

Dipl.-Ing. (FH) Dipl.-Ing. Harald Enzinger

Behavioral Modeling and Digital Predistortion of Radio Frequency Power Amplifiers

Doctoral Thesis

in partial fulfillment of the requirements for the degree of
Doctor of Technical Sciences

submitted to

Graz University of Technology, Austria
Doctoral School of Information and Communications Engineering

Advisor and First Assessor/Examiner:

FH-Prof. Priv.-Doz. Dipl.-Ing. Dr.techn. Christian Vogel
FH Joanneum – University of Applied Sciences, Austria

Second Assessor/Examiner:

Prof. Dr. Thomas Eriksson
Chalmers University of Technology, Sweden

Third Assessor/Examiner:

Univ.-Prof. Dipl.-Ing. Dr.techn. Gernot Kubin
Graz University of Technology, Austria

Graz, January 2018

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

EIDESSTÄTTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

date

signature

Abstract

The radio frequency power amplifier (RF-PA) within a digital wireless transmitter is a critical component regarding both the energy consumption and the signal quality. Especially due to today's broadband multicarrier modulation methods that generate signals with high peak-to-average power ratio, it is very hard to construct RF-PAs that achieve good energy efficiency and fulfill the strict linearity requirements imposed by the standard. Because of this, the digital predistortion (DPD) of RF-PAs has become a key technique for implementing energy efficient, high data rate wireless transmitters.

This thesis investigates theoretical foundations and practical methods for the behavioral modeling and DPD of RF-PAs. The main contributions are a semi-physical model of the joint linearity-efficiency characteristics of RF-PAs, a detailed analysis of polynomial baseband models of RF-PAs focusing on the often neglected even-order terms in baseband, and a collection of practical methods for the dual-band DPD of RF-PAs.

The trade-off between the linearity and efficiency of RF-PAs is investigated based on the semi-physical RF-PA model. For this purpose, linearity and efficiency quantification methods are introduced and applied to the model. Furthermore, an overview on highly efficient RF-PA operation modes and efficiency enhancement methods is given.

A central result regarding polynomial baseband models is the first formal justification of even-order terms in baseband. This is achieved by deriving explicit passband-baseband pairs for the quasi-memoryless polynomial and the Volterra series, which show that even-order terms in baseband correspond to modified basis functionals in passband.

Another central result of the presented analysis of polynomial baseband models is the formulation and proof of the phase homogeneity requirement, which represents a necessary symmetry of all complex baseband models of time-invariant passband systems.

The practical methods for the dual-band DPD of RF-PAs include a method for dual-band crest factor reduction by clipping and error-filtering, a dual-band DPD model based on a vector-switched generalized memory polynomial, and a dual-band DPD training algorithm based on the indirect and direct learning architectures. These methods were implemented in MATLAB and validated at the 2017 DPD competition at the IEEE International Microwave Symposium, reaching the first place within the competition.

Abstract in German (Kurzfassung)

Der Hochfrequenz-Leistungsverstärker (radio frequency power amplifier, RF-PA) innerhalb eines digitalen drahtlosen Senders ist eine kritische Komponente, sowohl bezüglich des Energieverbrauchs, als auch bezüglich der Signalqualität. Insbesondere durch die heutigen breitbandigen Mehrträger-Modulationsverfahren, welche Signale mit hohem Spitzen-zu-Durchschnitts Leistungsverhältnis erzeugen, ist es sehr schwer RF-PAs zu konstruieren welche eine gute Energieeffizienz erreichen und die strengen Linearitätsvorgaben des Standards einhalten. Aus diesem Grund wurde die digitale Vorverzerrung (digital predistortion, DPD) von RF-PAs zu einer Schlüsselmethode für die Realisierung energieeffizienter, hochratiger, drahtloser Sender.

Diese Abschlussarbeit erforscht theoretische Grundlagen und praktische Methoden für die Verhaltensmodellierung und DPD von RF-PAs. Die hauptsächlichen Beiträge sind ein semi-physikalisches Modell des gemeinsamen Linearitäts- und Effizienz-Verhaltens von RF-PAs, eine detaillierte Analyse von polynomiellen Basisband-Modellen von RF-PAs, ausgerichtet auf die oft vernachlässigten geradzahligen Terme im Basisband und eine Sammlung praktischer Methoden zur Doppelband-DPD von RF-PAs.

Der Zielkonflikt zwischen der Linearität und Effizienz von RF-PAs wird anhand des semi-physikalischen RF-PA Modells untersucht. Zu diesem Zweck werden Methoden zur Linearitäts- und Effizienz-Quantifizierung eingeführt und auf das Modell angewandt. Weiters wird ein Überblick über hoch effiziente RF-PA Betriebsarten und Methoden zur Effizienzverbesserung gegeben.

Ein zentrales Resultat bezüglich polynomieller Basisband Modelle ist die erste formale Rechtfertigung von geradzahligen Termen im Basisband. Dies wird erreicht durch die Herleitung expliziter Übertragungsband-Basisband Paare für das Quasi-Gedächtnislose Polynom und die Volterra Reihe, welche eine Korrespondenz zwischen geradzahligen Termen im Basisband und modifizierten Basisfunktionalen im Übertragungsband zeigen.

Ein weiteres zentrales Resultat der vorgestellten Analyse polynomieller Basisband Modelle ist die Formulierung und der Beweis der Phasen-Homogenitäts-Anforderung, welche eine notwendige Symmetrie von allen komplexwertiger Basisband Modellen von zeit-invarianten Übertragungsband Systemen darstellt.

Die praktischen Methoden zur Doppelband DPD von RF-PAs beinhalten eine Methode zur Doppelband Scheitelfaktorreduzierung durch Signalwertbegrenzung und Fehlerfilterung, ein Doppelband DPD Modell basierend auf einem vektorgeschalteten generalisiertem Gedächtnis-Polynom und einen Doppelband DPD Trainingsalgorithmus, basierend auf der indirekten und direkten Lernarchitektur. Diese Methoden wurden in MATLAB realisiert und bei dem DPD Wettbewerb beim IEEE International Microwave Symposium in 2017 validiert, wobei der erste Platz im Wettbewerb erreicht wurde.

Acknowledgment

A doctoral study is, like most things in life, a social experience. I deeply appreciate the chance of making this experience, which connected me with many intelligent, friendly, and inspiring people. In the following, I would like to express my sincere gratitude to the people who supported me during my doctoral studies and made this thesis possible.

First of all, I thank my doctoral advisor FH-Prof. Dr. Christian Vogel for his constant support, guidance, and scientific advise. It means a lot to me that he gave me the chance to conduct scientific research in such an interesting field. Foremost, however, his personal and professional support over the last five years was outstanding.

I also thank Dr. Karl Freiberger for being an extraordinary co-worker. The scientific discussions during our doctoral studies helped me a lot to develop my ideas and the non-scientific discussions were welcome distractions from the day-to-day work.

During my doctoral studies, I worked for the Telecommunications Research Center Vienna, and the Signal Processing and Speech Communications (SPSC) Laboratory at Graz University of Technology. In this context, I thank the head of the SPSC lab and co-assessor of this thesis, Univ.-Prof. Dr. Gernot Kubin for his support and helpful feedback, and for creating such a creative and pleasant working environment.

I also thank Prof. Dr. Thomas Eriksson from Chalmers University of Technology for agreeing to be the external co-assessor of this thesis and for his willingness to come to Graz to be part of the board of examiners at the final exam of my doctoral studies.

This thesis and my doctoral study were only possible due to the long-lasting project collaboration with Intel Connected Home Division Villach. In this context, I thank Dr. Steffen Trautmann and Alexander Kahl for their support and helpful feedback, and for giving me the freedom to work on both theoretical and practical problems.

The research leading to this thesis has received funding from the Austrian Research Promotion Agency FFG under the project numbers 835187 and 4718971. In this context, I thank the Austrian tax payers who provided the money for this funding.

Beside of the pleasant professional environment, I am very grateful to the people who form an important part of my private life. I thank my mother, who always was there for me and supported me throughout my whole life. And last, but definitely not least, I thank my girlfriend Julia for all the joy that her company brings into my life.

Preface

The main purpose of a doctoral thesis is to demonstrate the author's ability to conduct independent scientific research. Apart from this, however, a doctoral thesis can fulfill additional purposes that are aimed either at the author or the potential reader.

For the author, i.e., for me, the process of writing went along with a reflection of my own work over the last five years. It forced me to make a selection of the central results and to present them in a self-consistent way. This gave me a new perspective on my own work and inspired me to formulate and prove the phase homogeneity requirement, which is a previously unpublished result that was developed during the writing process.

For the potential reader, i.e., for you, this thesis may provide an introduction to the field of research, theoretical insights, or a collection of practical methods. My goal was to write this thesis as modular as possible, so that the individual chapters and appendices can be read independently from each other. I am aware that the audience of doctoral theses is in general very limited, but nevertheless, I hope that this thesis finds its way to the one or the other interested reader who finds something useful in it.

Contents

1	Introduction	1
1.1	Digital Predistortion of Wireless Transmitters	1
1.2	Motivation and Scope	3
1.3	Outline and Contributions	4
1.4	Related Doctoral Theses	5
2	The Linearity-Efficiency Trade-off	6
2.1	Introduction	6
2.2	A Joint Linearity-Efficiency Model	7
2.2.1	Fourier Series of the Drain Current	9
2.2.2	From Drain Current to Linearity and Efficiency	13
2.3	Evaluating Linearity and Efficiency	17
2.3.1	Signal Characterization	19
2.3.2	Linearity Quantification	21
2.3.3	Efficiency Quantification	23
2.3.4	Application Example	24
2.4	Highly Efficient RF Power Amplifiers	26
2.4.1	Classes of Operation	26
2.4.2	Efficiency Enhancement	28
2.5	Summary and Conclusion	29
3	Even-Order Terms in Polynomial Baseband Models	30
3.1	Introduction	30
3.2	Polynomial Models without Memory	33
3.2.1	Spectral Analysis	33
3.2.2	The Chebyshev Transform	34
3.2.3	From Passband to Baseband	36
3.2.4	From Baseband to Passband	37
3.2.5	Magnitude Power Functions	38
3.2.6	The Quasi-Memoryless Model	46
3.2.7	Phase Homogeneity	50
3.2.8	Discussion	51
3.3	Polynomial Models with Memory	53
3.3.1	The Volterra Series	53
3.3.2	The Baseband Volterra Series	55
3.3.3	The Baseband Volterra Series with Even-Order Terms	57
3.3.4	The Generalized Memory Polynomial	61
3.3.5	Phase Homogeneity	62
3.3.6	Discussion	64
3.4	Summary and Conclusion	65

4	Dual-Band Digital Predistortion of an Envelope-Tracking Power Amplifier	66
4.1	Introduction	66
4.1.1	Measurement Setup	67
4.1.2	Signal and Scoring	68
4.1.3	Linearization Architecture	70
4.2	Crest Factor Reduction	71
4.2.1	Structure of the Crest Factor Reducer	71
4.2.2	Design of the Crest Factor Reducer	72
4.3	Digital Predistortion	73
4.3.1	Structure of the Digital Predistorter	73
4.3.2	Training of the Digital Predistorter	76
4.4	Measurement Results	78
4.5	Summary and Conclusion	79
5	Summary and Conclusion	80
5.1	Summary of Results and Contributions	80
5.2	Comparison with Related Doctoral Theses	81
	Appendices	82
A	Complex Signals	83
B	Modulation and Demodulation	85
C	Wirtinger Calculus	88
D	Linear Least Squares	93
E	Synchronization	96
F	Combinatorics	105
G	Indefinite Integrals	107
H	Definite Integrals	109
	Bibliography	116

List of Figures

1.1	Block diagram of a digital wireless transmitter with CFR and DPD.	2
2.1	Circuit model of a radio frequency power amplifier.	7
2.2	Drain current characteristics of a transistor.	7
2.3	Mapping of the gate voltage to the drain current.	9
2.4	Spectral components of the drain current over the conduction angle.	12
2.5	Mapping of the drain current to the drain voltage.	13
2.6	Power, linearity, and efficiency characteristics of RF-PA models.	16
2.7	Effect of a nonlinearity on single-carrier and multi-carrier signals.	18
2.8	Magnitude distributions of single-carrier and multi-carrier signals.	20
2.9	Linearity quantification by measuring frequency-dependent error.	22
2.10	AM-AM characteristics of a memoryless DPD system.	24
2.11	EVM, ACPR, and average drain efficiency over output backoff.	25
3.1	PSDs produced by monomial basis functions.	33
3.2	Odd-symmetric and even-symmetric magnitude power functions.	39
3.3	Eigenvalues of the Chebyshev transform over its order and harmonic index.	42
3.4	PSDs produced by one-sided magnitude power functions.	43
3.5	PSDs produced by odd-symmetric magnitude power functions.	44
3.6	PSDs produced by even-symmetric magnitude power functions.	45
3.7	Quadrature model of a quasi-memoryless passband system.	48
3.8	Symmetric regions in the second and third order Volterra kernels.	54
3.9	Implementation of the GMP based on lookup tables.	61
4.1	Measurement setup of the 2017 DPD competition.	67
4.2	PSDs and waveforms of the dual-band OFDM signal.	68
4.3	Block diagram of the linearization architecture.	70
4.4	Block diagram of the dual-band CFR structure.	71
4.5	PSDs and linearity metrics after the dual-band CFR.	72
4.6	Block diagram of the dual-band DPD structure.	73
4.7	Kernel structure of the modified GMP.	74
4.8	Block diagram of the 2D-vector-switched model and its modeling regions.	75
4.9	Block diagrams of the indirect and direct learning architectures.	76
4.10	Linearization performance in the frequency and time domain.	78
C.1	Plot of a real-valued function in the complex plane and its gradient.	92
E.1	System models for the delay and the complex gain correction.	96
E.2	Example of the discrete-time delay estimation and correction.	103

List of Tables

3.1	Quasi-memoryless polynomial with odd-order terms only.	49
3.2	Quasi-memoryless polynomial with non-negative real-order terms.	49
3.3	Coefficient scaling factors of the quasi-memoryless polynomial.	49
4.1	Score specification of the 2017 DPD competition.	69
F.1	Summary of basic formulas from enumerative combinatorics.	106
F.2	Examples of basic problems from enumerative combinatorics.	106

List of Acronyms

AC	Alternating Current or Alternating Component
ACPR	Adjacent Channel Power Ratio
ADC	Analog-to-Digital Converter
AM-AM	Amplitude Modulation to Amplitude Modulation
AM-PM	Amplitude Modulation to Phase Modulation
CFR	Crest Factor Reduction
DAC	Digital-to-Analog Converter
DC	Direct Current or Direct Component
DPD	Digital Predistortion
EVM	Error Vector Magnitude
GMP	Generalized Memory Polynomial
IQ	In-Phase and Quadrature-Phase
LNA	Low Noise Amplifier
MOSFET	Metal-Oxid-Semiconductor Field-Effect Transistor
MSE	Mean Square Error
NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency-Division Multiplexing
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
PSD	Power Spectral Density
RF	Radio Frequency
RMS	Root Mean Square
RX	Receiver
TX	Transmitter

1

Introduction

Digital wireless communication has become an ubiquitous technology. Driven by data-hungry applications, there is an ongoing trend to higher data rates, larger bandwidths, and higher accuracies. On top of this, there is a strive for reduced energy consumption, motivated by the operational costs of basestations and the operational time of battery-powered devices. A critical component regarding both the energy consumption and the signal quality of wireless transmitters is the radio frequency power amplifier (RF-PA).

This thesis presents analytical results and practical methods for the behavioral modeling and digital predistortion (DPD) of RF-PAs, which are key techniques for the design and implementation of energy efficient, high data rate wireless transmitters.

1.1 Digital Predistortion of Wireless Transmitters

DPD is a form of digital enhancement of analog circuits [164]. As such, it requires an in-depth understanding of the analog impairments and their effects on the overall system performance. The dominating impairment targeted by the DPD of a wireless transmitter is the nonlinear distortion produced by the RF-PA, which causes

1. a spectral broadening of the transmitted signal, and
2. a reduction of the transmitted signal's modulation accuracy.

The spectral broadening can lead to interference with nearby channels and is therefore restricted by wireless standards and regulatory agencies. The modulation accuracy affects the achievable data rate and is therefore also specified by wireless standards. Today's high data rate wireless standards typically demand high linearity and generate signals with high peak-to-average power ratio (PAPR) [153]. This creates challenging requirements for RF-PAs, especially if the energy consumption must be kept low. To address these challenges, DPD is often included in modern wireless transmitters [110].

The working principle of DPD is to add a nonlinearity to the digital baseband processing of a wireless transmitter that predistorts the desired transmitted signal with the aim to pre-compensate the nonlinearity of the RF-PA. For this purpose, a complex baseband model of the inverse RF-PA behavior must be identified from a signal that is measured at the RF-PA output, demodulated, and processed by the digital baseband processor.

The block diagram of a digital wireless transmitter with DPD is shown in Figure 1.1.

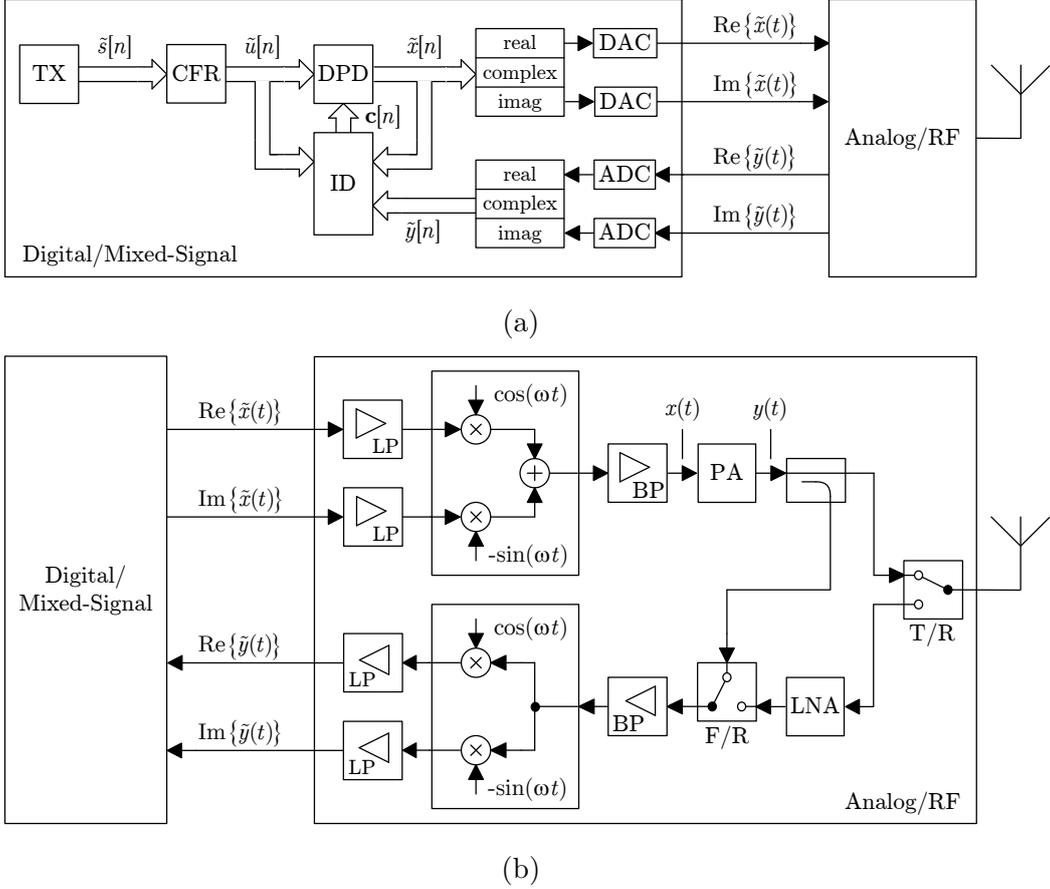


Figure 1.1: Block diagram of a digital wireless transmitter with CFR and DPD.

Components in the Digital/Mixed-Signal Section

The digital/mixed-signal section in Figure 1.1 (a) contains the following blocks:

- TX** The transmit (TX) signal generator block produces the discrete-time complex baseband signal, which represents the desired modulation of the RF carrier¹.
- CFR** The crest factor reduction (CFR) block reduces or limits the crest factor of the source signal, which is a synonym for its peak-to-average power ratio (PAPR).
- DPD** The DPD block predistorts the crest factor reduced baseband signal with the aim to pre-compensate the nonlinear distortion produced by the RF-PA.
- ID** The identification (ID) block adapts the nonlinearity of the DPD block based on the signal from the RF-PA output and the DPD input/output signals.

¹Unless otherwise noted, the symbol \sim is used to indicate complex baseband signals.

The interface between the digital/mixed-signal section and the analog/RF section is realized by two pairs of analog baseband signals, representing the real and imaginary parts of the respective complex baseband signals. In the transmit path, the analog baseband signals are generated by digital-to-analog converters (DACs) and in the receive path, the digital baseband signals are acquired by analog-to-digital converters (ADCs).

Components in the Analog/RF Section

The analog/RF section in Figure 1.1 (b) contains filters, mixers, and amplifiers. The mixers are arranged in a quadrature configuration to implement frequency translation based on complex modulation and demodulation. Introductions to complex signals and their application in modulators and demodulators are given in Appendices A and B. Before and after frequency translation, the signals pass through lowpass (LP) and bandpass (BP) amplifiers/filters that are designed to drive the following blocks and to suppress undesired signal components like images or harmonics. The power amplifier (PA) and low noise amplifier (LNA) are connected to the antenna by a transmit/receive (T/R) switch. In Figure 1.1 (b), the receive path is reused as a feedback path for DPD training by using a directional coupler and a feedback/receive (F/R) switch. For highly accurate linearization, however, dedicated DPD feedback paths are commonly used [239].

1.2 Motivation and Scope

Motivation of this Thesis

DPD is challenging, because it requires the combined knowledge of RF/microwave system design and nonlinear signal processing. Research on DPD is of high practical relevance, because advanced DPD methods can provide crucial competitive advantages in the highly contested wireless communications market. Despite the large amount of research on DPD [100, 101, 110, 121, 238], there are still aspects that call for further investigations. Central research questions that motivate this thesis are:

- What are the limitations of linearity and efficiency of RF-PAs?
- What are the foundations of complex baseband models of RF-PAs?
- Which methods give the best results in practical DPD applications?

Scope of this Thesis

The scope of this thesis is the analysis and design of digital signal processing methods for the behavioral modeling and DPD of non-switched RF-PAs with wideband signals. Related fields of research, beyond the scope of this thesis, are the digital enhancement of data converters [216], frequency converters [11], and frequency synthesizers [159], the design of full-digital wireless transceivers [176], and the circuit design of RF-PAs [56].

1.3 Outline and Contributions

Thesis Outline

This thesis consists of five chapters and eight appendices.

Chapter 1 gives an introduction to the field of research and the thesis.

Chapter 2 investigates the linearity-efficiency trade-off of RF-PAs based on a novel semi-physical model and presents an overview of linearity and efficiency quantification methods and highly efficient RF-PA operation modes and architectures. The main contribution is the joint linearity-efficiency model, which was published in [74, 80].

Chapter 3 investigates polynomial baseband models of RF-PAs with a focus on the often neglected even-order terms in baseband. The main contributions are explicit passband-baseband pairs of polynomial baseband models with even-order terms and the formulation and proof of the phase homogeneity requirement. The derivations of passband-baseband pairs with even-order terms were published in [77, 79].

Chapter 4 presents practical methods for CFR and DPD, exemplified by the dual-band linearization of an envelope tracking RF-PA. This chapter is based on the winning solution of the 2017 student design competition "PA Linearization through DPD" at the IEEE International Microwave Symposium, which was published in [81].

Chapter 5 summarizes the main results and contributions of the thesis.

The appendices provide introductions to mathematical concepts and proofs of statements encountered in this thesis. Appendix E is based on the survey published in [76].

Personal Contributions within this Thesis

This thesis is based on five publications [76, 77, 79–81] with co-authorship of Karl Freiberger, Gernot Kubin and Christian Vogel. The ideas, derivations, implementations, and formulations of these publications are primarily my own work. Contributions of my co-authors consist of scientific discussions and the review of manuscripts.

Personal Contributions beyond this Thesis

During my doctoral studies I published two additional works [78, 82] that are related, but outside the scope of this thesis. The work in [78] is supplemented with source code [75].

Furthermore, I contributed to the work of Karl Freiberger on optimization-based DPD identification [98] and error vector magnitude estimation [95–97].

Finally, I wrote the MATLAB program "bibget" [73], which simplifies the creation of BibTeX databases, by downloading bibliographic information from IEEE Xplore.

1.4 Related Doctoral Theses

Many doctoral theses have been written on the subject of DPD. Below, I summarize three of them. In Chapter 5, I compare them with my own thesis.

[63] Lei Ding, 2004, Georgia Institute of Technology

This thesis presents nonlinear baseband models and parameter identification methods for the DPD of wireless transmitters [64, 68, 70]. The work in [70] received the highest number of citations by presenting simulation results that demonstrate the good performance of a memory polynomial predistorter, identified by linear least squares fitting based on the indirect learning architecture. The feasibility of [70] is demonstrated by an implementation on a digital signal processor [67]. Furthermore, this thesis presents the first systematic investigation of even-order terms in polynomial baseband models [62] and digital compensation methods for the linear frequency response of two-stage frequency converters and the frequency-dependent IQ mismatch of complex modulators [65].

[204] Peter Singerl, 2006, Graz University of Technology

This thesis investigates nonlinear baseband modeling based on the Volterra series with the aim to reduce the number of parameters. For this purpose, the baseband Volterra series is approximated by orthogonal polynomials in the frequency domain [202]. To reduce the sampling rate during DPD identification, it is proposed to identify a model of the RF-PA at the Nyquist rate of the input signal and to use interpolation to make this model useable for the design of a DPD [201]. The performance of linearization by DPD is simulated using the indirect learning architecture [199] and a special low-rate identification architecture [200]. Furthermore this thesis presents a method for the identification of memory polynomials based on frequency-dependent dual-tone responses [203].

[211] Ali Soltani Tehrani, 2012, Chalmers University of Technology

This thesis presents a measurement-based comparison of RF-PA behavioral models that demonstrates the good complexity-accuracy trade-off of the generalized memory polynomial model [212]. Furthermore, it proposes to model long-term memory effects of RF-PAs by behavioral models with dynamic parameters [213] and shows that the same method can also be used to improve the performance of DPD systems regarding bursty communication signals. Finally this thesis presents an experimental setup for emulating the closed-loop parameter adaptation of DPD systems by using measurement devices that only allow the block-by-block generation and acquisition of signal vectors.

2

The Linearity-Efficiency Trade-off

Designing RF-PAs that are linear and efficient is challenging. In practice, there is a trade-off between the linearity and efficiency of RF-PAs. A key method to circumvent this trade-off is to use DPD for linearizing highly efficient, but nonlinear RF-PAs.

In this chapter, I investigate the linearity-efficiency trade-off, discuss metrics for linearity and efficiency quantification, and outline highly efficient RF-PA architectures. The main contribution is the joint linearity-efficiency model, which was published in [74, 80].

2.1 Introduction

The quest for efficient RF-PAs started in the early days of radio broadcasting [38]. Since then, the discipline of transmitter design has changed a lot, but the trade-off between linearity and efficiency remains [140]. Earl McCune explains the linearity-efficiency trade-off as a direct consequence of Ohm's law, combined with the operation principle of a transistor [152]. The qualitative analysis in [152] is an important foundation for understanding the design problems of RF-PAs. However, to further explore the linearity-efficiency trade-off, also quantitative methods are needed. An example for a quantitative method is the efficiency analysis of the ideal class A and class B operation of RF-PAs, which can, for example, be found in a book by Steve Cripps [53]. The main results of this analysis are the well-known peak efficiencies of 50% for class A and 78.5% for class B. A similar analysis can be applied with respect to linearity, but with the piecewise-linear transistor model that is typically used for efficiency analysis [53], the result is perfect linearity for class A and class B, which is clearly too simplistic. In [132], Lawrence Kushner uses a piecewise-quadratic transistor model for the efficiency analysis, but he does not evaluate its effect on the linearity. In [53], Steve Cripps proposes a piecewise-cubic transistor model, but he does not use it for linearity analysis either.

In this chapter, I present the derivation of a joint linearity-efficiency model of RF-PAs using both the piecewise-linear and the more realistic piecewise-cubic transistor model. Afterwards, I introduce linearity and efficiency quantification methods and demonstrate their application based on the previously derived RF-PA model. Finally, I conclude with an overview of highly efficient RF-PA operation modes and architectures.

2.2 A Joint Linearity-Efficiency Model

The following derivation is based on a generic single-ended RF-PA circuit, which is shown in Figure 2.1 and which is also used in [53]. This circuit consists of a transistor, an input and an output matching network, an inductor connecting the drain of the transistor to a fixed supply voltage, a signal generator at the input and a fixed load at the output. Due to the inductive coupling between the supply and the drain node of the transistor, the DC voltage¹ at the drain is equal to the supply voltage. For the gate node, we assume that its DC voltage can be adjusted by a bias circuit, which is not shown in Figure 2.1.

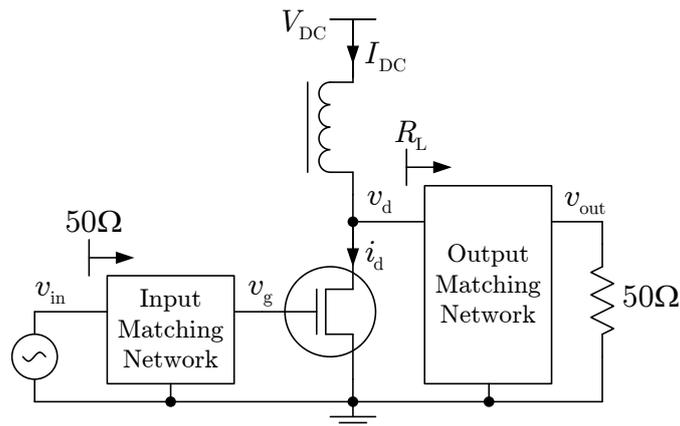


Figure 2.1: Circuit model of a radio frequency power amplifier.

The large-signal behavior of a transistor is typically modeled by a voltage-controlled current source with characteristics like in Figure 2.2. The blue and red lines represent different levels of approximation. For the following analysis we would like to use a very simple transistor model that still includes the dominating sources of nonlinearity.

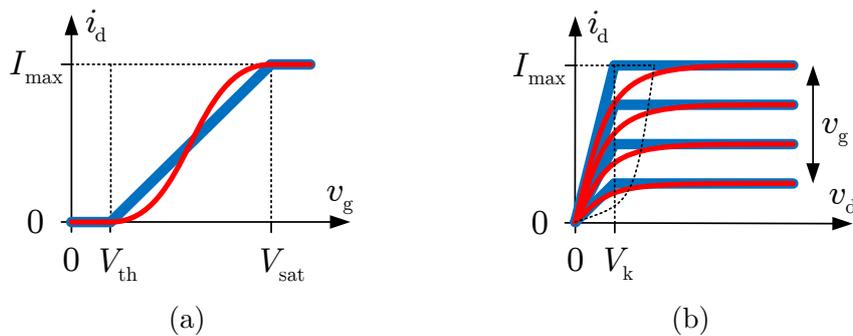


Figure 2.2: Drain current characteristics of a transistor.

¹Historically, the terms AC and DC refer to alternating current and direct current, respectively. In this thesis, I follow the common practice of using these terms in a broader sense for all types of signals. In this sense, AC can be read as alternating component and DC can be read as direct component.

In the classical efficiency analysis [53], the drain current is modeled by the blue lines in Figure 2.2, which represent a piecewise-linear model. The only nonlinearity considered is the break-point at the threshold voltage V_{th} in Figure 2.2 (a). The break-point at the saturation voltage V_{sat} in Figure 2.2 (a) is not considered, since it is assumed that the gate voltage v_g always stays below V_{sat} . The break-point at the knee voltage V_k in Figure 2.2 (b) is avoided by assuming that the knee voltage V_k is zero.

