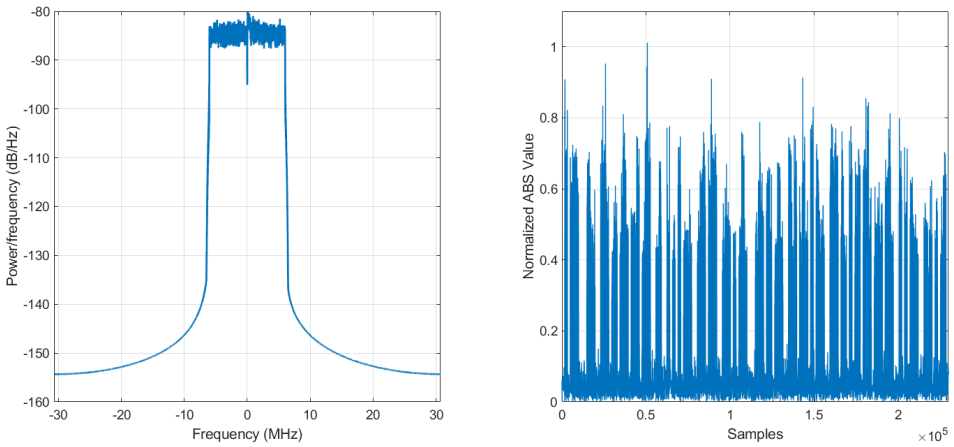

Figure 5.28: Spectrum and absolute value of the time-domain waveform of the test signal used in the Chalmers Weblab.

with the numerical increase. This property provides a mathematical stability to the designers to choose the mesh density in a real-case DPD implementation.

On the contrary, the mesh density is limited to a low value when considering I-Q mesh-selecting, because the storage of a normal server is not capable to contain a very large mesh. As depicted in Fig. 5.31 and Fig. 5.32, by using Memory I-Q mesh-selecting it is possible to achieve a MR of 200 with less than 2 dB performance loss, considering a mesh density value. The stability when using high density values is also as good like with the Memory ABS mesh-selecting, which further illustrates the mathematical robustness of this method.

A critical trade-off for I-Q memory MeS is the insufficiency of the RAM storage for a given computer, thus it can not reach as much mesh density as the ABS MeS. Apart from the tremendous RAM consuming, I-Q enhancement provide an excellent performance with very low matrix radio. Thus, in summary:

- If the storage space is limited, ABS MeS provides higher robustness with higher mesh density.

- If the storage space is not a limitation, the memory I-Q approach makes the MeS able to achieve lower matrix ratios with the same DPD performance and robustness.

- If an extreme low matrix ratio is required, the I-Q enhancement is the first priority.

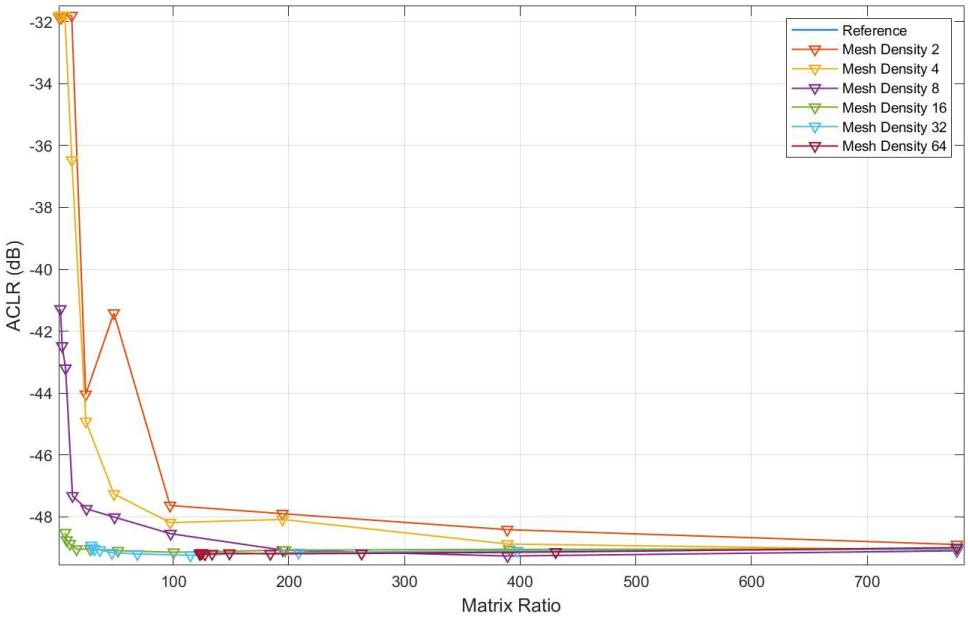

Figure 5.29: ACLR Performance of ABS Memory Mesh Selecting

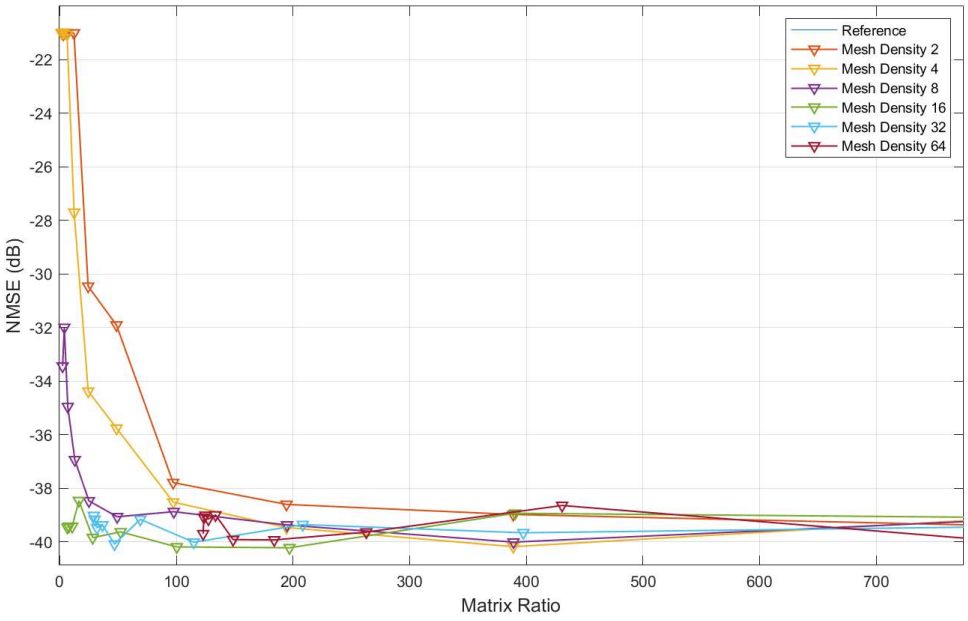

Figure 5.30: NMSE Performance of ABS Memory Mesh Selecting

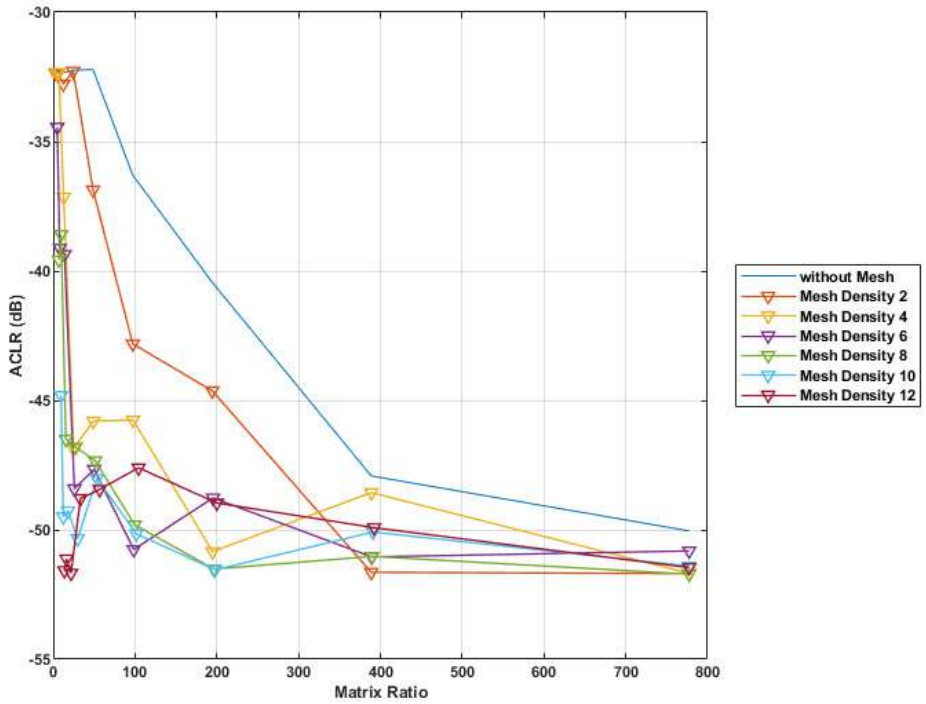

Figure 5.31: ACLR Performance of I-Q Memory Mesh Selecting

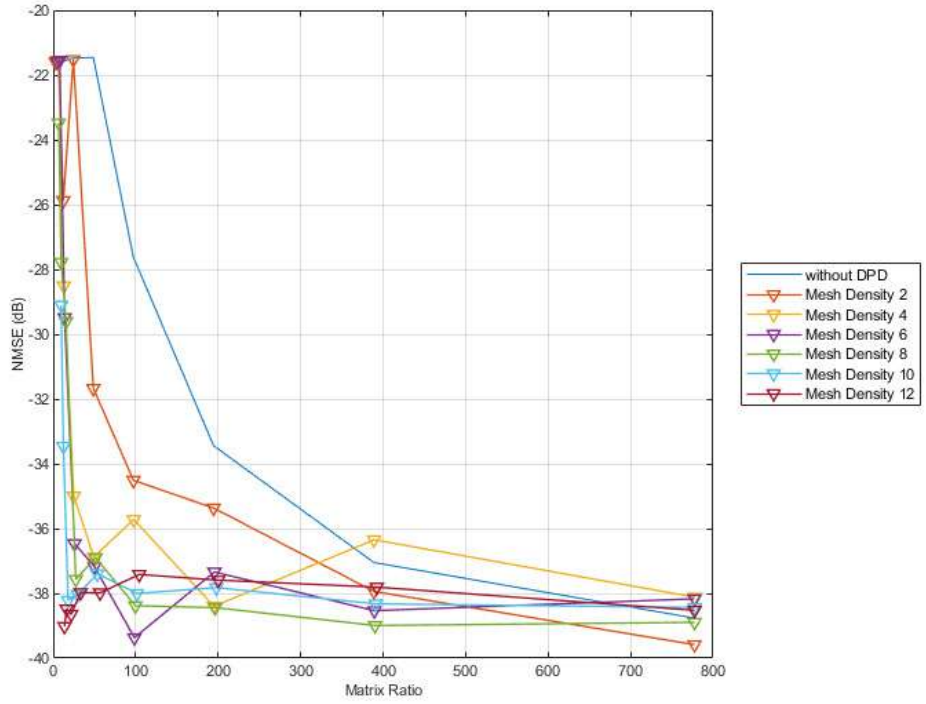

Figure 5.32: NMSE Performance of I-Q Memory Mesh Selecting

5.5.3 Combination of Memory I-Q Mesh-Selecting with Dimensionality Reduction Techniques

The following experimental results were obtained using the LMBA test-bench depicted in Fig. 5.13 and described in section 5.4.3 of this Chapter.

In a first approach we carried out a comparison of the different mesh-selecting (MeS) methods proposed in section 5.3.3 and their capability to reduce the number of samples while maintaining acceptable PA behavioral modeling capabilities. For PA behavioral modeling, we considered a test signal composed by 4 non contiguous channels of 64 QAM modulated LTE-20 signals spread in 200 MHz instantaneous bandwidth. Fig. 5.33 shows the identification performance when reducing the number of samples (or equations) for different 6 mesh-selecting techniques and the case of no mesh-selecting, i.e., considering removing simply consecutive samples (CoS). The 6 MeS techniques are: i) ABS (or AM-AM) memoryless MeS; ii) I-Q memoryless MeS; iii) ABS memory MeS; iv) I-Q memory Mes; v) Memory-Y ABS MeS and vi) Memory-Y I-Q MeS. This comparison is for PA behavioral modeling when considering the 200MHz signal and a GMP polynomial model with 139 coefficients. As observed in Fig. 5.33, focusing in the I-Q signal rather than in the AM-AM characteristic and including memory improves the reduction capabilities (i.e., NSME vs. number of samples used).

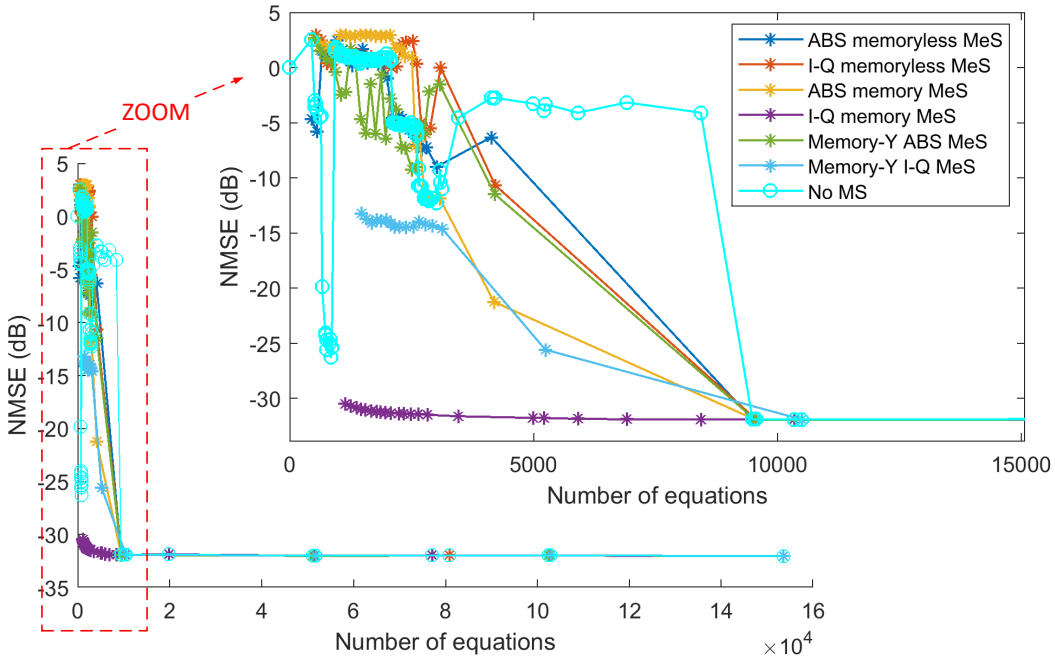

Figure 5.33: Comparison of mesh-selecting methods

In addition, the robustness (or resilience) of the model estimation when we remove samples is much more significant with the memory I-Q mesh-selecting method than by removing consecutive

samples. When removing consecutive samples without any specific criteria we end up removing samples from high amplitude values (which are less probable) and consequently the identification suffers from ill-conditioning degrading the overall NMSE. As an example, Fig. 5.34 shows two AM-AM plots when considering a) I-Q with memory mesh-selecting (MeS) and b) consecutive samples (CoS) selection. We have considered here a selection of 1,136 out of a 153,600 samples.

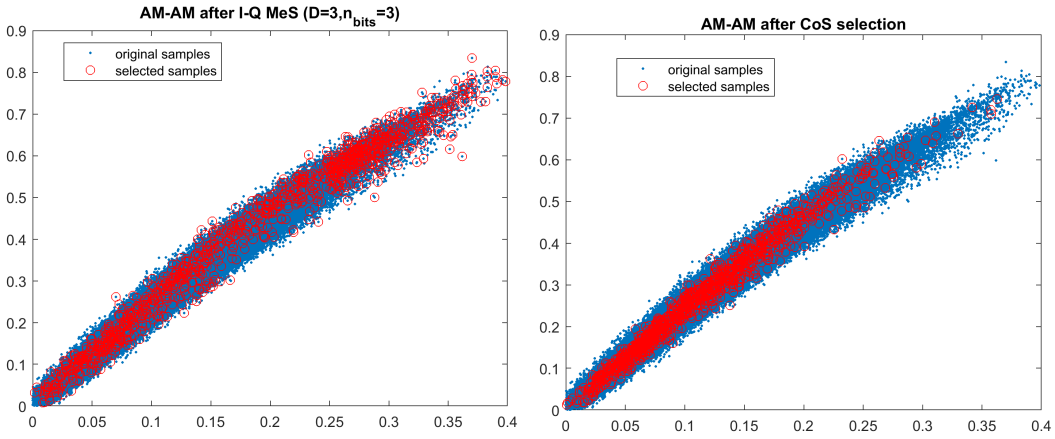

Figure 5.34: Comparison of Memory I-Q MeS vs. Consecutive sample selection.

Now, focusing on the Memory I-Q Mes (since it is the one providing better reduction performance) and in order to show the advantages of mesh selecting when properly combined with dimensionality reduction techniques, we conducted two tests: one in the context of PA behavioral modeling and another for DPD linearization.

Fig. 5.35 shows the NMSE between the measured and the modeled data when considering a GMP behavioral model, described in (3.8) section 3.1.1 of Chapter 3. The nonlinear functions associated to the memory and cross-memory terms in the GMP model can be particularized by polynomials (as in the original definition) or by interpolated look-up tables, such as, for example, B-splines. The NMSE of identification when using the GMP model implemented with polynomials and considering 139 coefficients is shown in Fig. 5.35-top, while Fig. 5.35-bottom shows the NMSE of identification for the GMP model implemented with B-splines of degree 3 and considering 189 coefficients. As previously observed, unlike discarding consecutive samples (CoS) or selecting samples with the AM-AM memoryless MeS method [Wan15a], by using the Memory I-Q MeS method (with memory D=3 and n_{bis}=3 in (5.10)) it is possible to maintain the NMSE value below -30 dB even when significantly reducing the number of required equations (samples) from an original number of 153,600 identification samples. The advantage of the Memory I-Q MeS is more evident when considering the GMP B-splines model [Bar14], since due to its piece-wise nature every segment requires a minimum number of samples to guarantee a proper estimation. Piece-wise models can benefit from higher locality, but at the same time it is necessary to guarantee a minimum number of samples for each 'piece' or segment. This is

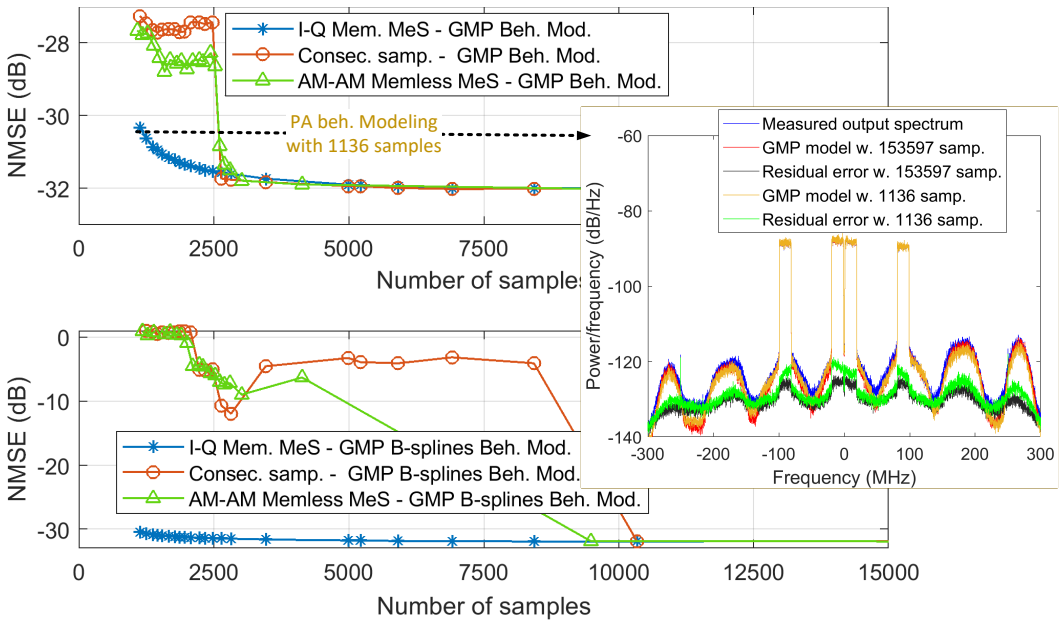

Figure 5.35: PA behavioral modeling: NMSE vs. identification samples.

the reason why the necessity of the proposed I-Q with memory mesh-selecting method is more evident when considering GMP-B-splines for PA behavioral modeling.