In the joint linearity-efficiency analysis presented below, we use two transistor models. The drain current in Figure 2.2 (a) is modeled either by a piecewise-linear or by a piecewise-cubic model, represented by the blue and red lines in Figure 2.2 (a). We do not restrict the range of the gate voltage, thus we consider both break-points at V_{th} and V_{sat} . For the output characteristic in Figure 2.2 (b) we use a piecewise-linear model like in [53]. The use of a more complicated model for the output characteristic would require the solution of nonlinear differential equations, which model the interaction between the drain voltage and the drain current. This is beyond the scope of the presented analysis.

To model the characteristics in Figure 2.2 (a), we define normalized input and output variables s_{in} and s_{out} using a coordinate system where the point $(v_g, i_d) = (V_{\text{th}}, 0)$ corresponds to $(s_{\text{in}}, s_{\text{out}}) = (0, 0)$ and the point $(v_g, i_d) = (V_{\text{sat}}, I_{\text{max}})$ corresponds to $(s_{\text{in}}, s_{\text{out}}) = (1, 1)$. Mapping v_g to i_d consists of three steps. First, v_g is normalized by

$$s_{\text{in}} = \frac{v_g - V_{\text{th}}}{V_{\text{sat}} - V_{\text{th}}}. \quad (2.1)$$

Then, s_{in} is mapped to s_{out} , by using either the piecewise-linear model given by

$$s_{\text{out}} = \begin{cases} 0 & s_{\text{in}} < 0 \\ s_{\text{in}} & 0 \leq s_{\text{in}} \leq 1 \\ 1 & s_{\text{in}} > 1 \end{cases} \quad (2.2)$$

or by using the piecewise-cubic model given by

$$s_{\text{out}} = \begin{cases} 0 & s_{\text{in}} < 0 \\ 3s_{\text{in}}^2 - 2s_{\text{in}}^3 & 0 \leq s_{\text{in}} \leq 1 \\ 1 & s_{\text{in}} > 1 \end{cases} \quad (2.3)$$

Finally, s_{out} is converted to the drain current by

$$i_d = I_{\text{max}} s_{\text{out}}. \quad (2.4)$$

The cubic polynomial in (2.3) was initially proposed in [53] and has several properties that make it well suited for the present application. First, it is simple, which is important for analytic derivations. Second, it produces a continuous piecewise model because its value at $s_{\text{in}} = 0$ is zero, its value at $s_{\text{in}} = 1$ is one and its first derivatives at $s_{\text{in}} = 0$ and $s_{\text{in}} = 1$ are zero. Third, it is also physically justified since for small values of s_{in} , it is dominated by the quadratic term, like the square-law model of MOSFETs.

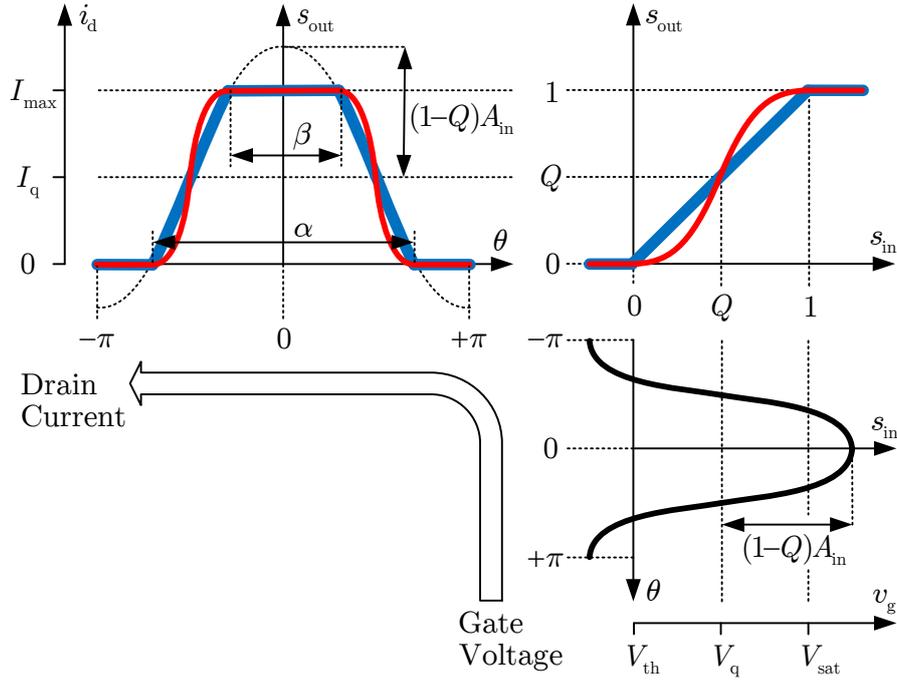


Figure 2.3: Mapping of the gate voltage to the drain current.

2.2.1 Fourier Series of the Drain Current

The first step in deriving the joint linearity-efficiency model is to derive a Fourier series of the drain current waveform resulting from a large-amplitude sinusoidal input signal at the gate. The mapping of the gate voltage to the drain current using the characteristics in Figure 2.2 (a) is illustrated in Figure 2.3. We define the input signal by

$$s_{\text{in}}(\theta) = Q + (1 - Q)A_{\text{in}} \cos(\theta) \quad (2.5)$$

where Q is the quiescent point, A_{in} is the input amplitude and $\theta = \omega t$ is an angle in radians representing the progress of time. We defined (2.5) in such a way that for $A_{\text{in}} = 1$ and $Q < 1$ the maxima of s_{in} are equal to one. If we send the input signal through one of the transfer functions in the upper right of Figure 2.3, we obtain the respective output signal $s_{\text{out}}(\theta)$ in the upper left of Figure 2.3. Due to clipping, we can discriminate three regions of the output signal. For $s_{\text{in}} < 0$ the output signal is constantly zero, for $s_{\text{in}} > 1$ the output signal is constantly one and for $0 \leq s_{\text{in}} \leq 1$ the output signal is either equal to the input signal, or it is a distorted version of the input signal. In the time domain, these regions are specified by the angles α and β . The conduction angle α specifies the region where the input and the output signal are above zero. The saturation angle β specifies the region where the input signal is above one and the output signal is equal to one. By setting (2.5) equal to zero (or one) and solving for θ , we can derive formulas for the conduction angle α and the saturation angle β as functions of A_{in} and Q .

The formula for the conduction angle is given by

$$\alpha = \begin{cases} 0 & Q \leq 0 \ \& \ A_{\text{in}} \leq |Q_{\text{r}}| \\ 2\pi & Q > 0 \ \& \ A_{\text{in}} \leq |Q_{\text{r}}| \\ 2 \arccos\left(\frac{Q_{\text{r}}}{A_{\text{in}}}\right) & \text{otherwise} \end{cases} \quad (2.6)$$

with $Q_{\text{r}} = Q/(Q - 1)$ and the formula for the saturation angle is given by

$$\beta = \begin{cases} 0 & Q \leq 1 \ \& \ A_{\text{in}} \leq 1 \\ 2\pi & Q > 1 \ \& \ A_{\text{in}} \leq 1 \\ 2 \arccos\left(\frac{1}{A_{\text{in}}}\right) & \text{otherwise} \end{cases} \quad (2.7)$$

The drain current $i_{\text{d}}(\theta)$ can be described for $-\pi \leq \theta \leq +\pi$ by

$$i_{\text{d}}(\theta) = \begin{cases} I_{\text{max}} & 0 < |\theta| \leq \frac{\beta}{2} \\ I_{\text{max}} s_{\text{out}}(\theta) & \frac{\beta}{2} < |\theta| \leq \frac{\alpha}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (2.8)$$

Since the drain current is 2π -periodic and even-symmetric in θ , it can be expressed for arbitrary values of θ by the Fourier series

$$i_{\text{d}}(\theta) = I_{\text{DC}} + \sum_{k=1}^{\infty} I_k \cos(k\theta) \quad (2.9)$$

where the k -th harmonic current amplitudes are given by

$$I_k = \frac{2}{\pi} \int_0^{\pi} i_{\text{d}}(\theta) \cos(k\theta) \, d\theta \quad (2.10)$$

and the DC current is given by

$$I_{\text{DC}} = \frac{I_0}{2}. \quad (2.11)$$

Inserting (2.8) into (2.10) we obtain

$$I_k = \frac{2}{\pi} \left[\int_0^{\frac{\beta}{2}} I_{\text{max}} \cos(k\theta) \, d\theta + \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} I_{\text{max}} s_{\text{out}}(\theta) \cos(k\theta) \, d\theta \right] \quad (2.12)$$

and dividing (2.12) by I_{max} results in

$$\frac{I_k}{I_{\text{max}}} = \frac{2}{\pi} \left[\int_0^{\frac{\beta}{2}} \cos(k\theta) \, d\theta + \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} s_{\text{out}}(\theta) \cos(k\theta) \, d\theta \right]. \quad (2.13)$$

To express (2.13) in terms of A_{in} and Q , we insert expressions for $s_{\text{out}}(\theta)$ into (2.13) using either the piecewise-linear (2.2) or the piecewise-cubic (2.3) transistor model.

Using the piecewise-linear transistor model with $\frac{\beta}{2} \leq \theta \leq \frac{\alpha}{2}$, or equivalently, $0 \leq s_{\text{in}} \leq 1$, we have $s_{\text{out}}(\theta) = s_{\text{in}}(\theta)$, with $s_{\text{in}}(\theta)$ defined in (2.5). Inserting (2.5) into (2.13), we get

$$\begin{aligned} \frac{I_k}{I_{\text{max}}} = \frac{2}{\pi} & \left[\int_0^{\frac{\beta}{2}} \cos(k\theta) \, d\theta \right. \\ & + Q \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos(k\theta) \, d\theta \\ & \left. + (1-Q)A_{\text{in}} \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos(\theta) \cos(k\theta) \, d\theta \right]. \end{aligned} \quad (2.14)$$

Using the piecewise-cubic transistor model with $\frac{\beta}{2} \leq \theta \leq \frac{\alpha}{2}$, or equivalently, $0 \leq s_{\text{in}} \leq 1$, we have $s_{\text{out}}(\theta) = 3s_{\text{in}}^2(\theta) - 2s_{\text{in}}^3(\theta)$ with $s_{\text{in}}(\theta)$ defined in (2.5). Inserting (2.5) into the cubic polynomial, expanding the terms and collecting equal powers of $\cos(\theta)$ we get

$$\begin{aligned} s_{\text{out}}(\theta) &= 3[Q + (1-Q)A_{\text{in}} \cos(\theta)]^2 \\ &\quad - 2[Q + (1-Q)A_{\text{in}} \cos(\theta)]^3 \\ &= 3[Q^2 + 2Q(1-Q)A_{\text{in}} \cos(\theta) + (1-Q)^2 A_{\text{in}}^2 \cos^2(\theta)] \\ &\quad - 2[Q^3 + 3Q^2(1-Q)A_{\text{in}} \cos(\theta) \\ &\quad \quad + 3Q(1-Q)^2 A_{\text{in}}^2 \cos^2(\theta) + (1-Q)^3 A_{\text{in}}^3 \cos^3(\theta)] \\ &= (3Q^2 - 3Q^3) \\ &\quad + [(6Q - 6Q^2)(1-Q)A_{\text{in}}] \cos(\theta) \\ &\quad + [(3 - 6Q)(1-Q)^2 A_{\text{in}}^2] \cos^2(\theta) \\ &\quad - [2(1-Q)^3 A_{\text{in}}^3] \cos^3(\theta). \end{aligned} \quad (2.15)$$

Inserting the final result of (2.15) into (2.13), we obtain

$$\begin{aligned} \frac{I_k}{I_{\text{max}}} = \frac{2}{\pi} & \left[\int_0^{\frac{\beta}{2}} \cos(k\theta) \, d\theta \right. \\ & + (3Q^2 - 2Q^3) \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos(k\theta) \, d\theta \\ & + (6Q - 6Q^2)(1-Q)A_{\text{in}} \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos(\theta) \cos(k\theta) \, d\theta \\ & + (3 - 6Q)(1-Q)^2 A_{\text{in}}^2 \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos^2(\theta) \cos(k\theta) \, d\theta \\ & \left. - 2(1-Q)^3 A_{\text{in}}^3 \int_{\frac{\beta}{2}}^{\frac{\alpha}{2}} \cos^3(\theta) \cos(k\theta) \, d\theta \right]. \end{aligned} \quad (2.16)$$

Solution formulas for the integrals in (2.14) and (2.16) are given in Appendix G.

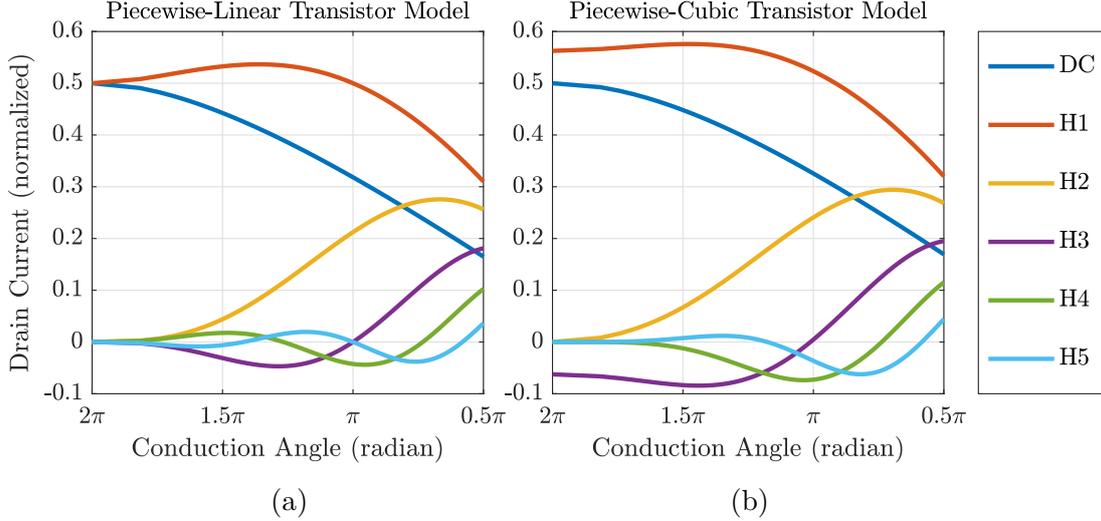


Figure 2.4: Spectral components of the drain current depending on the conduction angle. The plots show I_{DC}/I_{\max} (DC) and I_k/I_{\max} for $k \in \{1, 2, 3, 4, 5\}$ (H1-H5).

For simulating the spectral components of the drain current at various input amplitudes A_{in} and quiescent points Q , I have written a MATLAB program [74] that implements the mapping of A_{in} and Q to the conduction angle α and saturation angle β given by (2.6) and (2.7) and the computation of (2.14) and (2.16) using the integral formulas in Appendix G. Using this program, I simulated the spectral components of the drain current for $A_{\text{in}} = 1$ and $-10 \leq Q \leq 0.5$ and plotted the results over the conduction angle α as shown in Figure 2.4. The characteristics in Figure 2.4 illustrate the operation modes of RF-PAs, which are classified in the following way [53, 83]:

Class A This class is defined by a conduction angle of $\alpha = 2\pi$, obtained by $Q = 0.5$.

At $\alpha = 2\pi$, Figure 2.4 (a) shows a sinusoidal drain current, consisting of the DC and the first harmonic. Figure 2.4 (b) additionally shows a third harmonic.

Class AB This class is defined by a conduction angle of $\pi < \alpha < 2\pi$, obtained by $0 < Q < 0.5$. From Figure 2.4, we see that reducing the conduction angle below class A reduces the DC, but also increases the harmonics.

Class B This class is defined by a conduction angle of $\alpha = \pi$, obtained by $Q = 0$. At $\alpha = \pi$, Figure 2.4 (a) shows a reduced DC with respect to class A and even-order harmonics. Figure 2.4 (b) shows the presence of all harmonics.

Class C This class is defined by a conduction angle of $0 < \alpha < \pi$, obtained by $Q < 0$. From Figure 2.4, we see that reducing the conduction angle below class B reduces the DC, but also reduces the first harmonic and increases the higher harmonics.

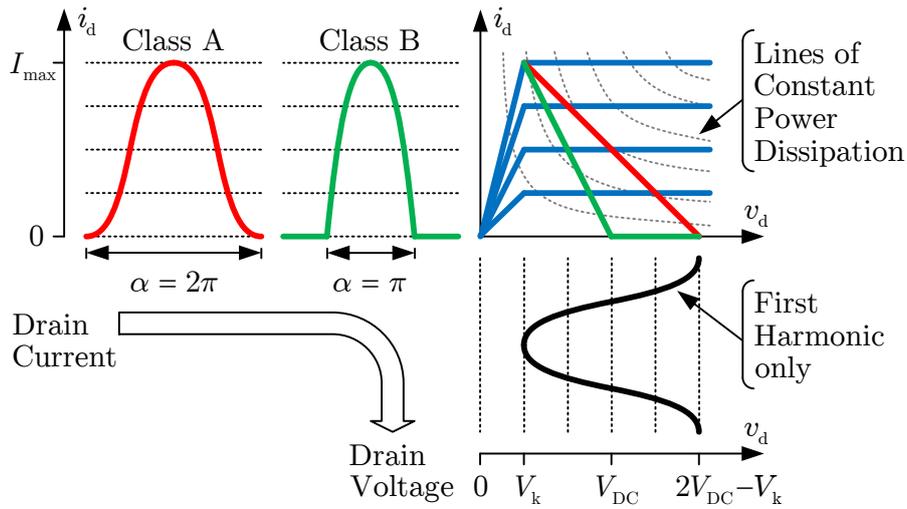


Figure 2.5: Mapping of the drain current to the drain voltage.

2.2.2 From Drain Current to Linearity and Efficiency

To model the linearity and efficiency of RF-PAs, we relate the drain current to the drain voltage and derive expressions for the DC power consumption and the RF output power. Simple relations can be found if the following two assumptions are made [53]:

Ideal transistor output The transistor behaves like an ideal current source with no feedback from the drain voltage to the drain current, as long as $v_d > V_k$.

Ideal matching network The output matching network is lossless, blocks DC, has a real impedance R_L at the first harmonic and zero impedance at higher harmonics.

The first assumption allows us to directly map the drain current to the drain voltage by neglecting any feedback within the transistor, as long as the drain voltage stays above the knee voltage. The second assumption results in a sinusoidal drain voltage, since the drain current harmonics are shorted by the output matching network.

The mapping of the drain current to the drain voltage for the piecewise-linear transistor model and the operation modes class A and class B is illustrated in Figure 2.5. At class A, the drain current is sinusoidal and the current-voltage trajectory is the straight line shown in red. At class B, the drain current is not sinusoidal and the current-voltage trajectory is the kinked line shown in green. The instantaneous power dissipation is equal to the product of drain voltage and drain current. Lines of constant power dissipation are hyperbolas, shown in gray. By comparing the current-voltage trajectories for class A and class B, we see that the path for class B moves along lower power dissipation, so we can expect a higher energy efficiency of class B compared to class A.

The drain voltage amplitude is the product of the first harmonic drain current amplitude and the load resistance R_L at the first harmonic. To maximize the output power obtained with full-scale input magnitude, we set the load resistance to

$$R_L = \frac{V_{\text{DC}} - V_k}{I_{1,\text{full}}}, \quad (2.17)$$

where V_{DC} is the supply voltage, V_k is the knee voltage and $I_{1,\text{full}}$ is the first harmonic drain current amplitude at $A_{\text{in}} = 1$. With this resistance, called the load line matching resistance [53], the drain voltage amplitude is maximized, subject to $v_d > V_k$. Since $I_{1,\text{full}}$ depends on the bias point, also the load line matching resistance depends on the bias point. In the following, we express resistances and powers relative to reference values obtained with the piecewise-linear transistor model, zero knee voltage, and class A or class B bias point. In this setup, the load line matching resistance is given by

$$R_{\text{ref}} = \frac{V_{\text{DC}}}{0.5 I_{\text{max}}}. \quad (2.18)$$

and the RF output power is obtained by multiplying the drain voltage amplitude V_{DC} by the drain current amplitude $0.5 I_{\text{max}}$, divided by two, resulting in

$$P_{\text{ref}} = \frac{V_{\text{DC}} I_{\text{max}}}{4}. \quad (2.19)$$

The DC power consumption is given by $P_{\text{DC}} = V_{\text{DC}} I_{\text{DC}}$. Dividing this equation by (2.19), we obtain the normalized DC power consumption by

$$\frac{P_{\text{DC}}}{P_{\text{ref}}} = 4 \frac{I_{\text{DC}}}{I_{\text{max}}}. \quad (2.20)$$

The RF output power is given by $P_{\text{RF}} = \frac{1}{2} I_1^2 R_L$. Dividing this equation by (2.18) and (2.19), we obtain the normalized RF output power by

$$\frac{P_{\text{RF}}}{P_{\text{ref}}} = 4 \left(\frac{I_1}{I_{\text{max}}} \right)^2 \frac{R_L}{R_{\text{ref}}}. \quad (2.21)$$

The normalized load resistance in (2.21) is chosen such that we get load line matching for the considered transistor model and bias point². The normalized load resistance that implements load line matching is obtained by dividing (2.17) by (2.18), resulting in

$$\frac{R_L}{R_{\text{ref}}} = \frac{1 - \frac{V_k}{V_{\text{DC}}}}{2 \frac{I_{1,\text{full}}}{I_{\text{max}}}}. \quad (2.22)$$

²In [80], we used a fixed load resistance for all bias points. To get a fairer comparison, we now use the actual load line matching resistance. The MATLAB program in [74] has been updated accordingly.

Equations (2.20) and (2.21) are implemented in the MATLAB program [74] that simulates the spectral components of the drain current. Using this program, I simulated the DC and RF powers for $0 \leq A_{\text{in}} \leq 2$ and $Q \in \{0, 0.14, 0.5\}$, resulting in the characteristics in Figure 2.6 (a,b). The results in Figure 2.6 (a) were obtained with the piecewise-linear transistor model and zero knee voltage, which is the setup used in conventional efficiency analysis [53]. The results in Figure 2.6 (b) were obtained with the piecewise-cubic transistor model and the non-zero knee voltage $V_k = 0.1 V_{\text{DC}}$.

The idealized characteristics in Figure 2.6 (a) show very simple behavior for class A and class B for $A_{\text{in}} \leq 1$. In both class A and class B, the RF power increases quadratically with the input amplitude, or linearly with the input power. In class A, the DC power is constant at twice the full-scale RF power and in class B, the DC power increases linearly with the input amplitude, or in a square-root relationship with the input power.

To better see the linearity and efficiency behavior, we can represent the DC and RF power characteristics in two additional forms. For this purpose, we define the normalized output amplitude A_{out} and the drain efficiency η_{drain} by

$$A_{\text{out}} = \sqrt{\frac{P_{\text{RF}}}{P_{\text{ref}}}} \quad \text{and} \quad \eta_{\text{drain}} = \frac{P_{\text{RF}}}{P_{\text{DC}}}. \quad (2.23)$$

The linearity is represented by the mapping $A_{\text{in}} \mapsto A_{\text{out}}$, which is called the amplitude modulation to amplitude modulation (AM-AM) characteristic and shown in Figure 2.6 (c) (d). The efficiency is represented by the mapping $A_{\text{in}} \mapsto \eta_{\text{drain}}$, which is called the instantaneous drain efficiency characteristic and shown in Figure 2.6 (e) (f).

With the piecewise-linear transistor model and $A_{\text{in}} \leq 1$, we obtain perfect linearity for class A and class B, and a kinked AM-AM characteristic for class AB, as shown in Figure 2.6 (c). With the piecewise-cubic transistor model, we obtain the more realistic linearity behavior of Figure 2.6 (d), showing gain compression for class A, a mixture of gain expansion and gain compression for class B, and good linearity for class AB.

The good linearity of class AB can be explained by a compensation of two nonlinearities within the power amplifier. First, there is current clipping, when the gate voltage drops below the threshold voltage and second, there is the progressive increase in drain current, when the gate voltage rises above the threshold voltage. At class AB, these nonlinearities compensate each other, producing an overall linear transfer characteristic. [53, p. 67]

From Figure 2.6 (e) (f), we see that the efficiency of the two setups is similar. With the piecewise-linear transistor model and zero knee voltage, we obtain the well known full-scale efficiencies of 50% for class A and 78.5% for class B, as shown in Figure 2.6 (e). With the piecewise-cubic transistor model and non-zero knee voltage, we obtain similar characteristics, as shown in Figure 2.6 (f). The similar efficiency of the two setups is caused by two opposite effects: The piecewise-cubic transistor model slightly increases the efficiency compared to the piecewise-linear transistor model and the non-zero knee voltage slightly decreases the efficiency compared to the zero knee voltage.

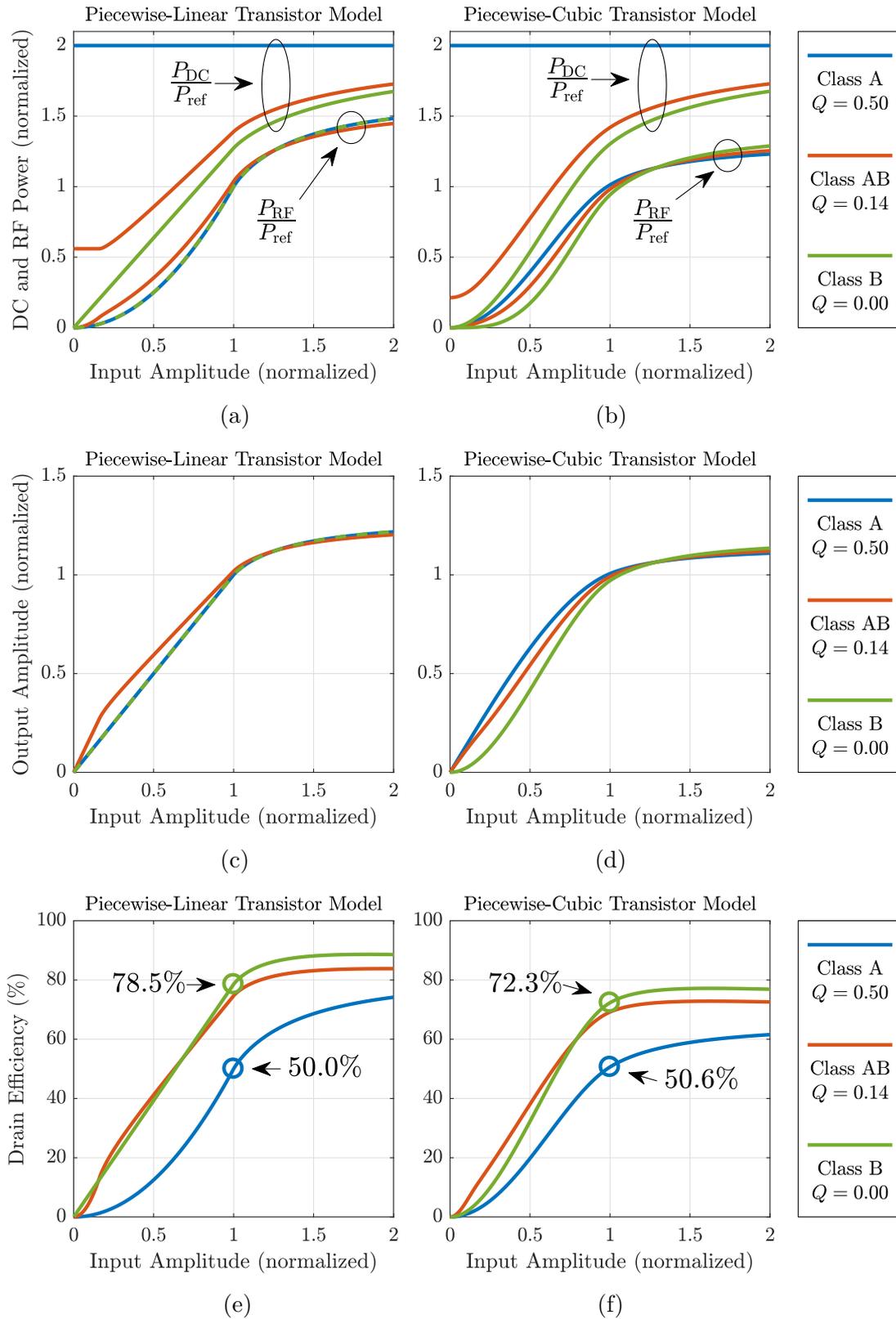


Figure 2.6: Power, linearity, and efficiency characteristics of RF-PA models.

2.3 Evaluating Linearity and Efficiency

In the previous section, we evaluated the linearity and efficiency of the RF-PA model by looking at its AM-AM characteristic and its instantaneous drain efficiency characteristic, respectively. These characteristics represent the steady state response of the system to a sinusoidal RF input signal. In an actual communication system, however, the RF-PA is not operated with a sinusoidal RF input signal, but with a complex modulated RF input signal. In the following, I demonstrate that depending on the type of complex modulation, the same RF-PA nonlinearity can have very different effects on the signal quality. For this purpose, I use the memoryless class AB nonlinearity from Figure 2.6 (d) and evaluated its effect on three different signal types, which are

- single-carrier quadrature phase shift keying (Q-PSK),
- single-carrier 64 point quadrature amplitude modulation (64-QAM), and
- orthogonal frequency division multiplexing (OFDM).

I generated the respective signals in complex baseband with a bandwidth of 20 MHz and a sampling rate of 200 MHz. For the single-carrier modulation, I used a root-raised-cosine pulse shaping filter with a rolloff factor of 0.5 and 50 filter taps. I used 16 samples per symbol, corresponding to a symbol rate of 12.5 MHz and a symbol length of 80 ns. For the OFDM signal³, I used a subcarrier spacing of 10 kHz and 1800 subcarriers. The used subcarrier spacing corresponds to a raw symbol length of 100 μ s, which is extend by a guard interval of 1/4 times the raw symbol length. The guard interval is added in front of each symbol by circular extension. To smooth the symbol transitions, I used a raised-cosine transition in the time domain, occupying 1/10 of the guard interval.

The effect of the class AB nonlinearity on the PSD and the constellation points of the respective signals is shown in Figure 2.7. Note that different levels of input backoff (IBO) were used, as indicated in the title of the plots to get similar magnitudes of nonlinear distortion. Here, the input backoff is defined as the reduction of the root-mean-square (RMS) magnitude of the complex baseband input signal with respect to the full-scale magnitude of one. For PSD estimation, I used Welch's method [230] with an overlap of 50 % and a Nuttall window [168, Fig. 15]. The signals had a length of 100 kSample and the fast Fourier transform was computed from segments of 1 kSample, corresponding to a bin spacing of 200 kHz and 199 averages. To obtain the constellation points of the single-carrier signals, I used a root-raised-cosine filter with the same specification as in the signal generator, followed by a linear 10-tap equalizer, which operates in the time domain at the symbol rate and is identified by data-aided least squares fitting. To obtain the constellation points of the OFDM signal, I used a linear frequency domain equalizer, also identified by data-aided least squares fitting. In the plot of the received constellation points in Figure 2.7 (f), the constellation points of all subcarriers are shown.

³A good introduction to OFDM is given in [236, Ch. 4]. For the history of OFDM, see [228].

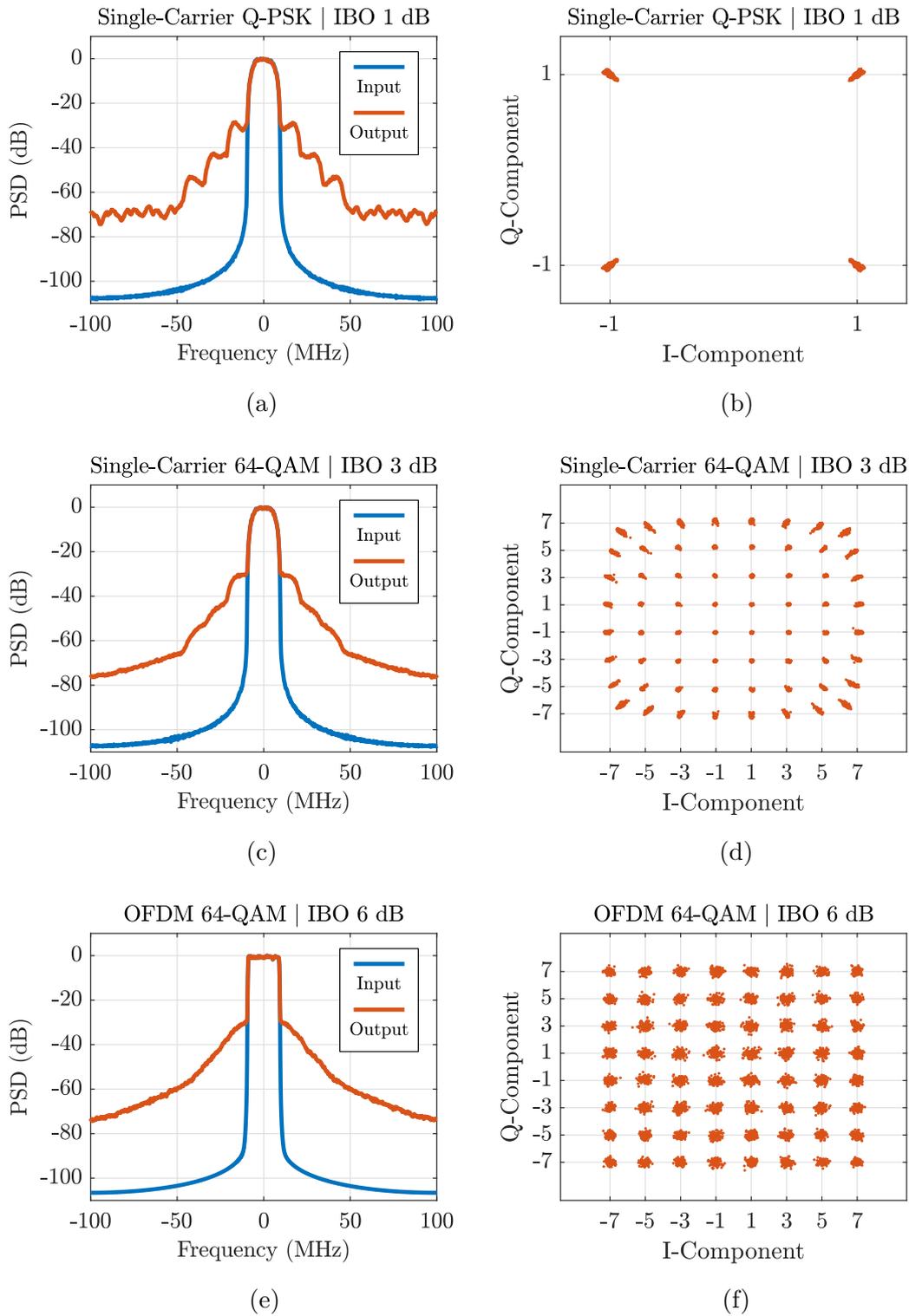


Figure 2.7: Effect of a nonlinearity on single-carrier and multi-carrier signals.

By comparing the PSDs in Figure 2.7, we notice that the out-of-band distortion strongly depends on the signal type. At the single-carrier Q-PSK signal, we see a step-wise decay of the out-of-band distortion, whereas at the other signals, the decay is smoother. PSDs like in Figure 2.7 can also be computed analytically based on a polynomial approximation of the nonlinearity and the statistics of the input signal [242–244, 246].

By looking at the constellation points in Figure 2.7, we notice two effects: First, there is a noise-like scattering around the mean of the respective constellation points and second, there is a distortion of the mean constellation points of the single-carrier 64-QAM signal. The noise-like scattering can be explained by missing orthogonality between the transmitted symbols. At the single-carrier signals, the impaired orthogonality leads to intersymbol interference and at the OFDM signal, it leads to intercarrier interference.

2.3.1 Signal Characterization

From Figure 2.7, we see that different signals require different levels of backoff to get similar levels of nonlinear distortion. These different requirements can be explained by the different magnitude statistics of the respective signals. To quantify the magnitude statistics of a complex baseband signal $x(t)$, we model it as a complex random process $X(t)$, whose magnitude follows a probability density function (PDF), defined by

$$\text{PDF}(a) = f(a) \quad \text{where} \quad \int_{a_1}^{a_2} f(a) da = \mathbb{P}\{a_1 \leq |X(t)| \leq a_2\}, \quad (2.24)$$

where $\mathbb{P}\{\cdot\}$ is the probability operator. Based on $x[n] = x(nT_s)$, where T_s is the sampling period and $n \in \mathbb{Z}$ with $1 \leq n \leq N$ are the sample indices, the magnitude PDF of $X(t)$ can be estimated by a kernel density estimator [221], given by

$$\text{PDF}(a) = \frac{1}{N} \sum_{n=1}^N K_h(a - |x[n]|), \quad (2.25)$$

where $K_h(a) = K(a/h)/h$ is the scaled kernel function, $K(\cdot)$ is the kernel function, which integrates to one and h is the kernel bandwidth. I estimated the PDFs of the signals from above, using a Gaussian kernel and a bandwidth of $h = 0.1$, resulting in the PDF estimates shown in Figure 2.8 (a). To better see the peak statistics, I converted the PDFs to complementary cumulative distribution functions (CCDFs), defined by

$$\text{CCDF}(a) = \mathbb{P}\{|X(t)| > a\} = 1 - \int_{-\infty}^a \text{PDF}(a) da, \quad (2.26)$$

which are shown in Figure 2.8 (b). For comparison, I also included the PDF and CCDF of a constant envelope signal in Figure 2.8. In general, the more the curves deviate from the constant envelope case, the more sensitive is the signal to nonlinear distortion.

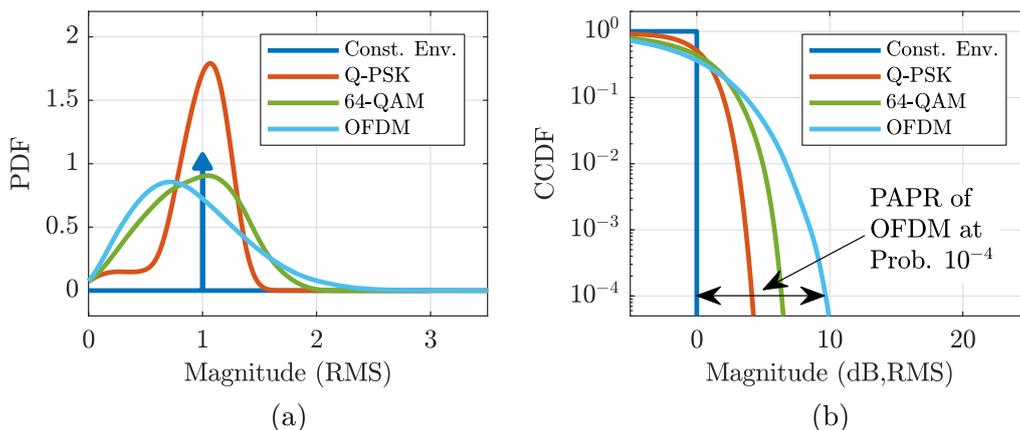


Figure 2.8: Magnitude distributions of single-carrier and multi-carrier signals.

A widely used metric to quantify the sensitivity of a signal to nonlinear distortion is the peak to average power ratio (PAPR). For a complex random process $X(t)$, the PAPR in decibel (dB) can be computed with respect to a peak probability p by the formula

$$\text{PAPR}(p) = 10 \log_{10} \left(\frac{(\text{CCDF}^{-1}(p))^2}{\mathbb{E}\{|X(t)|^2\}} \right), \quad (2.27)$$

where $\text{CCDF}^{-1}(\cdot)$ is the inverse magnitude CCDF and $\mathbb{E}\{\cdot\}$ is the expectation operator. If the magnitude CCDF is given like in Figure 2.8 (b), the PAPR can be read off directly from the plot, as shown for the OFDM signal, resulting in $\text{PAPR}(10^{-4}) \approx 10\text{dB}$.

Beside of the stochastic definition in (2.27), there is also a deterministic definition of the PAPR. For a signal $x(t)$, defined for $0 \leq t \leq T$, the deterministic PAPR is given by

$$\text{PAPR} = 10 \log_{10} \left(\frac{\max |x(t)|^2}{\text{mean } |x(t)|^2} \right). \quad (2.28)$$

If $x(t)$ is a realization of an ergodic random process $X(t)$, the deterministic PAPR of $x(t)$ for $T \rightarrow \infty$ is equal to the stochastic PAPR of $X(t)$ for $p \rightarrow 0$. However, for finite T , the deterministic PAPR strongly depends on the specific signal realization. Additionally, since (2.28) only captures the magnitude of the highest peak, the deterministic PAPR is only weakly correlated with the sensitivity to nonlinear distortion [20, 21]. To address the shortcomings of (2.28), several alternative metrics have been proposed, like the cubic metric [1, 61], the excess power [20], the variance of instantaneous power [21], and the amplifier-oriented metrics [198]. However, due to the conceptual and computational simplicity of (2.28), most methods aiming to decrease a signals sensitivity to nonlinear distortion are still designed with the objective to decrease its deterministic PAPR [60].

2.3.2 Linearity Quantification

To quantify the linearity of an RF-PA, we have to send an evaluation signal through it and quantify the nonlinear distortion within the output signal. Since nonlinear distortion strongly depends on the signal type, every linearity metric must be specified with respect to the signal type that is used for its evaluation. For quantifying the nonlinear distortion within a signal, there are several metrics.

The normalized mean square error (NMSE) is defined by

$$\text{NMSE} = 10 \log_{10} \left(\frac{\text{mean } |y_s(t) - x(t)|^2}{\text{mean } |x(t)|^2} \right), \quad (2.29)$$

where $x(t)$ is the input signal of the device under test (DUT) and $y_s(t)$ is the synchronized output signal. Synchronization is required, since the DUT typically causes a time delay and a complex gain between the input signal $x(t)$ and the output signal $y(t)$ of the DUT, which should not affect the linearity metric. For MSE optimal synchronization of the time delay and the complex gain, the methods in Appendix E can be used.

A possible disadvantage of the NMSE is that it does not discriminate between the in-band and the out-of-band error. Such a discrimination can be obtained by using the error vector magnitude (EVM) for quantifying the inband error and the adjacent channel power ratio (ACPR) for quantifying the out-of-band error. The EVM is defined by

$$\text{EVM} = 10 \log_{10} \left(\frac{\text{mean } |s_{\text{rx}}[n, k] - s_{\text{tx}}[n, k]|^2}{\text{mean } |s_{\text{tx}}[n, k]|^2} \right), \quad (2.30)$$

where $s_{\text{tx}}[n, k]$ are the transmitted symbols at symbol index n and subcarrier index k and $s_{\text{rx}}[n, k]$ are the received symbols at the respective indices after linear equalization. Since EVM operates on the symbol level, it measures only the inband error and since it includes a linear equalizer, it is insensitive to linear, frequency-dependent error.

An illustrative example of the difference between NMSE and EVM is shown in Figure 2.9. Figure 2.9 (a) is based on the same setup as Figure 2.7 (e) (f), which is an OFDM signal, sent through a memoryless class AB nonlinearity. Additionally to the PSDs of the input signal $x(t)$ and the output signal $y(t)$, the plot shows the PSD of the error signal $e_y(t) = y_s(t) - x(t)$ and the EVM at each of the 1800 subcarriers. With the memoryless nonlinearity, the PSD of the error signal is equal to the EVM in the in-band region, and it is equal to the spectral regrowth in the out-of-band region. Figure 2.9 (b) shows the results of a similar simulation with a linear FIR filter⁴ added in front of the class AB nonlinearity. In this case, the PSD of the error signal $e_y(t)$ is not equal to the EVM any more. The difference between the error signal and the EVM is caused by the linear frequency-dependent error, which is contained in the error signal $e_y(t)$, but removed by the equalizer within the EVM analyzer.

⁴The filter had an impulse response of $h[n] = (1, 0.9, \dots, 0.1)^T$ at a sampling rate of 200 MHz.

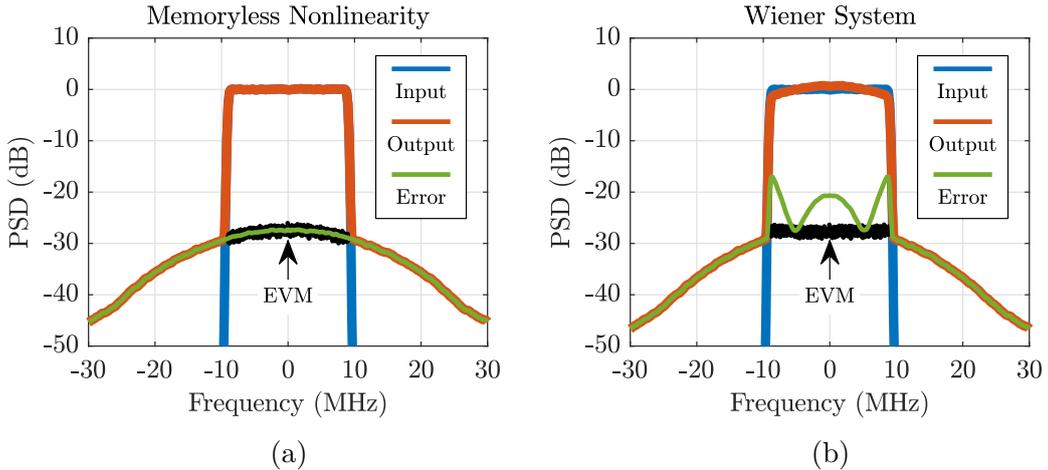


Figure 2.9: Linearity quantification by measuring frequency-dependent error.

A metric for the out-of-band error, is the adjacent channel power ratio (ACPR)

$$\text{ACPR} = 10 \log_{10} \left(\frac{\max \left(P_y(f_0 - B, B), P_y(f_0 + B, B) \right)}{P_y(f_0, B)} \right), \quad (2.31)$$

where f_0 is the center frequency of the transmission band, B is the bandwidth of the transmission band, and $P_y(f_0, B)$ is the frequency-dependent output power given by

$$P_y(f_0, B) = \int_{f_0 - \frac{B}{2}}^{f_0 + \frac{B}{2}} S_y(f) df, \quad (2.32)$$

where $S_y(f)$ is the PSD of the output signal $y(t)$. The ACPR defined in (2.31) uses the adjacent band with the higher distortion power to quantify the linearity. Alternatively, one can specify separate ACPR values for the lower and the upper adjacent band, or similar measures for other frequency bands. Some communication standards also specify a spectral mask, either absolute in mW/Hz, or relative to the inband PSD.

Measuring ACPR is simpler than measuring EVM, because ACPR only requires power measurements, whereas EVM requires a wideband receiver. Alternative methods to estimate the EVM with lower complexity, higher accuracy, or higher flexibility have been investigated in [94]. These alternatives include a variant of the noise power ratio (NPR), called the error power ratio (EPR) [95, 97], the swept error power ratio (SWEEP) [94], the subtraction of linearly correlated components EVM (SLIC-EVM) [96], and a method for excluding IQ mismatch in EVM measurements of multicarrier signals [94].

2.3.3 Efficiency Quantification

To quantify the efficiency of an RF-PA, we have to send an evaluation signal through it, measure the relevant power metrics and compute the efficiency as a ratio of the power metrics. Similar to linearity quantification, the result depends on the evaluation signal. If we use an unmodulated carrier, the result is called the instantaneous efficiency, which is typically represented as a function of the carrier amplitude, like we have seen in Figure 2.6 (e) (f). If we use a modulated carrier, we get the practically more relevant average efficiency, which can be represented as a function of the input backoff.

There are three ways how the power metrics of an RF-PA can be combined to an efficiency metric. Above, we have already used the drain efficiency, which is defined by

$$\eta_{\text{drain}} = \frac{P_{\text{RF}}}{P_{\text{DC}}}, \quad (2.33)$$

where P_{RF} is the RF output power and P_{DC} is the DC input power. A possibly relevant power metric that is not included in the drain efficiency is the RF input power P_{in} . If we include P_{in} in the numerator, we get the power added efficiency (PAE), defined by

$$\eta_{\text{PAE}} = \frac{P_{\text{RF}} - P_{\text{in}}}{P_{\text{DC}}}. \quad (2.34)$$

If we include P_{in} in the denominator, we get the overall efficiency, defined by

$$\eta_{\text{overall}} = \frac{P_{\text{RF}}}{P_{\text{DC}} + P_{\text{in}}}. \quad (2.35)$$

For typical RF-PAs which use a common-source topology, the overall efficiency is the physically justified version, because the RF input power is dissipated at the gate of the transistor [152, Part 4]. Nevertheless, the power added efficiency is also widely used. If the power gain $G_{\text{pwr}} = P_{\text{RF}}/P_{\text{in}}$ is high, the efficiency metrics defined above give approximately the same result. For evaluating the average efficiency of the joint linearity-efficiency model presented above, we have to use the average drain efficiency defined in (2.33), because the model does not include the power gain of the RF-PA. Including the power gain requires additional assumptions on the transistor technology.

For model-based evaluation of the average efficiency, we have to compute the expected values of the numerator and the denominator. For the drain efficiency, this gives

$$\eta_{\text{drain,avg}} = \frac{\mathbb{E}\{P_{\text{RF}}\}}{\mathbb{E}\{P_{\text{DC}}\}}. \quad (2.36)$$

where P_{RF} and P_{DC} are interpreted as random variables. The evaluation of (2.36) can be done in the ensemble domain or in the time domain. In the ensemble domain, we require power characteristics like in Figure 2.6 (a) (b) and the magnitude PDF of the input signal. In the time domain, we require the signal waveforms of P_{RF} and P_{DC} . The result of (2.36) is in general not equal to the expected value of (2.33). This is only the case for class A amplifiers, where P_{DC} is independent of the input signal [169].

2.3.4 Application Example

In the following, I demonstrate the linearity and efficiency quantification methods based on a simulation example using the class AB model of Figure 2.6 (b) (d) (f). Similar as before, we use the three signal types of Figure 2.7. For linearity quantification, we use the EVM and ACPR and for efficiency quantification, I use the average drain efficiency.

We evaluate the RF-PA model with and without ideal memoryless DPD. The AM-AM characteristics of the PA, the DPD, and the linearized system are shown in Figure 2.10 (a). The DPD implements the inverse PA characteristic within the invertible region of the PA. Beyond the invertible region, the DPD output is kept constant. Figure 2.10 (b) shows the gain instead of the output magnitude. This enables us to see the gain variation at low input magnitudes, which is perfectly compensated by the DPD.

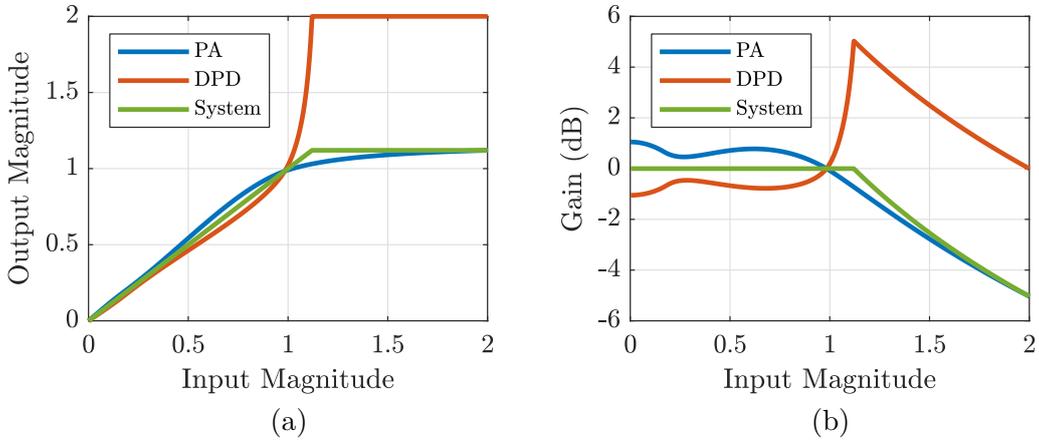
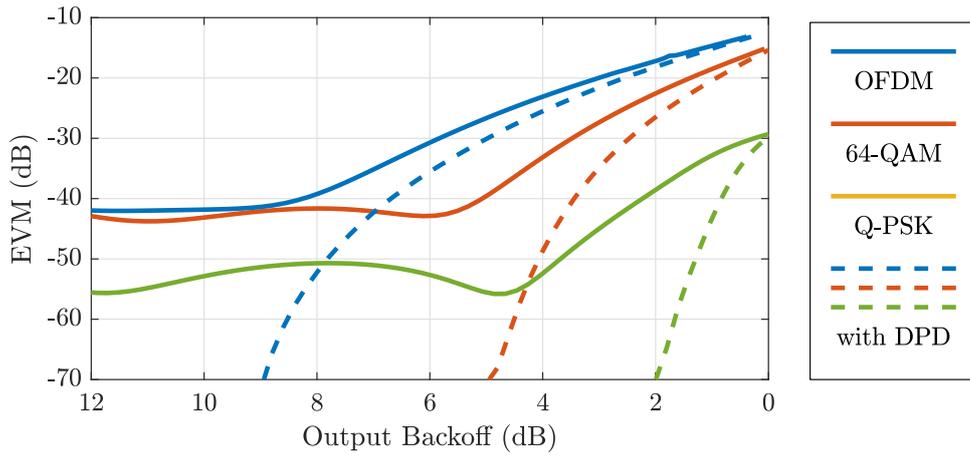
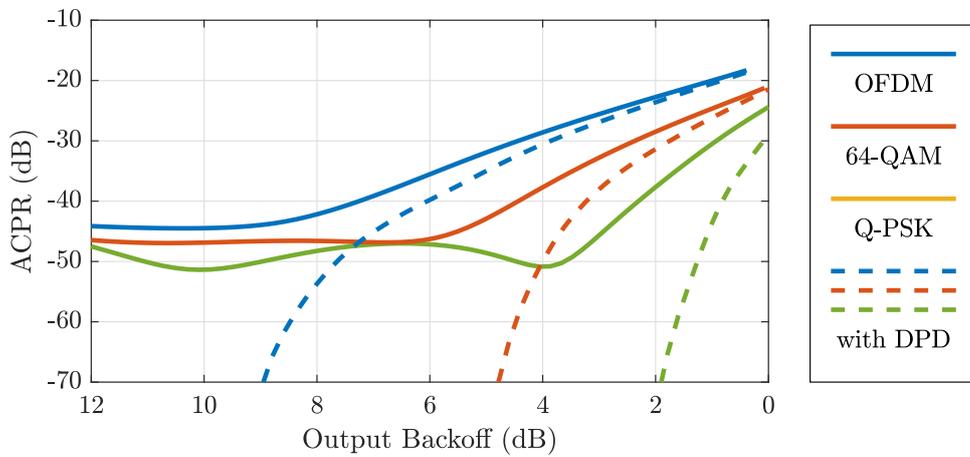


Figure 2.10: AM-AM characteristics of a memoryless DPD system.

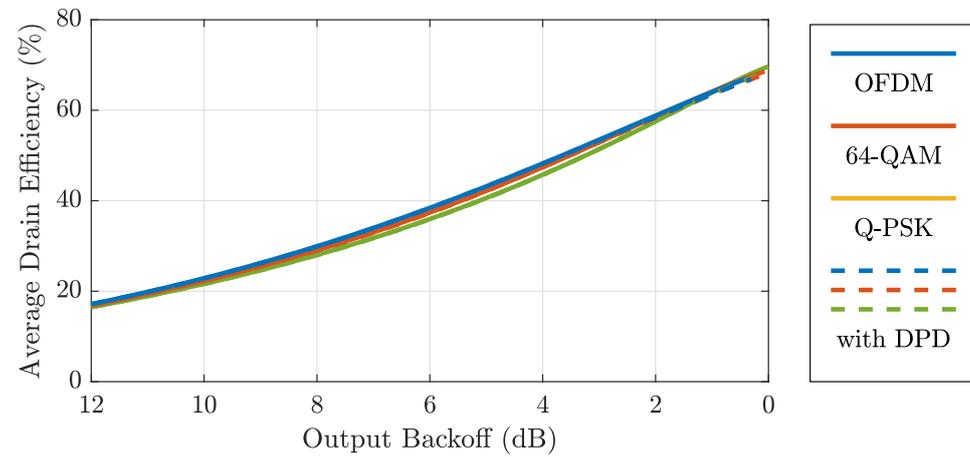
The simulation results are shown in Figure 2.11, which shows the EVM, the ACPR, and the average drain efficiency as functions of the output backoff (OBO). Here, OBO is defined as the reduction of the RMS output magnitude with respect to the full-scale magnitude of one. As can be expected from the previous results in Figure 2.7, we see in Figure 2.11 that the OFDM signal requires the largest backoff for a given level of linearity. At high backoff, the linearity is limited by the gain variation of the RF-PA. The average efficiency is mainly defined by the backoff and shows only slight variations with respect to different signal types or the use of a DPD. Based on the results in Figure 2.11, we can predict the required backoff and the average drain efficiency for a given signal and linearity requirement. For example, if we use an OFDM signal and require an EVM and ACPR better than -50 dB, we have to use a DPD, operated at an output backoff of at least 8 dB, resulting in a maximum average drain efficiency of approximately 30 %.



(a)



(b)



(c)

Figure 2.11: EVM, ACPR, and average drain efficiency over output backoff.

2.4 Highly Efficient RF Power Amplifiers

The design of highly efficient RF power amplifiers has been the subject of intensive research since the early days of radio transmission [38]. A concise overview of this subject is given in papers by Frederick Raab et al. [185, 186] and Earl McCune [152]. An in-depth treatment of the subject is given in the excellent book by Steve Cripps [53] and its extension [52]. Books with a focus on switched and harmonic tuned RF-PAs have been written by Andrei Grebennikov et al. [107] and Paolo Colantonio et al. [46].

In the following, I present an overview of methods for implementing highly efficient RF-PAs. The aim is to illustrate how circuit design, system design, and digital signal processing can be combined to push the boundaries of the linearity-efficiency trade-off.

2.4.1 Classes of Operation

Beside of the classification based on the conduction angle, as listed on page 12, there are other classes of operation which offer improved efficiency, often at the cost of linearity.

Class D In this class, the transistor is used as a switch instead of a current source. This enables a theoretical efficiency of 100 %, since in the conducting state the voltage across the transistor is zero and in the non-conducting state the current through the transistor is zero. The practical efficiency is limited by the transition between the states, where both the voltage and the current are non-zero. Class D amplifiers can be implemented in a voltage mode or a current mode architecture [107, Ch. 2]. For audio applications, class D is commonly used [117], but for RF transmitters, class D is less common, since switching losses increase with frequency [118]. Fundamental limits of switching RF-PAs are discussed in [157]. When class D is used within RF transmitters, it is often called class S [152, Part 2].

Class E In this class, the transistor is used as a switch, similar to class D, but with a resonant matching network at the output that ensures zero voltage switching and zero derivative voltage switching [205]. By imposing these conditions on the voltage waveform, the transition time between the switching states becomes less critical than in class D. Furthermore, the parasitic drain-source capacitance can be utilized as part of the resonant output matching network. Class E was introduced by Nathan and Alan Sokal in 1975 [205], but has roots in earlier Russian literature [106]. An efficiency analysis of class E can be found in [179]. Improved design equations are given in [206]. There is also an inverse class E that implements zero current switching and zero derivative current switching [165].

Class F In this class, the efficiency of an RF-PA is improved by overdrive and harmonic tuning [53, Ch. 6]. The load network of a class F amplifier implements shorts at even harmonics and open circuits at odd harmonics [107, Ch. 3]. This shapes the voltage waveform towards a square-wave, which decreases the overlap of simultaneous high current and high voltage, and therefore improves the efficiency. There is also an inverse class F, where the load network implements open circuits at even harmonics and shorts at odd harmonics [107, Ch. 4]. This shapes the current waveform towards a square-wave, which also improves the efficiency. In [237], class F and inverse class F are compared, resulting in the conclusion that inverse class F can reach a higher efficiency. A theory that relates class C, class E, and class F is presented in [183]. Architectures for implementing class F are discussed in [99].

Class G In this class, the efficiency of a PA is improved by using several constant supply voltages that are switched according to the input signal [117]. The switching can be implemented either by a serial or a parallel architecture [143]. Class G was introduced for audio applications [85] but can also be used for RF envelope modulation [220, 241]. An efficiency analysis of class G can be found in [181]. A concept similar to Class G is average power tracking (APT). The difference between class G and APT is that at class G the supply switching is based on the instantaneous signal, whereas at APT it is based on the average power.

Class H In this class, the efficiency of a PA is improved by a variable supply voltage that is controlled according to the input signal [117]. Class H was introduced for audio applications [86], but a similar concept called envelope tracking (ET) is used for RF applications. The difference between class H and ET is that at class H the supply control is based on the instantaneous signal, whereas at ET it is based on the magnitude of the equivalent baseband signal.

Class J In this class, a linear and wideband RF-PA is implemented with significant reactive components in the fundamental and second harmonic load impedances, which produces a non-sinusoidal voltage waveform [53, Ch. 4]. By proper selection of the harmonic impedances, a continuous range of linear operation modes from class J* to class J can be realized that includes class AB/B as a special case and features the same output power and efficiency as class AB/B [56]. Class J can be interpreted as an extension of the design space of class AB/B amplifiers at high frequencies where a short circuit termination of the second harmonic impedance is hard to realize due to the parasitic drain-source capacitance of the transistor. Implementation examples of class J can be found in [8, 161, 240].

Class S This is a synonym for class D designed for radio frequencies. Implementation examples of class S can be found in [148, 171, 231]. Earlier literature uses the term class S for switched RF envelope modulators [124, 128, 185, 186].

2.4.2 Efficiency Enhancement

The methods outlined below can be used to improve the average efficiency of linear RF-PAs, or to make nonlinear RF-PAs useable with non-constant envelope signals.

Doherty

The Doherty architecture, named after its inventor William Doherty [71], uses two RF-PAs in parallel, called the main amplifier and the peaking amplifier. Depending on the input signal magnitude, the peaking amplifier actively load modulates the main amplifier. An efficiency analysis is given in [182]. Historical reviews and current trends can be found in [35, 107, 174]. Digitally enhanced Doherty RF-PAs are discussed in [59].

Outphasing

The outphasing architecture, invented by Henri Chireix [43], combines the outputs of two RF-PAs operated with constant envelope signals. Amplitude modulation is implemented by modulating the phase difference between the two signals. Later, outphasing was re-introduced as linear amplification by nonlinear components (LINC) [47]. The efficiency of outphasing is analyzed in [180] and overviews of such systems are given in [16, 24].

Dynamic Power Supply

A dynamic power supply can be applied in different forms, ranging from average power tracking (APT) and envelope tracking (ET) [12, 222, 223] to direct polar (DP) [156], which is a modern form of envelope elimination and restoration (EER), also known as the Kahn method [119]. The important difference between APT/ET and DP is that APT/ET improves the efficiency of a linear RF-PA, whereas DP utilizes an efficient but nonlinear RF-PA for amplifying non-constant envelope signals [154, 155] [152, Part 6].

Dynamic Load Modulation

Dynamic load modulation was proposed by Frederick Raab for implementing direct polar modulation with a class E power amplifier and an electronically tunable matching network [184]. Recently, it was also used for the efficiency enhancement of tuned RF-PAs [9, 167, 194]. The generation of the tuning signal and the DPD of dynamic load modulation transmitters are discussed in [36]. A recent book on this topic is [224].

Pulse Modulation

Methods like pulse width modulation (PWM) [82] or delta sigma modulation (DSM) [172] can be used to encode non-constant envelope signals for the amplification by a nonlinear RF-PA. Such methods can be applied on the envelope [72, 114] or on the modulated carrier [178, 219]. Out-of-band components after the RF-PA must be removed by a linear filter, which degrades the overall efficiency. This degradation is quantified by the coding efficiency [42, 115]. Pulse modulation is essential for all-digital RF transmitters [176].

Linearization

Linearization is often required for highly efficient RF-PAs. Basically there are three types of linearization: feedback, feedforward, and predistortion [185]. The methods of feedback and feedforward linearization were invented by Harold Black in the 1920s [27] and are discussed in [124]. Today, predistortion is the most common method for the linearization of RF-PAs [121]. For very wideband systems, analog predistorsers are used [120], whereas for cellular systems, DPD dominates [109]. Recent books on DPD and digitally enhanced wireless transmitters are [101, 144, 238].

2.5 Summary and Conclusion

Summary

In this chapter, I investigated the linearity-efficiency trade-off of RF-PAs based on a memoryless semi-physical model of their joint linearity-efficiency behavior. This model is derived from a generic RF-PA circuit using either a piecewise-linear transistor model, or a more realistic piecewise-cubic transistor model. By using the piecewise-cubic transistor model, a realistic joint linearity-efficiency model of a generic RF-PA was obtained.