For DPD linearization purposes we considered a signal composed by two LTE-20 MHz carriers separated 20 MHz with 10.5 dB of PAPR centered the RF frequency around 2 GHz and exciting the LMBA providing 34 dBm of mean output power with 25% drain power efficiency. Fig. 5.36 shows the out-of-band linearity levels obtained (in terms of ACPR) for different number of samples used in the identification/adaptation subsystem.

The DPD coefficients of the GMP behavioral model (108 coefficients) in the DPD forward path have been extracted in the adaptation subsystem considering two solvers: the QR-LS algorithm and the PCA-DPLS algorithm described in [Pha19b]. This technique, based in the combination of the PCA transformation with the dynamic PLS (DPLS) extraction of components, is used to dynamically estimating and updating the DPD coefficients. The PCA-DPLS approach is equivalent to a canonical correlation analysis (CCA) updating solution, which is optimal in the sense of generating components with maximum correlation. The number of DPD coefficients used to update or estimate the DPD functions is significantly reduced by using the PCA-PLS algorithm [Pha19b].

As shown in Fig. 5.36, independently on the type of LS solver used, the mesh selecting method (with D=3 and n_{bis}=3) allows significantly reducing the number of samples (in comparison to removing consecutive samples) and still presenting acceptable ACPR values. Table 5.1 shows the minimum number of samples required for both LS solvers, to meet the ACPR threshold of -45 dBc when considering the proposed mesh selecting (MeS) method and simply discarding

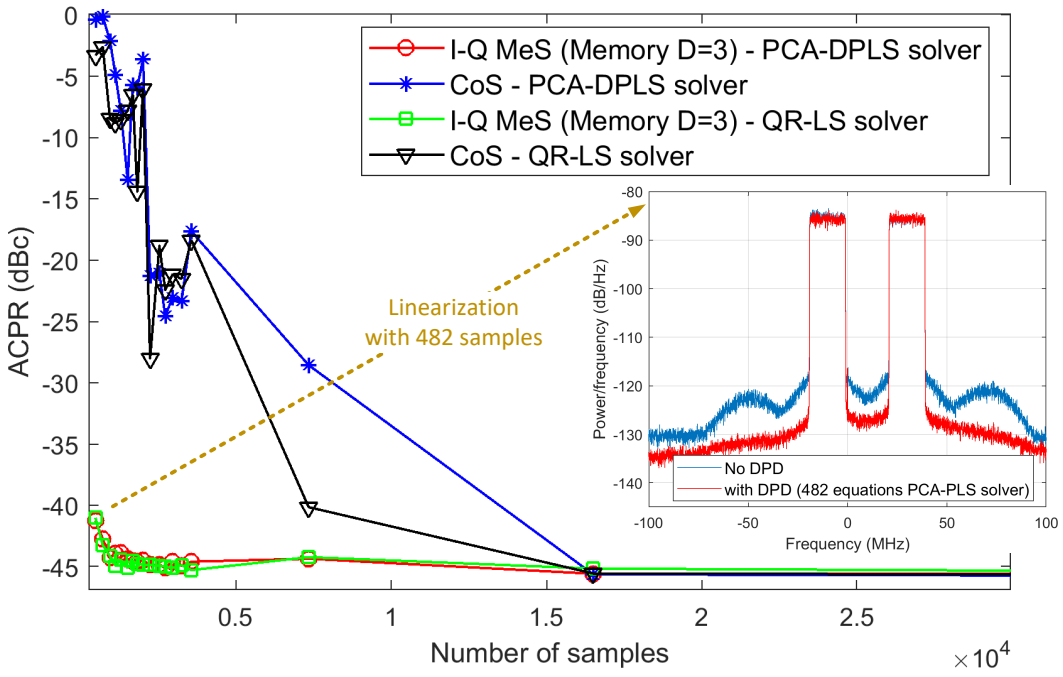

Figure 5.36: DPD linearization: ACPR values vs. identification samples.

consecutive samples (CoS).

Table 5.1: DPD Performance Comparison.

Solver (GMP 108 coeff.)	Technique/ # Samp.	Coeff. ID.	NMSE [dB]	ACPR [dBc]	SP* [%]
No DPD	– / 259,823	–	-25.9	-33.7	100
DPD QR-LS	CoS / 16,513	108	-36.4	-45.5	6.4
DPD QR-LS	MeS / 1,522	108	-36.1	-45.1	0.6
DPD PCA-DPLS	CoS / 16,513	10	-36.2	-45.6	6.4
DPD PCA-DPLS	MeS / 2,731	10	-36.1	-45.1	1.1

*SP is the percentage of selected samples with respect to the total.

Properly combining the proposed mesh selecting method with feature extraction techniques such as the PCA-DPLS (e.g., only 10 coefficients with 1.1% of the total available samples are required to meet the -45 dBc of ACPR) the computational complexity in the identification subsystem can be reduced by a factor of 65 in comparison to using the classical QR-LS and consecutive samples selection (e.g., 120 coefficients with 6.4% of the total available samples are required to meet the -45 dBc of ACPR).

5.5.4 Memory I-Q Mesh-Selecting for Training Artificial Neural Networks for DPD Linearization

The first set of results were obtained using the Chalmers Weblab (described in section 5.4.2 of this Chapter) with a 40 MHz OFDM-like test signal. A feed-forward ANN composed by 4 hidden layers with the following neuron configuration per layer [20, 20, 10, 10] was used for DPD linearization purposes. The original number of samples available to train the ANN was 230,160 samples. By applying the memory I-Q MeS, it is possible to reduce the number of training samples by a given reduction factor (RDF) without significantly degrading the performance of the trained DPD. The number of training samples are approximately the original number (i.e., 230,160 samples) divided by the reduction factor. Fig. 5.37 shows the results obtained when training the ANN considering several reduction factors using the memory I-Q MeS technique and simply removing consecutive samples.

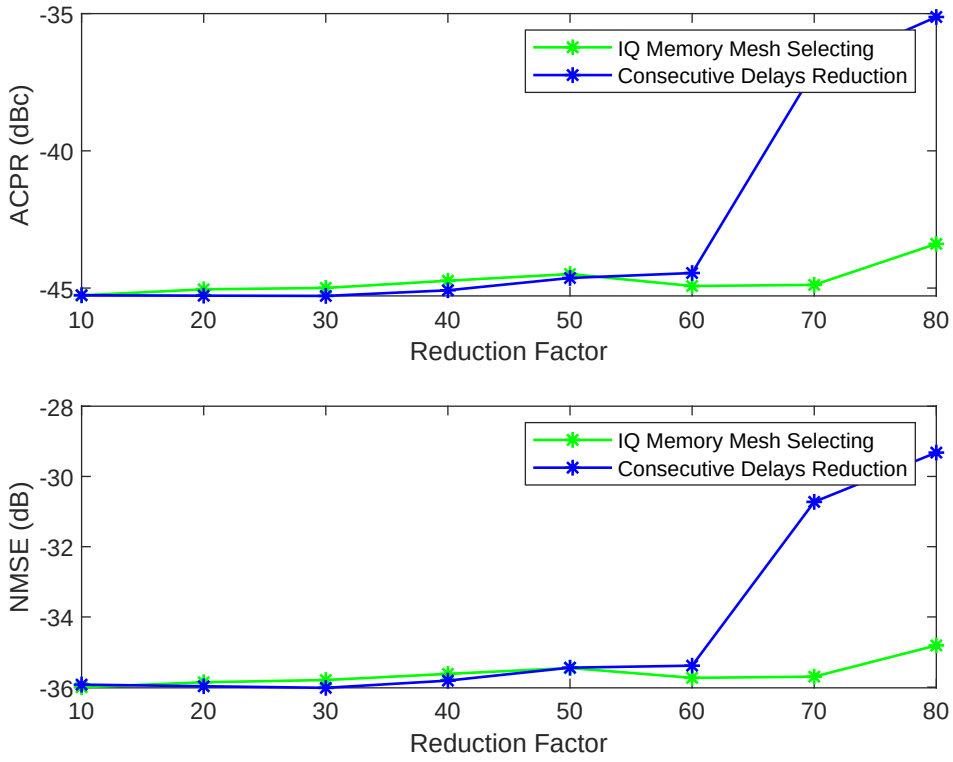

Figure 5.37: NMSE and ACPR after DPD linearization using ANN for different reduction factors.

As shown in Fig. 5.37 for reduction factors higher than 60, the memory I-Q MeS is more robust against NMSE and ACPR degradation than simply remove consecutive samples. Another contribution of MeS method is that the performance of DPD is degraded in a more stable curve in comparison to the CoS methods. The reason behind this effect is because the MeS method is able to keep the distribution and statistic sufficiency of the original signal by maintaining the

multi-histogram distribution. Instead, with the CoS method removing some relevant samples may cause a significant degradation by losing the statistic sufficiency suddenly.

In addition, by considering a reduction factor of 70 (taking into account both I-Q MeS and CoS approaches), thus the number of training samples is approximately 3K as 230,160/70=3,288 samples. We compare the evolution of the NMSE and ACPR when updating the DPD along several iterations with different data, which is illustrated in Fig. 5.38. It is obvioused that MeS provides a better and more stable performance than the CoS.

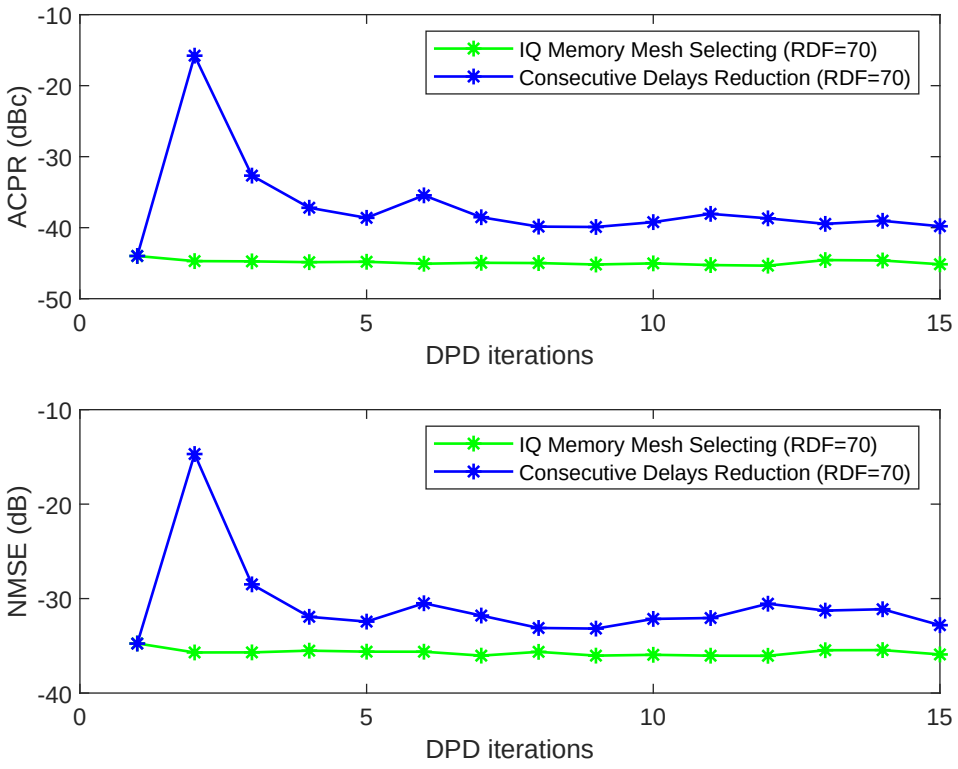

Figure 5.38: Evolution of the ACPR and NMSE for different ANN DPD iterations, considering a RDF of 70 with memory I-Q MeS and CoS.

Further more, the same feed-forward ANN configuration [20, 20, 10, 10] was used to linearize the LMBA described in section 5.4.3 of this Chapter. With the ANN DPD it is possible to meet the linearity specs (i.e., ACPR<-45 dBc) when driving the LMBA with a test signal composed by 4 non contiguous channels of 64 QAM modulated LTE-20 MHz signals spread in 200 MHz instantaneous bandwidth, as shown in Fig. 5.39.

In order to linearize the LMBA when operated with such wide-band signal, the amount of data required at each update iteration is quite important. However, by using the memory I-Q mesh-selecting algorithm it is possible to reduce the number of samples and still maintain the linearization performance of the ANN DPD to meet the targeted linearity specs. In Fig. 5.40 we can see the time required for the ANN to update the coefficients considering different reduction

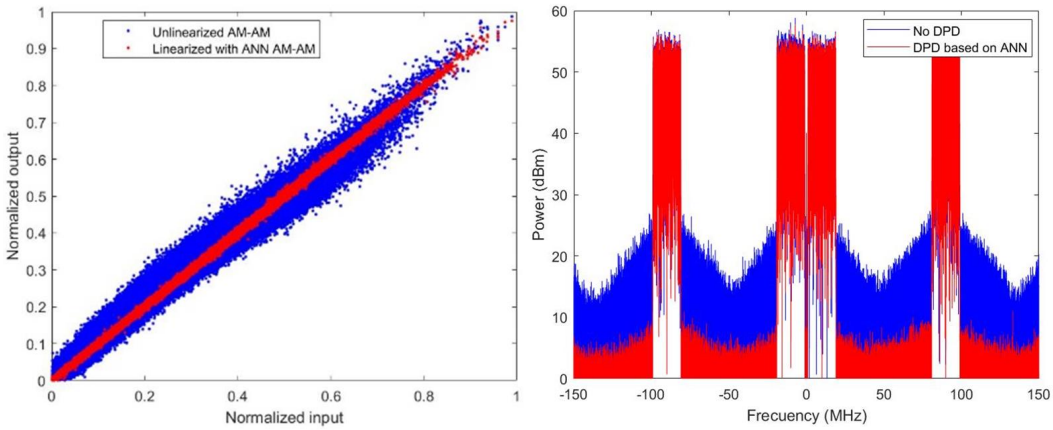

Figure 5.39: AM-AM and Output Power Spectra before and after ANN DPD considering a 200MHz multi-carrier signal.

factors (RF), while Fig. 5.41 shows the degradation suffered in terms of NMSE and ACPR when reducing the number of samples for the training.

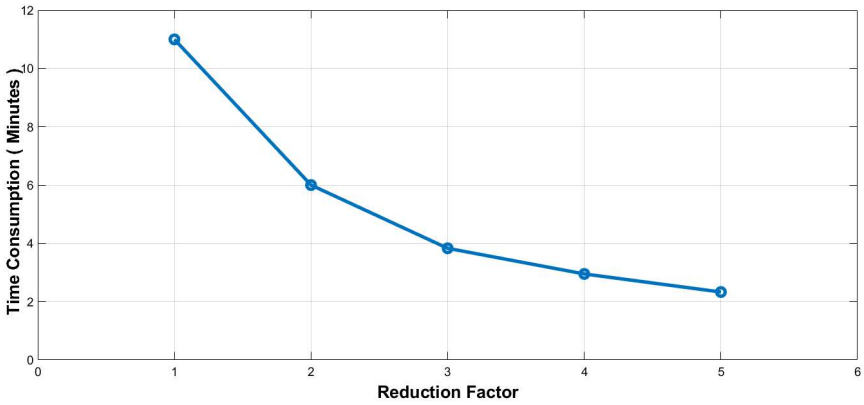

Figure 5.40: ANN DPD training time for different reduction factors

As observed, thanks to the memory I-Q MeS method it is possible to trade-off the computational complexity (and time) and the linearity performance, in terms of ACPR and NMSE degradation.

5.6 Conclusion

In this Chapter we have presented several mesh-selecting approaches that resulted from modifying the original AM-AM (or ABS) memoryless mesh-selecting method presented in Chapter 4, by considering memory or the I-Q components to build the mesh. In addition, the PA output signal can also be included to generate the mesh. After a comparison of the different mesh-selecting

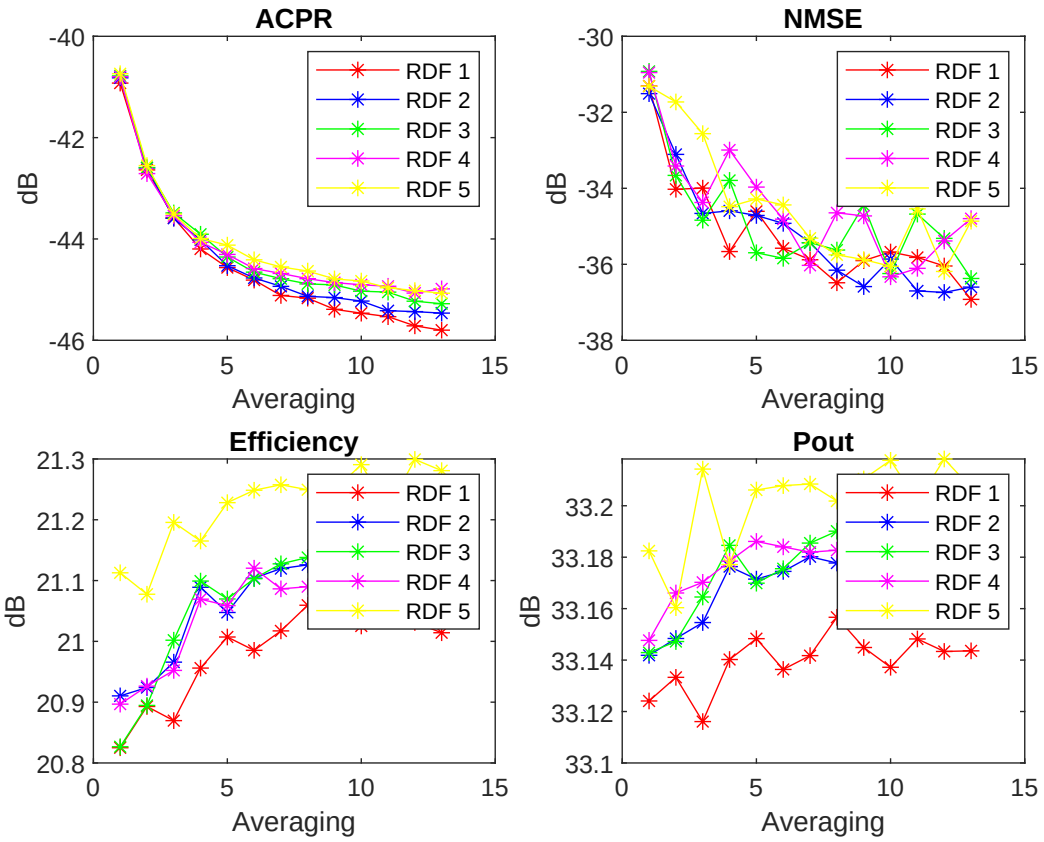

Figure 5.41: Evolution of the ACPR, NMSE, Power Efficiency and Output Power along several update iterations when considering different reduction factors of the training samples.

techniques, we have concluded that by properly choosing the mesh-density and memory depth, the memory I-Q mesh-selecting method presented superior trade-off between the sample reduction capabilities and the memory storage requirements.