Furthermore, I presented an overview of linearity and efficiency quantification methods, illustrated by an application example of the joint linearity-efficiency model, and an overview of operation modes and architectures for highly efficient RF-PAs.

Conclusion

Linearity and efficiency are two important, but often conflicting design requirements for RF-PAs. A main insight with respect to this linearity-efficiency trade-off is that highly efficient amplification requires circuit-level nonlinearity, but this does not necessarily restrict the overall system linearity. By using highly efficient RF-PA architectures in combination with DPD, it is possible to push the linearity-efficiency trade-off to regions that could not be reached without the help of digital signal processing.

3

Even-Order Terms in Polynomial Baseband Models

The large majority of RF-PA behavioral models and digital predistorters is based on polynomial baseband models, ranging from memoryless and quasi-memoryless complex baseband polynomials to the complex baseband Volterra series. Central questions with respect to these models are, which types of basis functionals are mathematically justified, and which subsets of basis functionals provide a good accuracy-complexity trade-off.

In this chapter, I investigate these questions with a focus on the often neglected even-order terms in the baseband model. The main contributions are the explicit relations between polynomial passband and baseband models, which were published in [77, 79], and the formulation and proof of the phase homogeneity requirement.

3.1 Introduction

Polynomials are widely used throughout science and engineering, especially for the tasks of interpolation and approximation. A formal basis for these tasks is provided by the mathematical discipline of approximation theory, which studies the approximation of functions by simpler functions [215]. With respect to RF-PA behavioral modeling, we can apply approximation theory to represent a given AM-AM characteristic by a finite set of input and output values and use polynomial interpolation in between these points. Furthermore, we can also identify an AM-AM characteristic from the measured input and output signals of an RF-PA by synchronizing them and fitting a memoryless complex baseband polynomial [245] to the synchronized signals¹. If we look at the AM-AM scattering of the synchronized signals, we will notice a certain amount of dispersion in the measured data [34]. This dispersion consists of random components, like thermal noise, on the one hand, and deterministic components, like distortion caused by memory effects, on the other hand. If we increase the signal bandwidth, the distortion caused by memory effects typically increases and the fitting error of a memoryless polynomial becomes larger. Because of this, the behavioral modeling and DPD of RF-PAs excited by wideband signals typically requires polynomial models with memory.

¹For synchronization, see Appendix E. For least-squares model fitting, see Appendix D.

Polynomial models with memory are typically based on the Volterra series. The Volterra series is a combination of a polynomial and a finite impulse response (FIR) filter. It is named after the Italian mathematician Vito Volterra, who developed the theory of analytical functionals in the 1880s [218]. Volterra’s theory was subsequently applied to nonlinear systems modeling by Norbert Wiener, who published his results in 1958 [233].

The theories of Volterra and Wiener were presented and further developed in the books by Martin Schetzen [196] and Wilson Rugh [191]. A more recent book on this topic was written by V. John Mathews and Giovanni Sicuranza [151]. Overview articles can be found in [18, 44, 45, 195]. A proof that the Volterra series is a universal approximator for continuous nonlinear systems with fading memory was developed by Stephen Boyd and Leon Chua [31]. A historical overview is given in [196, App. B]. In [78], we showed that Horner’s method for polynomial evaluation can also be applied to the Volterra series, enabling the computation of output samples with one multiplication per coefficient.

All resources on the Volterra series discussed so far consider the modeling of nonlinear systems with real-valued input and output signals. At wireless communication systems, however, it is much more efficient to model the signals and systems in the complex baseband representation [15]. The complex baseband Volterra series was first derived by Sergio Benedetto et al. in 1979 [22]. In this work, the authors showed that

- only odd-order terms of the passband Volterra series produce output centered at the carrier frequency, which is relevant for the complex baseband model, and
- the p -th order terms in the baseband Volterra series are products of p delayed signal terms, where $\frac{p+1}{2}$ terms are not conjugated and $\frac{p-1}{2}$ terms are conjugated.

These results indicate that the complex baseband Volterra series requires a structure that is different from the real-valued Volterra series. The memoryless and quasi-memoryless baseband polynomials are special cases of the complex baseband Volterra series and therefore require this structure as well [245]. Consequently, many authors restricted polynomial baseband models to odd-order terms only [109, 112, 116, 131, 160, 232].

In 1999, Claudio Rey showed that by including second order terms in the magnitude and phase polynomials of a quasi-memoryless analog predistorter, its performance could be improved [188]. The memory polynomial [125] and generalized memory polynomial [163] models were also formulated without the restriction to odd-order terms.

The first systematic investigation of even-order terms in polynomial baseband models was conducted by Lei Ding and G. Tong Zhou [62, 63, 69]. In these publications, the authors showed by simulations and measurements that even-order terms in memory polynomial based RF-PA models and DPDs can improve the modeling and linearization accuracy. However, no formal proof was presented that justifies the use of even-order terms, therefore an apparent contradiction with the formal results from [22] remained.

Another investigation of even-order terms in polynomial baseband models was conducted by Eduardo Lima, Telmo Cunha, and José Pedro [57, 58, 142]. In these publications, the authors propose new variants of the baseband Volterra series that include even-order terms. They show by measurements that the proposed models can achieve higher accuracy for RF-PA behavioral modeling than conventional baseband Volterra models. The theoretical justifications for the new models are primarily based on analogy with memoryless models and physical intuition and do not constitute a formal proof.

In [141], the same authors formulate symmetry requirements for complex baseband models and apply them to an RF-PA model based on neural networks. The proposed requirements are inspired by the work of Nelson Blachman [25, 26], but no formal proof is presented that extends these requirements to models with memory.

The work of Nelson Blachman is also the basis of an investigation by Jacques Sombrin, who studied polynomial models that are non-analytic at the origin [207, 208]. In these works, the author shows that the proposed models can reproduce the measured behavior of nonlinear passband devices that could not be explained by conventional polynomials². Furthermore, the author shows that the even or odd symmetry of the proposed basis functions is not necessarily equal to the even or odd nonlinear order of the respective basis function³. This is an important insight for a theoretical foundation of even-order terms in polynomial baseband models, but was not further investigated by the author.

Further investigations on even-order terms in polynomial baseband models can be found in [87, 137, 147, 162]. An intuitive discussion by Steve Cripps is given in [55].

In summary, up till now, many authors have investigated the effects of even-order terms in polynomial baseband models. The motivation for my own work on this topic is the lack of a formal proof that explicitly shows how even-order terms in polynomial baseband models can be derived from their corresponding passband models.

In this chapter, I present such formal proofs for the memoryless and quasi-memoryless polynomials and the Volterra series. The proofs for the models without memory are based on the Chebyshev transform, introduced by Nelson Blachman in 1971 [26]. The proof for the Volterra series is an extension of the work by Sergio Benedetto et al. from 1979 [22]. In addition to these passband-baseband transformation pairs, I formulate and prove the phase homogeneity requirement, which is a symmetry that applies to all complex baseband models of time-invariant passband systems. This symmetry extends the concept of even- and odd-symmetric functions to complex baseband models and provides a simple guideline for designing complex baseband models of RF-PAs.

²If a two-tone signal is sent through a polynomial, the power of the p -th order intermodulation products will increase with the power of the input signal. In the small signal region, this increase has a slope of p dB/dB. It is a common misconception that this must hold for actual RF-PAs too [54].

³Jacques Sombrin uses the mathematically more accurate term *degree* for the exponent of a monomial. We use the term *order* as a synonym for *degree*, since it is more common in the related literature. Further information on the usage and the meaning of *degree* and *order* can be found in [229]

3.2 Polynomial Models without Memory

In the first part of this chapter, I investigate polynomial models without memory. I start with an intuitive approach to the spectral analysis of memoryless nonlinear distortion. Afterwards, I introduce the Chebyshev transform [26], apply it to the analysis of memoryless models, extend the results to quasi-memoryless models, and formulate the phase homogeneity requirement. I conclude with a discussion of the derived results.

3.2.1 Spectral Analysis

According to the Weierstrass theorem [215, Ch. 6], a memoryless nonlinearity can be approximated within a finite interval with arbitrary accuracy by a polynomial. Therefore we can analyze the spectral properties of memoryless nonlinear distortion by studying the spectral properties of a passband signal sent through a polynomial, given by

$$y(t) = \sum_{p=1}^P \alpha_p x^p(t), \quad x(t) = \text{Re} \{ \tilde{x}(t)e^{j\omega t} \} = \frac{1}{2} (\tilde{x}(t)e^{j\omega t} + \tilde{x}^*(t)e^{-j\omega t}), \quad (3.1)$$

where α_p are real coefficients⁴, $x(t)$ is a real passband signal, $\tilde{x}(t)$ is its equivalent complex baseband signal⁵, and $y(t)$ is the real output signal. The spectrum of $y(t)$ is a linear combination of the spectra of $x^p(t)$. Since multiplication in the time domain is convolution in the frequency domain [170, Ch. 2.5], the spectrum of $x^p(t)$ can be obtained by p repeated convolutions of the spectrum of $x(t)$. An example is shown in Figure 3.1.

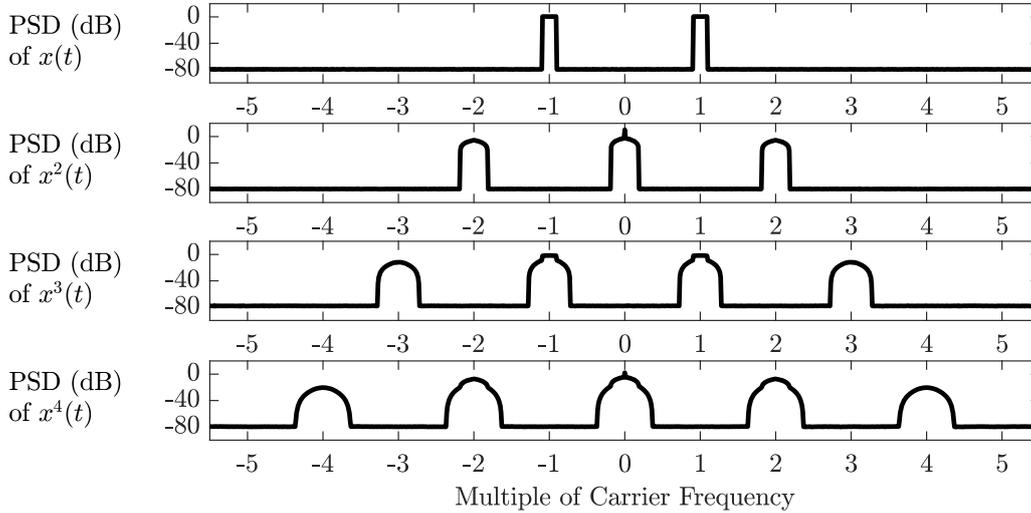


Figure 3.1: PSDs produced by monomial basis functions.

⁴For mathematical simplicity, we exclude the constant DC offset, modeled by the coefficient α_0 .

⁵The relation between passband and equivalent baseband signals is discussed in Appendix B

Figure 3.1 shows that the PSD of a nonlinear distorted passband signal consists of spectral components at integer multiples of the carrier frequency, each having a larger bandwidth than the original input signal. Typically, only the spectral components close to the carrier frequency are relevant, since the other components are sufficiently suppressed by the output matching network or the antenna [173]. Figure 3.1 also shows that only the odd-order terms of the passband polynomial create output close to the carrier frequency. If (3.1) is further evaluated by using the binomial theorem, an equivalent baseband model can be derived, which only includes odd-order terms [23, Example 2.18].

In the following, I present an alternative derivation of this equivalent baseband model based on the Chebyshev transform. An advantage of the Chebyshev transform is that it is not limited to polynomial models. More importantly, however, we will see that the inverse Chebyshev transform allows us to derive the passband model that corresponds to a memoryless baseband polynomial with even-order terms.

3.2.2 The Chebyshev Transform

Let's consider an alternative formulation of (3.1) given by

$$y(t) = f(x(t)), \quad x(t) = a(t) \cos(\omega t + \varphi(t)), \quad (3.2)$$

where $f(\cdot)$ is an arbitrary real-valued function. Nelson Blachman showed in [25] that the spectral properties of $y(t)$ can be analyzed by representing $y(t)$ as a Fourier series with time-varying coefficients. For this purpose, we define $\theta(t) = \omega t + \varphi(t)$ and note that the output signal can be represented as $y(t) = F(a(t), \theta(t))$, where $F(a, \theta)$ is the memoryless two-dimensional function $F(a, \theta) = f(a \cos(\theta))$. Since $F(a, \theta)$ is periodic in θ , it can be expanded into a Fourier series along this dimension⁶, resulting in⁷

$$f(a \cos(\theta)) = \frac{1}{2} \hat{f}_0(a) + \sum_{k=1}^{\infty} \hat{f}_k(a) \cos(k\theta), \quad (3.3)$$

where the coefficients are functions of a , given by

$$\hat{f}_k(a) = \frac{2}{\pi} \int_0^{\pi} f(a \cos(\theta)) \cos(k\theta) d\theta. \quad (3.4)$$

According to (3.3), $\hat{f}_k(\cdot)$ describes the mapping of the input amplitude a to the amplitude modulation of the k -th harmonic of the output signal $y(t)$. In the following, we call $\hat{f}_k(\cdot)$ the k -th order envelope nonlinearity. From (3.3), we see that a memoryless nonlinear system only produces AM-AM nonlinearity. For modeling AM-PM nonlinearity, we have to extend the model to the quasi-memoryless model, which is discussed in Section 3.2.6.

⁶In [82], we used this method for the analytical description of carrier-based pulse width modulation.

⁷(3.3) and (3.4) exploit the fact that $f(a \cos(\theta))$ is even-symmetric in θ for arbitrary functions $f(\cdot)$.

The term Chebyshev transform was proposed by Nelson Blachman in [26], where he showed that the Fourier series in θ , given by (3.3) and (3.4) can also be represented by a Chebyshev series⁸ in x . Substituting $\theta = \arccos(x/a)$ in (3.3) and (3.4), we obtain

$$f(x) = \frac{1}{2}\hat{f}_0(a) + \sum_{k=1}^{\infty} \hat{f}_k(a) T_k(x/a), \quad (3.5)$$

where the coefficients are still functions of a , given by

$$\hat{f}_k(a) = \frac{2}{\pi} \int_{-a}^{+a} \frac{f(x) T_k(x/a)}{\sqrt{a^2 - x^2}} dx, \quad (3.6)$$

and $T_k(\cdot)$ are the Chebyshev polynomials of the first kind, defined by

$$T_k(x) = \cos(k \arccos(x)) \quad \text{for } |x| \leq 1. \quad (3.7)$$

The mapping $f(x) \mapsto \hat{f}_k(a)$ described by (3.6), or equivalently by (3.4), is an integral transform. Due to its explicit representation in (3.6), Nelson Blachman called this transform the k -th order Chebyshev transform [26]. Studying the properties of the Chebyshev transform is of practical relevance, since it explicitly relates a memoryless passband nonlinearity with the corresponding baseband nonlinearity. In [26], Nelson Blachman presented a table of k -th order Chebyshev transforms of elementary functions and derived analytic formulas for the inverse k -th order Chebyshev transforms. In the following, we make use of the inverse first-order Chebyshev transform, given by

$$f(x) = \frac{1}{2} \int_0^{\pi/2} \left(\hat{f}_1(x \cos(\theta)) + \hat{f}'_1(x \cos(\theta)) x \cos(\theta) \right) d\theta + f_{\text{even}}(x), \quad (3.8)$$

where $\hat{f}'_1(\cdot)$ is the first derivative of $\hat{f}_1(\cdot)$ and $f_{\text{even}}(\cdot)$ is any even-symmetric function. The non-uniqueness of the inverse first-order Chebyshev transform indicates that $\hat{f}_1(\cdot)$ is not sufficient to fully describe $f(\cdot)$. Indeed, based on (3.6) and the symmetry of the Chebyshev polynomials, given by $T_k(-x) = (-1)^k T_k(x)$, it can be shown that the even-order envelope nonlinearities $\hat{f}_k(\cdot), k \in \mathbb{N}_{\text{even}}$ are only affected by the even-symmetric part of $f(\cdot)$, the odd-order envelope nonlinearities $\hat{f}_k(\cdot), k \in \mathbb{N}_{\text{odd}}$ are only affected by the odd-symmetric part of $f(\cdot)$, and in general, the k -th order envelope nonlinearities must have the same symmetry as the k -th order Chebyshev polynomials, given by

$$\hat{f}_k(-a) = (-1)^k \hat{f}_k(a). \quad (3.9)$$

The symmetry in (3.9) is significant for our discussion on even-order terms in polynomial baseband models. From (3.9) it follows that the even-order envelope nonlinearities must be even-symmetric functions and the odd-order ones must be odd-symmetric functions. The first-order envelope nonlinearity must therefore be an odd-symmetric function.

⁸The Chebyshev series is discussed in [215, Ch. 3]

3.2.3 From Passband to Baseband

In the following, we use the Chebyshev transform to derive the equivalent baseband model $\tilde{f}(\cdot) : \tilde{x}(t) \mapsto \tilde{y}(t)$ for the passband model $f(\cdot) : x(t) \mapsto y(t)$, given by

$$y(t) = f(x(t)), \quad \text{where} \quad f(x) = \sum_{p=1}^P \alpha_p f_p(x), \quad \text{with} \quad f_p(x) = x^p. \quad (3.10)$$

Since the Chebyshev transform in (3.4) is linear in $f(\cdot)$, we can transform the polynomial basis functions $f_p(\cdot)$ individually and recombine them with the coefficients α_p afterwards. By inserting the polynomial basis functions $f_p(x) = x^p$ into (3.4), we obtain

$$\hat{f}_{p,1}(a) = \frac{2}{\pi} \int_0^\pi a^p \cos^p(\theta) \cos(\theta) d\theta \quad (3.11)$$

$$= \left(\frac{2}{\pi} \int_0^\pi \cos^{p+1}(\theta) d\theta \right) a^p. \quad (3.12)$$

Evaluating the integral in (3.12), as discussed in Appendix H, we get

$$\hat{f}_{p,1}(a) = \begin{cases} \lambda_p a^p & p \text{ is odd,} \\ 0 & p \text{ is even,} \end{cases} \quad \text{where} \quad \lambda_p = \frac{1}{2^{p-1}} \binom{p}{\frac{p-1}{2}}. \quad (3.13)$$

By recombining the basis functions $\hat{f}_{p,1}(\cdot)$ with the coefficients α_p , we get

$$\hat{f}_1(a) = \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \alpha_p \lambda_p a^p. \quad (3.14)$$

According to (3.3), the passband output signal can be written as

$$y(t) = \frac{1}{2} \hat{f}_0(a(t)) + \sum_{k=1}^{\infty} \hat{f}_k(a(t)) \cos(k\omega t + k\varphi(t)). \quad (3.15)$$

The complex modulation of the k -th harmonic is therefore given by

$$\tilde{y}_k(t) = \hat{f}_k(a(t)) e^{jk\varphi(t)}, \quad (3.16)$$

where the special case $k = 1$ represents the equivalent baseband model $\tilde{f}(\cdot) : \tilde{x}(t) \mapsto \tilde{y}(t)$. By inserting (3.14) into (3.16), we get for $k = 1$ the equivalent baseband model

$$\tilde{y}(t) = \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \tilde{\alpha}_p \tilde{x}(t) |\tilde{x}(t)|^{p-1}, \quad (3.17)$$

where the baseband coefficients $\tilde{\alpha}_p = \lambda_p \alpha_p$ are defined for odd p only.

3.2.4 From Baseband to Passband

In the following, we investigate the effect of including even-order terms in (3.17) by removing the restriction to odd-order terms and transforming the resulting model back to passband. The baseband model with both odd- and even-order terms is given by

$$\tilde{y}(t) = \sum_{p=1}^P \tilde{\alpha}_p \tilde{x}(t) |\tilde{x}(t)|^{p-1}. \quad (3.18)$$

By pulling out the phase term from (3.18) and removing it, we obtain the polynomial model of the first-order envelope nonlinearity, given by

$$\hat{f}_1(a) = \sum_{p=1}^P \tilde{\alpha}_p a |a|^{p-1}. \quad (3.19)$$

To transform $\hat{f}_1(a)$ to the passband nonlinearity $f(x)$, we use the inverse first-order Chebyshev transform defined by (3.8). Since (3.8) is linear in $\hat{f}_1(\cdot)$, we can transform the basis functions of (3.19) individually and recombine them with the coefficients $\tilde{\alpha}_p = \lambda_p \alpha_p$ afterwards. The basis functions and their first derivatives are given by

$$\hat{f}_{p,1}(a) = a |a|^{p-1}, \quad (3.20)$$

$$\hat{f}'_{p,1}(a) = p |a|^{p-1}. \quad (3.21)$$

By inserting (3.20) and (3.21) into (3.8), we obtain

$$f_p(x) = \frac{1}{2} \int_0^{\pi/2} x \cos(\theta) |x|^{p-1} |\cos(\theta)|^{p-1} \\ + p |x|^{p-1} |\cos(\theta)|^{p-1} x \cos(\theta) d\theta + f_{p,\text{even}}(x). \quad (3.22)$$

By noting that $|\cos(\theta)| = \cos(\theta)$ within $0 \leq \theta \leq \pi/2$, we can simplify (3.22) to

$$f_p(x) = \left(\frac{p+1}{2} \int_0^{\pi/2} \cos^p(\theta) d\theta \right) x |x|^{p-1} + f_{p,\text{even}}(x). \quad (3.23)$$

Evaluating the integral in (3.23), as discussed in Appendix H, we get

$$f_p(x) = \frac{1}{\lambda_p} x |x|^{p-1} + f_{p,\text{even}}(x). \quad (3.24)$$

where the scaling factor λ_p is a generalization of λ_p in (3.13), given by

$$\lambda_p = \frac{2}{\sqrt{\pi}} \frac{\Gamma\left(\frac{p+2}{2}\right)}{\Gamma\left(\frac{p+3}{2}\right)}, \quad (3.25)$$

where $\Gamma(\cdot)$ is the gamma function, defined in (H.24) in Appendix H.

After recombining the passband basis functions in (3.24) with the coefficients $\tilde{\alpha}_p = \lambda_p \alpha_p$, and setting the arbitrary even-symmetric functions $f_{p,\text{even}}(\cdot)$ to zero, we obtain

$$y(t) = \sum_{p=1}^P \alpha_p x(t) |x(t)|^{p-1}. \quad (3.26)$$

By comparing (3.26) with the conventional passband polynomial in (3.10), we notice that the passband models differ with respect to their basis functions. The model in (3.10) uses basis functions of the form x^p , whereas the model in (3.26) uses basis functions of the form $x|x|^{p-1}$. For real-valued x , we have $x|x|^{p-1} = x^p$ for odd p . Furthermore, we have $x|x|^{p-1} = \text{sign}(x)x^p$ for even p . This means that (3.10) and (3.26) have identical odd-order terms, but the even-order terms of (3.26) are odd-symmetric functions.

3.2.5 Magnitude Power Functions

The result of the previous section motivates us to reconsider the choice of basis functions in the passband model. The set of conventional polynomial basis functions is given by

$$\mathcal{B}_{\text{poly}} = \{x^p \mid p \in \mathbb{N}_0\}, \quad (3.27)$$

where \mathbb{N}_0 is the set of non-negative integers. Restricting p to non-negative integers is reasonable, since it ensures that the basis functions are smooth⁹ and real inputs produce real outputs. For negative p , the basis functions x^p are not smooth, since there is a discontinuity at the origin. For non-integer p , there is also a discontinuity at the origin, and furthermore, real negative inputs produce imaginary outputs.

In the following, we investigate a new set of basis functions that is also smooth and produces real output for real input¹⁰. We define this new set by

$$\mathcal{B}_{\text{mag,pwr}} = \{\text{sign}^q(x) |x|^p \mid q \in \{1, 2\}, p \in \mathbb{R}_{\geq 0}\}, \quad (3.28)$$

where $\mathbb{R}_{\geq 0}$ is the set of non-negative real numbers. We call this new set of basis functions *magnitude power functions*. They are parameterized by a binary symmetry parameter q and a non-negative real order parameter p . The conventional polynomial basis functions are obtained from (3.28) by setting $p \in \mathbb{N}_0$ and using an integer of the same parity for q . Based on q , we can separate (3.28) into two subsets, given by

$$\mathcal{B}_{\text{odd}} = \{\text{sign}(x) |x|^p \mid p \in \mathbb{R}_{\geq 0}\}, \quad \mathcal{B}_{\text{even}} = \{|x|^p \mid p \in \mathbb{R}_{\geq 0}\}. \quad (3.29)$$

The subsets defined by (3.29) separate the magnitude power functions into even-symmetric and odd-symmetric basis functions, which can clearly be seen from Figure 3.2.

⁹A function $f(x)$ is smooth, if it belongs to the differentiability class C^∞ . The differentiability class C^m includes all functions for which the derivatives $f'(x), f''(x), \dots, f^{(n)}(x)$ exist and are continuous.

¹⁰The restriction to real input and output signals is reasonable for modeling real passband systems.

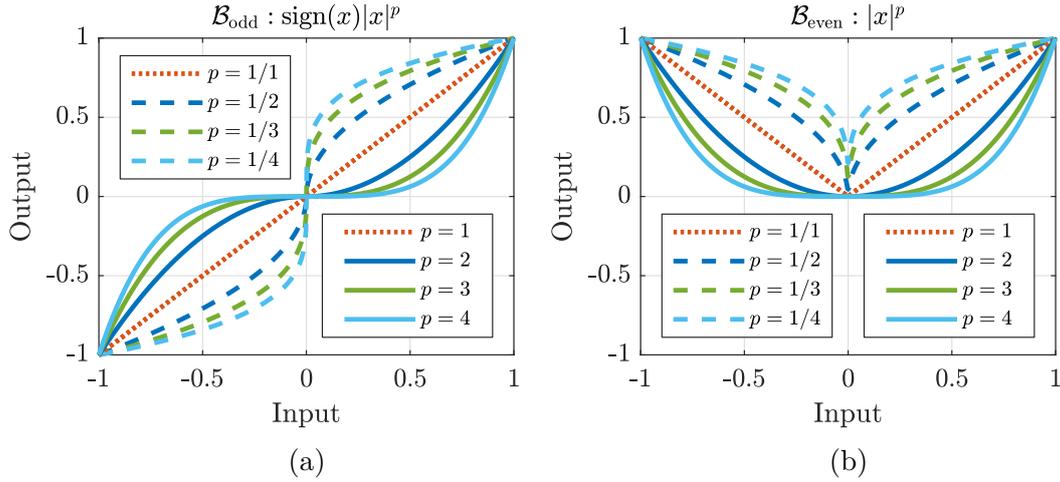


Figure 3.2: Odd-symmetric and even-symmetric magnitude power functions.

Differentiability of Polynomial Basis Functions and Magnitude Power Functions

To analyze the differentiability of the polynomial basis functions, we note that

$$\frac{d^n}{dx^n}(x^p) = \begin{cases} 0 & p \in \mathbb{N}_0 \text{ and } n > p, \\ (p)_n x^{p-n} & \text{otherwise,} \end{cases} \quad \text{where } (p)_n = \prod_{q=0}^{n-1} (p-q). \quad (3.30)$$

From (3.30), it follows that the n -th derivatives of the functions in (3.27) are either scaled versions of the functions in (3.27) (for $n \leq p$), or they are zero (for $n > p$). In both cases the derivatives exist and are continuous, so the functions in (3.27) are smooth.

To analyze the differentiability of the magnitude power functions in (3.28), we note that for $x > 0$, they are identical to x^p and for $x < 0$ they are identical to $\pm(-x)^p$. Possible discontinuities can therefore only occur at $x = 0$. If $p \in \mathbb{N}_0$ and $p + q \in \mathbb{N}_{\text{even}}$, the magnitude power functions are identical to the polynomial basis functions and therefore they are smooth. If $p \in \mathbb{N}_0$ and $p + q \in \mathbb{N}_{\text{odd}}$, the magnitude power functions have the form $\text{sign}(x)x^p$, where the sign change at the origin produces a discontinuity, if the function x^p or one of its derivatives is not equal to zero at the origin. By inspection of (3.30), we note that $f^{(n)}(0) \neq 0$ for $p = n$. Therefore, for $p \in \mathbb{N}_0$ and $p + q \in \mathbb{N}_{\text{odd}}$, the magnitude power functions are in the continuity class C^{p-1} . If we have $p \in \mathbb{R}_{\geq 0}$ and $p \notin \mathbb{N}_0$, there is a discontinuity at the origin in the n -th derivative of the magnitude power functions, if the exponent of x^{p-n} in (3.30) is negative. This is the case for the derivative of order $n = \lceil p \rceil$. Therefore, for $p \in \mathbb{R}_{\geq 0}$ and $p \notin \mathbb{N}_0$, the magnitude power functions are in the continuity class $C^{\lceil p \rceil - 1}$. We can summarize our analysis by

$$\text{sign}^q(x)|x|^p \in \begin{cases} C^\infty & p \in \mathbb{N}_0, p + q \in \mathbb{N}_{\text{even}}, \\ C^{\lceil p \rceil - 1} & \text{otherwise.} \end{cases} \quad (3.31)$$

Chebyshev Transform of One-Sided Magnitude Power Functions

A useful property of the magnitude power functions is that they are eigenfunctions¹¹ of the Chebyshev transform. We can show this by deriving the Chebyshev transform of the one-sided magnitude power functions¹², given by

$$f_p^{(+)}(x) = H(x) |x|^p, \quad \text{where} \quad H(x) = \begin{cases} 1 & x > 0, \\ 0 & x \leq 0. \end{cases} \quad (3.32)$$

The Chebyshev transform of the symmetric magnitude power functions in (3.28) immediately follows from the Chebyshev transform of (3.32) and symmetry considerations.

By inserting (3.32) into (3.4), we obtain

$$\hat{f}_{p,k}^{(+)}(a) = \frac{2}{\pi} \int_0^\pi H(a \cos(\theta)) |a \cos(\theta)|^p \cos(k\theta) d\theta. \quad (3.33)$$

For $a \geq 0$, we have

$$H(a \cos(\theta)) |a \cos(\theta)|^p = \begin{cases} a^p \cos^p(\theta) & 0 < \theta < \pi/2, \\ 0 & \pi/2 < \theta < \pi. \end{cases} \quad (3.34)$$

Therefore, we can simplify (3.33) to

$$\hat{f}_{p,k}^{(+)}(a) \Big|_{a>0} = \frac{2}{\pi} \int_0^{\pi/2} a^p \cos^p(x) \cos(k\theta) d\theta \quad (3.35)$$

$$= \left(\frac{2}{\pi} \int_0^{\pi/2} \cos^p(x) \cos(k\theta) d\theta \right) a^p. \quad (3.36)$$

The part of $\hat{f}_{p,k}^{(+)}(a)$ for $a < 0$ follows from the symmetry stated in (3.9).

Solving the integral in (3.36), as discussed in Appendix H, we get

$$\hat{f}_{p,k}^{(+)}(a) = \frac{1}{2} \lambda_{p,k} \text{sign}^k(a) |a|^p \quad (3.37)$$

where the scaling factor $\lambda_{p,k}$ is given by

$$\lambda_{p,k} = \frac{1}{2^{p-1}} \binom{p}{\frac{p-k}{2}} \quad \text{where} \quad \binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(x-y+1)\Gamma(y+1)}. \quad (3.38)$$

The formula in (3.38) generalizes λ_p in (3.13) and (3.25) to values of k other than one.

¹¹A function $f(\cdot)$ is an eigenfunction of an operator $\mathcal{F} : f(\cdot) \mapsto \hat{f}(\cdot)$, if it is invariant up to a scaling factor λ under the operator \mathcal{F} , i.e., we have $\hat{f}(\cdot) = \lambda f(\cdot)$. The factor λ is the corresponding eigenvalue.

¹²The Chebyshev transform of (3.32) was presented by Nelson Blachman in [26], but without derivation. The closely related Chebyshev transform of (3.29) was presented by Jacques Sombrin in [207].

Chebyshev Transform of Symmetric Magnitude Power Functions

The symmetric magnitude power functions given by

$$f_p^{(q)}(x) = \text{sign}^q(x) |x|^p \quad (3.39)$$

can be written as a linear combination of a one-sided magnitude power function $f_p^{(+)}(x)$ and a mirrored one-sided magnitude power function $f_p^{(+)}(-x)$ by combining them to

$$f_p^{(q)}(x) = f_p^{(+)}(x) + (-1)^q f_p^{(+)}(-x). \quad (3.40)$$

Since the Chebyshev transform is linear, also the Chebyshev-transformed symmetric magnitude power functions can be written as such a linear combination, given by

$$\hat{f}_{p,k}^{(q)}(a) = \hat{f}_{p,k}^{(+)}(a) + (-1)^q \hat{f}_{p,k}^{(+)}(-a). \quad (3.41)$$

The symmetry stated in (3.9), allows us to simplify (3.41) to

$$\hat{f}_{p,k}^{(q)}(a) = \hat{f}_{p,k}^{(+)}(a) + (-1)^{q+k} \hat{f}_{p,k}^{(+)}(a). \quad (3.42)$$

If the parity of q and k is identical, their sum is even, the term $(-1)^{q+k}$ simplifies to one, and the function $\hat{f}_{p,k}^{(q)}(a)$ simplifies to $2\hat{f}_{p,k}^{(+)}(a)$. On the other hand, if the parity of q and k is different, their sum is odd, the term $(-1)^{q+k}$ simplifies to minus one, and the function $\hat{f}_{p,k}^{(q)}(a)$ simplifies to zero. Therefore, we can write (3.42) as

$$\hat{f}_{p,k}^{(q)}(a) = \begin{cases} \lambda_{p,k} \text{sign}^q(a) |a|^p & q+k \in \mathbb{N}_{\text{even}}, \\ 0 & \text{otherwise,} \end{cases} \quad (3.43)$$

where $\lambda_{p,k}$ is defined in (3.38). By comparing (3.39) and (3.43), we clearly see that the magnitude power functions are eigenfunctions of the Chebyshev transform. The eigenvalues are $\lambda_{p,k}$ if the parity of q and k is identical, and they are zero otherwise.

As special cases of the transformation pair (3.39) and (3.43), we have the transformation pair for the odd-symmetric magnitude power functions, given by

$$f_p^{(1)}(x) = \text{sign}(x) |x|^p \quad \Leftrightarrow \quad \hat{f}_{p,k}^{(1)}(a) = \begin{cases} \lambda_{p,k} \text{sign}(a) |a|^p & k \in \mathbb{N}_{\text{odd}}, \\ 0 & \text{otherwise,} \end{cases} \quad (3.44)$$

and for the even-symmetric magnitude power functions, we have

$$f_p^{(2)}(x) = |x|^p \quad \Leftrightarrow \quad \hat{f}_{p,k}^{(2)}(a) = \begin{cases} \lambda_{p,k} |a|^p & k \in \mathbb{N}_{\text{even}}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.45)$$

Eigenvalue Function of the Chebyshev Transform

The scaling factor $\lambda_{p,k}$ defined in (3.38) is a two-argument function, which maps the order $p \in \mathbb{R}_{\geq 0}$ and the harmonic index $k \in \mathbb{Z}$ to a real-valued scalar. A plot of this function with a generalized harmonic index $k \in \mathbb{R}$ is shown in Figure 3.3.

Figure 3.3 (a) shows that for $p = 0$, the factor $\lambda_{p,k}$ follows a sinc function along k . In (H.53) of Appendix H, we verify that for $p = 0$, we have $\lambda_{0,k} = 2 \sin(\frac{k}{2} \pi) / (\frac{k}{2} \pi)$.

Figure 3.3 (b) shows the magnitude of $\lambda_{p,k}$ in a logarithmic scale. This plot clearly shows the zeros of $\lambda_{p,k}$, which occur at $k = \pm(p + 2n)$ with $n \in \mathbb{N}$.

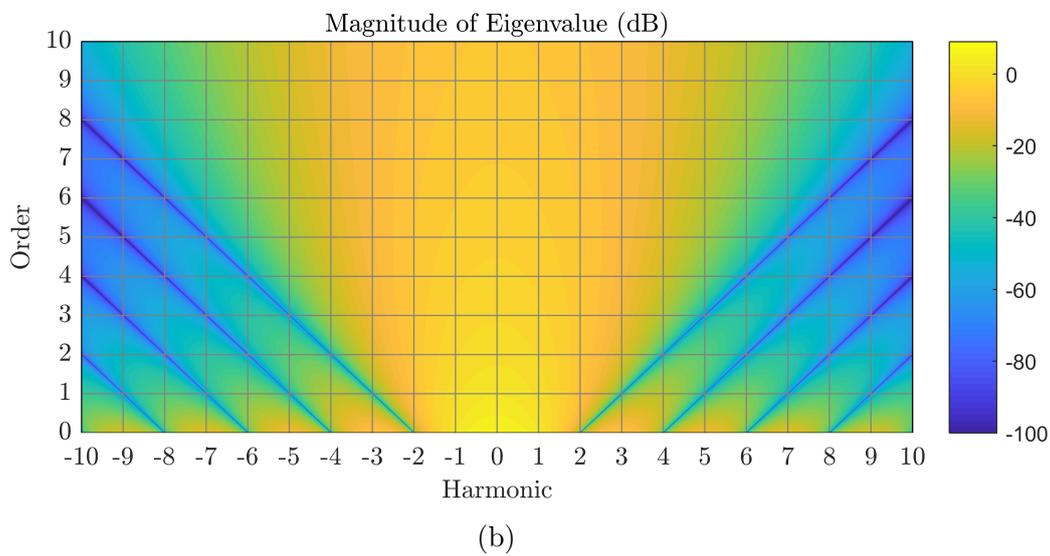
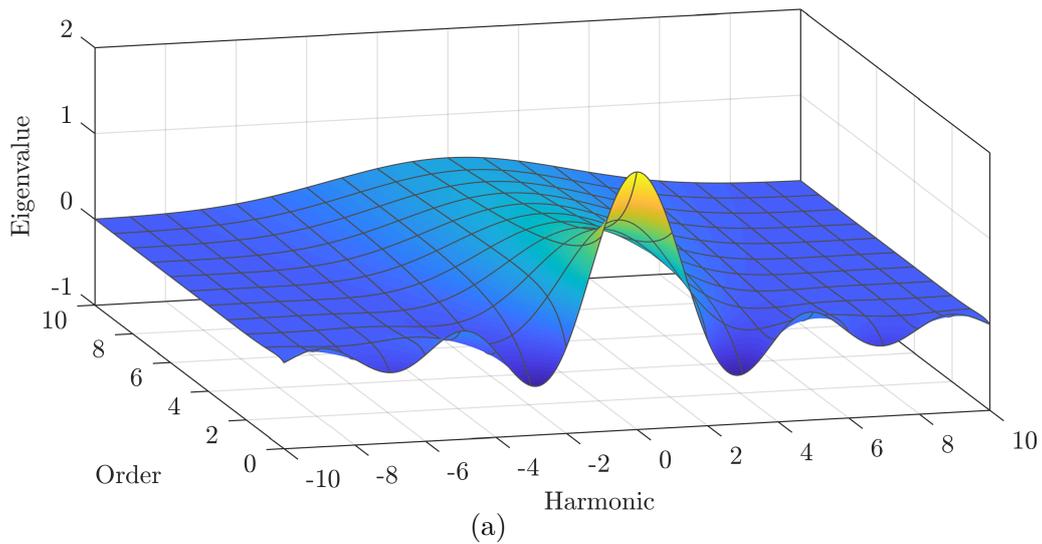
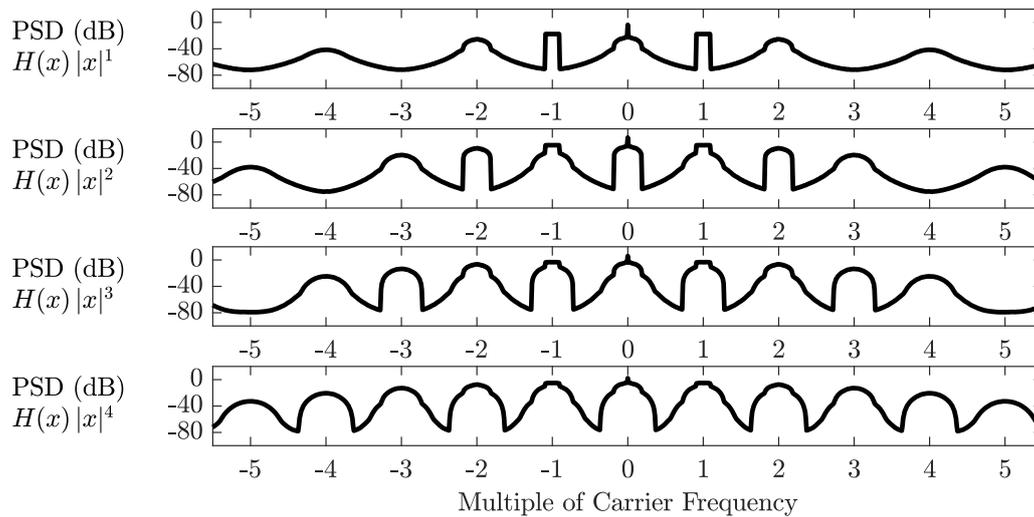


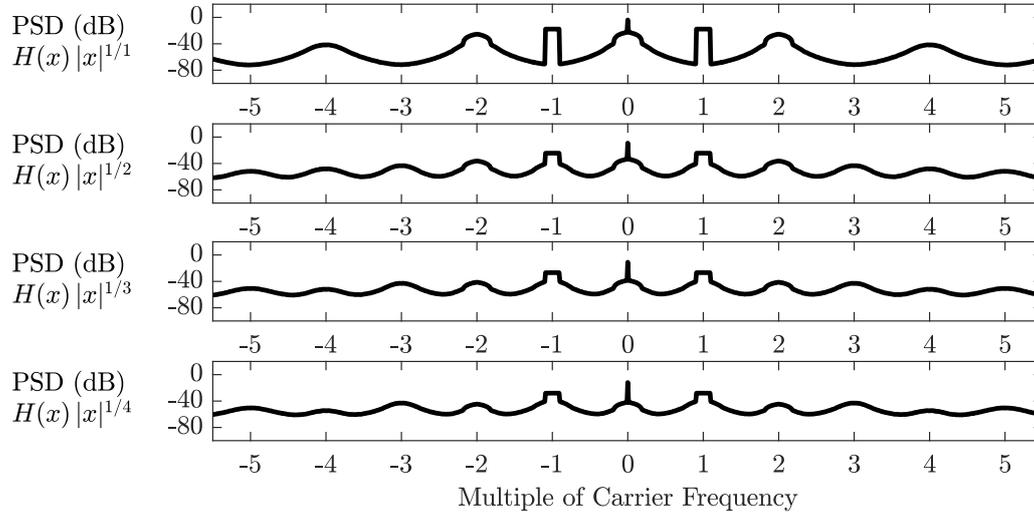
Figure 3.3: Eigenvalues of the Chebyshev transform over its order and harmonic index.

Spectral Analysis of One-Sided Magnitude Power Functions

To illustrate the spectral characteristics of magnitude power functions, I repeated the simulations of Figure 3.1 with this new type of basis function. The results for the one-sided magnitude power functions are shown in Figure 3.4. From these plots, we clearly see that the one-sided magnitude power functions produce output at all harmonics, except for $k = \pm(p + 2n)$, $n \in \mathbb{N}$, where the eigenvalue $\lambda_{p,k}$ is zero.



(a)



(b)

Figure 3.4: PSDs produced by one-sided magnitude power functions.

Spectral Analysis of Odd-Symmetric Magnitude Power Functions

The spectral characteristics of the odd-symmetric magnitude power functions are shown in Figure 3.5. If we compare these plots with Figure 3.4, we notice that the spectral output at odd harmonics is identical. The spectral output at even harmonics, however, is zero. For odd-integer orders, the odd-symmetric magnitude power functions reduce to polynomial basis functions and the spectra are therefore the same as in Figure 3.1.

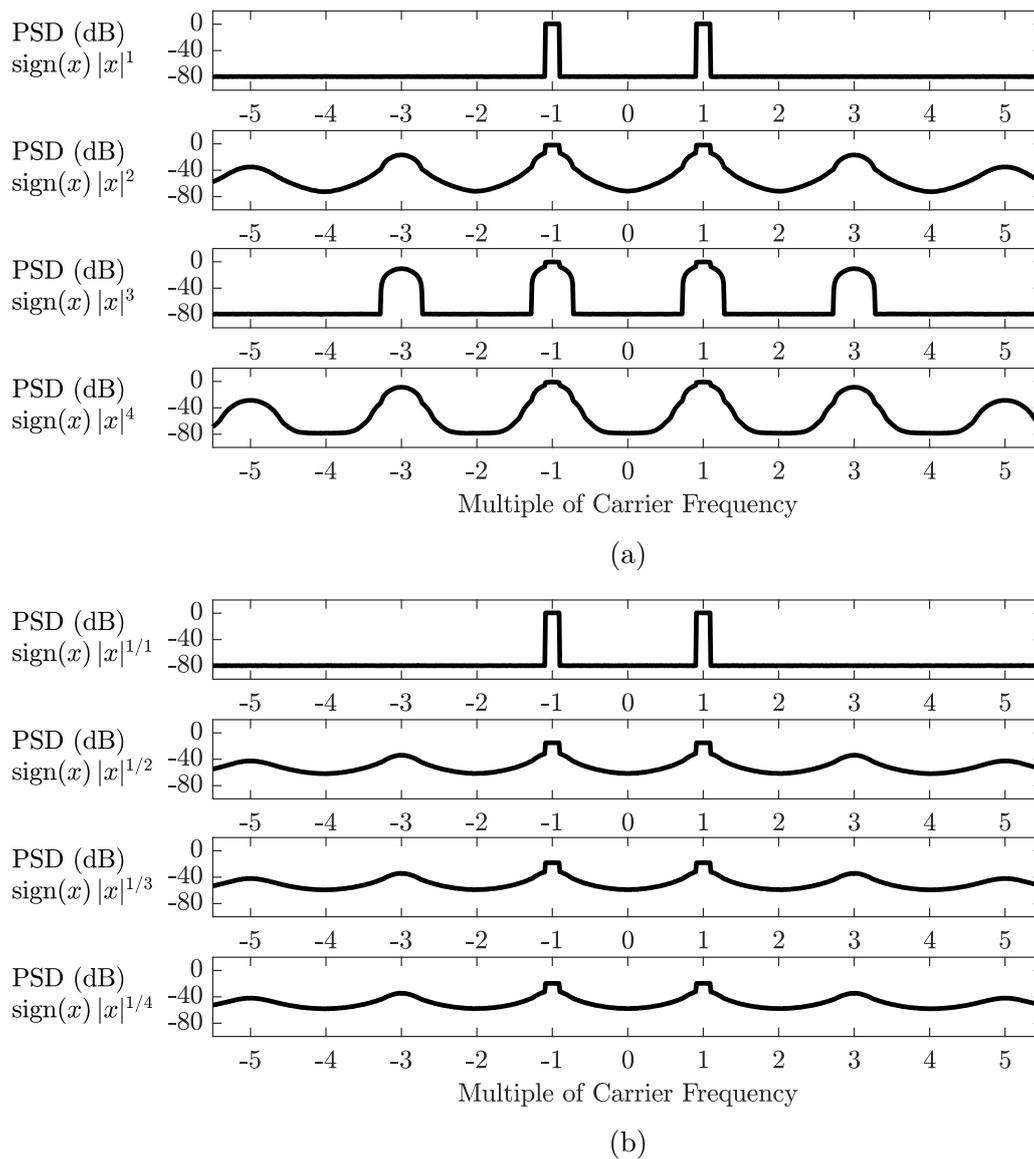
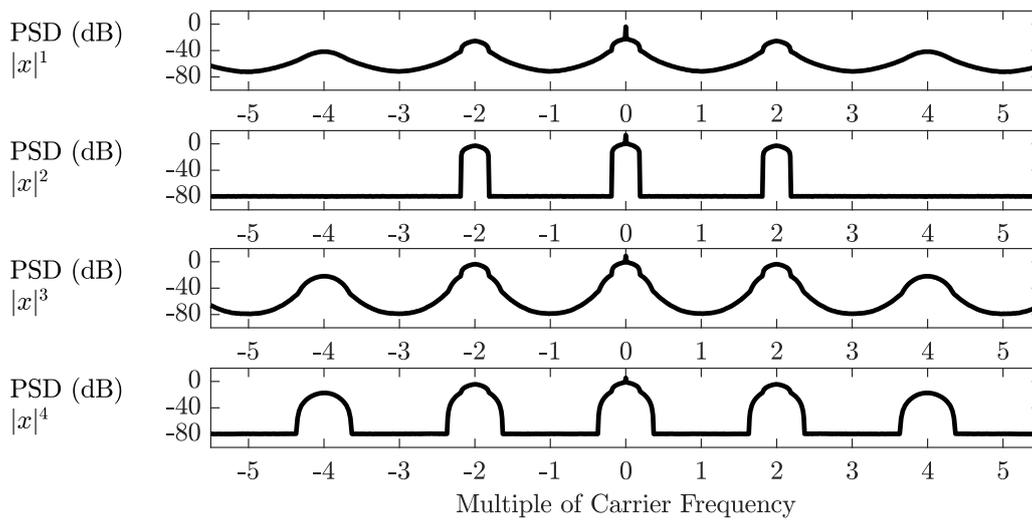


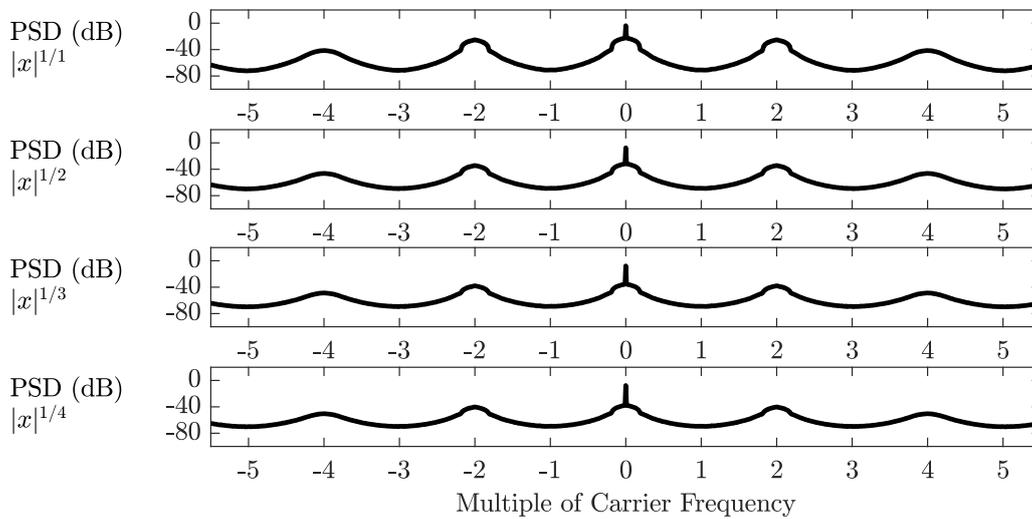
Figure 3.5: PSDs produced by odd-symmetric magnitude power functions.

Spectral Analysis of Even-Symmetric Magnitude Power Functions

The spectral characteristics of the even-symmetric magnitude power functions are shown in Figure 3.5. The spectral output at even harmonics is identical to the spectral output of the one-sided magnitude power functions. At odd harmonics, the spectral output is zero. For even-integer orders, the even-symmetric magnitude power functions reduce to polynomial basis functions and therefore, the spectra are the same as in Figure 3.1.



(a)



(b)

Figure 3.6: PSDs produced by even-symmetric magnitude power functions.

3.2.6 The Quasi-Memoryless Model

Above, I have presented an investigation of memoryless models and showed that they can model AM-AM, but cannot model AM-PM. The coefficients of memoryless baseband polynomials are real, since they are obtained from the real coefficients of memoryless passband polynomials by multiplication with real scaling factors.

In the following, I show that AM-PM can be modeled by complex coefficients in the baseband polynomial and derive the corresponding passband model, consisting of a Hilbert transformer and two memoryless passband nonlinearities [122]. The resulting model is called quasi-memoryless, since the baseband representation is memoryless, but the passband representation is not memoryless due to the Hilbert transformer. With respect to even-order terms in polynomial baseband models, the following analysis shows that the results derived for memoryless models also apply to quasi-memoryless models.

Modeling AM-PM by Complex Coefficients in Baseband

According to (3.3), the output of a memoryless nonlinearity, driven by a modulated carrier input signal $x(t) = a(t) \cos(\omega t + \varphi(t))$ can be described by the Fourier series

$$y(t) = \frac{1}{2} \hat{A}_0(a(t)) + \sum_{k=1}^{\infty} \hat{A}_k(a(t)) \cos(k\omega t + k\varphi(t)), \quad (3.46)$$

where $\hat{A}_k(\cdot)$ is the k -th order envelope nonlinearity, producing AM-AM distortion¹³. A model that also includes AM-PM can be formulated by generalizing (3.46) to

$$y(t) = \frac{1}{2} \hat{A}_0(a(t)) + \sum_{k=1}^{\infty} \hat{A}_k(a(t)) \cos(k\omega t + k\varphi(t) + \hat{\Phi}_k(a(t))), \quad (3.47)$$

where $\hat{\Phi}_k(\cdot)$ is the AM-PM nonlinearity of the k -th harmonic. By using the identity $\text{Re}(e^{j\theta}) = \cos(\theta)$, we can represent (3.47) in polar form by

$$y(t) = \frac{1}{2} \hat{A}_0(a(t)) + \sum_{k=1}^{\infty} \text{Re} \left(\hat{A}_k(a(t)) e^{j\hat{\Phi}_k(a(t))} e^{j(k\omega t + k\varphi(t))} \right). \quad (3.48)$$

By applying the conversion formulas in Appendix A to the term $\hat{A}_k(a(t)) e^{j\hat{\Phi}_k(a(t))}$ in (3.48), the Fourier series can also be written in Cartesian form by

$$y(t) = \frac{1}{2} \hat{I}_0(a(t)) + \sum_{k=1}^{\infty} \text{Re} \left(\left(\hat{I}_k(a(t)) + j\hat{Q}_k(a(t)) \right) e^{j(k\omega t + k\varphi(t))} \right), \quad (3.49)$$

where $\hat{I}_k(\cdot)$ and $\hat{Q}_k(\cdot)$ are the k -th order in-phase and quadrature-phase nonlinearities.

¹³Here, we introduce $\hat{A}_k(\cdot)$ as a synonym for $\hat{f}_k(\cdot)$ to emphasize the AM-AM effect.

The k -th harmonic modulation of the polar form in (3.48) is given by

$$\tilde{y}_k(t) = \hat{A}_k(a(t)) e^{j\hat{\Phi}_k(a(t))} e^{jk\varphi(t)}, \quad (3.50)$$

and the k -th harmonic modulation of the Cartesian form in (3.49) is given by

$$\tilde{y}_k(t) = \left(\hat{I}_k(a(t)) + j\hat{Q}_k(a(t)) \right) e^{jk\varphi(t)}. \quad (3.51)$$

A polynomial baseband model is obtained from (3.50) or (3.51), if we set $k = 1$ and approximate the nonlinearities $\hat{A}_1(a)$, $\hat{\Phi}_1(a)$ or $\hat{I}_1(a)$, $\hat{Q}_1(a)$ by a power series in $a(t)$.

If we use the Cartesian form in (3.51) and approximate the nonlinearities $\hat{I}_1(a)$ and $\hat{Q}_1(a)$ with the odd-symmetric magnitude power functions defined in (3.29), we obtain

$$\tilde{y}(t) = \left(\sum_{p=1}^P \tilde{\alpha}_p a(t) |a(t)|^{p-1} + j \sum_{p=1}^P \tilde{\beta}_p a(t) |a(t)|^{p-1} \right) e^{j\varphi(t)} \quad (3.52)$$

$$= \left(\sum_{p=1}^P (\tilde{\alpha}_p + j\tilde{\beta}_p) a(t) |a(t)|^{p-1} \right) e^{j\varphi(t)} \quad (3.53)$$

$$= \sum_{p=1}^P \tilde{c}_p \tilde{x}(t) |\tilde{x}(t)|^{p-1} \quad \text{with} \quad \tilde{c}_p = \tilde{\alpha}_p + j\tilde{\beta}_p, \quad (3.54)$$

which shows that the effect of AM-PM can be modeled by generalizing the memoryless complex baseband polynomial in (3.26) by complex-valued coefficients.

Modeling AM-PM by Two Memoryless Nonlinearities in Passband

To show that it is justified to use the odd-symmetric magnitude power functions as basis functions for $\hat{I}_1(a)$ and $\hat{Q}_1(a)$, we show that the effect of AM-AM and AM-PM can also be modeled by two memoryless nonlinearities in passband. This allows us to re-use the Chebyshev transformation pairs from above to derive explicit passband-baseband transformation pairs for the quasi-memoryless polynomial model.

We start the derivation by re-writing the Fourier series in (3.49) as

$$y(t) = \frac{1}{2} \hat{I}_0(a(t)) + \sum_{k=1}^{\infty} \hat{I}_k(a(t)) \cos(k\omega t + k\varphi(t)) - \sum_{k=1}^{\infty} \hat{Q}_k(a(t)) \sin(k\omega t + k\varphi(t)). \quad (3.55)$$

From (3.55), we see that the quasi-memoryless model can be represented by a linear combination of two memoryless models that are excited by a modulated cosine carrier $x_i(t)$ (in-phase) and a modulated sine carrier $x_q(t)$ (quadrature-phase), given by

$$x_i(t) = +\text{Re}(\tilde{x}(t) e^{j\omega t}) = +a(t) \cos(\omega t + \varphi(t)), \quad (3.56)$$

$$x_q(t) = -\text{Im}(\tilde{x}(t) e^{j\omega t}) = -a(t) \sin(\omega t + \varphi(t)). \quad (3.57)$$

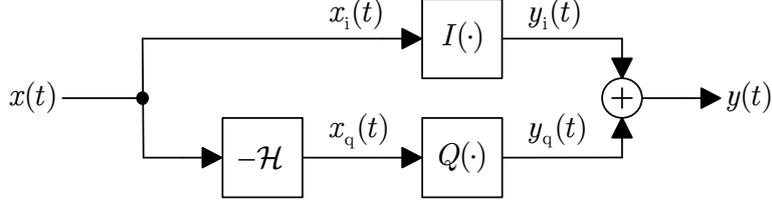


Figure 3.7: Quadrature model of a quasi-memoryless passband system.

The signals in (3.56) and (3.57) can be generated by using a Hilbert transformer as it is shown in Figure 3.7. The Hilbert transform \mathcal{H} applies a phase delay of 90° on the carrier, but does not modify the baseband signals $a(t)$ and $\varphi(t)$. If the signal is narrowband, \mathcal{H} can be approximated by a delay of $\Delta t = T_c/4$, where T_c is the carrier period.

The output of the memoryless passband nonlinearities $I(\cdot)$ and $Q(\cdot)$ in Figure 3.7 can be described by the Fourier series of the in-phase component

$$y_i(t) = \frac{1}{2}\hat{I}_0(a(t)) + \sum_{k=1}^{\infty} \hat{I}_k(a(t)) \cos(k\omega t + k\varphi(t)), \quad (3.58)$$

and the Fourier series of the quadrature-phase component

$$y_q(t) = \frac{1}{2}\hat{Q}_0(a(t)) + \sum_{k=1}^{\infty} \hat{Q}_k(a(t)) \cos\left(k\omega t + k\varphi(t) + k\frac{\pi}{2}\right). \quad (3.59)$$

In (3.59), we use the facts that $-\sin(\theta) = \cos(\theta + \frac{\pi}{2})$ and a phase shift of φ at the input of a memoryless model is equal to a phase shift of $k\varphi$ at the k -th harmonic output. Due to the phase shift of $k\frac{\pi}{2}$ in (3.59), we can re-write the quadrature-component as

$$\begin{aligned} y_q(t) = & \frac{1}{2}\hat{Q}_0(a(t)) + \sum_{\substack{k=1 \\ k \text{ is odd}}}^{\infty} (-1)^{\frac{k+1}{2}} \hat{Q}_k(a(t)) \sin(k\omega t + k\varphi(t)) \\ & + \sum_{\substack{k=2 \\ k \text{ is even}}}^{\infty} (-1)^{\frac{k}{2}} \hat{Q}_k(a(t)) \cos(k\omega t + k\varphi(t)). \end{aligned} \quad (3.60)$$

At odd harmonics, the signals in (3.58) and (3.60) form an actual quadrature pair consisting of a cosine and sine carrier such that AM-PM can be modeled. At even harmonics, however, both components are modulated cosine carriers such that AM-PM cannot be modeled. The complex modulation of the odd k -th harmonic is described by

$$\tilde{y}_k(t) = \left(\hat{I}_k(a(t)) + j(-1)^{\frac{k+3}{2}} \hat{Q}_k(a(t)) \right) e^{jk\varphi(t)} \quad \text{for odd } k. \quad (3.61)$$

For $k = 1, 5, 9, \dots$, (3.61) is identical to (3.51), which shows that for these harmonics, the passband model in Figure 3.7 produces exactly the same AM-AM and AM-PM effects that were assumed in (3.47). For $k = 3, 7, 11, \dots$, the sign of the phase modulation $\hat{\Phi}_k(\cdot)$ in (3.47) must be changed such that Figure 3.7 and (3.47) become equivalent models.

Summary of Transformation Pairs

Above, we have seen that the passband model in Figure 3.7 results in the k -th harmonic baseband model given by (3.61). The baseband nonlinearities $\hat{I}_k(\cdot)$ and $\hat{Q}_k(\cdot)$ in (3.61) are the Chebyshev transforms of the passband nonlinearities $I(\cdot)$ and $Q(\cdot)$. Therefore, the results derived for memoryless models also apply to quasi-memoryless models and it is justified to use the magnitude power functions with arbitrary order $p \in \mathbb{R}_{\geq 0}$ in quasi-memoryless polynomial baseband models, as we have done in (3.52).

A summary of passband-baseband transformation pairs for the quasi-memoryless model is given in the tables below. Table 3.1 shows the conventional transformation pair, which includes odd-order terms only. Table 3.2 shows the generalized transformation pair, where the set \mathcal{P} may include any non-negative real number. Table 3.3 lists the scaling factors λ_p for the first ten integer orders, computed by (H.40) in Appendix H.

PB	$y(t) = \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \alpha_p x_i^p(t) + \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \beta_p x_q^p(t) \quad \text{where} \quad \begin{cases} x_i(t) = +\text{Re}(\tilde{x}(t) e^{j\omega t}) \\ x_q(t) = -\text{Im}(\tilde{x}(t) e^{j\omega t}) \end{cases}$
BB	$\tilde{y}(t) = \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \tilde{c}_p \tilde{x}(t) \tilde{x}(t) ^{p-1} \quad \text{where} \quad \tilde{c}_p = \lambda_p (\alpha_p + j\beta_p), \quad \lambda_p = \frac{1}{2^{p-1}} \binom{p}{\frac{p-1}{2}}$

Table 3.1: Quasi-memoryless polynomial with odd-order terms only.

PB	$y(t) = \sum_{p \in \mathcal{P}} \alpha_p x_i(t) x_i(t) ^{p-1} + \sum_{p \in \mathcal{P}} \beta_p x_q(t) x_q(t) ^{p-1} \quad \begin{array}{l} x_i(t) \text{ and } x_q(t) \\ \text{as in Table 3.1} \end{array}$
BB	$\tilde{y}(t) = \sum_{p \in \mathcal{P}} \tilde{c}_p \tilde{x}(t) \tilde{x}(t) ^{p-1} \quad \text{where} \quad \tilde{c}_p = \lambda_p (\alpha_p + j\beta_p), \quad \lambda_p = \frac{2}{\sqrt{\pi}} \frac{\Gamma\left(\frac{p+2}{2}\right)}{\Gamma\left(\frac{p+3}{2}\right)}$

Table 3.2: Quasi-memoryless polynomial with non-negative real-order terms.

p	1	2	3	4	5	6	7	8	9	10
λ_p	1	$\frac{1}{\pi} \frac{8}{3}$	$\frac{3}{4}$	$\frac{1}{\pi} \frac{32}{15}$	$\frac{5}{8}$	$\frac{1}{\pi} \frac{64}{35}$	$\frac{35}{64}$	$\frac{1}{\pi} \frac{512}{315}$	$\frac{63}{128}$	$\frac{1}{\pi} \frac{1024}{693}$

Table 3.3: Coefficient scaling factors of the quasi-memoryless polynomial.