In order to prove the possibility of further reducing the computational complexity in the DPD adaptation subsystem, the memory I-Q mesh-selecting method has been combined with a feature extraction dimensionality reduction technique. Considering the example of the GMP-based DPD linearization of the LMBA, excited with a 60 MHz test signal, the computational complexity in the identification subsystem has been reduced by a factor of 65 when using the memory I-Q MeS combined with the PCA-DPLS solver in comparison to using the classical QR-LS solver and consecutive samples selection.

Finally, the memory I-Q MeS has been proved to be of crucial interest when training ANN for DPD purposes. In an adaptive DPD approach where the ANN has to be periodically trained, the memory I-Q MeS can be used to trade-off the required training time and the linearity performance, in terms of ACPR and NMSE degradation.

Chapter 6

Automatic Tuning to Find the Basis Functions of a Digital Predistorter

6.1 Introduction

An underfitted model lacks of essential coefficients in the model description, while an overfitted model contains more parameters than the model really needs. Both underfitted and overfitted models tend to misrepresent the training data and will therefore have poor predictive performance.

One of most demanding part of the DPD design is to choose a proper DPD behavioral model and then properly select the required basis functions, which assuming a polynomial implementation, depends on the number of memory terms and order of the polynomials chosen for the model. This procedure not only requires large knowledge of the DPD algorithm and experience dealing with DPD linearization, but sometimes also to deal with non-straightforward relations. Most of the times, the trail and error is the most popular method, but is time consuming and provides no guarantee of the optimality of the solution found, since it is sometimes not possible to try manually all the possible combinations. Additionally, our preconceptions may limit the testing coverage and ignore some possible good results.

In the field of DPD linearization, dimensionality reduction techniques are used with a double objective. On the one hand to ensure a proper, well-conditioned parameter identification and, on the other hand, to reduce the number of coefficients to be estimated and thus relaxing the computational complexity and memory requirements of a hardware implementation. As previously discussed, several approaches have been proposed targeting both robust identification and model order reduction such as LASSO [Wis08], the Ridge regression [Gua12b], the sparse Bayesian learning (SBL) algorithm [Pen16a] or the orthogonal matching pursuit (OMP) [Rei15].

In this Chapter we propose a slightly different approach, where machine learning is applied to

trade-off the search time and the performance of the optimum configuration found. Therefore, the use of a global optimization method is proposed to select the optimum DPD parameter configuration in the whole search area. Experimental results using the UPC Weblab are presented and discussed.

6.2 Adaptive Lipschitz Optimisation (adaLIPO)

In this Chapter we have considered the use of the adaptive Lipschitz optimisation (adaLIPO) algorithm to determine the best parameter configuration (e.g., memory terms, non-linear order) of a DPD behavioral model. A brief description of the adaLIPO algorithm is given in the following.

The smoothness-based approach to global optimisation assumes that the system presents some regularity with respects to the input. In particular, the use of the Lipschitz constant (the bound of the first derivative of a Lipschitz function, i.e., a continuous function limited in how fast it can change) in [Shu72, S.A72], played a key role in the development of many efficient global optimisation algorithms. The adaLIPO algorithm proposed in [Mal17] is oriented to exploit the global smoothness of the unknown function for global optimisation and, according to the authors, it can achieve faster rates of convergence on globally smooth problems than the previously known methods.

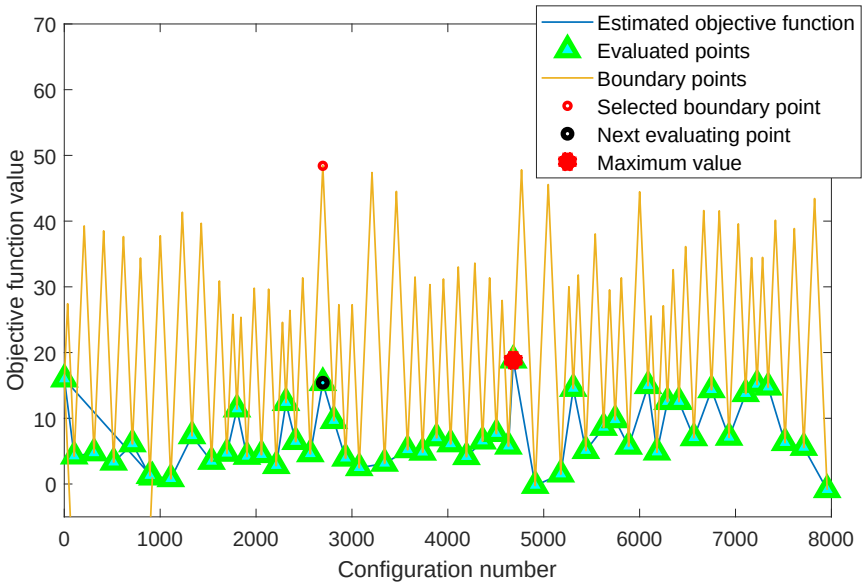

Figure 6.1: Graphical example of the adaLIPO algorithm.

The principles of the adaLIPO algorithm are summarized in the example depicted in Fig. 6.1, consisting in the optimisation of a one-dimensional function. The blue line is the given function to be optimised. The basics of the adaLIPO algorithm consist in maintaining a piece-wise upper

bound of the given function to be optimized, according to the up-to-date Lipschitz constant. At the beginning, for few iterations, the Lipschitz constant and the bound (the orange line) are estimated. Then, as shown in Fig. 6.1, the algorithm finds the maximum value of the upper bound (the red circle) and obtains its real value (the black circle) after the evaluation of the given function. In the next iteration, the up-to-date Lipschitz constant and the bound will be updated and using a Bernoulli distribution, among all cross points of the upper bound, the next point will be selected. The optimizing procedure will converge after several iterations.

6.3 Overview on the Tuning Parameters of the GMP Behavioral Model

As described in Chapter 3, the Generalized Memory Polynomial (GMP) proposed by Morgan et al. in [D. 06], is one to the most popular PA behavioral models because it allows to cover several cross-memory interactions.

For convenience, let us review again the input-output relationship in the DPD taking into account a GMP model

$$\begin{aligned} x[n] = & \sum_{i=0}^{N_a-1} \sum_{p=0}^{P_a-1} \alpha_{pi} \cdot u[n-\tau_i^a] \left| u[n-\tau_i^a] \right|^p \\ & \sum_{j=1}^{M_b} \sum_{i=0}^{N_b-1} \sum_{p=1}^{P_b-1} \beta_{pij} \cdot u[n-\tau_i^b] \left| u[n-\tau_i^b-\tau_j^b] \right|^p \\ & \sum_{j=1}^{M_c} \sum_{i=0}^{N_c-1} \sum_{p=1}^{P_c-1} \gamma_{pij} \cdot u[n-\tau_i^c] \left| u[n-\tau_i^c+\tau_j^c] \right|^p \end{aligned} \tag{6.1}$$

The first term is the expression of the memory polynomial behavioral model, with P_a being the nonlinear order and N_a the memory depth. In general, for wireless communications applications with fractional bandwidth (defined as the ratio between the signal bandwidth and the centre frequency of operation) significantly smaller than 1, only odd order terms of the polynomial are considered to characterize the intermodulation distortion (the effect of harmonic distortion is considered negligible since it can be filtered). However, in the global optimization search with the adaLIPO algorithm we will consider both odd and even terms for the polynomial orders. The second and third terms of (6.1) introduce both negative and positive cross-term delays to the memory polynomial model, with P_b and P_c being the polynomial orders and N_b, N_c, M_b, M_c relates to memory depths. The coefficients describing the model are α_{pi}, β_{pij}, while τ^a, τ^b and τ^c (with $\tau \in \mathbb{Z}$ and $\tau_0 = 0$) are the most significant non-consecutive delays of the input signal $u[n]$ that better contribute to characterize memory effects. Thus, the total number of coefficients of

the GMP model is $O = P_a N_a + P_b N_b M_b + P_c N_c M_c$. Theoretically, each of these parameters can be tuned independently in the DPD design within reasonable ranges.

As discussed in Chapter 3 the GMP behavioral model is a simplification of the general Volterra series, and can be seen as a good trade-off between complexity and accuracy. The number of possible combinations for defining the optimum number of nonlinear orders and memory parameters is huge and is highly dependent on the specific PA and transmitted signal.

In the following we will use the adaLIPO global optimization algorithm to determine the optimum combination of parameters to meet the linearity specifications with the minimum number of coefficients possible. In order to favor certain combinations we will have to properly define a cost function that prioritizes the solutions that uses the minimum number of coefficients with the maximum linear output power possible.

6.4 Experimental Results

The following experimental results were obtained using the UPC Weblab depicted in Fig. 5.11 and described in section 5.4.1 of Chapter 5. The test signal was an LTE 10 MHz signal. Taking into account the bandwidth expansion that experience the predistorted signals, the bandwidth of the transmitted signal was determined by the bandwidth limitations of the Zynq-AD936 platform (up to 56 MHz bandwidth with a clock rate of 61.14 MSa/s).

In the global search, a part from the GMP parameters, an internal gain (K_{Gain} that controls the input back-off of operation) is considered into the tuning area. The range of each value is listed below, with a total number of possible combinations around $5.8 \cdot 10^6$.

- Polynomial order $P_a \in [1,\ 13]$, with $P_a \subset \mathbb{N}$,.
- Polynomial order $P_b \in [1,\ 13]$, with $P_b \subset \mathbb{N}$,.
- Polynomial order $P_c \in [1,\ 13]$, with $P_c \subset \mathbb{N}$,.
- Memory length $N_a \in [1,\ 5]$, with $N_a \subset \mathbb{N}$,.
- Memory length $N_b \in [1,\ 5]$, with $N_b \subset \mathbb{N}$,.
- Memory length $N_c \in [1,\ 5]$, with $N_c \subset \mathbb{N}$,.
- Memory length $M_a \in [1,\ 3]$, with $M_a \subset \mathbb{N}$,.
- Memory length $M_b \in [1,\ 3]$, with $M_b \subset \mathbb{N}$,.
- Internal Gain $K_{Gain} \in [0.8,\ 1.3]$, with $K_{Gain} \subset \mathbb{Q}$, with one fractional digit.

The adaLIPO algorithm will therefore search for the best DPD parameters configuration taking into account a given score (or cost) function. The score function that we want to maximize, i.e., the higher the better, is defined in (6.2),

$$\begin{aligned} \boldsymbol{Score} = \lambda_1 \cdot (-\text{NMSE} + \text{NMSE}_{th}) + \lambda_2 \cdot \max\big((-\text{ACPR} + \text{ACPR}_{th}), 0\big) \\ +\lambda_3 \cdot \max\big(\text{Max N}^\circ \text{ coeff} - \text{N}^\circ \text{ coeff GMP}), 0\big) \cdot \alpha + \lambda_4 \cdot (\text{P}_{out} - \text{P}_{out,th}) \cdot \alpha \end{aligned} \tag{6.2}$$

with $\alpha = \max\big((-\text{ACPR} + \text{ACPR}_{th}), 0\big)$. As observed, in the definition of the score function we have tried to find a balance among linearity performance (NMSE and ACPR), computational complexity (N° of coefficients) and power efficiency (indirectly, through the mean output power). Consequently we have included the NMSE (the lower the better), the ACPR (the lower the better but it is necessary to at least meet the threshold on -45 dBc), the number of coefficients (the lower the better but cannot exceed the *Max N° coeff* threshold) and the mean output power (the higher the better). The score function in (6.2) includes some thresholds and some weights to give more importance to certain figures of merit. Therefore, for example, unless we meet the ACPR specs (determined by the threshold ACPR_{th}) the benefits of using few coefficients or delivering high mean output power are not taken into account. The specific weights and thresholds used in this particular adaLIPO search are:

- Weights: $\lambda_1 = 1$; $\lambda_2 = 20$; $\lambda_3 = 0.5$; $\lambda_4 = 5$.
- Thresholds: NMSE_{th}=-30 dB; ACPR_{th}=-45 dBc; Max N° coeff=100; $\text{P}_{out,th}$=30 dBm.

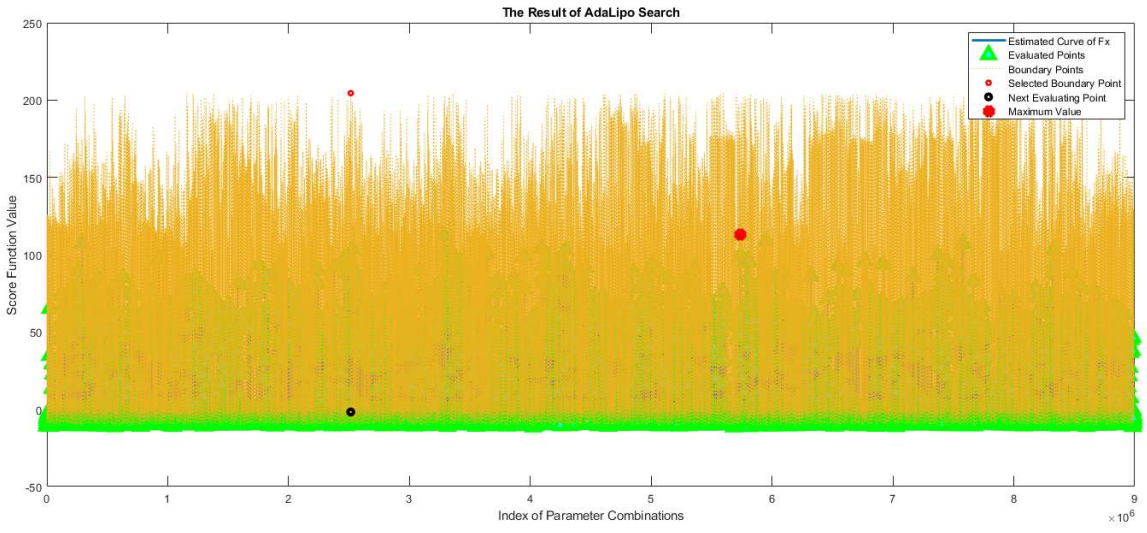

Figure 6.2: Results of the AdaLIPO search on the best configuration of the GMP model for DPD linearization.

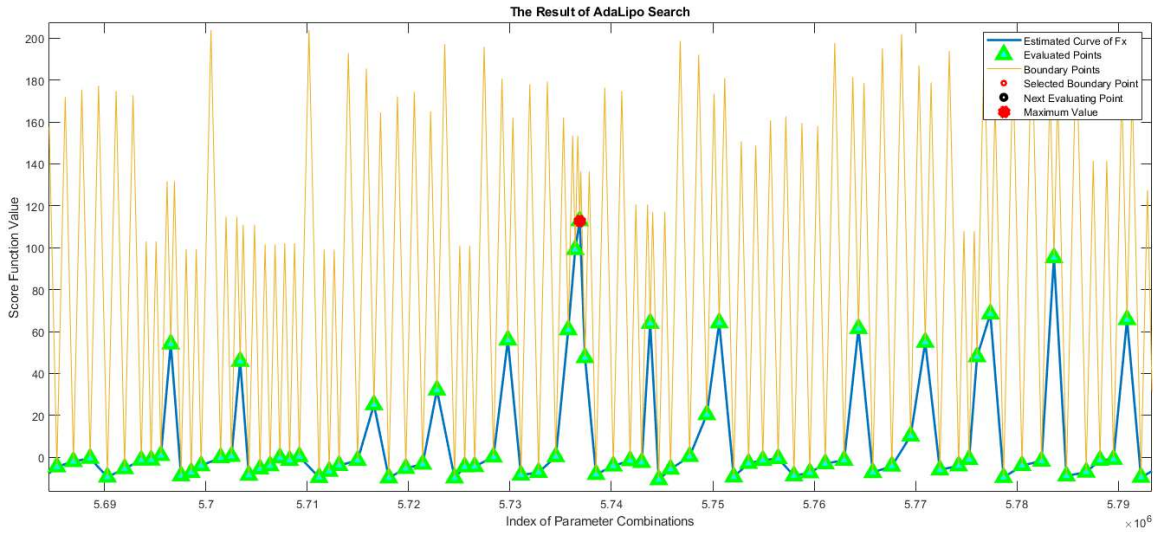

Figure 6.3: Detail of the adaLIPO search on the best GMP model configuration for DPD linearization

Fig. 6.2 and its zoomed version in Fig. 6.3 show the score values of the different configurations considered in the adaLIPO search. Table 6.1 lists the top best configurations found by the adaLIPO search. Note that some of the configurations providing very good score (and thus finding a good trading-off between linearity, computational complexity and power efficiency) are not obvious, and they wouldn't have been conceived in a manual search. Fig. 6.4 and Fig. 6.5 show the AM-AM and AM-PM characteristics and the output power spectra, respectively, before and after DPD linearization when considering one of top best configurations of the GMP model DPD reported in Table 6.1.

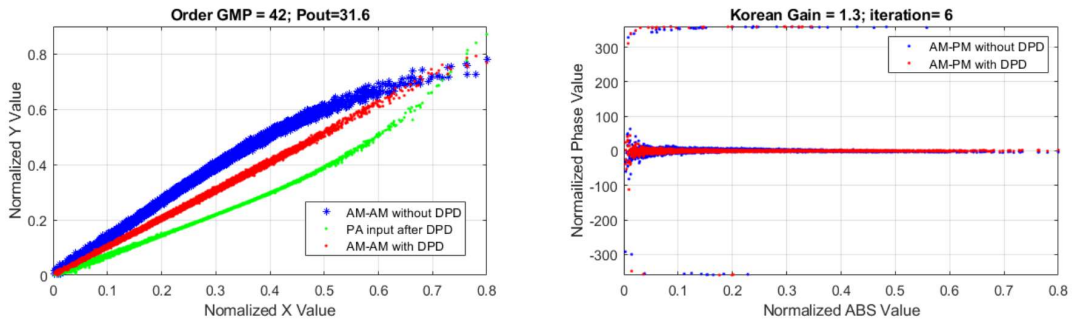

Figure 6.4: The AM-AM and AM-PM characteristics using of one of top best configurations of the GMP model DPD.