3.2.7 Phase Homogeneity

Above we have focused on memoryless and quasi-memoryless polynomial models and showed in much detail that even-order terms in baseband are mathematically justified. Beside of polynomial models, however, there are also other types of quasi-memoryless models, like the ones listed in [91]. Applying a detailed mathematical analysis on every new model is clearly a big effort. It is therefore desirable to have a simple requirement that tells us whether a baseband model of an RF-PA is mathematically justified.

If we apply no restriction on the model structure, a quasi-memoryless baseband model can be any memoryless mapping from one complex signal to another, given by

$$\tilde{f}_k(\cdot) : \tilde{x}(t) \mapsto \tilde{y}_k(t). \quad (3.62)$$

Mappings of the form (3.62) also include mappings that act differently on the real and imaginary part, like $\tilde{y}_1(t) = a \operatorname{Re}(\tilde{x}(t)) + j b \operatorname{Im}(\tilde{x}(t))$. Such a model is reasonable for modeling the IQ mismatch of a quadrature modulator, but it is not reasonable for modeling the nonlinear distortion of an RF-PA. For an RF-PA, it is common to assume distortion in form of AM-AM and AM-PM, resulting in a model structure given by

$$\tilde{f}_k(\cdot) : \tilde{x}(t) \mapsto \tilde{y}_k(t) = \hat{F}_k(|\tilde{x}(t)|) e^{jk\angle\tilde{x}(t)}, \quad (3.63)$$

where the function $\hat{F}_k(\cdot)$ may be represented in polar or Cartesian form by

$$\hat{F}_k(a) = \hat{A}_k(a) e^{j\hat{\Phi}_k(a)} = \hat{I}_k(a) + j \hat{Q}_k(a). \quad (3.64)$$

The functions $\hat{I}_k(\cdot)$ and $\hat{Q}_k(\cdot)$ are k -th order Chebyshev transforms, which must fulfill the symmetry in (3.9). Consequently, the AM-AM and AM-PM must fulfill

$$\hat{A}_k(-a) = (-1)^k \hat{A}_k(a), \quad \hat{\Phi}_k(-a) = \hat{\Phi}_k(a). \quad (3.65)$$

If we investigate the symmetry of $\tilde{f}_k(\cdot)$ in (3.63), we find that

$$\tilde{f}_k(e^{j\xi} \tilde{x}) = e^{jk\xi} \tilde{f}_k(\tilde{x}), \quad (3.66)$$

which can be interpreted as a generalization of (3.9), where the factor -1 is replaced by $e^{j\xi}$, with $\xi \in \mathbb{R}$. The symmetry in (3.66) is similar to k -th order homogeneity¹⁴, but with the important restriction that the magnitude of the scaling factor $e^{j\xi}$ is one. It seems therefore appropriate to call (3.66) k -th order phase homogeneity.

In Section 3.3.5, I prove that k -th order phase homogeneity is not limited to quasi-memoryless models that model AM-AM and AM-PM, but it is a general requirement for complex baseband models of time-invariant passband systems.

¹⁴See for example [196, Ch. 12], [191, Ch. 1], or [151, p. 27].

3.2.8 Discussion

Below, I summarize insights from the analysis presented above and discuss theoretical and practical implications of non-negative real orders in polynomial baseband models.

Order and Symmetry

A central insight of the analysis presented above is the importance of discriminating between the order of a term and the symmetry of a basis function. If we use polynomial basis functions, the order of the term is identical to the symmetry of the basis function, but with magnitude power basis functions, this is in general not true. In contrary, we have seen that the presence of even-order terms in baseband models can be explained by the odd symmetry of the corresponding magnitude power basis functions.

Intermodulation Products

If we send a real two-tone signal with tones at $f_c \pm f_{\text{env}}$ through a nonlinear passband system, the output near f_c consists of tones at $f_c \pm k f_{\text{env}}$ with $k \in \mathbb{N}_{\text{odd}}$. The integer k is called the order of the intermodulation product. The fact that only odd-order intermodulation products occur near the carrier frequency is related to the fact that first-order envelope nonlinearities must be odd-symmetric functions. It is important to discriminate between the order of an intermodulation product and the order of a term.

Numerical Properties

Basis functions with low order terms have in general better numerical properties than basis functions with high order terms. This is important for model identification, where the condition number of the regression matrix should be kept low and also for hardware implementation, where the bit-width of arithmetic operations should be kept low.

Spectral Properties

A unique property of the polynomial basis functions x^p with $p \in \mathbb{N}_0$ is that their output is bandlimited to p times the bandwidth of the input signal. This property is related to the smoothness of the polynomial basis functions. If magnitude power basis functions are used that cannot be reduced to polynomial basis functions, the output is not bandlimited. An important special case of not-bandlimited basis functions are the odd-symmetric basis functions $\text{sign}(x)|x|^p$ with $p \in \mathbb{N}_{\text{even}}$, which correspond to even-order terms in polynomial baseband models. If such even-order terms are included, it is important to use sufficient oversampling to keep aliasing below the desired dynamic range.

Analytic Properties

A function is analytic, if it can be locally represented by a convergent Taylor series. Smoothness is a necessary, but not sufficient condition for analyticity. Complex analytic functions are also called holomorphic. Holomorphicity is rather restrictive, since such functions must fulfill the Cauchy-Riemann conditions, outlined in Appendix C.

The polynomial basis functions x^p with $p \in \mathbb{N}_0$ are smooth, analytic, and holomorphic. The real magnitude power functions $\text{sign}^q(x) |x|^p$, that cannot be reduced to polynomial basis functions are not smooth and therefore not analytic or holomorphic.

The complex magnitude power functions $\tilde{x} |\tilde{x}|^{p-1}$ with $p \in \mathbb{R}_{\geq 0}$, $p \neq 1$, $\tilde{x} \in \mathbb{C}$ are not holomorphic. To see this, we write them as $\tilde{x} (\tilde{x} \tilde{x}^*)^{(p-1)/2}$ and note that holomorphic functions must be independent of their complex-conjugate argument [129, p. 8].

Physical Justification

It is interesting to note that the semi-physical RF-PA model of Chapter 2 is dominated by a second-order term. The passband nonlinearity in (2.3) shows square-law behavior for small inputs, which matches the physical model of MOSFETs. For gate voltages below the threshold voltage, the drain current is zero, similar to the one-sided magnitude power functions. In an actual transistor, there is no discontinuity at the threshold voltage. However, it seems reasonable that the overall behavior can be approximated well, regardless of the discontinuity at the origin of the magnitude power functions.

Completeness of Odd-Order Terms

Since every function consists of an odd-symmetric and an even-symmetric part and the symmetry of polynomial basis functions is identical to the parity of their order, we can state the following corollary to the Weierstrass theorem: Every odd-symmetric function can be approximated within a finite interval with arbitrary accuracy by a polynomial with odd-order terms only. The reason why even-order terms are nevertheless practically relevant is that they can provide higher accuracy by using only a few terms.

Fitting of Polynomial Baseband Models

The coefficients of quasi-memoryless polynomial baseband models can be identified from complex baseband input and output signals by using the complex polynomial baseband model in Table 3.2 and the method of least squares described in Appendix D.

If polynomial models of the AM-AM and AM-PM are fitted separately, the resulting behavioral model is in general not optimal in the signal domain. An AM-AM polynomial that is fitted onto $|\tilde{x}| \mapsto |\tilde{y}|$ is additionally a biased estimate, since complex noise on \tilde{y} is not averaged out completely, but increases the estimated small signal gain.

3.3 Polynomial Models with Memory

In the first part of this chapter, I focused on the simplest case, which is the memoryless polynomial, to establish a solid theoretical basis for even-order terms in memoryless and quasi-memoryless polynomial baseband models. For wideband communication signals, however, nonlinear models with memory are required. Since linearity in the coefficients greatly simplifies the identification of the coefficients, most of these models are based on the baseband Volterra series, which was first derived in [22].

In the following, I present an introduction to the Volterra series and derive its equivalent baseband version, similar as in [22], but by using a more compact vector notation. Afterwards, I apply insights from the analysis of memoryless models to derive a generalized baseband Volterra series with even-order terms. Finally, I prove that phase homogeneity is a necessary symmetry of all baseband models of time-invariant passband systems.

3.3.1 The Volterra Series

The Volterra series maps a real input signal $x(t)$ to a real output signal $y(t)$ by

$$y(t) = \sum_{p=1}^P \underbrace{\int \cdots \int}_{p\text{-fold integral}} h_p(\tau_1, \dots, \tau_p) \prod_{i=1}^p x(t - \tau_i) d\tau_1 \dots d\tau_p, \quad (3.67)$$

where $h_p(\tau_1, \dots, \tau_p)$ are the p -th order Volterra kernels¹⁵, which are functions of the p time-lag variables τ_1 to τ_p . The structure of (3.67) becomes more clear, if we explicitly write the first few terms of the series, which are given by

$$\begin{aligned} y(t) &= \int h_1(\tau_1) x(t - \tau_1) d\tau_1 \\ &+ \int \int h_2(\tau_1, \tau_2) x(t - \tau_1) x(t - \tau_2) d\tau_1 d\tau_2 \\ &+ \int \int \int h_3(\tau_1, \tau_2, \tau_3) x(t - \tau_1) x(t - \tau_2) x(t - \tau_3) d\tau_1 d\tau_2 d\tau_3 \\ &+ \dots \end{aligned} \quad (3.68)$$

If the Volterra kernels are zero, except for $\tau_1 = 0, \tau_2 = 0, \dots, \tau_p = 0$, the Volterra series reduces to the memoryless polynomial. On the other hand, if the Volterra kernels are zero, except for $h_1(\tau_1)$, the Volterra series reduces to a linear convolution.

The kernels of order $p \geq 2$ can be restricted to a subspace of \mathbb{R}^p without loss of generality, since all permutations of the time lags τ_1, \dots, τ_p produce the same input product. An illustration of the symmetric regions for $p = 2$ and $p = 3$ is shown in Figure 3.8.

¹⁵For mathematical simplicity, we exclude the constant DC offset, modeled by the scalar kernel h_0 .

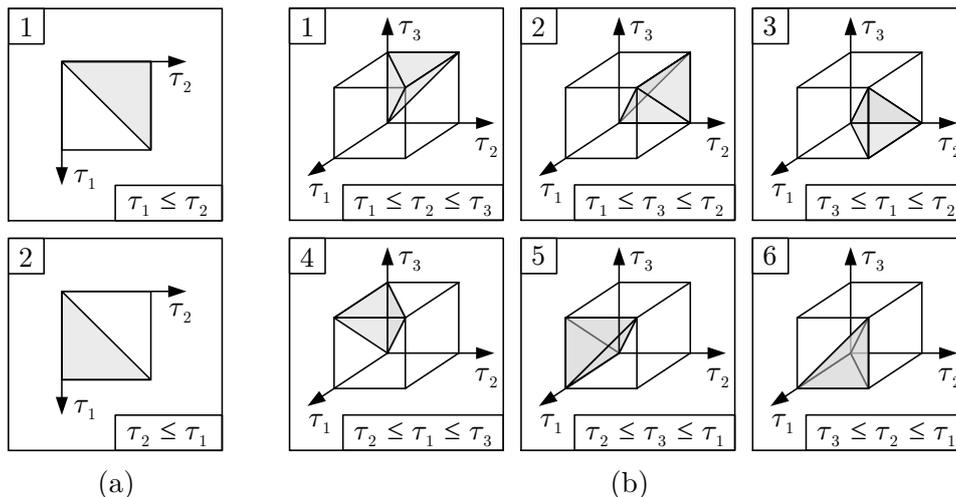


Figure 3.8: Symmetric regions in the second and third order Volterra kernels.

Discrete-Time Implementation

The discrete-time version of the continuous-time Volterra series in (3.67) is given by

$$y[n] = \sum_{p=1}^P \underbrace{\sum_{m_1=0}^M \sum_{m_2=m_1}^M \cdots \sum_{m_p=m_{p-1}}^M}_{p\text{-fold sum}} h_p[m_1, \dots, m_p] \prod_{i=1}^p x[n - m_i], \quad (3.69)$$

where the time lags τ_1 to τ_p have been replaced by the sample lags m_1 to m_p . In (3.69), we have restricted the sample lags to $0 \leq m_1 \leq m_2 \leq \dots \leq m_p \leq M$, corresponding to the causal kernel of memory depth M and the first symmetric regions in Figure 3.8.

Computational Complexity

The computational complexity of (3.69) can be quantified by the number of coefficients, multiplied by the number of operations that are required for each of the coefficients. By restricting the summation to one of the symmetric regions in the kernel, the number of coefficients in the p -th order kernel is reduced from $C_p = (M + 1)^p$ to

$$C_p = \binom{M + 1}{p} = \binom{M + p}{p}, \quad (3.70)$$

which computes the number of p -multicombinations out of $M + 1$ possible sample lags¹⁶. A direct implementation of (3.69) requires p multiplications per coefficient. However, if Horner's method for polynomial evaluation is used, Volterra filtering can be implemented with only one multiplication per coefficient [78]. Examples in C are available from [75].

¹⁶Here, we use the definition $\binom{n}{k} = \binom{n+k-1}{k}$, which expresses the number of multicombinations in a similar way as the binomial coefficient does for combinations. For further details, see Appendix F.

3.3.2 The Baseband Volterra Series

To derive the baseband Volterra series, we introduce a compact notation by defining the p -dimensional time lag vector $\boldsymbol{\tau}_p = [\tau_1, \tau_2, \dots, \tau_p]^T$, such that (3.67) can be written as

$$y(t) = \sum_{p=1}^P \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \prod_{i=1}^p x(t - \tau_i) d\boldsymbol{\tau}_p. \quad (3.71)$$

Furthermore, we separate (3.71) into three parts, given by¹⁷

$$y(t) = \sum_{p=1}^P y_p(t), \quad y_p(t) = \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \Psi_p(t, \boldsymbol{\tau}_p) d\boldsymbol{\tau}_p, \quad \Psi_p(t, \boldsymbol{\tau}_p) = \prod_{i=1}^p x(t - \tau_i). \quad (3.72)$$

The modulated carrier input signal is given by

$$x(t) = \operatorname{Re} \left(\tilde{x}(t) e^{j\omega t} \right) = \frac{1}{2} \tilde{x}(t) e^{+j\omega t} + \frac{1}{2} \tilde{x}^*(t) e^{-j\omega t}, \quad (3.73)$$

where $\tilde{x}(t)$ is the complex baseband signal. An equivalent form of (3.73) is given by

$$x(t) = \sum_{c \in \mathbb{B}} \hat{x}_{(t)}^{(c)}, \quad \hat{x}_{(t)}^{(c)} = \left(\frac{1}{2} \tilde{x}(t) e^{j\omega t} \right)^{(*c)}, \quad \mathbb{B} = \{0, 1\}. \quad (3.74)$$

Inserting the 2-term sum in (3.74) into the basis functional $\Psi_p(t, \boldsymbol{\tau}_p)$ in (3.72), we get

$$\Psi_p(t, \boldsymbol{\tau}_p) = \prod_{i=1}^p \sum_{c \in \mathbb{B}} \hat{x}_{(t-\tau_i)}^{(c)} = \sum_{\mathbf{c}_p \in \mathbb{B}^p} \prod_{i=1}^p \hat{x}_{(t-\tau_i)}^{(c_i)}, \quad (3.75)$$

where we expanded the p -term product of 2-term sums into a 2^p -term sum of p -term products. Inserting the latter into the integral in (3.72), we get

$$y_p(t) = \sum_{\mathbf{c}_p \in \mathbb{B}^p} \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \prod_{i=1}^p \hat{x}_{(t-\tau_i)}^{(c_i)} d\boldsymbol{\tau}_p. \quad (3.76)$$

The integral in (3.76) produces the same result for all permutations of elements in \mathbf{c}_p . We can therefore restrict the sum to vectors \mathbf{c}_p consisting of q zeros followed by $p - q$ ones and multiply each term by the number of equivalent permutations¹⁸, resulting in

$$y_p(t) = \sum_{q=0}^p \binom{p}{q} \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \prod_{i=1}^p \hat{x}_{(t-\tau_i)}^{(i>q)} d\boldsymbol{\tau}_p. \quad (3.77)$$

¹⁷Here, we introduce the symbol $\Psi_p(t, \boldsymbol{\tau}_p)$ for the input product, which we call the basis functional.

¹⁸This reduction is similar to the binomial theorem and can be illustrated by the fact that the sum of the p -th row of Pascal's triangle (starting at $p = 0$ with the one-element row consisting of 1) is 2^p

The index q in (3.77) represents the number of not-conjugated terms within the input products. The difference of not-conjugated terms and conjugated terms is given by the integer $k = q - (p - q) = 2q - p$. Substituting $q = \frac{p+k}{2}$ in (3.77), we get

$$y_p(t) = \sum_{\substack{k=-p \\ \text{steps of 2}}}^{+p} \binom{p}{\frac{p+k}{2}} \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \prod_{i=1}^p \dot{x}_{(t-\tau_i)}^{(i > \frac{p+k}{2})} d\boldsymbol{\tau}_p. \quad (3.78)$$

Inserting the definition of $\dot{x}_{(t)}^{(c)}$ from (3.74) into (3.78) and summing over p , we get

$$y(t) = \sum_{p=1}^P \sum_{\substack{k=-p \\ \text{steps of 2}}}^{+p} \int_{\mathbb{R}^p} \frac{1}{2^p} \binom{p}{\frac{p+k}{2}} h_p(\boldsymbol{\tau}_p) \times \prod_{i=1}^{\frac{p+k}{2}} \tilde{x}(t - \tau_i) e^{+j\omega(t-\tau_i)} \prod_{l=\frac{p+k}{2}+1}^p \tilde{x}^*(t - \tau_l) e^{-j\omega(t-\tau_l)} d\boldsymbol{\tau}_p. \quad (3.79)$$

The exponentials in (3.79) can be combined to $e^{j(k\omega t + \varphi_{p,k}(\boldsymbol{\tau}_p))}$, which allows us to write

$$y(t) = \sum_{p=1}^P \sum_{\substack{k=-p \\ \text{steps of 2}}}^{+p} \frac{1}{2} \tilde{y}_{p,k}(t) e^{jk\omega t}, \quad \tilde{y}_{p,k}(t) = \int_{\mathbb{R}^p} \tilde{h}_{p,k}(\boldsymbol{\tau}_p) \tilde{\Psi}_{p,k}(t, \boldsymbol{\tau}_p) d\boldsymbol{\tau}_p, \quad (3.80)$$

where the harmonic kernels and the harmonic basis functionals are given by

$$\tilde{h}_{p,k}(\boldsymbol{\tau}_p) = h_p(\boldsymbol{\tau}_p) \lambda_{p,k}(\boldsymbol{\tau}_p), \quad \tilde{\Psi}_{p,k}(t, \boldsymbol{\tau}_p) = \prod_{i=1}^{\frac{p+k}{2}} \tilde{x}(t - \tau_i) \prod_{l=\frac{p+k}{2}+1}^p \tilde{x}^*(t - \tau_l), \quad (3.81)$$

and the complex scaling factors of the Volterra kernels are given by

$$\lambda_{p,k}(\boldsymbol{\tau}_p) = \frac{1}{2^{p-1}} \binom{p}{\frac{p+k}{2}} \exp \left(j\omega \left(-\sum_{i=1}^{\frac{p+k}{2}} \tau_i + \sum_{l=\frac{p+k}{2}+1}^p \tau_l \right) \right). \quad (3.82)$$

The expression for $y(t)$ in (3.80) is a sum of modulated harmonics that produces a spectrum like in Figure 3.1 on page 33. The baseband Volterra series models the complex modulation of the first harmonic. Defining $\tilde{h}_p(\boldsymbol{\tau}_p) = \tilde{h}_{p,1}(\boldsymbol{\tau}_p)$, we may write it as

$$\tilde{y}(t) = \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \int_{\mathbb{R}^p} \tilde{h}_p(\boldsymbol{\tau}_p) \underbrace{\prod_{i=1}^q \tilde{x}(t - \tau_i)}_{q\text{-fold product}} \underbrace{\prod_{l=q+1}^{2q-1} \tilde{x}^*(t - \tau_l)}_{(q-1)\text{-fold product}} d\boldsymbol{\tau}_p, \quad q = \frac{p+1}{2}. \quad (3.83)$$

Note that (3.83) contains only odd-order terms and that the basis functionals contain a number of not-conjugated terms that exceeds the number of conjugated terms by one.

Discrete-Time Implementation

The discrete-time version of the baseband Volterra series in (3.83) is given by

$$\begin{aligned} \tilde{y}[n] = & \sum_{\substack{p=1 \\ p \text{ is odd}}}^P \underbrace{\sum_{m_1=0}^M \sum_{m_2=m_1}^M \cdots \sum_{m_q=m_{q-1}}^M}_{q\text{-fold sum, } q=\frac{p+1}{2}} \underbrace{\sum_{m_{q+1}=0}^M \sum_{m_{q+2}=m_{q+1}}^M \cdots \sum_{m_{2q-1}=m_{2q-2}}^M}_{(q-1)\text{-fold sum, } q=\frac{p+1}{2}} \\ & \times \tilde{h}_p[m_1, \dots, m_p] \prod_{i=1}^q \tilde{x}[n - m_i] \prod_{l=q+1}^{2q-1} \tilde{x}^*[n - m_l], \end{aligned} \quad (3.84)$$

where the time lags τ_1 to τ_p have been replaced by sample lags m_1 to m_p . The kernel symmetry of (3.84) is different from that of (3.69), since only sample lag permutations within not-conjugated terms or within conjugated terms produce the same output. The p -fold sum is therefore separated into a q -fold sum, describing the sample lags of the not-conjugated terms, and a $(q-1)$ -fold sum, describing the sample lags of the conjugated terms. The overall number of summations stays the same, since $q + (q-1) = p$.

Computational Complexity

The computational complexity of (3.84) is higher than that (3.69), because there are more non-redundant coefficients in the p -th order kernel and each of them requires p complex multiplications. The number of coefficients in the odd p -th order kernel is

$$C_p = \binom{M+1}{q} \binom{M+1}{q-1} = \binom{M+q}{q} \binom{M+q-1}{q-1}, \quad \begin{cases} q = \frac{p+1}{2}, \\ p \in \mathbb{N}_{\text{odd}}. \end{cases} \quad (3.85)$$

The number of multiplications can be reduced by reusing methods similar to [78], but the potential for complexity reduction is less, because of the reduced symmetry.

3.3.3 The Baseband Volterra Series with Even-Order Terms

Based on our investigation of memoryless polynomial baseband models, we expect that a baseband Volterra series with even-order terms can be derived, if we assume a passband Volterra series with modified basis functionals, given by

$$\Psi_p(t, \boldsymbol{\tau}_p) = \begin{cases} \prod_{i=1}^p x(t - \tau_i) & p \text{ is odd,} \\ \prod_{i=1}^{p-1} x(t - \tau_i) |x(t - \tau_p)| & p \text{ is even.} \end{cases} \quad (3.86)$$

Above, we have already shown that the odd-order passband basis functionals in (3.86) correspond to the odd-order baseband basis functionals in (3.81). In the following, I present a similar derivation for the even-order passband basis functionals in (3.86).

Factorization of the Modified Even-Order Volterra Terms

According to (3.72), the p -th order output of the passband Volterra series is given by

$$y_p(t) = \int_{\mathbb{R}^p} h_p(\boldsymbol{\tau}_p) \Psi_p(t, \boldsymbol{\tau}_p) d\boldsymbol{\tau}_p. \quad (3.87)$$

If we assume that p is even, and define $r = p - 1$, which is odd, we may separate the p -fold integral in (3.87) into a one-fold integral and an r -fold integral, given by

$$y_p(t) = \int_{\mathbb{R}} \int_{\mathbb{R}^r} h_p(\boldsymbol{\tau}_r, \tau_p) \Psi_r(t, \boldsymbol{\tau}_r) |x(t - \tau_p)| d\boldsymbol{\tau}_r d\tau_p. \quad (3.88)$$

An equivalent way of writing (3.88) is given by

$$y_p(t) = \int_{\mathbb{R}} f_r(t, \tau_p) g(t, \tau_p) d\tau_p, \quad \text{where} \quad \begin{cases} f_r(t, \tau_p) = \int_{\mathbb{R}^r} h_p(\boldsymbol{\tau}_r, \tau_p) \Psi_r(t, \boldsymbol{\tau}_r) d\boldsymbol{\tau}_r, \\ g(t, \tau_p) = |x(t - \tau_p)|. \end{cases} \quad (3.89)$$

In (3.89), we have factorized the even p -th order output of the modified Volterra series into an odd r -th order Volterra term and a delayed magnitude term. In the following, we expand these terms into two harmonic series and then combine them to one harmonic series of the even p -th order output of the modified Volterra series.

Harmonic Series of the Odd-Order Volterra Terms

According to (3.80), we can represent the odd r -th order Volterra terms by

$$f_r(t, \tau_p) = \sum_{\substack{k_1=-r \\ \text{steps of 2}}}^{+r} \frac{1}{2} \tilde{f}_{r,k_1}(t, \tau_p) e^{jk_1\omega t}, \quad (3.90)$$

where the complex modulation of the harmonics is given by

$$\tilde{f}_{r,k}(t, \tau_p) = \int_{\mathbb{R}^r} \tilde{h}_{p,k}(\boldsymbol{\tau}_r, \tau_p) \tilde{\Psi}_{r,k}(t, \boldsymbol{\tau}_r) d\boldsymbol{\tau}_r. \quad (3.91)$$

Harmonic Series of the Delayed Magnitude Term

According to (3.3) and (3.45), we can represent the delayed magnitude term by

$$g(t, \tau_p) = \sum_{\substack{k_2=-\infty \\ k_2 \text{ is even}}}^{+\infty} \tilde{g}_{k_2}(t, \tau_p) e^{jk_2\omega t}, \quad (3.92)$$

where the complex modulation of the harmonics is given by

$$\tilde{g}_k(t, \tau_p) = \lambda_{1,k} |\tilde{x}(t - \tau_p)| e^{jk(\angle \tilde{x}(t - \tau_p) - \omega \tau_p)}. \quad (3.93)$$

Harmonic Series of the Modified Even-Order Volterra Terms

Inserting (3.90) and (3.92) into (3.89), we obtain

$$y_p(t) = \sum_{\substack{k_1=-r \\ \text{steps of 2}}}^{+r} \sum_{\substack{k_2=-\infty \\ k_2 \text{ is even}}}^{+\infty} \frac{1}{2} \tilde{y}_{p,k_1,k_2}(t, \tau_p) e^{j(k_1+k_2)\omega t}, \quad (3.94)$$

where the complex modulation of the harmonics is given by

$$\tilde{y}_{p,k_1,k_2}(t) = \int_{\mathbb{R}} \tilde{f}_{r,k_1}(t, \tau_p) \tilde{g}_{k_2}(t, \tau_p) d\tau_p. \quad (3.95)$$

To transform (3.94) into a harmonic series over a single harmonic index k , we define $k = k_1 + k_2$ and apply the substitution $k_2 = k - k_1$. This changes the sum over even k_2 into a sum over odd k . The change of parity follows from the fact that p is even, which implies that r and k_1 are odd. After the substitution, we have

$$y_p(t) = \sum_{\substack{k=-\infty \\ k \text{ is odd}}}^{+\infty} \frac{1}{2} \tilde{y}_{p,k}(t, \tau_p) e^{jk\omega t}, \quad (3.96)$$

where the complex modulation of the harmonics is given by

$$\tilde{y}_{p,k}(t) = \sum_{\substack{k_1=-r \\ \text{steps of 2}}}^{+r} \int_{\mathbb{R}} \tilde{f}_{r,k_1}(t, \tau_p) \tilde{g}_{k-k_1}(t, \tau_p) d\tau_p. \quad (3.97)$$

By inserting (3.91) and (3.93) into (3.97), we get

$$\tilde{y}_{p,k}(t) = \sum_{\substack{k_1=-r \\ \text{steps of 2}}}^{+r} \int_{\mathbb{R}^p} \tilde{h}_{p,k}^{(k_1)}(\tau_p) \tilde{\Psi}_{p,k}^{(k_1)}(t, \tau_p) d\tau_p, \quad (3.98)$$

where the baseband kernels are given by

$$\tilde{h}_{p,k}^{(k_1)}(\tau_p) = h_p(\tau_p) \lambda_{r,k_1}(\tau_r) \lambda_{1,k-k_1} e^{-j\omega(k-k_1)\tau_p}, \quad (3.99)$$

and the baseband basis functionals are given by

$$\tilde{\Psi}_{p,k}^{(k_1)}(t, \tau_p) = \tilde{\Psi}_{r,k_1}(t, \tau_r) |\tilde{x}(t - \tau_p)| e^{j(k-k_1)\angle \tilde{x}(t-\tau_p)}. \quad (3.100)$$

The harmonic series in (3.96) shows that, as expected, the modified even-order passband basis functionals produce output at odd harmonics. The modulation of these harmonics is described by the even p -th order Volterra integrals in (3.98), which contain several even-order basis functionals, given by (3.100), and indexed by the odd integer k_1 .

Baseband Volterra Series with Even-Order Terms

If we want a baseband Volterra series that exactly reproduces the first harmonic output of the passband Volterra series with the modified basis functionals defined in (3.86), we have to use all baseband basis functionals defined in (3.100). However, since our goal is not an exact reproduction of a given passband model, but a baseband model with a good complexity-accuracy trade-off, we are free to select a subset of the basis functionals. Inserting the odd-order baseband basis functional from (3.81) on page 56 into (3.100) and setting $k_1 = k$, such that the exponential term vanishes, we get for even orders

$$\tilde{\Psi}_{p,k}(t, \boldsymbol{\tau}_p) = \prod_{i=1}^{\frac{r+k}{2}} \tilde{x}(t - \tau_i) \prod_{l=\frac{r+k}{2}+1}^r \tilde{x}^*(t - \tau_l) |\tilde{x}(t - \tau_p)|. \quad (3.101)$$

By extending the odd-order baseband Volterra series in (3.83) with the even-order terms in (3.101), we can formulate the baseband Volterra series with even-order terms by

$$\tilde{y}(t) = \sum_{p=1}^P \int_{\mathbb{R}^p} \tilde{h}_p(\boldsymbol{\tau}_p) \underbrace{\prod_{i=1}^q \tilde{x}(t - \tau_i)}_{\substack{q\text{-fold product} \\ q = \lfloor \frac{p+1}{2} \rfloor}} \underbrace{\prod_{l=q+1}^{2q-1} \tilde{x}^*(t - \tau_l)}_{\substack{(q-1)\text{-fold product} \\ q = \lfloor \frac{p+1}{2} \rfloor}} \underbrace{|\tilde{x}(t - \tau_p)|}_{\substack{\text{only if} \\ p \text{ is even}}} d\boldsymbol{\tau}_p, \quad (3.102)$$

Discrete-Time Implementation

The discrete-time version of (3.102) is given by

$$\begin{aligned} \tilde{y}[n] = & \sum_{p=1}^P \underbrace{\sum_{m_1=0}^M \sum_{m_2=m_1}^M \cdots \sum_{m_q=m_{q-1}}^M}_{q\text{-fold sum, } q = \lfloor \frac{p+1}{2} \rfloor} \underbrace{\sum_{m_{q+1}=0}^M \sum_{m_{q+2}=m_{q+1}}^M \cdots \sum_{m_{2q-1}=m_{2q-2}}^M}_{(q-1)\text{-fold sum, } q = \lfloor \frac{p+1}{2} \rfloor} \underbrace{\sum_{m_p=0}^M}_{\substack{\text{only if} \\ p \text{ is even}}} \\ & \times \tilde{h}_p[m_1, \dots, m_p] \underbrace{\prod_{i=1}^q \tilde{x}[n - m_i]}_{\substack{q\text{-fold product} \\ q = \lfloor \frac{p+1}{2} \rfloor}} \underbrace{\prod_{l=q+1}^{2q-1} \tilde{x}^*[n - m_l]}_{\substack{(q-1)\text{-fold product} \\ q = \lfloor \frac{p+1}{2} \rfloor}} \underbrace{|\tilde{x}[n - m_p]|}_{\substack{\text{only if} \\ p \text{ is even}}}, \end{aligned} \quad (3.103)$$

where we have used the floor operator $\lfloor \cdot \rfloor$ to ensure integer values of q for even p and integrated an additional sum over the sample lag m_p that is only included if p is even.