Table 6.1: Linearization results with the LTE 10 MHz bandwidth signal

Opt. Config.	ACLR (dBc)	NSME (dB)	No.of Coeff	Optput Power (dBm)	Total Score
PA=4, PB=2, PC=1, LA=3, LB=3, LC=4, MB=1, MC=3, KG=1.3	-48.0	-31.1	38	31.39	99.1
PA=1, PB=4, PC=10, LA=2, LB=2, LC=1, MB=1, MC=1, KG=1.3	-48.7	-29.2	42	31.22	108.2
PA=4, PB=2, PC=1, LA=3, LB=3, LC=4, MB=1, MC=3, KG=1.3	-46.1	-29.6	21	31.37	67.0
PA=5, PB=1, PC=3, LA=2, LB=1, LC=5, MB=2, MC=1, KG=1.3	-46.4	-29.7	29	31.41	71.1
PA=5, PB=1, PC=10, LA=2, LB=4, LC=1, MB=1, MC=1, KG=1.3	-47.9	-28.3	42	31.56	92.2
PA=6, PB=3, PC=9, LA=4, LB=2, LC=1, MB=2, MC=1, KG=1.3	-46.9	-30.2	70	31.34	59.5
PA=8, PB=4, PC=1, LA=1, LB=2, LC=1, MB=2, MC=1, KG=1.3	-48.7	-30.1	40	31.37	110.6
PA=10, PB=4, PC=1, LA=1, LB=1, LC=1, MB=2, MC=3, KG=1.3	-47.1	-31.7	38	31.42	81.9
PA=11, PB=3, PC=2, LA=2, LB=3, LC=1, MB=2, MC=3, KG=1.3	-47.1	-30.5	62	31.49	68.0

6.5 Conclusion

In this Chapter, we used a global optimization algorithm, adaLIPO, to find the best parameter configuration of a GMP behavioral model for DPD. In our particular context, the adaLIPO algorithm is capable to find the best configuration among $5.8 \cdot 10^6$ possible combinations constituting the searching space. The definition of the score function is of key importance since prioritizes the different figures of merit (FOM) under evaluation. In our particular case, the linearity performance (evaluated in terms of NMSE and ACPR), the model complexity (evaluated in terms of number of coefficients) and the power efficiency (indirectly evaluated in terms of mean output power) are the FOMs weighted to define the adaLIPO score function.

The proposed approach has been validated through experimental results obtained with the

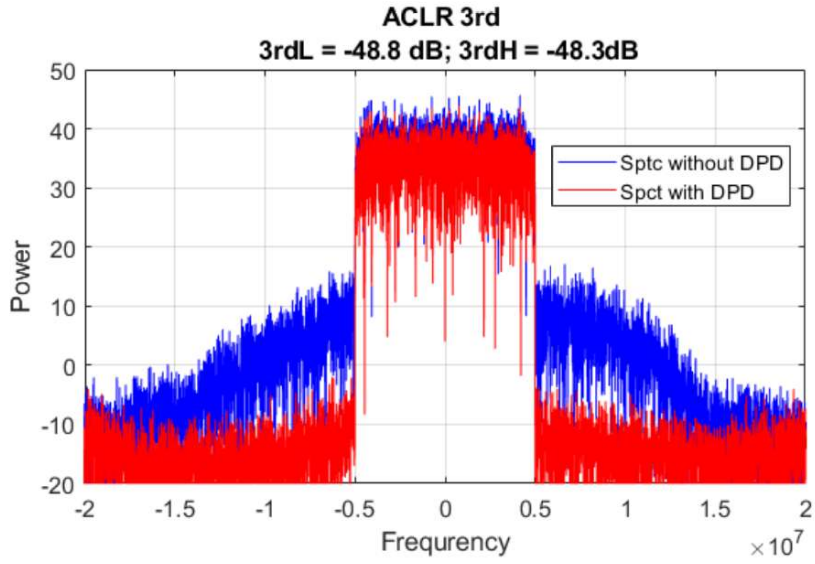

Figure 6.5: Linearized and unlinearized spectra using one of top best configurations of the GMP model DPD.

UPC Weblab. In our particular case, some of the best configurations found by the adaLIPO algorithm to linearize a class-J PA operated with a LTE-10 MHz signal were able to deliver around 31.4 dBm of mean output power, with approximately -48 dBc of ACPR and -31 dB of NMSE, with a total number of GMP parameters around 40. Besides, having a look at the top best configurations we can observe that are quite different among each other, which shows how the adaLIPO search can find configurations that may have never been considered in a manual tuning.

Chapter 7

Auto-Tuning and Digital Predistortion of Power Amplifiers with Several Free-Parameters

7.1 Introduction

In this Chapter we present an auto-tuning approach for dual-input PAs, using a combination of global optimization search algorithms and adaptive linearization in the optimization of a multiple-input PA. The objective is to exploit the extra degrees of freedom provided by dual-input topologies to enhance the power efficiency figures along wide signal bandwidths and high PAPR values, while being compliant with the linearity requirements. By using heuristic search global optimization algorithms, such as the simulated annealing (SA) or the adaLIPO algorithm (described in section 6.2 in Chapter 6) it is possible to find the best parameter configuration for PA biasing, signal calibration and DPD linearization to help mitigating the inherent trade-off between linearity and power efficiency.

All power amplifier architectures based on active load modulation, such as Doherty, LMBA and outphasing, rely on the non-linear interaction between multiple transistors to enhance the average efficiency in presence of modulated signals with large dynamic range. While these architectures can be designed with a single RF input to simplify their use in a transmitter, there are benefits in maintaining separate inputs controlled by different upconverter chains. For example, some of the record bandwidth Doherty PAs have separate inputs [Bat11, And13], and the advantages of dual-input Doherty compared to single input have been explored in specific studies [Pia18]. This does not mean that single-input Doherty PAs with good bandwidth do not exist, see for example [Gio14, Rub18]. However, the additional degrees of freedom offered by the separate inputs can be used to optimize the performance on the same or larger bandwidth, or to improve other performance metrics such as linearity and average efficiency [Dar16, Kal19].

Similar considerations can be made for the LMBA, which was originally proposed for telecom applications as a dual-input structure [Qua18]. However, single-input solutions have also been proposed [Ped18], and critically compared to the dual-input case [Col18]. Also in this case the single input solutions are viable, with a clear advantage in terms of simplicity, but at the cost of compromised performance. The outphasing represents a very different case, since it requires in principle separate inputs that are with constant amplitude and phase modulated only. This is best achieved with separated inputs; however, some attempt has been made to realise single-input narrowband outphasing circuits [Bar15] that demonstrated comparable performance to dual-input cases.

Focussing on dual-input PAs, it is reasonable to state that evaluating the performance of a set of free-parameters often requires experimental cross-validations with significant computational cost and time, especially when the search space is vast. The idea behind global optimization is to find the optimum output value (i.e., the globally best solution in the presence of multiple local optima) of an unknown function with limited evaluations. Several techniques have been proposed in the literature [Hor02] to find the most suitable set of parameters among large tunable ranges. Among the exact methods we can find, for example, Bayesan search algorithms, branch and bound algorithms, adaptive stochastic search methods or successive approximation methods; while among the heuristic methods we can find, for example, evolution strategies (e.g., genetic algorithms), the tabu search or the simulated annealing [Hor02].

In this Chapter we propose an auto-tuning approach to take advantage of the possibilities given by the extra degrees of freedom in dual-input PAs by using the simulated annealing and the adaLIPO heuristic search approaches. The objective is to find the best configuration for PA biasing, signal calibration and digital predistortion linearization that guarantees the linearity specifications, in terms of Normalized Mean Squared Error (NMSE), Error Vector Magnitude (EVM) and Adjacent Channel Leakage Ratio (ACLR), and maximizes power efficiency of dual-input PAs.

7.2 Description of the LMBA and its Free-Parameters

7.2.1 Multiple-Input Power Amplifier Architectures

All the aforementioned PA architectures based on active load modulation (Doherty, LMBA and outphasing) and separate inputs can be visualised, by generalisation, as the block diagram of Fig. 7.1. The amplifier has N RF inputs, a drain or collector bias (or, in some cases, more than one at different rail voltages), and M different gate voltages to control the stages independently. A typical example of PA with independent gate voltages is the Doherty, where the Main is biased in class AB, and the auxiliary in class C. The instantaneous amplitude and phase of each input,

as well as the M gate bias voltages, can be controlled and adjusted separately, allowing for a large number of degrees of freedom that can be exploited to optimize a target Figure of Merit (FoM).

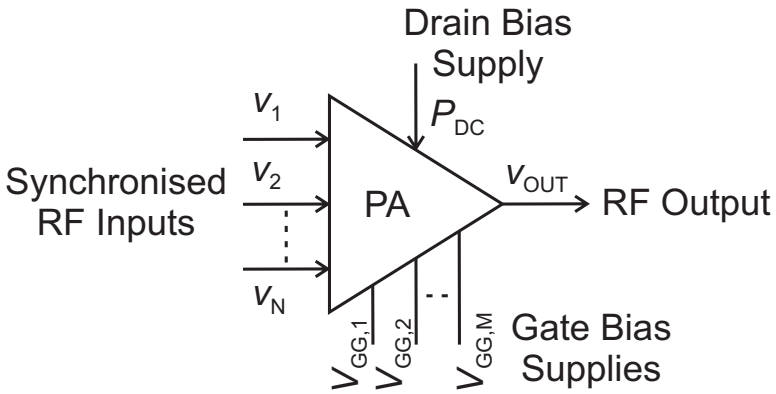

Figure 7.1: General block diagram of a multiple-input power amplifier.

The LMBA presented in [Qua18] is used as Device Under Test (DUT). Its block diagram is shown in Fig. 7.2. There are two separate RF inputs; v_1 controls the balanced amplifier (BPA) pair, based on two CGH40025F transistors from Wolfspeed, biased in class AB with $V_{GG,1}$ at -2.8 V corresponding to 80 mA of quiescent drain current; v_2 controls the Control Signal Power (CSP) amplifier, also based on a CGH40025F, and biased in class C, with $V_{GG,2}$ left as a free parameter within the range of DC voltages -5.5 V to -3.5 V. The matching networks and the output hybrid couplers are based on soft-board microstrip networks, with surface mount device (SMD) capacitors and resistors for the by-pass and stabilisation networks. An off-the-shelf hybrid coupler is used on the input. The circuit is mounted on an aluminium fixture, and SMA connectors coaxial launchers are used for the RF ports.

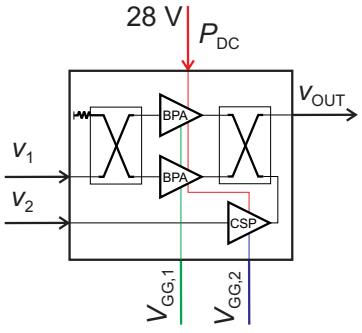

Figure 7.2: Block diagram of the LMBA used as DUT in this paper and described in detail in [Qua18].

The continuous wave (CW) measurements reported in [Qua18] showed, a maximum power exceeding 63 W, and a 8 dB back-off efficiency exceeding 39%, over the 1.7–2.5 GHz frequency

range. Modulated signal measurements were also performed with 5 MHz and 20 MHz channel LTE signals, showing the 'linearisability' (which indicates whether the PA is able or easy to be linearized in practical implementation) of the LMBA under these conditions. Both sets of measurements were performed with a manual search for the optimum amplitude, phase and bias settings. In particular, the relative phase was maintained at a constant offset that led to a good compromise between output power and back-off efficiency, while the relative amplitude was following a square relation between the BPA and CSP inputs [Qua18].

The relative amplitude between the two inputs will be now defined through a shaping function, where it will be possible to tune two degrees of freedom. Similarly, the relative phase will be also a parameter to be optimized.

7.2.2 Free-Parameters of the Dual-Input PA

The DUT used to validate the proposed tuning approach is the LMBA presented in [Qua18]. In order to maximize the power efficiency of the LMBA when operated with wideband signals presenting high PAPR while meeting at the same time the in-band and out-of-band linearity specifications, we have considered some free-parameters, which when properly tuned, should help to mitigate the inherent linearity versus power efficiency trade-off.

The FOMs that we are considering to define the cost or objective function and determine the values of the free-parameters are: the ACLR, the PA power efficiency (η), the NMSE and the EVM. The out-of-band (defined in terms of ACLR) and the in-band (defined in terms of EVM) linearity specifications must be met (being compliant with the communications standard) while the power efficiency is a figure of merit that justifies the election of one topology instead of another.

The parameters to be tuned are schematically depicted in Fig. 7.3 and listed in the following:

- Shaping functions parameters; Offset percentage (OP) and degree of the root p.
- Relative phase (ψ_{rel}) between the BPA and CSP signals.
- The DC gate voltage of the CSP amplifier, $V_{\mathrm{GG},2}$.
- The maximum PAPR (max_PAPR) in dB, of the complex baseband signal ($u[n]$) to be sent.
- The baseband gain (GainBB), which controls the mean input power and thus the input back-off (IBO).

As depicted in Fig. 7.3, by applying some CFR technique (such as, peak cancellation in [Lop14]) it is possible to limit the maximum PAPR. Consequently, the input back-off can

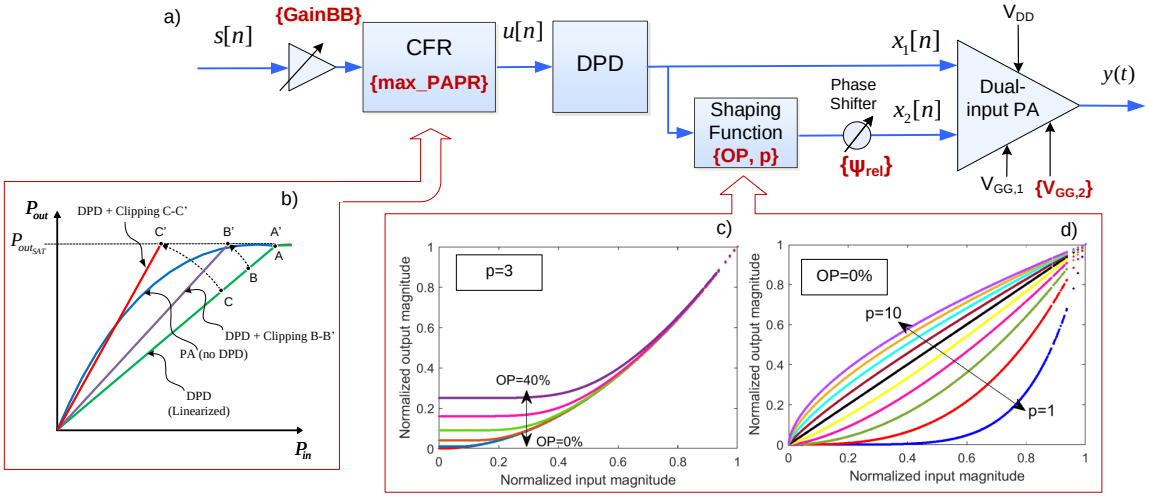

Figure 7.3: Simplified block-diagram showing the degrees of freedom of the dual-input PA system including DPD and CFR.

be reduced (by increasing GainBB) and thus operating closer to compression, as graphically described in Fig.7.3-b).

The BPA signal is defined directly as $x_1[n] = x[n]$. The CSP signal $x_2[n]$ is generated by using a shaping function previously employed in envelope tracking (dynamic supply modulation) and outphasing (dynamic load modulation) applications [Gil19], because it provides two degrees of freedom. One of the parameters can be tuned to allow some level of detroughing (preventing the signal from dropping to zero), while the other parameters controls the shape of the AM/AM characteristic. More specifically, the CSP signal $x_2[n]$ is defined as

$$x_2[n] = x_{sf}[n]\ e^{i\psi_{rel}} \tag{7.1}$$

with ψ_{rel} being the relative phase (in radians) between the BPA and the CSP signals; and where the signal after the shaping function $x_{sf}[n]$ is defined as

$$x_{sf}[n] = A_s[n]\ K_0\ e^{i\phi_x} \tag{7.2}$$

where $K_0 = \frac{max\{|x[n]|\}}{max\{A_s[n]\}}$, $\phi_x = phase\{x[n]\}$ and the amplitude relation between signals is characterised by the following shaping function:

$$A_s[n] = \left((x_{min})^6 + (|x[n]|)^6\right)^{1/p} \tag{7.3}$$

with p being the degree of the root (p^{th} root) and the lower bound x_{min} defined as

$$x_{min} = max\{|x[n]|\}\ OP \tag{7.4}$$

where OP is the offset percentage, defining the threshold for the detouring function. The input-output characteristics of the shaping function when sweeping the parameters OP and p are depicted in Fig.7.3-c) and Fig.7.3-d).

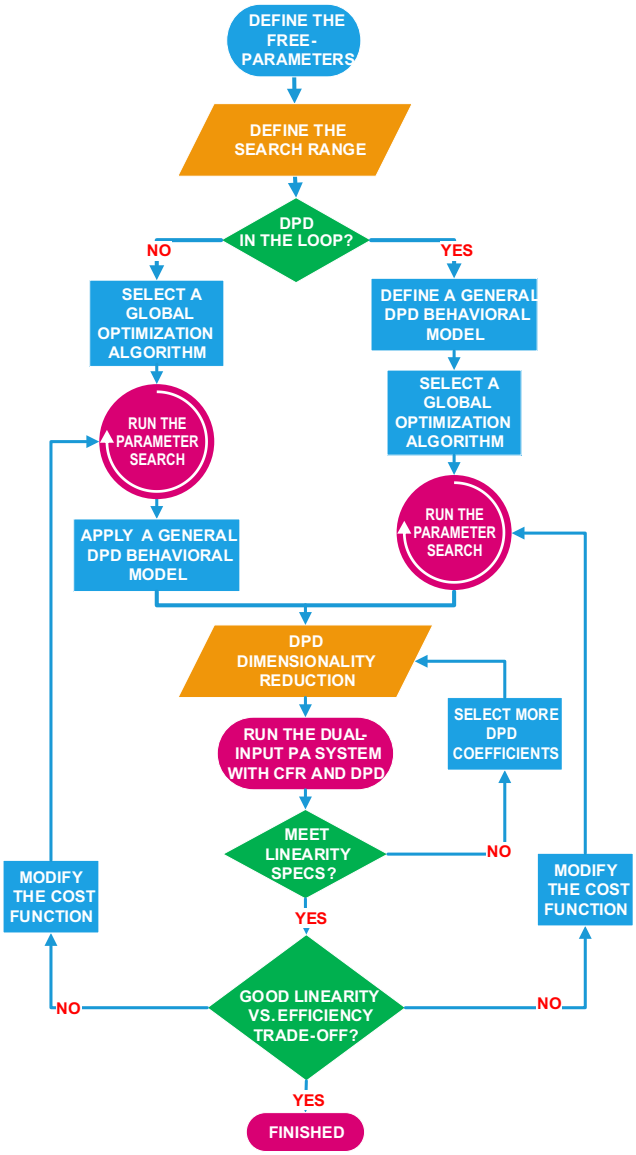

Figure 7.4: Flowchart of the proposed auto-tuning technique for dual-input PAs.

7.3 Description of the Dual-Input PA Auto-tuning Approach

7.3.1 Dual-Input PA Tuning Approach

The tuning approach proposed to configure the dual-input PA is summarized in the flowchart in Fig. 7.4. More specifically, the steps are described in the following:

i. Define the degrees of freedom (free-parameters) to be tuned. Typically, device and system parameters that have an impact on the linearity v.s. efficiency trade-off.

ii. Define the tuning range of free-parameters (upper and lower bounds). Typically, some preliminary tests, or information about the DUT, are necessary to determine this range.

iii. Decide whether to include the DPD in the optimization process. If included, the final linearity specs can be targeted inside the optimization algorithm. If not included, lower linearity specs can be targeted assuming that a later application of DPD will be able to meet system requirements. When considering wide-band signals, where the linearity specifications will be more difficult to meet, it is better to include the DPD in the optimization process. This way we can avoid solutions where the power efficiency is optimum but then the linearity levels (mainly in terms of ACLR) cannot be met (even with DPD) without significantly degrading the original power efficiency figures. When including the DPD in the optimization, the behavioural model needs to be oversized to linearize the dual-input power amplifier under significantly different operation modes. Then, once the optimum configuration is fixed, model order reduction techniques can be applied to the DPD to reduce the number of required parameters.

iv. Choose the optimization algorithm and design the cost (or objective) function. In this cost function, all the FOMs should appear weighted according to their importance. Additionally, some thresholds values for each FOM can be also be defined to further penalize not meeting the desired specifications. This is an important feature when dealing with mandatory system requirements such as ACLR limits.

v. Configure the DUT characterisation and capture input-output data searching for the parameters values until the cost function threshold is achieved.

vi. Carry out an off-line model order reduction of the DPD behavioural model. A feature selection technique, such as the orthogonal matching pursuit (OMP), is used to reduce the number of parameters of the DPD behavioural model and ensure a well-conditioned estimation.

vii. Check the linearity specification after model reduction. If not satisfactory, go back to step 5 and increase the number of coefficients.

viii. Check the linearity vs. power efficiency trade-off which is obtained with the free-parameters found. If not satisfactory, go back to step 4 and re-define the cost function changing its weights and thresholds.

In the following we will provide a more in-depth description of the specific details involving each one of the steps of the proposed tuning approach. Similarly as in Chapter 6, we will use the adaLIPO algorithm for global optimization. In addition, the simulated annealing search algorithm will be also considered.

7.3.2 Digital Predistortion Linearization

Despite the fact that the principles of the closed-loop DPD have been already addressed in section 3.1.2 of Chapter 3, the DPD behavioral model used in this Chapter has not been previously introduced. Therefore, for completeness, we review again the principles of closed-loop DPD.

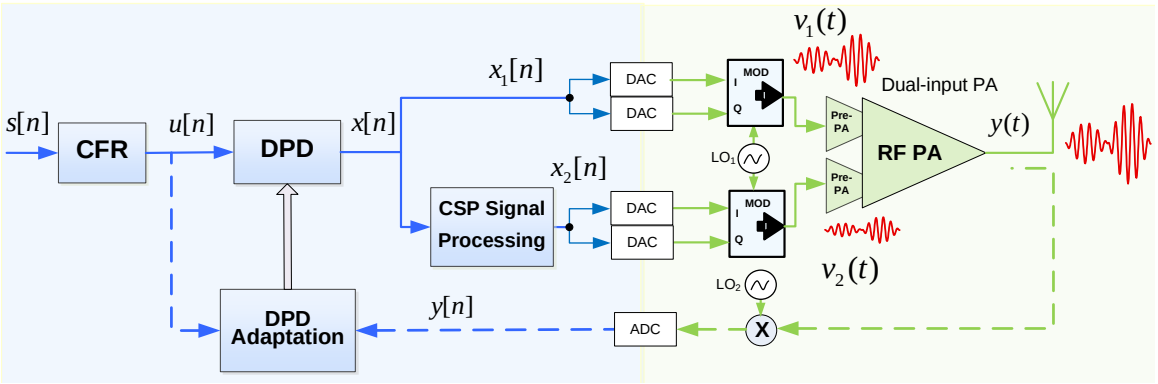

Figure 7.5: Block diagram of the dual-input PA with CFR and DPD linearization.

The block diagram of our closed-loop adaptive DPD architecture is shown in Fig 7.5. The input-output relationship at the DPD block can be described as

$$x[n] = u[n] - d[n] \tag{7.5}$$

where $x[n]$ is the signal at the output of the DPD block, $u[n]$ is the input signal and $d[n]$ is the distortion signal, described by the following simplified Volterra-based model,

$$\begin{aligned} d[n] = & \sum_{i=0}^{N_1-1} \alpha_i \; u[n-\tau_i^1] + \\ & \sum_{j=0}^{N_3-1} \sum_{i=0}^{j} \sum_{k=0}^{i} \beta_{kij} \; u[n-\tau_j^3] \; u[n-\tau_i^3] \; u^*[n-\tau_k^3] + \\ & \sum_{j=0}^{N_5-1} \sum_{i=0}^{j} \sum_{k=0}^{i} \sum_{l=0}^{k} \sum_{s=0}^{l} \gamma_{slkij} \; u[n-\tau_j^5] \; u[n-\tau_i^5] \\ & \qquad u^*[n-\tau_k^5] \; u[n-\tau_l^5] \; u^*[n-\tau_s^5] + \\ & \sum_{i=0}^{N_a-1} \sum_{p=0}^{\frac{P_a-7}{2}} \delta_{pi} \; u[n-\tau_i^a] \; \left|u[n-\tau_i^a]\right|^{2p+6} \end{aligned} \tag{7.6}$$

where τ^1, τ^3, τ^5 and τ^a (with $\tau^{1,3,5,a} \in \mathbb{Z}$ and $\tau_0^{1,3,5,a} = 0$) are the most significant sparse delays of the input ($u[n]$) that contribute to characterise memory effects. As it can be observed in (7.6), the first, third and fifth-order kernels have been included (limiting the number of combinations to avoid repetitions) and for higher odd-order non-linearities, a simple memory polynomial model

has been considered. Since we are targeting the linearization of wideband signals, we want to be able to capture as many cross-memory products as possible.

In a more compact notation, (7.5) can be rewritten as

$$x[n] = u[n] - \boldsymbol{\varphi}^T[n]\boldsymbol{w}[n] \tag{7.7}$$

where $\boldsymbol{w}[n] = \Big(w_1[n], \cdots, w_i[n], \cdots, w_M[n]\Big)^T$ is a vector of coefficients at time n with dimensions $M \times 1$, where M, is the order of the behavioural model, and $\boldsymbol{\varphi}^T[n] = \Big(\varphi_1[n], \cdots, \varphi_i[n], \cdots, \varphi_M[n]\Big)$ is the vector containing the basis functions $\varphi_i[n]$ (with $i = 1, \cdots, M$) following the Volterra-based model in (7.6). The mapping between the simplified Volterra-based model specific coefficients ($\alpha_i, \beta_{kij}, \gamma_{slkij}$ and δ_{pi}) in (7.6) and the general purpose DPD coefficients $w_i[n]$ in (7.7) is straightforward.

Finally, when expressed in matrix notation, (7.5) can be rewritten as

$$\boldsymbol{x} = \boldsymbol{u} - \boldsymbol{U}\boldsymbol{w} \tag{7.8}$$

where $\boldsymbol{x} = (x[0], \cdots, x[n], \cdots, x[N-1])^T$ and $\boldsymbol{u} = (u[0], \cdots, u[n], \cdots, u[N-1])^T$, with $n = 0, \cdots, N-1$, are the predistorted and input vectors respectively, and $\boldsymbol{U} = (\boldsymbol{\varphi}[0], \cdots, \boldsymbol{\varphi}[n], \cdots, \boldsymbol{\varphi}[N-1])^T$ is the $N \times M$ data matrix, with N being the number of samples and M being the number of basis functions or the order of the model.

Following a closed-loop error minimization technique as in [Bra, LB18], the coefficients can be extracted iteratively finding the least squares (LS) solution. At the i^{th} iteration (i.e., when considering buffers of N data samples) the coefficients are obtained as

$$\boldsymbol{w}^{i+1} = \boldsymbol{w}^i + \mu\Big(\boldsymbol{U}^H\boldsymbol{U}\Big)^{-1}\boldsymbol{U}^H\boldsymbol{e} \tag{7.9}$$

with μ ($0 \leq \mu \leq 1$) being the weighting factor and $\boldsymbol{e} = (e[0], \cdots, e[n], \cdots, e[N-1])^T$ is the $N \times 1$ vector of the error defined as

$$\boldsymbol{e} = \frac{\boldsymbol{y}}{G_0} - \boldsymbol{u} \tag{7.10}$$

where G_0 determines the desired linear gain of the PA, and where $\boldsymbol{y}$ and $\boldsymbol{u}$ are the $N \times 1$ vectors of the PA output and the transmitted input, respectively.

In addition, if we want to further simplify the number of basis functions defining our DPD model, we can apply some feature selection algorithm. As discussed in section 3.1.5 of Chapter 3, one of the algorithms used for dimensionality reduction is the orthogonal matching pursuit (OMP), a greedy algorithm for sparse approximation used in [Rei15, Pha18a] for model order reduction purposes.

7.3.3 Global Optimization Algorithms

In this Chapter we have considered the use of two heuristic search algorithms to determine the free-parameters of the dual-input PA, namely, the well-known simulated annealing (SA) and the

adaptive Lipschitz optimization (adaLIPO) algorithm. A brief description of the SA is given in the following, while a description of the adaLIPO algorithm is given in section 6.2 of Chapter 6.

Simulated Annealing (SA)

One of the most famous large scale heuristic searching method is the simulated annealing, which was first introduced by Kirkpatrick in 1983 [Kir83]. The SA method (named after a technique in metallurgy involving heating and controlled cooling of a material to increase the size of its crystals and reduce their defects) performs well in the case of large scale searching, and also has a good property of converging. Following the analogy with metallurgy, the slow cooling in the simulated annealing has to do with a slow decrease in the probability of accepting worse solutions as the solution space is explored. To find the global optimum solution, the algorithm has to be able to carry out an extensive search, that is the reason why accepting worse solutions is a fundamental property. Therefore, at each iteration, the algorithm randomly selects a solution and evaluates it, then decides the next move based on either one of two probabilities according to the quality of the new solution in comparison to the previous ones. During the search, the SA parameter named temperature (again, in analogy with metallurgy) is progressively decreased (until reaching the zero value) and the probabilities of moving to a better new solution and moving to a worse new solution updated accordingly.

7.4 Experimental Results

7.4.1 General Considerations

The following experimental results were obtained using the LMBA test-bench depicted in Fig. 5.13 and described in section 5.4.3 of Chapter 5.

The proposed auto-tuning approach for LMBA or dual-input PA systems was tested with LTE (OFDM-like) waveforms. In particular, we considered two types of test signals: i) a 64 quadrature amplitude modulation (QAM) 20 MHz bandwidth LTE signal (LTE-20) at 2 GHz RF frequency with 10.2 dB of PAPR, and ii) a non contiguous intra-band carrier-aggregated (CA) LTE system consisting in 4 channels of 64 QAM modulated LTE-20 signals (CA-4×LTE-20) spread in 200 MHz instantaneous bandwidth at 2 GHz RF frequency and a PAPR of 10.7 dB. To be noted, these signals are more demanding than the ones previously used to characterise the same DUT [Qua18].

Following the proposed procedure schematically described in Fig. 7.4, the first step is to define the free-parameters to be optimized. In order to show the difficulty of properly tuning the parameters defined in Section 7.2.2, Fig. 7.6 shows the evolution of the best-case and worst-case

ACLR, best-case and worst-case EVM, the NMSE, the output power and the power efficiency for a 200 MHz CA-4×LTE-20 test signal, when:

- Sweeping the relative phase (ψ_{rel}) but keeping $p = 3$, $OP = 0$ and $V_{\text{GG,2}} = -3.5\,\text{V}$.
- Sweeping the OP but keeping $p = 3$, $\psi_{rel} = 190^{\circ}$ and $V_{\text{GG,2}} = -3.5\,\text{V}$.
- Sweeping p but keeping, $OP = 0$, $\psi_{rel} = 190^{\circ}$ and $V_{\text{GG,2}} = -3.5\,\text{V}$.
- Sweeping $V_{\text{GG,2}}$ but keeping $p = 3$, $OP = 0$ and $\psi_{rel} = 190^{\circ}$.

For simplicity, no CFR has been considered and the input gain has been kept fixed. As observed in Fig. 7.6, by sweeping the values of one parameter and fixing the values of the rest, it is possible to evaluate the different FOMs and determine the best configuration for each one individually. However, by fixing some of their values, we are already limiting the search space and thus, there is no guarantee that the solution found for this specific set of parameters is a global optimum.

Therefore, to properly tune the free-parameters, the next step, according to the flow diagram in Fig. 7.4, is to define the search range (upper and lower bounds). This is empirically determined, and in our particular case, we considered the same search range for both test cases. The upper and lower bounds defined for the free-parameters under search were:

- Offset percentage, $OP = [0.01,\ 0.40]$. We empirically found (as an example, see Fig. 7.6) that for $OP > 0.4$, the linearity and efficiency performance was significantly degraded.
- degree of the root, $p = [1.0,\ 10.0]$. We empirically found that for $p > 10$, no significant variations are appreciated in the linearity performance.
- Relative phase, $\psi_{rel} = [0,\ 359]^{o}$.
- The gate voltage of the CSP amplifier, $V_{GG,2} = [-3.5,\ -5.5]V$. This provides a reasonable variation between a deep-class C condition that should favour efficiency, and a near-class B bias where linearity should improve.
- The maximum PAPR, $max\_PAPR = [7.0,\ 12.0]dB$. For PAPR values lower than 7 dB the EVM degradation resulted unacceptably high, while no CFR was applied for PAPR values higher than 11.5 dB.
- The baseband gain, $GainBB = [16.0,\ 19.0]$. This range of baseband gain values provides a variation of 1.5 dB to adjust the IBO.

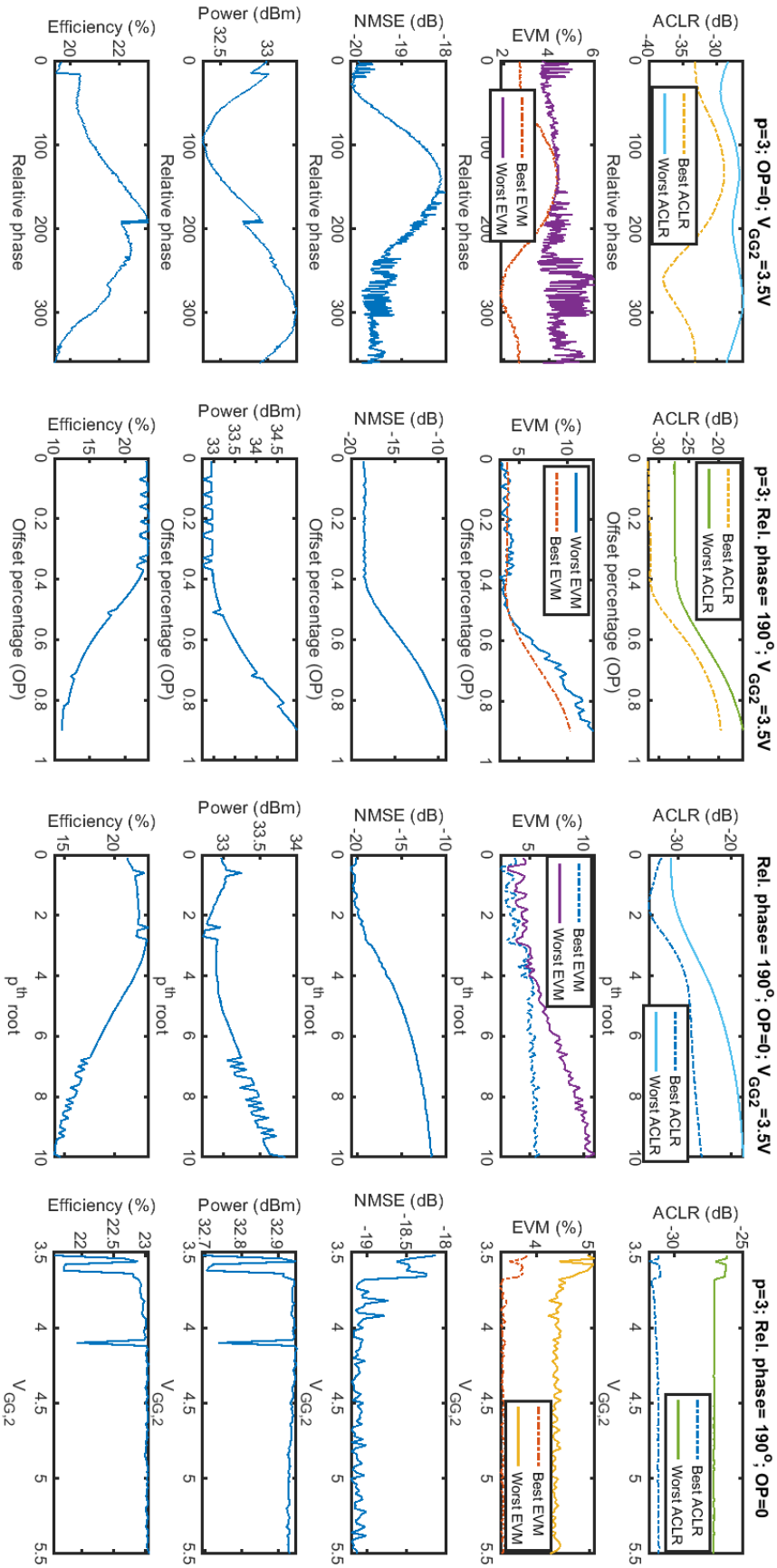

Figure 7.6: Evaluation of the FOMs when sweeping some parameters (degrees of freedom) individually.

At this point it is necessary to decide: i) if the DPD linearization needs to be included in the search procedure, ii) the optimization algorithm to be used, and iii) the FOMs, weights and thresholds of the cost function. As it will be shown in the next subsections, we considered to include DPD linearization in the search process only for the CA-4×LTE-20 test signal case. In addition, despite the fact that in both test cases we considered the use of both SA and adaLIPO algorithms to determine the values of the free-parameters, the cost functions used in each test-case were different.

7.4.2 Test case 1: 20 MHz Bandwidth LTE Signal (LTE-20)

In the case of amplifying a 20 MHz LTE signal, the required linearization is not that challenging (i.e., the linearity specs are easy to meet with DPD linearization after fixing the free-parameters). Consequently, we can conduct an optimization for selecting the optimum value of the free-parameters without the need to include DPD linearization in the search process.

The cost function for the LTE-20 case is defined in (7.11),

$$\begin{aligned} \boldsymbol{J} = (\eta_{th} - \eta)\lambda_{\text{eff}} + (\text{ACLR} - \text{ACLR}_{th})\lambda_{\text{ACLR}} \\ +(\text{NMSE} - \text{NMSE}_{th})\lambda_{\text{NMSE}} + (\text{EVM} - \text{EVM}_{th})\lambda_{\text{EVM}} \end{aligned} \tag{7.11}$$

As depicted in Table 7.1, with this configuration of weights, more importance is given to minimize the out-of-band distortion (i.e., ACLR) and maximize power efficiency, while the in-band distortion (i.e., NMSE and EVM) requirements are more relaxed, since they are easier to meet.

The results obtained when considering both SA and adaLIPO optimizations are listed in Table 7.1. As an example, Fig. 7.7 shows the solution found by the adaLIPO algorithm (out of 4.3507 10^{10} possible configurations) for the given objective or cost function. Note that the adaLIPO algorithm searches the maximum of the cost function, consequently, we have to change the sign of the cost function described in (7.11) to run the algorithm. In addition, taking into account that the weights of the cost function are multiplying the FOMs, the threshold values defined in this cost function have no real impact or penalization effect. In this particular case, they are simply included to create an offset for better interpreting the score value (i.e., positive score values correspond to configurations where most of the targeted thresholds are met).