Computational Complexity

The number of coefficients in the odd p -th order kernel is described by (3.85) on page 57. For even orders, the number of coefficients C_p is given by $C_p = C_{p-1} (M + 1)$.

3.3.4 The Generalized Memory Polynomial

The baseband Volterra series is very general, but the rapid growth of its kernel size limits its applicability to low orders and memory depths. The generalized memory polynomial (GMP), proposed by Dennis Morgan et al. in [163] is a model with reduced kernel size. In the GMP, the sample lags m_1 to m_p of the baseband Volterra series are reduced to the sample lags m_1 and m_2 for all orders. The sample lag m_1 corresponds to the first not-conjugated term and the sample lag m_2 applies to all other terms, resulting in¹⁹

$$\tilde{y}[n] = \sum_{p=1}^P \underbrace{\sum_{m_1=0}^M \sum_{m_2=0}^M}_{p=1:m_1=m_2} \tilde{h}_p[m_1, m_2] \tilde{x}[n - m_1] |\tilde{x}[n - m_2]|^{p-1}. \quad (3.104)$$

By rearranging the summations in (3.104), we can write the GMP as

$$\tilde{y}[n] = \sum_{m_1=0}^M \sum_{m_2=0}^M \tilde{x}[n - m_1] \tilde{g}_{m_1, m_2}(|\tilde{x}[n - m_2]|), \quad (3.105)$$

where the factor $\tilde{g}_{m_1, m_2}(\cdot)$ is a magnitude-dependent complex gain, given by

$$\tilde{g}_{m_1, m_2}(a) = \sum_{p=p_0[m_1, m_2]}^P \tilde{h}_p[m_1, m_2] |a|^{p-1}, \quad p_0[m_1, m_2] = \begin{cases} 1 & m_1 = m_2, \\ 2 & \text{otherwise.} \end{cases} \quad (3.106)$$

The formula in (3.105) can be mapped to an efficient lookup-table implementation, which is shown in Figure 3.9 and requires only $(M + 1)^2$ multiply-accumulate units²⁰.

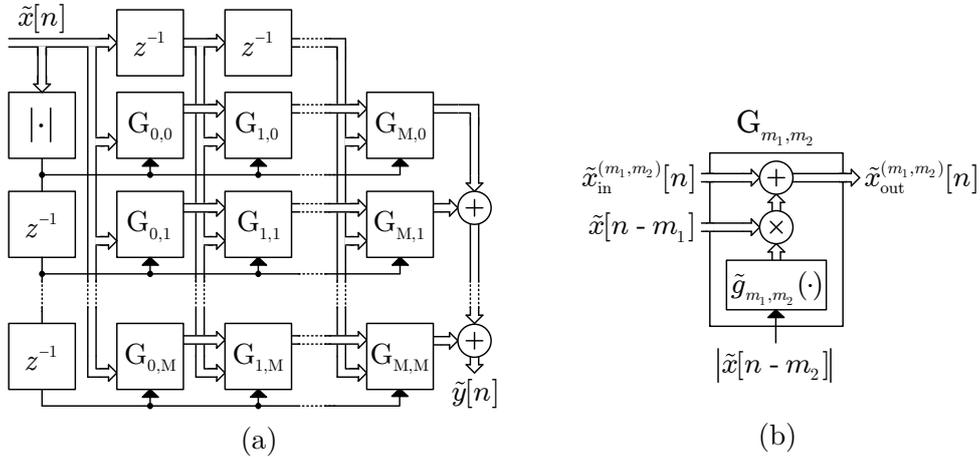


Figure 3.9: Implementation of the GMP based on lookup tables.

¹⁹The original formula in [163] describes the same kernel subspace, but uses different coordinates.

²⁰Further details on lookup table based RF-PA models can be found in [102]

3.3.5 Phase Homogeneity

Above, I have shown that it is mathematically justified to extend the conventional odd-order baseband Volterra series by even-order terms that are formed by extending the odd-order basis functionals by an additional magnitude term. However, the Volterra series is not the only possible foundation for an RF-PA model with memory, so it is interesting to ask, what requirement must be fulfilled by a complex baseband model such that it is mathematically justified for modeling an RF-PA in passband.

To answer this question, I return to the Fourier series analysis of a nonlinear distorted passband signal, as presented for memoryless systems in Section 3.2.2 on page 34. In the following, I extend this analysis to systems with memory and prove that phase homogeneity, as defined in (3.66) on page 50, is not limited to quasi-memoryless models, but applies to all equivalent baseband models of time-invariant passband systems.

Fourier Series of a Nonlinear Distorted Passband Signal

Let's consider a nonlinear distorted passband signal $y(t)$, described by

$$y(t) = \mathcal{N}\{x(t)\}, \quad x(t) = \operatorname{Re}\left(\tilde{x}(t) e^{j\omega t}\right), \quad (3.107)$$

where $\mathcal{N}\{\cdot\}$ is a time-invariant operator. This signal can also be described by

$$y(t) = Y(t, \tau) \Big|_{\tau=t} = \mathcal{N}\left\{ \operatorname{Re}\left(\tilde{x}(t) e^{j\omega\tau}\right) \right\} \Big|_{\tau=t}, \quad (3.108)$$

where $Y(t, \tau)$ is a memoryless function, mapping $t \in \mathbb{R}$ and $\tau \in \mathbb{R}$ to $y \in \mathbb{R}$.

Since the term $e^{j\omega\tau}$ in (3.108) is T -periodic in τ with $T = \frac{2\pi}{\omega}$, also the function $Y(t, \tau)$ is T -periodic in τ . Therefore, we can expand $Y(t, \tau)$ into the Fourier series

$$Y(t, \tau) = \frac{1}{2} \tilde{y}_0(t) + \sum_{k=1}^{\infty} \operatorname{Re}\left(\tilde{y}_k(t) e^{jk\omega\tau}\right) \quad (3.109)$$

where the Fourier coefficients are functions of t , given by

$$\tilde{y}_k(t) = \frac{2}{T} \int_{\tau_0}^{\tau_0+T} Y(t, \tau) e^{-jk\omega\tau} d\tau. \quad (3.110)$$

The Fourier series in (3.109) and (3.110) is based on the same method as the Fourier series in (3.3) and (3.4) on page 34, which lead to the definition of the Chebyshev transform. However, unlike to the memoryless case, where the term $f(a \cos(\theta))$ is even-symmetric in θ for arbitrary functions $f(\cdot)$, there is no similar symmetry of $Y(t, \tau)$ in τ . Therefore, we have to use the complex Fourier series in (3.109) and integrate over one complete period in (3.110). The only simplification follows from the fact that $Y(t, \tau)$ is real, which allows us to combine the positive and negative frequency terms in (3.109).

By setting $\tau = t$ in (3.109), we obtain the Fourier series of $y(t)$, given by

$$y(t) = \frac{1}{2} \tilde{y}_0(t) + \sum_{k=1}^{\infty} \operatorname{Re} \left(\tilde{y}_k(t) e^{jk\omega t} \right). \quad (3.111)$$

The terms $\tilde{y}_k(t)$ in (3.111) describe the complex modulation of the k -th harmonic output of $y(t)$. They can be related to $\tilde{x}(t)$ by inserting (3.108) into (3.110), which leads us to the definition of the time-invariant operators $\tilde{\mathcal{N}}_k$ with $k \in \mathbb{N}_0$, given by

$$\tilde{\mathcal{N}}_k : \tilde{x}(t) \mapsto \tilde{y}_k(t) = \frac{2}{T} \int_{\tau_0}^{\tau_0+T} \mathcal{N} \left\{ \operatorname{Re} \left(\tilde{x}(t) e^{j\omega\tau} \right) \right\} e^{-jk\omega\tau} d\tau. \quad (3.112)$$

Equation (3.112) describes very concisely, how a k -th harmonic baseband model $\tilde{\mathcal{N}}_k\{\cdot\}$ can be derived from a time-invariant passband model $\mathcal{N}\{\cdot\}$. By substituting $\mathcal{N}\{\cdot\}$ with a specific passband model, like a memoryless polynomial or a Volterra series, the respective k -th harmonic baseband model $\tilde{\mathcal{N}}_k\{\cdot\}$ can be derived. A practical difficulty of (3.112) is, however, that the resulting expressions become very long, which is a problem for manual derivations. Because of this, we have used more specific methods for deriving the complex baseband models presented in the previous sections of this chapter.

Proof of the Phase Homogeneity Requirement

In the following, we use (3.112) to prove the phase homogeneity requirement, given by

$$\tilde{\mathcal{N}}_k \left\{ e^{j\xi} \tilde{x}(t) \right\} = e^{jk\xi} \tilde{\mathcal{N}}_k \left\{ \tilde{x}(t) \right\}. \quad (3.113)$$

Inserting the term $e^{j\xi} \tilde{x}(t)$ into $\tilde{\mathcal{N}}_k\{\cdot\}$, defined in (3.112), we get

$$\tilde{\mathcal{N}}_k \left\{ e^{j\xi} \tilde{x}(t) \right\} = \frac{2}{T} \int_{\tau_0}^{\tau_0+T} \mathcal{N} \left\{ \operatorname{Re} \left(e^{j\xi} \tilde{x}(t) e^{j\omega\tau} \right) \right\} e^{-jk\omega\tau} d\tau. \quad (3.114)$$

By defining $\xi = \omega\nu$, we can combine the exponentials $e^{j\xi}$ and $e^{j\omega\tau}$ in (3.114), such that

$$\tilde{\mathcal{N}}_k \left\{ e^{j\xi} \tilde{x}(t) \right\} = \frac{2}{T} \int_{\tau_0}^{\tau_0+T} \mathcal{N} \left\{ \operatorname{Re} \left(\tilde{x}(t) e^{j\omega(\tau+\nu)} \right) \right\} e^{-jk\omega\tau} d\tau. \quad (3.115)$$

Next, we change the variable of integration in (3.115) from τ to μ by defining $\mu = \tau + \nu$. Applying the substitution $\tau = \mu - \nu$ to (3.115), we get

$$\tilde{\mathcal{N}}_k \left\{ e^{j\xi} \tilde{x}(t) \right\} = \frac{2}{T} \int_{\mu_0}^{\mu_0+T} \mathcal{N} \left\{ \operatorname{Re} \left(\tilde{x}(t) e^{j\omega\mu} \right) \right\} e^{-jk\omega(\mu-\nu)} d\mu. \quad (3.116)$$

Pulling out the factor $e^{jk\omega\nu}$ from (3.116), we get

$$\tilde{\mathcal{N}}_k \left\{ e^{j\xi} \tilde{x}(t) \right\} = e^{jk\omega\nu} \frac{2}{T} \int_{\mu_0}^{\mu_0+T} \mathcal{N} \left\{ \operatorname{Re} \left(\tilde{x}(t) e^{j\omega\mu} \right) \right\} e^{-jk\omega\mu} d\mu \quad (3.117)$$

$$= e^{jk\xi} \tilde{\mathcal{N}}_k \left\{ \tilde{x}(t) \right\}, \quad (3.118)$$

which completes the proof of the phase homogeneity requirement stated in (3.113).

3.3.6 Discussion

Below, I summarize insights from the analysis presented above and discuss practical implications for the complex baseband modeling of RF-PAs and wireless transmitters.

The Baseband Volterra Series

The large majority of RF-PA behavioral models and digital predistorters is based on the baseband Volterra series. Above, I have presented a concise derivation of the baseband Volterra series with odd-order terms, followed by a derivation of even-order terms in baseband that correspond to modified even-order terms in passband. The numerical advantages of even-order terms in baseband are illustrated by simulations in [77].

The Phase Homogeneity Requirement

The phase homogeneity requirement derived in this chapter is a simple criterion for the design of complex baseband models of time-invariant passband systems. It is easy to verify that the odd-order baseband Volterra functionals in (3.81) and the even-order baseband Volterra functionals in (3.100) fulfill the phase homogeneity requirement. However, phase homogeneity is not limited to Volterra models, but also applies to other structures for RF-PA modeling like piecewise models [247] or neural networks [141].

Modeling of IQ Modulators

The phase homogeneity requirement does not apply to IQ modulators, since they are not time-invariant passband systems. Linear frequency-dependent IQ mismatch can be modeled and compensated by widely-linear models [11, 65, 175]. Nonlinear impairments of IQ modulators or joint impairments of IQ modulators and RF-PAs can be modeled by combining the widely-linear and baseband Volterra models [10, 48]. A Volterra model for general complex-valued signals can be defined by combining two real-valued Volterra models [37], or by generalizing the baseband Volterra model of RF-PAs [49].

Selection of Basis Functionals

A practical RF-PA behavioral model or digital predistorter typically uses only a subset of the baseband Volterra series basis functionals. The comparison in [212] showed that the generalized memory polynomial [163] features a good complexity-accuracy trade-off. However, also many other model structures based on the baseband Volterra series have been proposed [50, 51, 100, 101, 248]. The search for the best basis functionals is basically a task of combinatoric optimization. An efficient heuristic for this task is the greedy algorithm fast orthogonal search [7, 127]. Other heuristics for selecting the best basis functionals in RF-PA models can be found in [40, 104, 187, 235].

3.4 Summary and Conclusion

Summary

In this chapter, I presented a formal mathematical analysis of even-order terms in polynomial baseband models. By deriving explicit passband-baseband transformation pairs for the memoryless, the quasi-memoryless, and the Volterra model with even-order terms in baseband, I provided proofs that these terms are mathematically justified.

A further result of this chapter is the formulation and proof of the phase homogeneity requirement, which is a necessary symmetry of all complex baseband models of time-invariant passband systems. This symmetry explains the specific form of basis functionals in polynomial baseband models and provides a simple guideline for the design of other structures for the complex baseband modeling of RF-PAs.

Conclusion

The simple question regarding the theoretical justification of even-order terms in polynomial baseband models finally lead to a detailed investigation of the mathematical principles of complex baseband models. During this investigation, we saw that the work of Nelson Blachman in [25, 26] actually provides a solid mathematical basis for the theoretical justification of even-order terms in memoryless polynomial baseband models. However, working out all the relevant details of even-order terms in polynomial baseband models still required considerable analysis, which is documented in this chapter.

A fundamental insight from this chapter is that complex baseband modeling is, at its core, a special type of Fourier series analysis. The central idea of this analysis is to formulate a distorted passband signal as a two-dimensional signal, where one dimension models the random nature of the baseband modulation and the second dimension models the periodicity of the carrier. This way, the Fourier series framework can be applied to any distorted passband signal, even if the modulation is random and the passband signal is not periodic. It is interesting to note that exactly the same technique can also be applied to the analysis of pulse width modulation systems, like in [82, 114, 178].

4

Dual-Band Digital Predistortion of an Envelope-Tracking Power Amplifier

The annual student design competition "PA Linearization through DPD" at the IEEE International Microwave Symposium is a good opportunity for the benchmarking of DPD methods. In 2017, the DPD competition took place for the fourth time and the aim was to linearize an envelope tracking RF-PA under concurrent dual-band operation.

In this chapter, I present the winning solution of the 2017 DPD competition, which was published in the IEEE Microwave Magazine [81]. The main contribution is a collection of practical methods for the DPD of RF-PAs that were tested successfully on a realistic measurement setup against seven international competitors.

4.1 Introduction

The final evaluation of methods for the DPD of RF-PAs requires measurements with a real RF-PA in a realistic RF measurement setup. The achievable linearization accuracy of such a setup depends strongly on the characteristics of the RF-PA and the impairments of the signal generation and acquisition devices. As a result, a meaningful comparison can only be obtained, if the structures and algorithms are compared directly with the same RF-PA and the same RF measurement devices. The need for such a direct, measurement-based comparison creates high efforts for researchers who want to compare newly developed DPD methods with the state-of-the-art. On the one hand, the implementation of a reliable RF measurement setup is not trivial, and on the other hand, the implementation of reference algorithms proposed by other researchers requires great care, since a wrongly implemented reference algorithm would lead to false conclusions.

A remedy against the often insufficient comparability of DPD research results and the high effort of implementing a reliable RF measurement setup is the "WebLab" [136]. The WebLab is a remotely accessible RF measurement setup, which allows researchers from all over the world to freely access a common, high quality DPD test bench. The WebLab was developed by Chalmers University of Technology for the first student design competition "PA Linearization through DPD", which took place in 2014. After the first DPD competition, the WebLab was kept online as a permanent free service.

Since its introduction in 2014, the DPD competition has been established as an annual event at the IEEE International Microwave Symposium. The participants must be undergraduate or doctoral students and they can form teams of up to four members. The number of participating teams was five in 2014, three in 2015, three in 2016, and eight in 2017. Each year, the winning team is invited to submit an article to IEEE Microwave Magazine. The winning solutions up till now are described in [14, 81, 123, 145].

In 2017, the DPD competition was organized by Prof. Pere L. Gilabert, Prof. Gabriel Montoro and Mr. David López-Bueno from the Polytechnic University of Catalonia with support by Hermann Boss from Rohde & Schwarz. The aim was to linearize an envelope-tracking RF-PA under concurrent dual-band excitation, as described below.

4.1.1 Measurement Setup

The measurement setup of the 2017 DPD competition was located in a laboratory of the Polytechnic University of Catalonia in Barcelona. All measurements before and during the competition were conducted remotely over the internet, based on the upload and download of signal vectors. An overview of the setup is shown in Figure 4.1¹.

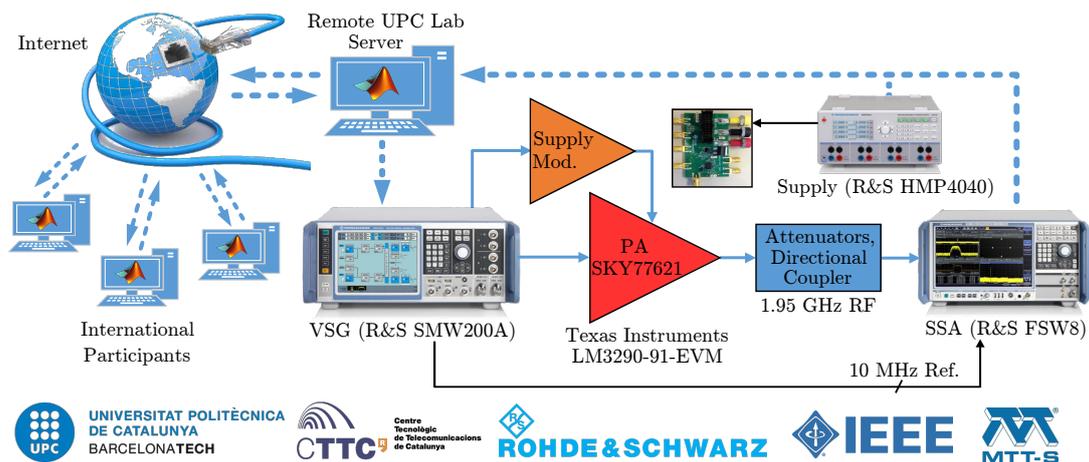


Figure 4.1: Measurement setup of the 2017 DPD competition.

The measurement setup in Figure 4.1 consists of a vector signal generator (VSG), a signal and spectrum analyzer (SSA), and a power supply from Rohde & Schwarz. The device-under-test is an evaluation board from Texas instruments, equipped with the supply modulator chips LM3290 and LM3291, and the RF-PA SKY77621 from Skyworks. For each measurement, the user must generate a complex baseband input signal and a real-valued supply signal and upload these signals to the server. After the measurement is completed, the user receives the corresponding complex baseband output signal.

¹Figure 4.1 is based on an illustration that was provided by Prof. Pere L. Gilabert.

4.1.2 Signal and Scoring

The MATLAB code for generating the complex baseband dual-band OFDM signal and the corresponding real-valued supply signal was provided by the organizers. The signals are generated at a sampling rate of 122.88 MHz, which is an integer multiple of the LTE subcarrier spacing of 15 kHz. Each signal contains 122880 samples, which corresponds to a duration of one millisecond. Plots of the signals are shown in Figure 4.2.

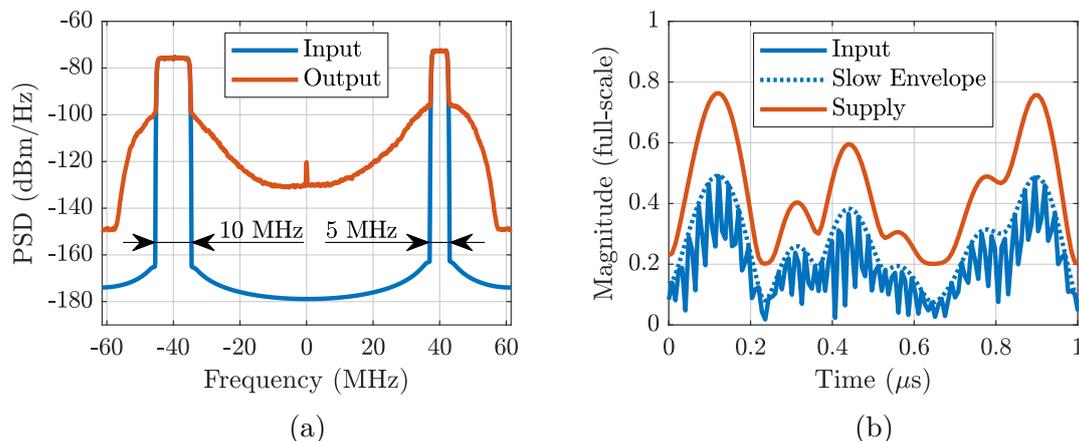


Figure 4.2: PSDs and waveforms of the dual-band OFDM signal.

Dual-Band OFDM Signal

The complex baseband input signal consists of two frequency bands with bandwidths of 10 MHz and 5 MHz and the same average power. The center frequencies of the two bands are separated by 80 MHz. The PSDs of the complex baseband input signal and the corresponding output signal without DPD are shown in Figure 4.2 (a). From this plot, we see that the RF-PA output signal contains high levels of out-of-band distortion. The time domain characteristics of the input signal are shown in Figure 4.2 (b). From this plot, we see that there are rapid variations in the signal's magnitude. These variations are a result of the concurrent dual-band nature of the input signal [189].

Supply Signal Generation

The supply signal is derived from the complex baseband input signal by a two-step process: First, a slow envelope signal is derived from the complex baseband input signal, and second, a memoryless shaping function is applied to the slow envelope signal to produce the supply signal. For generating the slow envelope signal, the concurrent dual-band input signal $u[n]$ is separated into two single-band signals $u_1[n]$ and $u_2[n]$ by filtering out the other frequency band. The slow envelope signal is computed by

$$u_{\text{env}}[n] = |u_1[n]| + |u_2[n]|. \quad (4.1)$$

Figure 4.2 (b) shows that this slow envelope tracks the peaks of the dual-band signal.

The purpose of the second step in the supply signal generation process is to map the slow envelope signal to a defined value range. For this purpose, it is normalized by

$$u_{\text{env,norm}}[n] = \frac{u_{\text{env}}[n]}{\max(u_{\text{env}}[n])}, \quad (4.2)$$

which ensures that the maximum of $u_{\text{env,norm}}[n]$ is one. The normalized signal is then sent through a memoryless shaping function to produce the supply signal, given by

$$z[n] = (z_{\text{min}}^6 + u_{\text{env,norm}}[n])^{1/6}. \quad (4.3)$$

This shaping function ensures that the minimum of the supply signal $z[n]$ is larger or equal to z_{min} . Figure 4.2 (b) shows that the produced supply signal always stays above $z_{\text{min}} = 0.2$. Alternative methods for the supply signal generation are summarized [133].

Score Computation

The MATLAB code for analyzing the RF-PA output signal and computing the score was provided by the organizers. The score is based on four criteria: adjacent channel power ratio (ACPR), normalized mean square error (NMSE), average output power, and average drain efficiency. An overview of the target values for the respective metrics and the weighting factors for the score computation is shown in Table 4.1.

	Target	Score	Conditions
ACPR	-45 dB	1 point/dB (*)	—
NMSE	-33 dB	0.5 point/dB (*)	(1)
Power	20 dBm	10 points/dBm (*)	(1) (2) (3)
Efficiency	—	1 point/percent	(1) (2) (3)
(*) Relative to target value. (1) ACPR target is achieved. (2) NMSE target is achieved. (3) Power difference between bands is within ± 0.5 dBm.			

Table 4.1: Score specification of the 2017 DPD competition.

From the weighting factors in Table 4.1, we see that the average output power has the highest influence on the score. However, due to the conditions on the score computation, the points for the average output power only count if the linearity requirements of ACPR and NMSE and the required power difference between the two bands are fulfilled. Failing in any of these requirements immediately leads to a score near to zero.

4.1.3 Linearization Architecture

A concurrent dual-band transmitter requires a different DPD architecture than a single-band transmitter. If the nonlinearity of a DPD is applied directly onto a discrete-time dual-band signal, the DPD output signal contains intermodulation products of the two frequency bands, which are typically located in higher Nyquist zones and therefore show up as aliasing components in the fundamental Nyquist zone. To prevent the generation of these intermodulation products, the two frequency bands must be separated and processed by a combination of two dual-input, dual-output, nonlinear systems.

To separate the two frequency bands, we transformed the dual-band signal into the frequency domain, applied frequency translation, filtering, and decimation in the frequency domain, and converted the respective single-band signals back to the time domain. For the single-band signals, we used a center frequency of zero and a sampling rate of 80 MHz. This allows to linearize the complete frequency range of the dual-band signal.

An overview of the used linearization architecture is shown in Figure 4.3. It follows the common practice of separating the digital processing before the PA into crest factor reduction (CFR) and digital predistortion (DPD). The purpose of CFR is to limit the peaks of the input signal, such that the operating range of the DPD is clearly defined. Within this range, the DPD ideally implements the inverse PA transfer characteristic, such that the PA output signal is a scaled replica of the CFR signal.

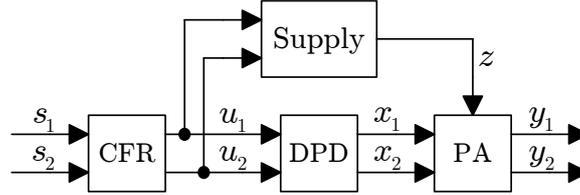


Figure 4.3: Block diagram of the linearization architecture.

Figure 4.3 shows that the supply signal is derived from the CFR signal. We chose the CFR signal as the basis for the supply signal generation, because the supply signal should track the instantaneous envelope of the PA output signal, which, after the DPD is operating, is very similar to the CFR signal. Initially, we also experimented with a DPD architecture where the supply signal is used as a third input to the DPD, as proposed in [103], but in the present setup, this did not improve the performance.

An important detail concerning the supply signal is that it reaches the PA over a different path than the RF input signal. Any delay mismatch between these paths deteriorates the linearity. To equalize the delay mismatch, we applied a time-shift on the supply signal and performed a line search over the time-shift parameter to optimize the ACPR.

4.2 Crest Factor Reduction

The large number of CFR methods outlined in [111] can be categorized into:

1. methods that add a certain amount of tolerable distortion to the signal, and
2. methods that do not add distortion, but modify the signal generation process.

Since for the competition, the signal generation was fixed, a method of the first type had to be used. Most methods of the first type use some sort of clipping and filtering. The clipping reduces the crest factor, and the filtering trades off more inband distortion for less out-of-band distortion. The reduction of out-of-band distortion is important, because clipping distortion is very broadband and most standards impose stricter limits on the out-of-band distortion levels as on the inband distortion levels.

4.2.1 Structure of the Crest Factor Reducer

For the competition, we used the dual-band CFR structure in Figure 4.4.

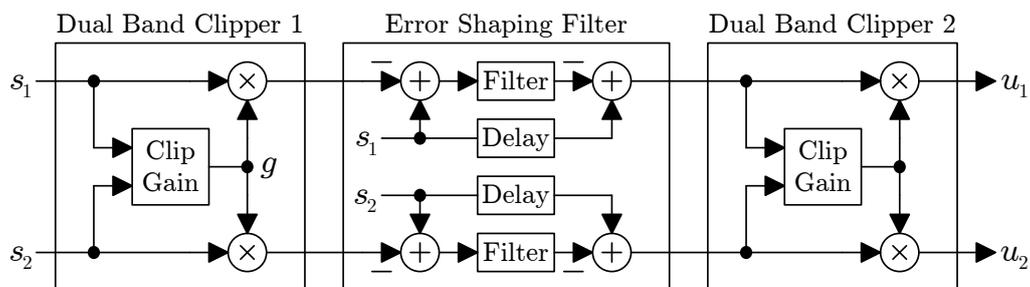


Figure 4.4: Block diagram of the dual-band CFR structure.

Figure 4.4 represents a combination of dual-band clipping [41] and error filtering [126]. Error filtering was also used by previous winners of the DPD competition [123, 145].

The main idea of dual-band clipping is that the peak magnitude of the dual-band signal can be limited by applying a time-varying gain on the single-band signals. The required gain depends on the peak-magnitude of the dual-band signal, which is estimated by

$$s_{\text{env}}[n] = |s_1[n]| + |s_2[n]|. \quad (4.4)$$

If $s_{\text{env}}[n]$ is larger than the predefined limit u_{max} , the magnitude must be reduced. The gain factor that implements the required magnitude reduction is given by

$$g[n] = \begin{cases} \frac{u_{\text{max}}}{s_{\text{env}}[n]} & s_{\text{env}}[n] > u_{\text{max}}, \\ 1 & \text{otherwise.} \end{cases} \quad (4.5)$$

After the first dual-band clipper, Figure 4.4 contains an error shaping filter, which computes the clipping error, applies a lowpass filter, and then recombines the filtered clipping error with the original signal. Applying the lowpass filter on the error signal instead of the clipped signal gives more flexibility for the filter design.

Due to the filtering, there is a certain amount of peak-regrowth. To get a clearly defined maximum magnitude, we added another clipper after the error shaping filter. By setting the clipping threshold of the second clipper to the same value as in the first one, the clipping distortion of the second clipper could be tolerated without further filtering.

4.2.2 Design of the Crest Factor Reducer

After defining the structure of the CFR, the clipping threshold and the lowpass filter had to be designed. The aim for this design is to achieve a low crest factor of the output signal while fulfilling the requirements of Table 4.1 with sufficient margin.

For the design of the lowpass filter, we used the function `fir2` from MATLAB, which designs a linear-phase finite impulse response (FIR) filter by using an inverse Fourier transform and a Hamming window. We specified that the filter should have a passband-width equal to the signal bandwidth, an out-of-band suppression of 40 dB, and a memory depth of 80 samples. To select the clipping threshold, we performed Monte Carlo simulations based on 100 random signal realizations. After several iterations, we selected a clipping threshold that produces an output crest factor of 8.6 dB.

The PSDs of the CFR input, output, and error signals are shown in Figure 4.5 (a). This plot shows that the filter shapes the clipping error, but does not affect the desired signal. From the results of the Monte Carlo simulation in Figure 4.5 (b), we see that there is sufficient margin to the ACPR limit of -45 dB and the NMSE limit of -33 dB.

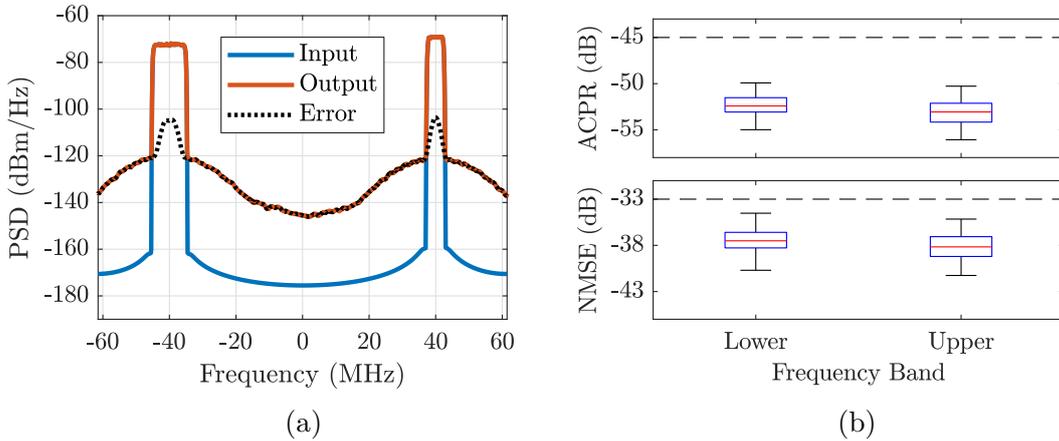


Figure 4.5: PSDs and linearity metrics after the dual-band CFR.