With the free-parameters found in Table 7.1 using both SA and adaLIPO optimization algorithms, we applied the DPD and we obtained the results listed in Table 7.2. To be noted, no CFR was applied, since both algorithms discard to apply CFR reduction. In addition, in Table 7.2 we can observe a triple compromise between the power efficiency, the linearity and the computational complexity. The power efficiency is around 31% (with less than 1 percentage point of variation) independently on the optimization method or the number of coefficients of the DPD, since the PA power efficiency is more sensitive to the chosen input power back-

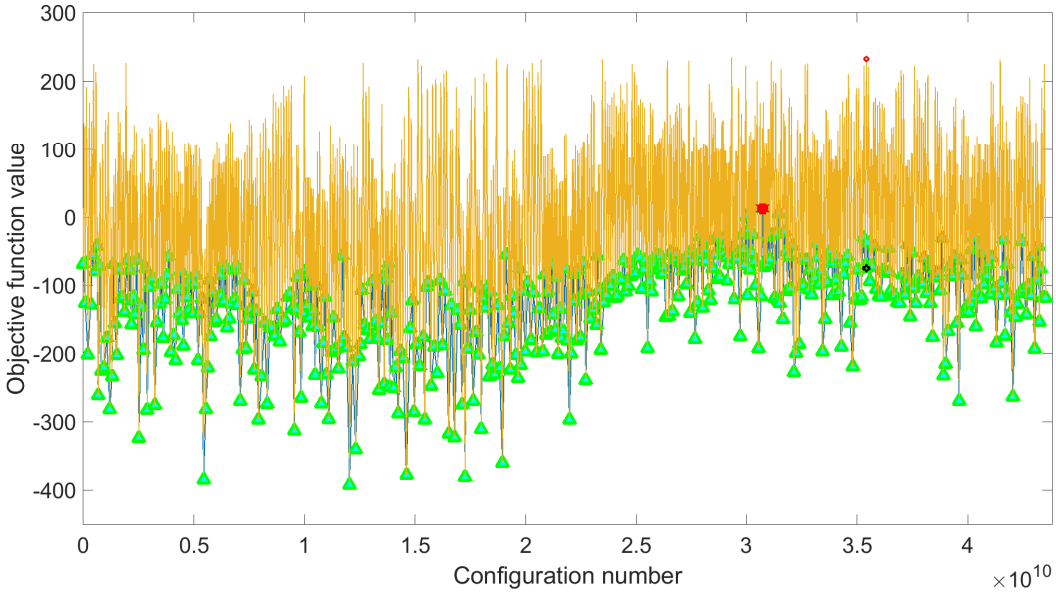

Figure 7.7: Example of the adaLIPO search for the LTE-20 signal test case.

Table 7.1: Parameters configuration for a LTE 20 MHz bandwidth signal

Opt. Config.	Threshold	Weight	Optim. values
Simulated Annealing (No DPD)	$\eta = 25\%$ ACLR =-45 dBc NSME=-26 dB EVM= 4 %	$\lambda_{\text{eff}} = 10$ $\lambda_{\text{ACLR}} = 7$ $\lambda_{\text{NMSE}} = 2$ $\lambda_{\text{EVM}} = 5$	$OP = 0.31$ $p = 6$ $\psi_{rel} = 245^o$ $maxPAPR = 10.8dB$ $V_{GG,1} = 5.4V$ $GainBB = 6.0$
adaLIPO (No DPD)			$OP = 0.26$ $p = 4.6$ $\psi_{rel} = 254^o$ $maxPAPR = 11dB$ $V_{GG,1} = 5.3V$ $GainBB = 6.1$

off. The linearity levels are easily met (e.g., the EVM after DPD is always below 1%), but we can trade-off the ACLR levels and the number of coefficients by using a dimensionality reduction method such as the OMP algorithm. Therefore, as depicted in Fig. 7.8 and listed in Table 7.2, we can meet the ACLR specifications (i.e., $ACLR < -45\ dBc$) with only 66 coefficients or, alternatively, achieving better spectral regrowth compensation by including more DPD coefficients (e.g., up to 108 coefficients) when considering the parameters' configuration found by the adaLIPO algorithm and listed in Table 7.1.

Table 7.2: linearization results with the LTE 20 MHz bandwidth signal

Opt. Config.	Worst ACLR (dBc)	NSME (dB)	Worst EVM (%)	Optput Power (dBm)	η (%)
SA config. without DPD	-38.7	-29.0	2.0	36.6	31.1
SA config. with DPD (108 coeff.)	-49.0	-37.6	0.8	36.6	30.3
SA config. with DPD (62 coeff.)	-48.3	-37.8	0.7	36.4	30.7
adaLIPO config. without DPD	-36.7	-27.5	2.3	36.2	31.0
adaLIPO config. with DPD (108 coeff.)	-53.4	-40.9	0.6	36.2	31.5
adaLIPO config. with DPD (66 coeff.)	-46.7	-38.5	0.7	36.2	31.8

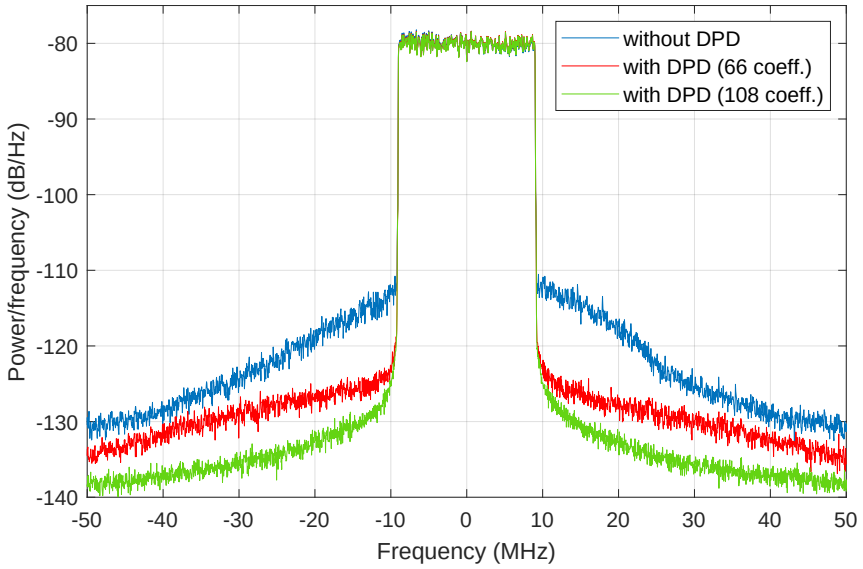

Figure 7.8: LTE-20 signal test case. Output power spectra before and after DPD linearization, when considering a DPD behavioural model with 66 and 108 coefficients, respectively.

7.4.3 Test case 2: 200 MHz Bandwidth CA-4×LTE-20 Signal

For the CA-4×LTE-20 test signal case, the attempts of optimization without DPD inclusion in the process did not lead to a configuration where the output signal was compliant with the ACLR and EVM thresholds. Therefore, we included DPD linearization to run the optimization search process. We wanted to make sure that the solution found resulted in a PA behaviour that could be later linearised with the 200 MHz instantaneous bandwidth signal. We considered a general Volterra-like behavioural model (as described in (7.6)) with a generic configuration that yielded to a DPD behavioural model with 592 coefficients. Another change compared to the 20 MHz LTE case was related to the definition of the cost function, where some thresholds were added together with the weights (this time defined as exponents) to not only emphasize the desired behaviour, but also to add further penalization in case of not meeting the linearity threshold values.

$$\begin{aligned} \boldsymbol{J} = (\eta_{th} - \eta)^{\lambda_{\text{eff}}} + (\text{ACLR} - \text{ACLR}_{th})^{\lambda_{\text{ACLR}}} \\ +(\text{NMSE} - \text{NMSE}_{th})^{\lambda_{\text{NMSE}}} + (\text{EVM} - \text{EVM}_{th})^{\lambda_{\text{EVM}}} \end{aligned} \tag{7.12}$$

The results obtained when considering both SA and adaLIPO optimizations including DPD are listed in Table 7.3. As an example, Fig. 7.9 and Fig. 7.10 show the evolution of free-parameter values and the evolution of the FOMs respectively, along 200 SA iterations. The values to which the free-parameters converged are shown in Fig. 7.9 and listed in Table 7.3.

With the free-parameters found in Table 7.3 using both SA and adaLIPO optimization algorithms, we applied CFR (defined by the $maxPAPR$ parameter) and DPD (using the Volterra-based DPD model in (7.6)) and we obtained the results showing the linearity vs. efficiency trade-off listed in Table 7.4. As it can be observed, even when the parameters' configuration differ between SA and adaLIPO, their performance is quite similar. For the 200 MHz instantaneous bandwidth signal tested, we can meet the out-of-band and in-band linearity specifications with a mean output power around 33 dBm and a power efficiency around 22%.

In addition, after applying the OMP algorithm for feature selection, we were able to reduce the number of coefficients of the DPD behavioural model up to 374 coefficients in the case of the SA configuration, and 364 coefficients in the case of the adaLIPO configuration, and still being compliant with the required linearity specifications. Fig. 7.11 shows the spectra of the 200 MHz instantaneous bandwidth CA-4×LTE-20 test signal before and after DPD linearization (considering the SA configuration in Table 7.4); while Fig. 7.12 depicts the AM-AM and AM-PM characteristics before and after DPD linearization. Note that in both cases CFR was applied to the original signal to limit the PAPR to 9.8 dB (as described in Table 7.3).

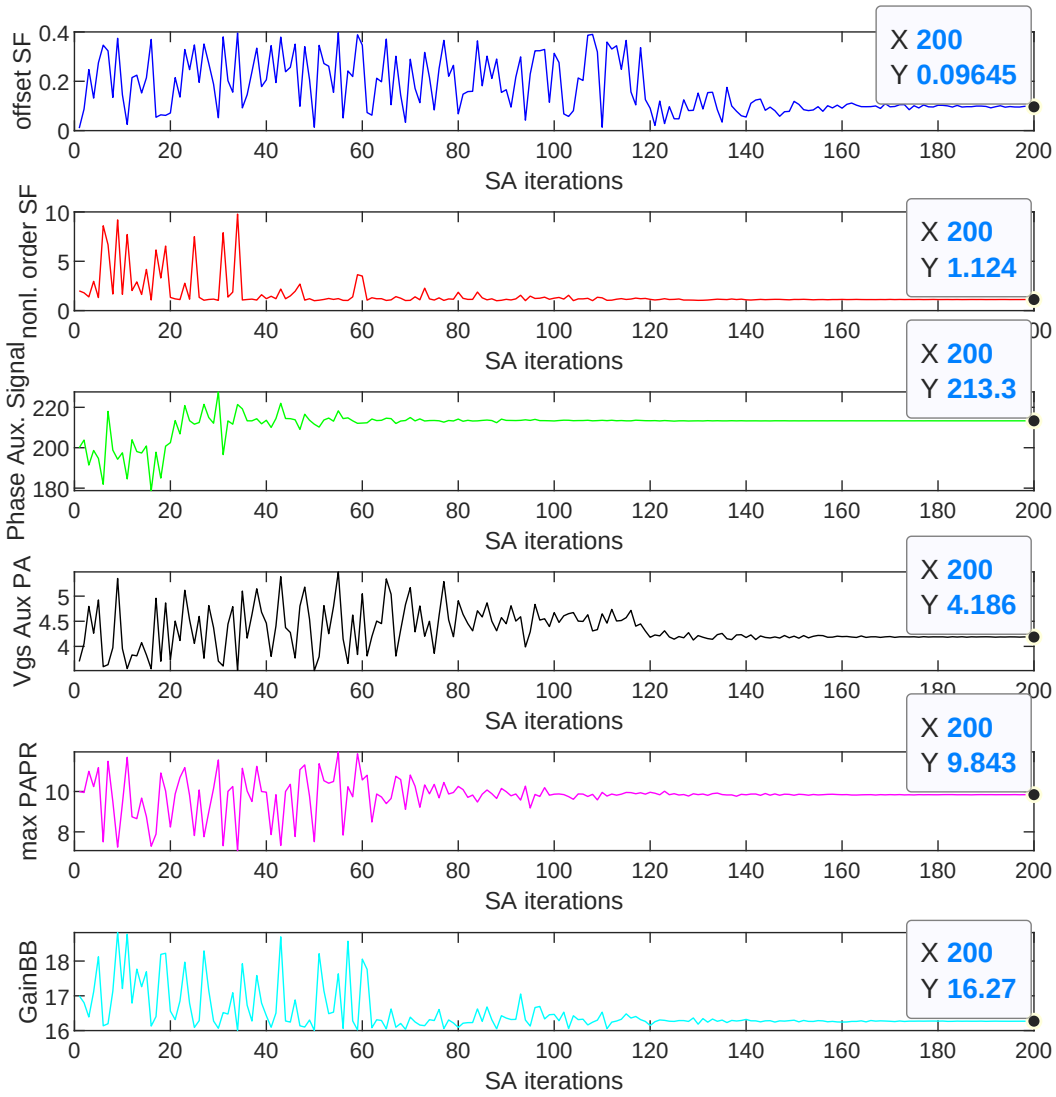

Figure 7.9: Evolution of the parameters' values for different SA iterations for the CA-4×LTE-20 signal test case.

Table 7.3: Parameters configuration for the CA-4×LTE-20 200 MHz bandwidth signal

Opt. Config.	Thresholds	Weight	Optim. values
Simulated Annealing (with DPD)	η= 19% ACLR =-45 dBc NSME=-30 dB EVM=1%	$\lambda_{\text{eff}} = 5$ $\lambda_{\text{ACLR}} = 5$ $\lambda_{\text{NMSE}} = 1$ $\lambda_{\text{EVM}} = 1$	$OP = 0.09$ $p = 1.12$ $\psi_{rel} = 213^o$ $maxPAPR = 9.8dB$ $V_{GG,1} = 4.2V$ $GainBB = 16.3$
adaLIPO (with DPD)			$OP = 0.02$ $p = 1.5$ $\psi_{rel} = 182^o$ $maxPAPR = 9.6dB$ $V_{GG,1} = 4.6V$ $GainBB = 16.3$

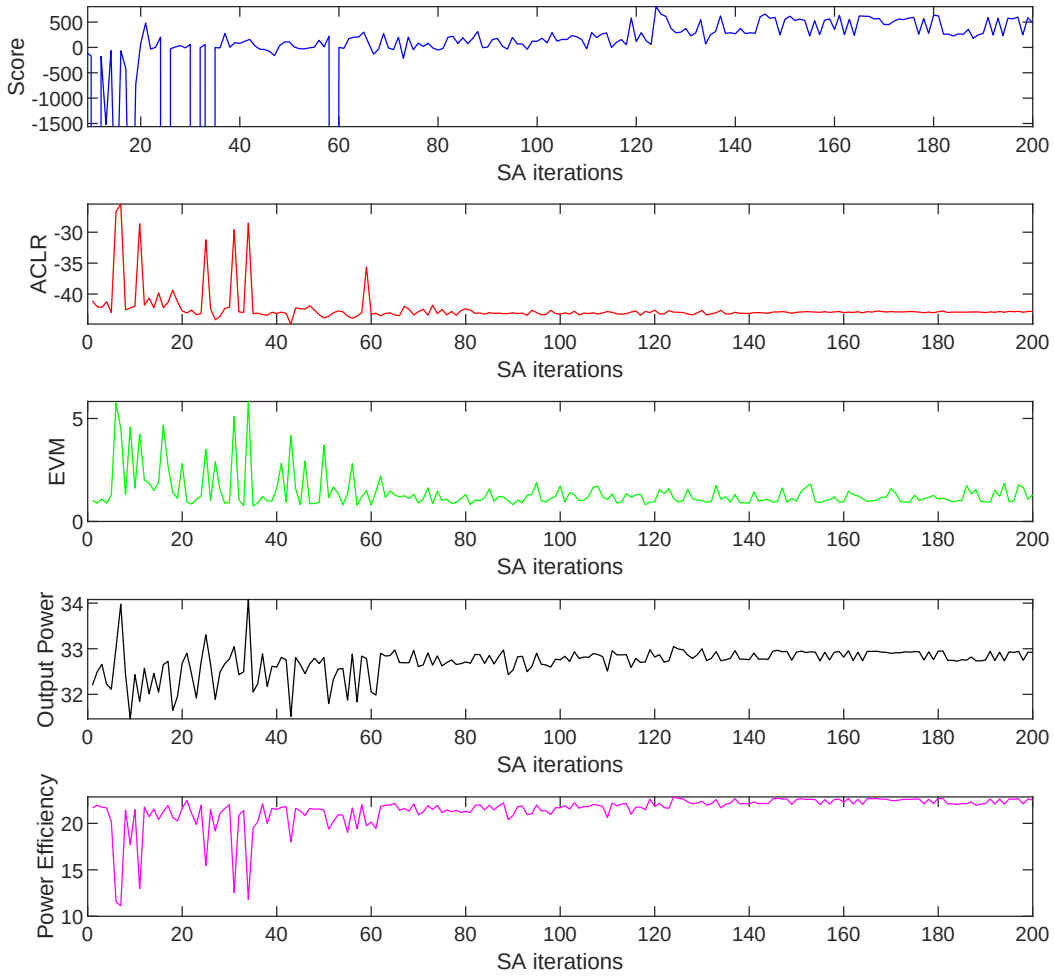

Figure 7.10: Evolution of the FOMs along different SA iterations for the CA-4×LTE-20 signal test case.

Table 7.4: linearization results with the CA-4×LTE-20 200 MHz bandwidth signal

Opt. Config.	Worst ACLR (dBc)	NSME (dB)	Worst EVM (%)	Optput Power (dBm)	η (%)
SA with CFR and without DPD	-30.3	-20.0	4.1	33.8	24.8
SA with CFR and with DPD	-45.2	-35.8	0.9	32.9	22.2
Adalipo with CFR and without DPD	-30.4	-20.1	4.3	33.7	24.7
Adalipo with CFR and with DPD	-45.1	-35.5	0.9	32.8	22.2

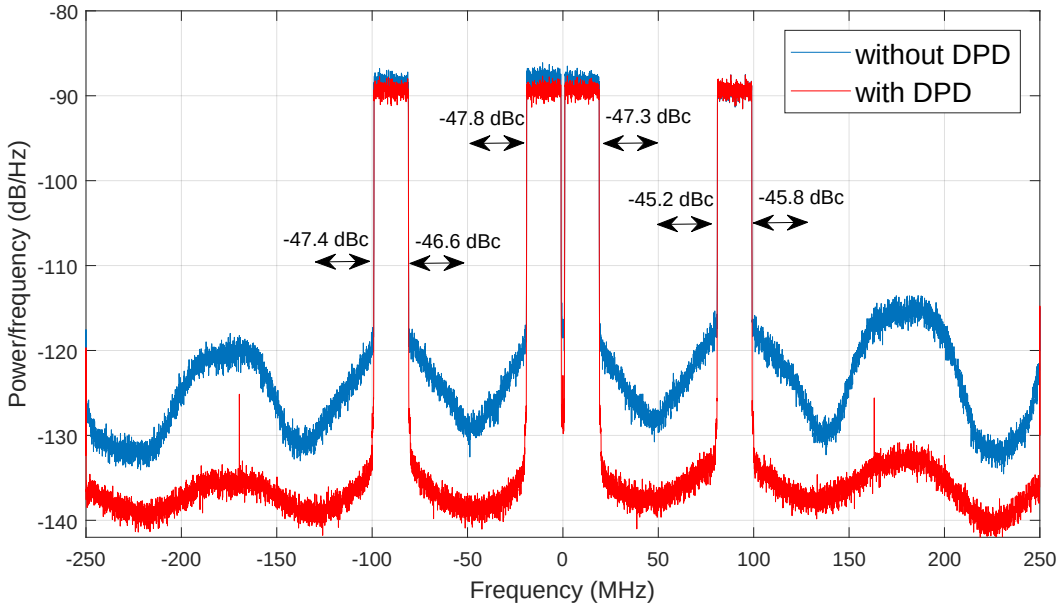

Figure 7.11: Output power spectra before and after DPD linearization for the CA-4×LTE-20 signal test case..

7.5 Conclusion

In this Chapter we proposed an approach to exploit at best dual-input PAs in terms of maximizing power efficiency along wide bandwidths while being compliant with the linearity specifications. The proposed technique relies on conducting a global optimization to find the optimum values of a set of key circuit and system level parameters that properly combined with DPD linearization and CFR techniques can find a good compromise for the inherent linearity vs. efficiency trade-off.

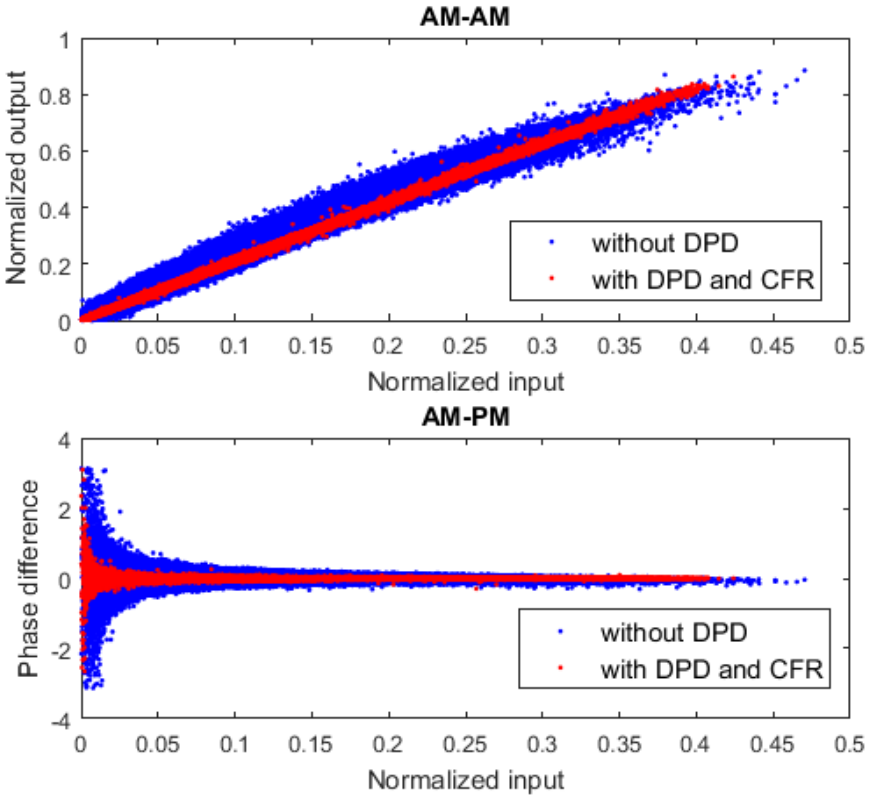

Figure 7.12: AM-AM and AM-PA before and after DPD for the CA-4×LTE-20 signal test case.

The proposed approach has been validated through experimental results. In our particular case, we have used the SA and adaLIPO heuristic search global optimization algorithms to find the best parameter configuration, taking into account two different test cases with different cost functions each one. By using a LMBA and after properly tuning the selected free-parameters, it was possible to achieve power efficiency values greater than 30% when considering the LTE-20 test signal. Moreover, up to 22% of mean power efficiency was obtained when considering the CA-4×LTE-20 test signal with 200 MHz instantaneous bandwidth. In both test-cases the peak-to-average power ratio (PAPR) of the signals was greater than 10 dB. The out-of-band linearity requirements ($ACLR < -45$ dBc) were met, and the error vector magnitude was kept always below 1%.

Despite obtaining different parameter's configurations depending on the type of heuristic search algorithm used, in both test cases (i.e., LTE-20 and CA-4×LTE-20 test cases) the linearization performance (in terms of ACLR and EVM) and power efficiency figures obtained were quite similar independently on the optimization algorithm used.

Chapter 8

Conclusions and Future Prospects

8.1 Conclusions

In the context of UAV wireless communications, the power consumption, size and weight of the payload is of significant importance. The power amplifier, present in every transmitter chain, can be designed to maximize its power efficiency or its linearity, but not both. Therefore, a way to deal with this inherent trade-off between linearity and power efficiency is the use of PA linearizers. Among them, DPD linearization is the preferred solution to both academia and industry, for its high flexibility and linearization performance. In order to save as many computational and power resources as possible, the implementation of an open-loop DPD results a very attractive solution for UAV applications. However, engineers have to deal with the complexity of DPD linearization since its both time and knowledge demanding.

The research presented in this Thesis was aimed to provide a fast engineering approach from several perspectives:

- To enhance the off-line training performance of open-loop DPD, which provides simplicity to the DPD implementation.
- To reduce the computational complexity for both the DPD design and real-case application.
- To automate the DPD tuning procedure by importing artificial intelligence to the experimental part.

The main contribution of this research consisted in providing two different methods for reducing the design and operating costs of an open-loop DPD, based on the analysis of the DPD function. The first method focuses on the input domain analysis, aiming to provide an efficient way to reduce the computational resources of the DPD adaptation function and enhance the robustness of the estimation. The second method involves the use of machine learning techniques

in the DPD design procedure to enlarge the capacity of the DPD algorithm when considering a high number of free-parameters to tune.

In particular, in Chapter 4 we presented a mesh-selecting method to find a small subset of the original test signal for DPD coefficients extraction while keeping the same accuracy. By using model order reduction based on the PCA theory combined with a proper selection of the equations through the AM-AM memoryless mesh-selecting method, we were able to significantly reduce the data matrix dimensions as well as to improve the matrix conditioning. The price to pay for reducing the computational complexity is some accuracy loss.

While in Chapter 5 we proposed and compared a series of improved mesh-selecting methods (resulted from modifying the original AM-AM memoryless mesh-selecting) to provide enhanced selecting performance. Focusing in the mesh-selecting method with better performance, the memory I-Q mesh-selecting method was combined with the PCA-DPLS feature extraction dimensionality reduction technique to allow a computational complexity reduction in the identification subsystem by a factor of 65, in comparison to using the classical QR-LS solver and consecutive samples selection. In addition, the memory I-Q mesh-selecting has been proved to be of crucial interest when training ANN for DPD purposes, since it can significantly reduce the ANN training time.

Finally, in Chapter 6 and Chapter 7 we proposed a machine learning based tuning method to assist the DPD design procedure. The adaLIPO global optimization algorithm, was used in Chapter 6 to find the best parameter configuration of a GMP behavioral model for DPD. While in Chapter 7 we proposed a methodology to conduct a global optimization search to find the optimum values of a set of key circuit and system level parameters that properly combined with DPD linearization and CFR techniques can exploit at best dual-input PAs in terms of maximizing power efficiency along wide bandwidths while being compliant with the linearity specifications.

With the achievements reported in this thesis, we have contributed to the long-term objective that consists in designing an artificial intelligence core algorithm (AICA) that will help researchers around the world test the new methods/models with all available weblabs, automatically. This way, the factory calibration of open-loop DPDs will be significantly simplified and available to SMEs around the globe which, in the field of avionics, could foster the adoption open-loop DPD linearization solutions. Given the LOS transmission characteristics in UAVs and the available behavioral models to account for the time-variant behavior of PAs due to temperature changes, a properly off-line tuned open-loop DPD would become a feasible solution in UAVs communications to extend the coverage and battery autonomy of UAVs. This idea will be further developed in the following and final section.

8.2 Future Prospects

The research work developed in this thesis can be extended in a number of directions as follows:

- **Mesh-Selecting Method with Arbitrary Distribution.** As shown in previous chapters, by keeping the original distribution of the multi-dimension histogram, the MeS provides a good selecting performance that keeps the DPD performance and reduces the computation complexity. This illustrates the importance of the original statistics of the DPD training signal.

  However, some other researches also point out that some supplementation is required to the original input signal statistical feature. The most famous approach is the piece-wise method provided by Kim at [Kim06]. By dividing the signal into high and low amplitude and predistorting via independent DPD functions, it operates on the foundation of the signal amplitude distribution. As the higher part of the signal gains high importance when applied a separate DPD function, the performance improved accordingly. Another recent research in [Kra20] proposes a method for sample selection based on a Genetically Optimised Histogram (GOH), which is a histogram of signal magnitudes optimized with respect to characteristics of the PA and of the transmitted signal.

  A undergoing work to enhance the MeS is to include a arbitrarily set multi-dimension histogram to improve the DPD performance. Unlike the traditional histogram optimized like in the GOH method or even manually, the optimized distribution will be obtained with a machine learning method.

- **Mesh-Selecting Method with Signal Similarity based Selecting.** As mentioned several times in this dissertation, the demanding of the storage is the major problem of the MeS method. The reason for keeping the tremendous large mesh inside the RAM storage is to find the way to map a new coming PA input sample to the coordinate position, i.e., the original objective of the MeS.

  A further work to improve the performance of MeS is to get rid off the original multi-dimensional histogram or find another approach to map the new signal into the original statistical feature. The information of the mesh could also be saved in the selection result of the PA input signal. Comparing the similarity between different signals, by means of the mean square error, would be a possibility and it is now under test. The experimental result would be illustrated in further publications.

- **A full AI assisting approach that includes the PA, DPD model and the DPD methods.** In its final format, multiple test-benches will be allowing access to the artificial intelligence core algorithm (AICA). Researchers will be also able to upload their new ideas of DPD models or linearization methods, as envisaged in Fig. 8.1. The AICA will test

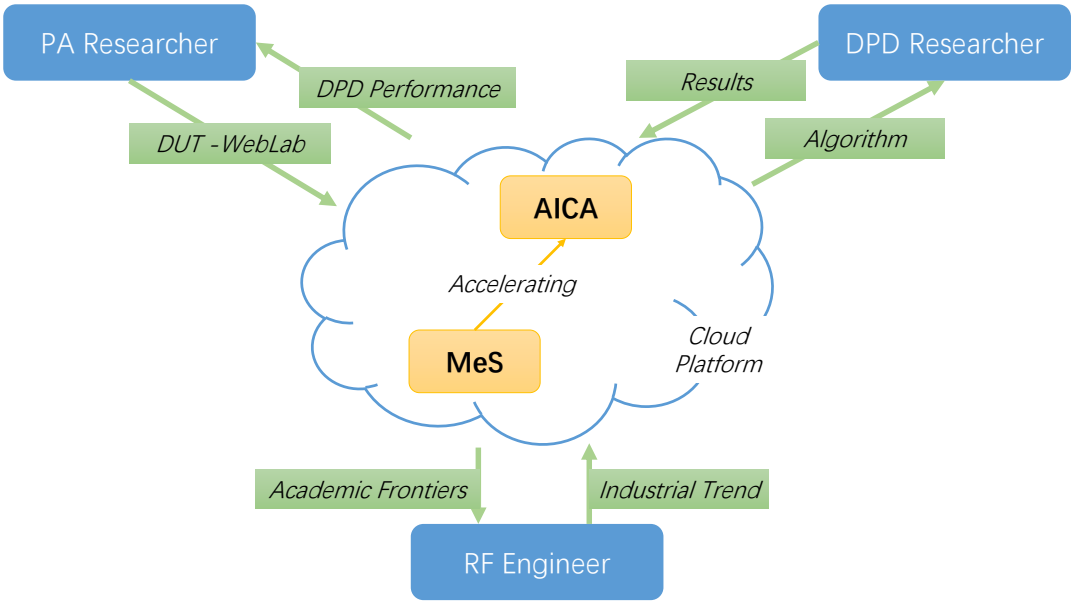

Figure 8.1: The structure of new DPD AI paradigm

the new methods with all PAs automatically and then will show the researcher a panoply of results evaluating the DPD linearization performance. PA designers and manufacturers will receive a suitable DPD solution that will be chosen among all the available academic publications automatically by AICA.

Bibliography

[ANA] "http://www.analog.com/en/products/rf-microwave/integrated-transceivers-transmitters-receivers/widebandtransceivers-ic/ad9364.htmlproduct-overview", .

[And13] C. M. Andersson, D. Gustafsson, J. Chani Cahuana, R. Hellberg, and C. Fager, "A 1–3-GHz Digitally Controlled Dual-RF Input Power-Amplifier Design Based on a Doherty-Outphasing Continuum Analysis", vol. 61, no. 10, pp. 3743–3752, Oct. 2013.

[Bar14] Filipe M. Barradas, Telmo R. Cunha, Pedro M. Lavrador, and Jose C. Pedro, "Polynomials and LUTs in PA behavioral modeling: A fair theoretical comparison", *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 12, pp. 3274–3285, dec 2014.

[Bar15] T. W. Barton, and D. J. Perreault, "Theory and Implementation of RF-Input Outphasing Power Amplification", vol. 63, no. 12, pp. 4273–4283, dec. 2015.

[Bar16] Taylor Barton, "Not just a phase: Outphasing power amplifiers", *IEEE Microwave Magazine*, vol. 17, no. 2, pp. 18–31, feb 2016.

[Bat11] Bathich, K. and Markos, A.Z. and Boeck, G., "Frequency Response Analysis and Bandwidth Extension of the Doherty Amplifier", vol. 59, no. 4, pp. 934 –944, Apr. 2011.

[Bel09] R. Beltran, F.H. Raab, and A. Velazquez, "Hf outphasing transmitter using class-e power amplifiers", *Microwave Symposium Digest, 2009. MTT '09. IEEE MTT-S International*, pp. 757–760, 2009.

[Bra] R. N. Braithwaite, "General principles and design overview of digital pre-distortion", *Digital Front-End in Wireless Communications and Broadcasting*, pp. 143–191, F. Luo, Ed. Cambridge, U.K.: Cambridge Univ. Press, 2011, ch. 6.

[Bra08] R.N. Braithwaite, "Wide bandwidth adaptive digital predistortion of power amplifiers using reduced order memory correction", *Microwave Symposium Digest, 2008 IEEE MTT-S International*, pp. 1517–1520, 2008.

[Bra11] R. Neil Braithwaite, "Adaptation of a digitally predistorted rf amplifier using selective sampling", *ZTE Communications*, , no. 9, pp. 3–12, set. 2011.

[Bra15] R. Neil Braithwaite, "Closed-loop digital predistortion (DPD) using an observation path with limited bandwidth", *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 2, pp. 726–736, feb 2015.

[BRI] "https://www.defensetech.org/2013/06/07/f-35a-launches-first-missile-in-test-flight/", .

[Cao17] Wenhui Cao, Yue Li, and Anding Zhu, "Magnitude-selective affine function based digital predistorter for RF power amplifiers in 5g small-cell transmitters", *2017 IEEE MTT-S International Microwave Symposium (IMS)*, IEEE, jun 2017.

[Cha] "http://dpdcompetition.com/rfweblab/", .

[Che17] Peng Chen, and Thomas J. Brazil, "Classifying load-pull contours of a broadband high-efficiency power amplifier using a support vector machine", *2017 Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*, IEEE, apr 2017.

[Chi35] H. Chireix, "High Power Outphasing Modulation", *Proceedings of the IRE*, vol. 23, no. 11, pp. 1370–1392, Nov. 1935.

[Col18] D. Collins, R. Quaglia, J. Powell, and S. Cripps, "Experimental characterization of a load modulated balanced amplifier with simplified input power splitter", *2018 Asia-Pacific Microwave Conference (APMC)*, pp. 461–463, Nov. 2018.

[Cox74] D. Cox, "Linear Amplification with Nonlinear Components", *IEEE Trans. on Communications*, vol. 22, no. 12, pp. 1942–1945, Dec 1974.

[D. 06] D. R. Morgan, Z. Ma et al., "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers", *IEEE Trans. on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[Dar16] Ramzi Darraji, Pedram Mousavi, and Fadhel M. Ghannouchi, "Doherty Goes Digital: Digitally Enhanced Doherty Power Amplifiers", *IEEE Microwave Magazine*, vol. 17, no. 8, pp. 41–51, aug. 2016.

[Den18] H. Deng, D. Zhou, D. Zhang, and Y. Zhang, "Analogue predistorter based on nonlinear crlhtl", *Electronics Letters*, vol. 54, no. 11, pp. 693–695, 2018.

[DJI] "http://www.dji.com/cn/phantom-4", .

[Doh36] W. H. Doherty, "A new high efficiency power amplifier for modulated waves", *Radio Engineers, Proceedings of the Institute of*, vol. 24, no. 9, pp. 1163–1182, 1936.

[El-12] M. El-Asmar, A. Birafane, M. Helaoui, A. B. Kouki, and F. M. Ghannouchi, "Analytical design methodology of outphasing amplification systems using a new simplified chireix combiner model", *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1886–1895, 2012.

[Flo17] Julieta Flomenbaum, *Design and linearization of an efficient class-E power amplifier using a test bench based on development boards*, Master thesis, Universitat Politècnica de Catalunya, 2017.

[Fra96] W.A. Frank, "Sampling requirements for volterra system identification", *IEEE Signal Processing Letters*, vol. 3, no. 9, pp. 266–268, sep 1996.

[Gao18] W. Gao, "Linearization of wideband wi-fi power amplifiers using rf analog memory predistortion", pp. 1–6, 2018.

[GE] "www.ge.com", .

[Gil08] P. L. Gilabert, A. Cesari, G. Montoro, E. Bertran, and J. M. Dilhac, "Multi Look-Up Table FPGA Implementation of a Digital Adaptive Predistorter for Linearizing RF Power Amplifiers with Memory Effects", *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 2, pp. 372–384, Feb. 2008.

[Gil11] P. L. Gilabert, G. Montoro, and E. Bertran, "Fpga implementation of a real-time narma-based digital adaptive predistorter", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 7, pp. 402–406, 2011.

[Gil12a] P.L. Gilabert, and G. Montoro, "Look-up table implementation of a slow envelope dependent digital predistorter for envelope tracking power amplifiers", *Microwave and Wireless Components Letters, IEEE*, vol. 22, no. 2, pp. 97–99, 2012.

[Gil12b] P.L. Gilabert, and G. Montoro, "Look-up table implementation of a slow envelope dependent digital predistorter for envelope tracking power amplifiers", *Microwave and Wireless Components Letters, IEEE*, vol. 22, no. 2, pp. 97–99, 2012.

[Gil13a] Pere L. Gilabert, Gabriel Montoro, David Lopez, Nikolaos Bartzoudis, Eduard Bertran, Miquel Payaro, and Alain Hourtane, "Order reduction of wideband digital predistorters using principal component analysis", *Microwave Symposium Digest (IMS), 2013 IEEE MTT-S International*, pp. 1–4, 2013.

[Gil13b] Pere L. Gilabert, Gabriel Montoro, David Lopez, and Jose A. Garcia, "3D Digital predistortion for dual-band envelope tracking power amplifiers", *Microwave Conference Proceedings (APMC), 2013 Asia-Pacific*, pp. 734–736, 2013.

[Gil15] Pere L. Gilabert, and Gabriel Montoro, "3-D distributed memory polynomial behavioral model for concurrent dual-band envelope tracking power amplifier linearization", *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 2, pp. 638–648, Feb. 2015.

[Gil19] Pere L. Gilabert, Gabriel Montoro, David Vegas, Nieves Ruiz, and Jose Angel Garcia, "Digital predistorters go multidimensional: DPD for concurrent multiband envelope tracking and outphasing power amplifiers", *IEEE Microwave Magazine*, vol. 20, no. 5, pp. 50–61, may 2019.

[Gil20] P. L. Gilabert, D. Vegas, Z. Ren, G. Montoro, J. R. P'erez-Cisneros, M. N. Ruiz, X. Si, and J. A. García, "Design and digital predistortion linearization of a wideband outphasing amplifier supporting 200 mhz bandwidth", *2020 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)*, pp. 46–49, 2020.

[Gio14] R. Giofré, L. Piazzon, P. Colantonio, and F. Giannini, "An Ultra-Broadband GaN Doherty Amplifier with 83% of Fractional Bandwidth", vol. 24, no. 11, pp. 775–777, Nov. 2014.

[Gua10] Lei Guan, and Anding Zhu, "Low-cost fpga implementation of volterra series-based digital predistorter for rf power amplifiers", vol. 58, no. 4, pp. 866–872, 2010.

[Gua12a] Lei Guan, Chao Yu, and Anding Zhu, "Bandwidth-constrained least squares-based model extraction for band-limited digital predistortion of RF power amplifiers", *2012 Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits*, IEEE, sep 2012.

[Gua12b] Lei Guan, and Anding Zhu, "Optimized low-complexity implementation of least squares based model extraction for digital predistortion of RF power amplifiers", *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 594–603, Mar. 2012.

[Gua12c] Lei Guan, and Anding Zhu, "Optimized low-complexity implementation of least squares based model extraction for digital predistortion of RF power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 594–603, mar 2012.

[Gui20] Estefanía Guillena, *Linealización de un Amplificador Balanceado con Modulación de Carga Mediante un Predistorsionador Digital Basado en Redes Neuronales para Comunicaciones en Vehículos Aereos no Tripulados*, Graduate thesis, Universitat Politècnica de Catalunya, 2020.

[Gum18] K. Gumber, and M. Rawat, "Low-cost rfinârfout predistorter linearizer for high-power amplifiers and ultra-wideband signals", *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2069–2081, 2018.

[Hay91] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1991.

[Hor02] Reiner Horst, *Introduction to Global Optimization (Nonconvex Optimization and Its Applications)*, Springer-Verlag, Berlin, Heidelberg, 2002.

[Isl17] A. Islam, P. J. Xia, H. Huang, and S. Boumaiza, "Constrained identification of rational functions for robust digital predistortion", pp. 166–168, 2017.

[Kal19] R. Kalyan, K. Rawat, and S. K. Koul, "A Digitally Assisted Dual-Input Dual-Band Doherty Power Amplifier With Enhanced Efficiency and Linearity]", vol. 66, no. 2, pp. 297–301, feb. 2019.

[Kat16] Allen Katz, John Wood, and Daniel Chokola, "The evolution of PA linearization: From classic feedforward and feedback through analog and digital predistortion", *IEEE Microwave Magazine*, vol. 17, no. 2, pp. 32–40, feb 2016.

[Kim01] J. Kim, and K. Konstantinou, "Digital Predistortion of Wideband Signals Based on Power Amplifier Model with Memory", *Electronics Letters*, vol. 37, pp. 1417–1418, Nov. 2001.

[Kim06] W.-J. Kim, K.-J. Cho, S. P. Stapleton, and J.-H. Kim, "Piecewise Pre-Equalized Linearization of the Wireless Transmitter With a Doherty Amplifier", *IEEE Trans. on Microwave Theory and Techniques*, vol. 54, no. 9, pp. 3469–3478, Sept. 2006.

[Kir83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science*, vol. 220, no. 4598, pp. 671–680, may 1983.

[Koe06] Heinz Koeppl, and Peter Singerl, "An efficient scheme for nonlinear modeling and predistortion in mixed-signal systems", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 12, pp. 1368–1372, dec 2006.

[Kra20] Jan Kral, Tomas Gotthans, Roman Marsalek, Michal Harvanek, and Markus Rupp, "On feedback sample selection methods allowing lightweight digital predistorter adaptation", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 6, pp. 1976–1988, jun 2020.

[Kwa13] A.K. Kwan, S.A. Bassam, M. Helaoui, and F.M. Ghannouchi, "Concurrent dual band digital predistortion using look up tables with variable depths", *Power Amplifiers for Wireless and Radio Applications (PAWR), 2013 IEEE Topical Conference on*, pp. 25–27, 2013.

[LB18] David Lopez-Bueno, Quynh Anh Pham, Gabriel Montoro, and Pere L. Gilabert, "Independent digital predistortion parameters estimation using adaptive principal component analysis", *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 12, pp. 5771–5779, dec. 2018.

[Li20] Yue Li, Xiaoyu Wang, and Anding Zhu, "Sampling rate reduction for digital predistortion of broadband RF power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 3, pp. 1054–1064, mar 2020.

[Liu14] Youjiang Liu, Jonmei J. Yan, Hayg-Taniel Dabag, and Peter M. Asbeck, "Novel technique for wideband digital predistortion of power amplifiers with an under-sampling ADC", *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 11, pp. 2604–2617, nov 2014.

[Liu15] Ying Liu, Wensheng Pan, Shihai Shao, and Youxi Tang, "A general digital predistortion architecture using constrained feedback bandwidth for wideband power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 5, pp. 1544–1555, may 2015.

[Lop14] David Lopez, Pere L. Gilabert, Gabriel Montoro, and Nikolaos Bartzoudis, "Peak cancellation and digital predistortion of high-order QAM wideband signals for next generation wireless backhaul equipment", *2014 International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC)*, IEEE, apr. 2014.

[Ma13] Y. Ma, Y. Yamao, Y. Akaiwa, and C. Yu, "Fpga implementation of adaptive digital predistorter with fast convergence rate and low complexity for multi-channel transmitters", *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 11, pp. 3961–3973, 2013.

[Ma14] Yuelin Ma, Yasushi Yamao, Yoshihiko Akaiwa, and Koji Ishibashi, "Wideband digital predistortion using spectral extrapolation of band-limited feedback signal", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 7, pp. 2088–2097, jul 2014.

[Mal17] Cédric Malherbe, and Nicolas Vayatis, "Global optimization of Lipschitz functions", *arXiv e-prints*, pag. arXiv:1703.02628, mar. 2017.

[Med18] A. Mediano, and F. J. Ortega-Gonzalez, "Class-e amplifiers and applications at mf, hf, and vhf: Examples and applications", *IEEE Microwave Magazine*, vol. 19, no. 5, pp. 42–53, 2018.

[Mol17] Albert Molina, Kannan Rajamani, and Kamran Azadet, "Concurrent dual-band digital predistortion using 2-D lookup tables with bilinear interpolation and extrapolation:

Direct least squares coefficient adaptation", *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 4, pp. 1381–1393, Apr. 2017.

[Mon10a] G. Montoro, P.L. Gilabert, E. Bertran, and J. Berenguer, "A method for real-time generation of slew-rate limited envelopes in envelope tracking transmitters", *RF Front-ends for Software Defined and Cognitive Radio Solutions (IMWS), 2010 IEEE International Microwave Workshop Series on*, pp. 1–4, 2010.

[Mon10b] G. Montoro, P.L. Gilabert, E. Bertran, and J. Berenguer, "A method for real-time generation of slew-rate limited envelopes in envelope tracking transmitters", *RF Front-ends for Software Defined and Cognitive Radio Solutions (IMWS), 2010 IEEE International Microwave Workshop Series on*, pp. 1–4, 2010.

[Mon11] G. Montoro, P. Gilabert, J. Berenguer, and E. Bertran, "Digital predistortion of envelope tracking amplifiers driven by slew-rate limited envelopes", *Microwave Symposium Digest (MTT), 2011 IEEE MTT-S International*, pp. 1–4, 2011.

[Mra12] Nizar Mrabet, Imaduddin Mohammad, Farouk Mkadem, Chiheb Rebai, and Slim Boumaiza, "Optimized hardware for polynomial digital predistortion system implementation", *2012 IEEE Topical Conf. on Power Amplifiers for Wireless and Radio Appl. (PAWR)*, pp. 83–84, IEEE, Jan. 2012.

[Mur06] S.D. Muruganathan, and A.B. Sesay, "A QRD-RLS-based predistortion scheme for high-power amplifier linearization", *IEEE Trans. Circuits Syst. II, Express Briefs*, vol. 53, no. 10, pp. 1108–1112, Oct. 2006.

[Ped18] P. H. Pednekar, W. Hallberg, C. Fager, and T. W. Barton, "Analysis and Design of a Doherty-Like RF-Input Load Modulated Balanced Amplifier", vol. 66, no. 12, pp. 5322–5335, Dec. 2018.

[Pen16a] Jun Peng, Songbai He, Bingwen Wang, Zhijiang Dai, and Jingzhou Pang, "Digital predistortion for power amplifier based on sparse bayesian learning", *IEEE Trans. on Circ. and Sys. II: Express Briefs*, vol. 63, no. 9, pp. 828–832, Sep. 2016.

[Pen16b] Raymond Pengelly, Christian Fager, and Mustafa Ozen, "Doherty's legacy: A history of the Doherty power amplifier from 1936 to the present day", *IEEE Microwave Magazine*, vol. 17, no. 2, pp. 41–58, feb. 2016.

[Pha18a] Quynh Anh Pham, David López-Bueno, Teng Wang, Gabriel Montoro, and Pere L. Gilabert, "Multi-dimensional LUT-based digital predistorter for concurrent dual-band envelope tracking power amplifier linearization", *Proc. 2018 IEEE Topical Conf. on RF/Microw. Power Amplifiers for Radio and Wireless Appl. (PAWR)*, pp. 47–50, Jan. 2018.

[Pha18b] Quynh Anh Pham, David Lopez-Bueno, Teng Wang, Gabriel Montoro, and Pere L. Gilabert, "Partial least squares identification of multi look-up table digital predistorters for concurrent dual-band envelope tracking power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 12, pp. 5143–5150, dec 2018.

[Pha19a] Quynh Anh Pham, David López-Bueno, Gabriel Montoro, and Pere L. Gilabert, "Dynamic selection and update of digital predistorter coefficients for power amplifier linearization", *Proc. 2019 IEEE Topical Conf. on RF/Microw. Power Amplifiers for Radio and Wireless Appl. (PAWR)*, pp. 1–4, Jan. 2019.

[Pha19b] Quynh Anh Pham, Gabriel Montoro, David Lopez-Bueno, and Pere L. Gilabert, "Dynamic selection and estimation of the digital predistorter parameters for power amplifier linearization", *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 3996–4004, oct 2019.

[Pha19c] Thi Quynh Anh Pham, *Contribution to Dimensionality Reduction of Digital Predistorter Behavioral Models for RF Power Amplifier Linearization*, phdthesis, Universitat Politènica de Catalunya, dec. 2019.

[Pia18] A. Piacibello, M. Pirola, V. Camarchia, C. Ramella, R. Quaglia, X. Zhou, and W. . Chan, "Comparison of S-Band Analog and Dual-Input Digital Doherty Power Amplifiers", *2018 48th European Microwave Conference (EuMC)*, pp. 1237–1240, Sep. 2018.

[Pop17] Zoya Popovic, "Amping up the PA for 5g: Efficient GaN power amplifiers with dynamic supplies", *IEEE Microwave Magazine*, vol. 18, no. 3, pp. 137–149, may 2017.

[Pop18] Zoya Popovic, and Jose A. Garcia, "Microwave Class-E Power Amplifiers: A Brief Review of Essential Concepts in High-Frequency Class-E PAs and Related Circuits", *IEEE Microwave Magazine*, vol. 19, no. 5, pp. 54–66, jul. 2018.

[Qi 16] Qi Cai, Wenquan Che, and Kaixue Ma, "A linear gan power amplifier using novel transistor based analog predistortion method", pp. 1–4, 2016.

[Qua18] R. Quaglia, and S. Cripps, "A Load Modulated Balanced Amplifier for Telecom Applications", *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 3, pp. 1328–1338, mar. 2018.

[Raa02] F.H. Raab, P. Asbeck, S. Cripps, P.B. Kenington, Z.B. Popovic, N. Pothecary, J.F. Sevic, and N.O. Sokal, "Power amplifiers and transmitters for rf and microwave", *Microwave Theory and Techniques, IEEE Transactions on*, vol. 50, no. 3, pp. 814–826, 2002.

[Rei15] Reina-Tosina, J. and Allegue-Martínez, M. and Crespo-Cadenas, C. and C. Yu, C., and Cruces, S., "Behavioral modeling and predistortion of power amplifiers under sparsity

hypothesis", *IEEE Trans. on Microw. Theory and Tech.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.

[RFB] "http://revistapesquisa.fapesp.br/en/2013/10/23/the-flight-of-the-falcon/", .

[RT15] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. C. Yu, and S. Cruces, "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis", *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.

[Rub18] J. J. Moreno Rubio, V. Camarchia, M. Pirola, and R. Quaglia, "Design of an 87% Fractional Bandwidth Doherty Power Amplifier Supported by a Simplified Bandwidth Estimation Method", vol. 66, no. 3, pp. 1319–1327, Mar. 2018.

[Rui16] M. N. Ruiz, A. L. Benito, J. R. Pérez-Cisneros, P. L. Gilabert, G. Montoro, and J. A. García, "Constant-gain envelope tracking in a uhf outphasing transmitter based on continuous-mode class-e gan hemt pas", *2016 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1–4, 2016.

[S.A72] S.A. Piyavskii, "An algorithm for finding the absolute extremum of a function", *USSR Computational Mathematics and Mathematical Physics*, vol. 12, no. 4, pp. 57–67, jan. 1972.

[Sal81] A. Saleh, "Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers", *IEEE Trans. on Communications*, vol. 29, no. 11, pp. 1715–1720, Nov 1981.

[Sch09] D. Scheurs, M. O'Droma, A. A. Goacher, and M. Gadringer (eds.), *RF Power Amplifier Behavioural Modeling*, Cambridge University Press, 2009.

[She16] D. J. Shepphard, J. Powell, and S. C. Cripps, "An efficient broadband reconfigurable power amplifier using active load modulation", *IEEE Microwave and Wireless Components Letters*, vol. 26, no. 6, pp. 443–445, 2016.

[Shu72] Bruno O. Shubert, "A Sequential Method Seeking the Global Maximum of a Function", *SIAM Journal on Numerical Analysis*, vol. 9, no. 3, pp. 379–388, 1972.

[TDO] "https://talk.dallasmakerspace.org/t/lily-camera-drone/3003", .

[Tom19] P. M. Tomé, F. M. Barradas, T. R. Cunha, and J. C. Pedro, "Hybrid analog/digital linearization of gan hemt-based power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 288–294, 2019.

[UAV] "http://www.afwing.com/baike/uav2/uav-2.htm", .

[Veg17] D. Vegas, F. Moreno, M. N. Ruiz, and J. A. García, "Efficient class-e power amplifier for variable load operation", *2017 Integrated Nonlinear Microwave and Millimetre-wave Circuits Workshop (INMMiC)*, pp. 1–3, 2017.

[Wan15a] Teng Wang, Pere L. Gilabert, and Gabriel Montoro, "Under-sampling effects and computational cost reduction in RF power amplifier behavioral modeling", *2015 10th European Microwave Integrated Circuits Conference (EuMIC)*, IEEE, sep 2015.

[Wan15b] Zhancang Wang, "Demystifying envelope tracking: Use for high-efficiency power amplifiers for 4G and beyond", *IEEE Microwave Magazine*, vol. 16, no. 3, pp. 106–129, apr. 2015.

[Wan16] Zonghao Wang, Wenhua Chen, Gongzhe Su, Fadhel M. Ghannouchi, Zhenghe Feng, and Yuanan Liu, "Low feedback sampling rate digital predistortion for wideband wireless transmitters", *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 11, pp. 3528–3539, nov 2016.

[Wan17] Ziming Wang, Lei Guan, and Ronan Farrell, "Undersampling observation-based compact digital predistortion for single-chain multiband and wideband direct-to-RF transmitter", *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 5274–5283, dec 2017.

[Wat18] Gavin T. Watkins, and Konstantinos Mimis, "How Not to Rely on Moore's Law Alone: Low-Complexity Envelope-Tracking Amplifiers", *IEEE Microwave Magazine*, vol. 19, no. 4, pp. 84–94, jun 2018.

[Wes96] Frank Westad, Klaus Diepold, and Harald Martens, "Qr-plsr: Reduced-rank regression for high-speed hardware implementation", *Journal Of Chemometrics*, vol. 10, pp. 439–451, 1996.

[WET] "http://www.wetouchsky.com/dji-drone-application-electricity-inspection-in-china/", .

[Wis08] D. Wisell, J. Jalden, and P. Handel, "Behavioral power amplifier modeling using the lasso", *2008 IEEE Inst. and Meas. Tech. Conf.*, pp. 1864–1867, May 2008.

[Wol10] N. Wolf, J. Mueller, and H. Klar, "Simple predistortion system for compensation of temperature dependent nonlinearity of power amplifiers", *2010 IEEE Radio and Wireless Symposium (RWS)*, pp. 152–155, 2010.

[Woo17] John Wood, "System-level design considerations for digital pre-distortion of wireless base station transmitters", *IEEE Trans. on Microw. Theory and Tech.*, vol. 65, no. 5, pp. 1880–1890, May 2017.

[WV2] "https://commons.wikimedia.org/wiki/file:wv-3 hurricane hunter at nas jacksonville in 1950s.jpg", .

[Yam96] K. Yamauchi, K. Mori, M. Nakayama, Y. Itoh, Y. Mitsui, and O. Ishida, "A Novel Series Diode Linearizer for Mobile Radio Power Amplifiers", *Proc. IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 831–834, June 1996.

[YOL] "https://pjreddie.com/darknet/yolo/", .

[Yu99] C. S. Yu, W. S. Chan, and W.-L. Chan, "1.9 GHz Low Loss Varactor Diode Pre-Distorter", *Electronics Letters*, vol. 35, pp. 1681–1682, Sept. 1999.

[Yu12] Chao Yu, Lei Guan, Erni Zhu, and Anding Zhu, "Band-limited volterra series-based digital predistortion for wideband RF power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 12, pp. 4198–4208, dec 2012.

[Yu17] Z. Yu, "An experimental evaluation of a digital predistortion system with thermal memory effects modeling", *2017 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1545–1548, 2017.

[Zha17] C. Zhao, H. Liu, Y. Wu, and K. Kang, "Analysis and design of cmos doherty power amplifier based on voltage combining method", *IEEE Access*, vol. 5, pp. 5001–5012, 2017.

[Zhu92] Y. M. Zhu, "Generalized sampling theorem", *IEEE Trans. on Circ. and Sys.âII: Analog and Dig. Signal Proc*, vol. 39, pp. 587â588, 1992.