4.3 Digital Predistortion

The design of a DPD can be divided into the following two steps:

1. designing the structure of the nonlinear model, and
2. designing the algorithm for training the nonlinear model.

4.3.1 Structure of the Digital Predistorter

The majority of DPD structures are linear in the coefficients, because this property allows the use of efficient training algorithms based on solving systems of linear equations. For concurrent dual-band systems, such a model can be written as

$$x_1[n] = \sum_{i=1}^{I_1} c_{1,i} \Phi_{1,i}\{u_1[n], u_2[n]\}, \quad (4.6)$$

$$x_2[n] = \sum_{i=1}^{I_2} c_{2,i} \Phi_{2,i}\{u_1[n], u_2[n]\}, \quad (4.7)$$

where $c_{1,i}$, $c_{2,i}$ are the coefficients, $\Phi_{1,i}\{\cdot\}$, $\Phi_{2,i}\{\cdot\}$ are the basis functionals, i enumerates the basis functionals, and I_1 , I_2 are the numbers of coefficients and basis functionals per frequency band. For the competition, we used basis functionals of the form

$$\Phi_{b,i}\{u_1[n], u_2[n]\} = u_b[n - m_0] |u_1[n - m_1]|^{p_1} |u_2[n - m_2]|^{p_2}, \quad (4.8)$$

which are inspired by the generalized memory polynomial (GMP) [163]. With this type of basis functionals, the top-level DPD structure can be drawn as in Figure 4.6 (a).

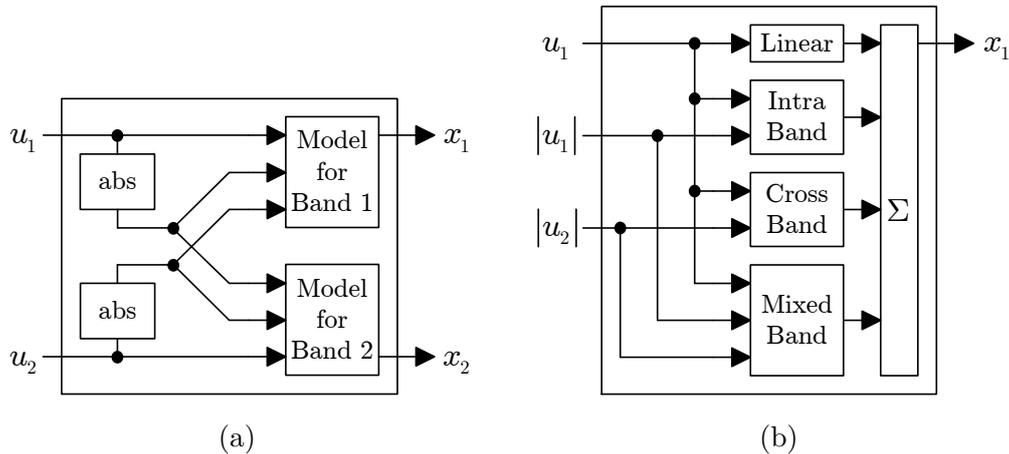


Figure 4.6: Block diagram of the dual-band DPD structure.

To further specify the model, we divided the basis functionals into four groups: 1) linear basis functionals where both p_1 and p_2 are zero, 2) intra-band and 3) cross-band basis functionals where either p_1 or p_2 is zero, and 4) mixed-band basis functionals where both p_1 and p_2 are larger than zero. Using this classification, the model for band one can be drawn like in Figure 4.6 (b). Below, I specify the behavior of each of these blocks in terms of equations. The model for band two has the same structure as the model for band one and is obtained by swapping the band indices of all variables.

The linear block contains a linear FIR filter given by

$$x_1^{(\text{linear})}[n] = \sum_{m=0}^M c_m u_1[n-m]. \quad (4.9)$$

The intra-band block contains a modified GMP, given by

$$\begin{aligned} x_1^{(\text{intra})}[n] = & \sum_{m=0}^{M_1} \sum_{p=1}^{P_1[m]} c_{m,p}^{(1)} u_1[n-m] |u_1[n-m]|^p \\ & + \sum_{k=1}^{K_2} \sum_{m=0}^{M_2[k]} \sum_{p=1}^{P_2[k,m]} c_{k,m,p}^{(2)} u_1[n-m] |u_1[n-m-k]|^p \\ & + \sum_{k=1}^{K_3} \sum_{m=0}^{M_3[k]} \sum_{p=1}^{P_3[k,m]} c_{k,m,p}^{(3)} u_1[n-m-k] |u_1[n-m]|^p. \end{aligned} \quad (4.10)$$

Compared to the original GMP [163], (4.10) contains the following modifications:

- The linear part is excluded, since it is modeled by a separate block.
- The third line of (4.10) is modified, such that only causal terms are included.
- The upper limits of the summations depend on the indices of previous sums.

The third modification allows the use of different orders for different memory depths to heavily prune the model without sacrificing too much accuracy. An illustration of the sample lags used by (4.10) with respect to m_0 and m_1 from (4.8) is shown in Figure 4.7.

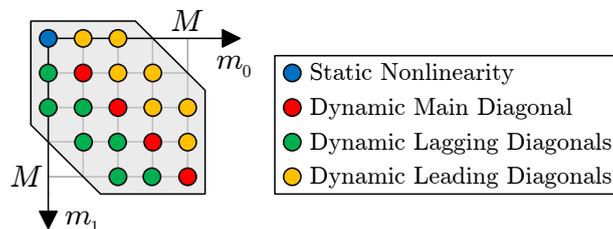


Figure 4.7: Kernel structure of the modified GMP.

The cross-band block contains the same model as the intra-band block. Its equation is obtained from (4.10) by using u_2 instead of u_1 inside the magnitude operators.

The mixed-band block contains a modified 2D memory polynomial [17], given by

$$x_1^{(\text{mixed})}[n] = \sum_{m=0}^M \sum_{p_1=1}^{P_1[m]} \sum_{p_2=1}^{P_2[m,p_1]} c_{m,p_1,p_2} u_1[n-m] |u_1[n-m]|^{p_1} |u_2[n-m]|^{p_2}. \quad (4.11)$$

To further increase the accuracy of DPD models, several authors have proposed piecewise models [6, 13, 166, 192, 247]. These models circumvent the problems of high-order basis functionals by using low-order basis functionals with higher locality. Most of them are based on splines, which enforce a certain level of continuity between the segments [135]. An exception is the vector-switched model [6] that does not enforce continuity. Dropping the continuity constraint has two advantages: 1) it simplifies the model formulation since any model can be used for the individual segments, and 2) it improves the numerical properties since the support of each basis functional is confined to the modeling region of the respective segment. Potential problems due to discontinuities have not been observed when the model is trained by least squares fitting with long signal vectors [6].

For the competition, we adapted the vector-switched technique to the dual-band case, which lead to the 2D-vector-switched model in Figure 4.8. Each of the regional models in Figure 4.8 (a) contains the dual-band DPD structure discussed above. The switching between the regional models is based on the magnitude of the current sample in each of the frequency bands. For mapping sample magnitudes to modeling regions, we used a vector quantizer with manually selected centroids as shown in Figure 4.8 (b). In the overall model, we included eight identical regional models, each having 72 coefficients per frequency band. By using identical regional models, the vector-switched model can be implemented efficiently by a single regional model with switched coefficients.

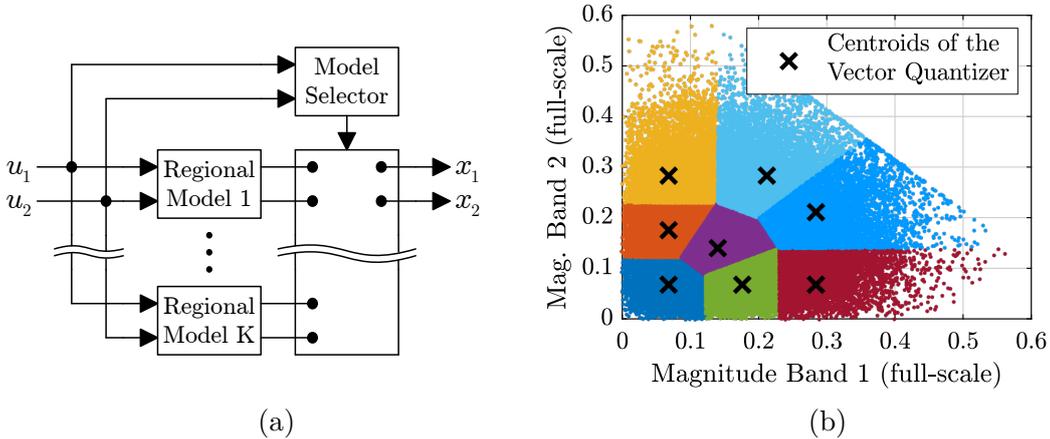


Figure 4.8: Block diagram of the 2D-vector-switched model and its modeling regions.

4.3.2 Training of the Digital Predistorter

Two architectures are commonly used to train digital predistorters: 1) indirect learning and 2) direct learning. Both originate from earlier work on adaptive control [177]. In the following, I give a short introduction to these architectures, followed by a description of the training algorithm for the 2D-vector-switched DPD used in the competition.

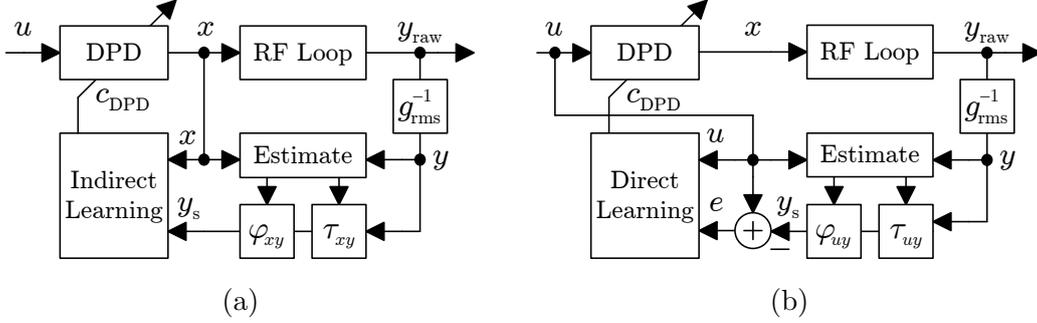


Figure 4.9: Block diagrams of the indirect and direct learning architectures.

Indirect Learning

In the indirect learning architecture of Figure 4.9 (a), the complex baseband input and output signals of the RF loop are used to train the DPD. Before the actual training, the output signal y_{raw} is synchronized to the input signal x . This consists of magnitude scaling, delay correction and phase correction. The magnitude scaling determines the output power of the linearized system. By using the RMS gain of the RF loop with deactivated DPD in all iterations of training, it is ensured that the output power with DPD converges to the same value as without DPD. The time delay τ_{xy} and the phase rotation φ_{xy} are newly estimated in each iteration based on the complex cross-correlation r_{xy} , as described in Appendix E. Using the synchronized signals x and y_s , we can set up a linear equation system for the post-inverse model of the RF loop, given by

$$\mathbf{x} = \mathbf{Y}_s \mathbf{c}_{\text{DPD}}, \quad (4.12)$$

where \mathbf{x} is the vector of input signal samples, \mathbf{Y}_s is the matrix formed by sending the synchronized output signal through the nonlinear basis functionals of the DPD model, and \mathbf{c}_{DPD} is the vector of model coefficients. In the indirect learning block of Figure 4.9 (a), the equation in (4.12) is set up and solved by the method of least squares, as described in Appendix D. Afterwards, the obtained coefficients are used in the DPD block, which maps the original input signal u to the predistorted input signal x_{DPD} by

$$\mathbf{x}_{\text{DPD}} = \mathbf{U} \mathbf{c}_{\text{DPD}}. \quad (4.13)$$

A serious drawback of indirect learning is that solving (4.12) finds the optimal post-inverse model, which is in general different from the optimal pre-inverse model.

Direct Learning

The goal of direct learning is to directly minimize the predistortion error by using an architecture like it is shown in Figure 4.9 (b). For synchronizing the output signal y_{raw} to the DPD input signal u , the same methods are used as with indirect learning. From the synchronized signals, the error signal is computed by

$$\mathbf{e} = \mathbf{u} - \mathbf{y}_s. \quad (4.14)$$

This error represents the deviation of y_{raw} from an ideal output signal, which is the original input signal u , multiplied by g_{rms} , and with arbitrary delay and phase rotation. To adapt the DPD, we require the deviation of x from the ideal predistorted signal. Since the RF loop introduces distortion, the deviations of x and y_{raw} from their respective ideal signals are in general different. However, this difference is often neglected, and the error signal is related to the error in the DPD coefficients by [32, 66, 108]

$$\mathbf{e} = \mathbf{U} \mathbf{c}_{\text{error}}. \quad (4.15)$$

In the direct learning block of Figure 4.9 (b), the equation in (4.15) is set up and solved by the method of least squares. Afterwards, the DPD coefficients are updated by

$$\mathbf{c}_{\text{DPD}}^{(i+1)} = \mathbf{c}_{\text{DPD}}^{(i)} + \mu \mathbf{c}_{\text{error}}^{(i)}, \quad (4.16)$$

where i is the iteration index and μ is the step size. The experimental evidence in [108] shows that despite the strong simplification required for (4.15), the resulting algorithm converges to better results than indirect learning. There are more advanced versions of direct learning, which do not rely on this simplification [113, 138, 139], however, it remains to be shown whether these algorithms actually converge to better solutions.

Training of the 2D-Vector-Switched DPD

With the 2D-Vector-Switched DPD structure, the training must be applied to each of the two frequency bands and to each of the eight modeling regions. This requires the solution of 16 equation systems per training iteration. In each training iteration, the coefficients can be updated either with indirect or with direct learning. For the competition, we used indirect learning in the first iteration and direct learning in subsequent iterations. For synchronizing the dual-band signals, we applied the single-band synchronization methods discussed in Appendix E separately to each frequency band. After synchronizing the signals, we set up a single equation system for each frequency band and subdivided this equation system into eight smaller equation systems, based on the magnitude of the current input sample in each frequency band. After solving these equation systems, we updated the coefficients of the respective regional models. During the competition, we used 19 iterations of training. After ten iterations of training, the linearization performance was already close to the final performance.

4.4 Measurement Results

At the competition, each team had 20 minutes for training. After this period, a new realization of the input signal was used to evaluate the score.

Our solution achieved a score of 71.8 points, which corresponds to

- an ACPR of -49.2 dB,
- an NMSE of -35.7 dB,
- an output power of 24.4 dBm, and
- a drain efficiency of 22.3 %.

Higher drain efficiencies, such as reported in [103], can be reached with a dynamic bias voltage, as discussed in [152, Part 6].

Plots of the linearization performance are shown in Figure 4.10. From Figure 4.10 (a), we see that after ten iterations of training, the DPD reduces the out-of-band distortion to levels close to the CFR-induced distortion. The scatter plot in Figure 4.10 (b) shows the AM-AM and AM-PM characteristics of the lower frequency band, before and after linearization. This plot shows that the strong dispersion, caused by memory effects and intermodulation of the two frequency bands, is well compensated by the DPD.

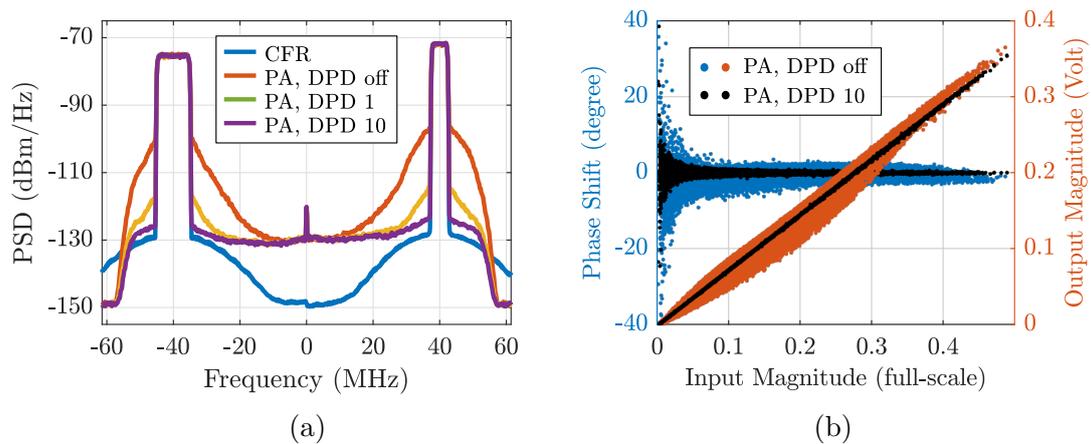


Figure 4.10: Linearization performance in the frequency and time domain.

4.5 Summary and Conclusion

Summary

In this chapter, I presented the winning solution of the 2017 DPD competition, held at the IEEE International Microwave Symposium. The main ingredients were:

- a dual-band CFR based on clipping and filtering,
- a dual-band DPD based on a modified GMP,
- the 2D-vector-switched technique, and
- direct learning with several iterations.

These ingredients combine several state-of-the-art methods for CFR and DPD with new extensions we have developed specifically for the 2017 DPD competition.

Conclusion

Over the last four years, the DPD competition provided reliable and objective evidence regarding best practices in the design of digital predistorters. By comparing the winning solutions [14, 81, 123, 145], it can be observed that three of them used CFR based on clipping and error filtering, and all of them used DPD training based on several iterations of the simple direct learning algorithm described above. The used DPD structures are more diverse, since they provide greater freedom for fine-tuning to the respective RF-PA and problem statement. However, a common feature of the last three winning solutions is that the DPD structures were based on some form of pruned baseband Volterra model, combined with some form of piecewise extension for higher locality.

5

Summary and Conclusion

In this chapter, I summarize the main results and contributions of this thesis and compare them with the related doctoral theses that are summarized in the first chapter.

5.1 Summary of Results and Contributions

Results

The main results of this thesis are:

1. A memoryless semi-physical RF-PA model.
2. Explicit passband-baseband pairs of polynomial models, including
 - a) the quasi-memoryless polynomial with even-order terms, and
 - b) the baseband Volterra series with even-order terms.
3. The phase homogeneity requirement for baseband models of RF-PAs.
4. A collection of practical methods for dual-band CFR and DPD.

Contributions

The results summarized above advance the state-of-the-art in the following aspects:

1. The RF-PA model extends the conventional efficiency analysis by a more accurate transistor model, enabling the joint modeling of linearity and efficiency.
2. The passband-baseband pairs of polynomial models provide a solid mathematical foundation for using even-order terms in polynomial baseband models.
3. The phase homogeneity requirement provides for the first time a simple justification for the specific structure of nonlinear complex baseband models of RF-PAs.
4. The CFR and DPD methods used for winning the 2017 DPD competition extend single-band models and methods for the use in concurrent dual-band systems.

5.2 Comparison with Related Doctoral Theses

A comparison of doctoral theses may include several aspects, like

1. the originality of the presented ideas,
2. the practical relevance of the presented results,
3. the clarity of the presentation, and
4. the number of publications and citations.

In the following, I exemplify these aspects based on the related doctoral theses, listed in Section 1.4 on page 5. Afterwards, I exemplify them with respect to my own thesis.

Discussion of Related Theses

Originality is rather subjective. In my opinion, it can be exemplified by Peter Singerl's method for fitting memory polynomial models based on dual-tone responses [203]. Model fitting based on single-tone [193] and dual-tone [130] responses was proposed before, but the originality of Peter Singerl's method is that the model fitting is formulated as a linear least squares problem, similar to conventional model fitting in the time domain.

Practical relevance can be exemplified by research results that present robust solutions to commonly encountered problems or give well-founded answers to commonly asked questions. In my opinion, the methods presented by Lei Ding in [65, 70] and the comparison presented by Ali Soltani Tehrani in [212] are of great practical relevance.

Clarity of presentation includes a well-designed thesis structure, a clear statement of the problems, solutions, and contributions, and a well-defined mathematical notation. In my opinion, the thesis of Lei Ding [63] fulfills these criteria very well.

The number of publications and citations is a quantitative measure for the scientific impact of a thesis. The thesis of Lei Ding [63] is the earliest in the present comparison and the corresponding publications received the highest number of citations.

Discussion of This Thesis

Personally, I think the most original idea within my own thesis is the formulation and proof of the phase homogeneity requirement. The results of highest practical relevance are probably the CFR and DPD methods of Chapter 4, since they were validated by measurements against several international competitors. The clarity of presentation is best judged by the reader, not by the author. This thesis is based on four conference papers [76, 77, 79, 80] and one magazine paper [81]. Additionally, I authored two conference papers [78, 82] and contributed to four co-authored papers [95–98].

Appendices

A

Complex Signals

Wireless communication systems exchange information by modulating the parameters of an RF carrier signal like its amplitude, frequency, or phase. Modern spectrally efficient communication systems jointly modulate the amplitude and phase, which is commonly described by a complex baseband signal. In this appendix, I discuss basic properties of such signals and show how to convert between the polar and Cartesian representation.

Polar and Cartesian Form of Complex Signals

A complex baseband signal $\tilde{x}(t)$ can be represented in polar form by

$$\tilde{x}(t) = a(t) e^{j\varphi(t)}, \quad (\text{A.1})$$

or it can be represented in Cartesian form by

$$\tilde{x}(t) = i(t) + jq(t). \quad (\text{A.2})$$

The real-valued signals $a(t)$, $\varphi(t)$, $i(t)$, and $q(t)$ are called the amplitude signal, the phase signal, the in-phase signal, and the quadrature-phase signal, respectively. If we modulate the complex baseband signal $\tilde{x}(t)$ onto a carrier of frequency f_c , as described in Appendix B, the signals $a(t)$ and $\varphi(t)$ directly represent the amplitude and phase of the modulated carrier signal. The signals $i(t)$ and $q(t)$ represent the amplitudes of the in-phase and quadrature-phase modulated carrier, which as a sum represent the modulated carrier signal. There are two important advantages of the Cartesian form:

- The frequency support of $i(t)$ and $q(t)$ does not exceed the frequency support of $\tilde{x}(t)$. This is especially important for wideband applications.
- The processing structures for $i(t)$ and $q(t)$ are identical. This simplifies their implementation and allows the matching of component values.

The polar form is only unique if the value ranges of $a(t)$ and $\varphi(t)$ are restricted. By defining $a(t) = |\tilde{x}(t)|$, we obtain $0 \leq a(t)$ and $\varphi_0 \leq \varphi(t) < \varphi_0 + 2\pi$. Alternatively, however, we can allow negative values of $a(t)$ and demand that $\varphi_0 \leq \varphi(t) < \varphi_0 + \pi$.

Conversion between Polar and Cartesian Form

The conversion from polar to Cartesian form is given by

$$i(t) = a(t) \cos(\varphi(t)), \quad (\text{A.3})$$

$$q(t) = a(t) \sin(\varphi(t)), \quad (\text{A.4})$$

and the conversion from Cartesian to polar form is given by

$$a(t) = \sqrt{i^2(t) + q^2(t)}, \quad (\text{A.5})$$

$$\varphi(t) = \text{atan2}(q(t), i(t)), \quad (\text{A.6})$$

where $\text{atan2}(\cdot)$ is the 2-input, 4-quadrant inverse tangent, defined by

$$\text{atan2}(q, i) = \begin{cases} \arctan\left(\frac{q}{i}\right) & i > 0, \\ \arctan\left(\frac{q}{i}\right) + \pi & i < 0 \text{ and } q \geq 0, \\ \arctan\left(\frac{q}{i}\right) - \pi & i < 0 \text{ and } q < 0, \\ +\frac{\pi}{2} & i = 0 \text{ and } q > 0, \\ -\frac{\pi}{2} & i = 0 \text{ and } q < 0, \\ \text{undefined} & i = 0 \text{ and } q = 0. \end{cases} \quad (\text{A.7})$$

An efficient method for implementing (A.3) to (A.6) is the coordinate rotation digital computer (CORDIC) algorithm [158, 217], which only requires comparison, bit shift, and add operations, converging to the true result with one bit per iteration.

Alternatively, we may use lookup tables of the functions $\sin(\cdot)$, $\cos(\cdot)$, $\sqrt{\cdot}$, and $\arctan(\cdot)$. For (A.3) and (A.4), we require one quarter period of a sinusoid. For (A.5), we ensure $|q(t)| \leq |i(t)|$ by a possible swap of $i(t)$ and $q(t)$ and compute the amplitude by

$$a(t) = |i(t)| \sqrt{1 + \left(\frac{q(t)}{i(t)}\right)^2}, \quad (\text{A.8})$$

which requires a lookup table of the square root function for inputs between 0 and 2. For (A.6), we may implement (A.7), by a lookup table of $\arctan(\cdot)$ for inputs between 0 and 1, since $\arctan(1/x) = \pi/2 - \arctan(x)$ for $x \geq 0$ and $\arctan(-x) = -\arctan(x)$.

A low complexity approximation of (A.5), which is especially relevant for the high speed implementation in a lookup-table-based DPD is given by [5, 89]

$$a(t) = \begin{cases} \alpha_1 A + \beta_1 B & A > K B, \\ \alpha_2 A + \beta_2 B & \text{otherwise,} \end{cases} \quad \text{where} \quad \begin{cases} A = \max(|i(t)|, |q(t)|), \\ B = \min(|i(t)|, |q(t)|). \end{cases} \quad (\text{A.9})$$

where α_1 , β_1 , α_2 , β_2 , and K are constants that may be selected such that (A.9) can be implemented using only comparison, bit shift, and add operations. For example with $\alpha_1 = 1$, $\beta_1 = 0$, $\alpha_2 = 7/8$, $\beta_2 = 1/2$, $K = 4$, the maximum error is 2.98 % [5, 89].

B

Modulation and Demodulation

RF-PAs are operated with real passband signals $x(t)$ that have a frequency support of $[-f_c - B/2, -f_c + B/2] \cup [+f_c - B/2, +f_c + B/2]$, where f_c is the carrier frequency, B is the bandwidth, $[f_1, f_2]$ is a frequency range defined by $\{f \in \mathbb{R} \mid f_1 \leq f \leq f_2\}$ and \cup is the union operator. Such signals can be represented equivalently by complex baseband signals $\tilde{x}(t)$ that have a frequency support of $[-B/2, +B/2]$. The mapping $\tilde{x}(t) \mapsto x(t)$ is called complex modulation and the mapping $x(t) \mapsto \tilde{x}(t)$ is called complex demodulation. In this appendix, I present a mathematical description of these processes.

Complex Modulation

At complex modulation, the complex baseband signal $\tilde{x}(t)$ is frequency-translated to the carrier frequency f_c by multiplying it with $e^{j\omega t}$, where $\omega = 2\pi f_c$, resulting in

$$\hat{x}(t) = \tilde{x}(t) e^{j\omega t}. \quad (\text{B.1})$$

The frequency support of $\hat{x}(t)$ is $[+f_c - B/2, +f_c + B/2]$. If $B < 2f_c$, the signal $\hat{x}(t)$ contains only positive frequency components, which means it is an analytic signal [19]. The real passband signal $x(t)$ is the real part of the analytic signal [150], given by¹

$$x(t) = \text{Re}(\tilde{x}(t) e^{j\omega t}). \quad (\text{B.2})$$

Inserting the polar form $\tilde{x}(t) = a(t) e^{j\varphi(t)}$ into (B.2), we get

$$x(t) = \text{Re}\left(a(t) e^{j(\omega t + \varphi(t))}\right) \quad (\text{B.3})$$

$$= a(t) \cos(\omega t + \varphi(t)). \quad (\text{B.4})$$

Inserting the Cartesian form $\tilde{x}(t) = i(t) + j q(t)$ into (B.2), we get

$$x(t) = \text{Re}\left((i(t) + j q(t)) (\cos(\omega t) + j \sin(\omega t))\right) \quad (\text{B.5})$$

$$= i(t) \cos(\omega t) - q(t) \sin(\omega t). \quad (\text{B.6})$$

¹From (B.2), it follows that $\max\{x(t)\} = \max\{|\tilde{x}(t)|\}$ and $\text{rms}\{x(t)\} = \text{rms}\{|\tilde{x}(t)|\}/\sqrt{2}$. As a result, the PAPR of $x(t)$ is 3 dB higher than the PAPR of $\tilde{x}(t)$. (PAPR is defined in (2.28))

Complex Demodulation using the Hilbert Transform

For complex demodulation, (B.2) must be inverted. For this purpose, we may reconstruct the analytic signal $\hat{x}(t)$ from the passband signal $x(t)$, and then multiply it with a complex exponential signal to frequency translate it back to baseband by

$$\tilde{x}(t) = \hat{x}(t) e^{-j\omega t}. \quad (\text{B.7})$$

To reconstruct $\hat{x}(t)$ from $x(t)$, we have to invert $x(t) = \text{Re}(\hat{x}(t))$. If $\hat{x}(t)$ is an arbitrary complex signal with non-zero imaginary part, the real part operation is not invertible since the information of the imaginary part is lost. However, since $\hat{x}(t)$ is an analytic signal, we can reconstruct its imaginary part from its real part. To see how this works, we derive expressions for the imaginary part, $\hat{x}(t) = \text{Im}(\hat{x}(t))$, similar as we did it for the real part in (B.3)-(B.6). Using the polar form of $\tilde{x}(t)$, we get

$$\hat{x}(t) = \text{Im} \left(a(t) e^{j(\omega t + \varphi(t))} \right) \quad (\text{B.8})$$

$$= a(t) \sin(\omega t + \varphi(t)). \quad (\text{B.9})$$

and using the Cartesian form of $\tilde{x}(t)$, we get

$$\hat{x}(t) = \text{Im} \left((i(t) + jq(t)) (\cos(\omega t) + j \sin(\omega t)) \right) \quad (\text{B.10})$$

$$= i(t) \sin(\omega t) + q(t) \cos(\omega t). \quad (\text{B.11})$$

By comparing (B.9) and (B.11) with (B.4) and (B.6), we notice that the expressions for $\hat{x}(t)$ differ from the expressions for $x(t)$ only by a phase shift of the carrier by 90 degree. An unmodulated carrier is phase shifted by 90 degree, if the signal is delayed by $\tau = T_c/4$, where $T_c = 1/f_c$ is the carrier period. To reconstruct $\hat{x}(t)$ from $x(t)$, we have to phase shift the modulated carrier, without delaying the modulating signal components $a(t)$, $\varphi(t)$, $i(t)$, or $q(t)$. This is achieved by the Hilbert transform, which is defined by

$$\mathcal{H} : x(t) \mapsto \hat{x}(t) = \mathcal{F}^{-1} \left\{ -j \text{sign}(f) \mathcal{F} \{ x(t) \} \right\}, \quad (\text{B.12})$$

where \mathcal{F} is the Fourier transform and \mathcal{F}^{-1} is the inverse Fourier transform. The Hilbert transform is a linear operator, which has a frequency response given by

$$H(f) = \begin{cases} +j = e^{+j\frac{\pi}{2}} & f < 0, \\ 0 & f = 0, \\ -j = e^{-j\frac{\pi}{2}} & f > 0. \end{cases} \quad (\text{B.13})$$

Using the Hilbert transform, we can reconstruct $\hat{x}(t)$ from $x(t)$ by

$$\hat{x}(t) = x(t) + j \mathcal{H} \{ x(t) \}. \quad (\text{B.14})$$

Complex Demodulation using an IQ Demodulator

Complex demodulation using the Hilbert transform is a useful mathematical concept, but it is not well suited for a practical implementation. A method that is better suited for a practical implementation is IQ demodulation. In an IQ demodulator, the signals $i(t)$ and $q(t)$ are obtained from the passband signal $x(t)$, by multiplying $x(t)$ with the RF carrier basis functions $\cos(\omega t)$ and $-\sin(\omega t)$, respectively, followed by lowpass filtering. In the following, we prove that this actually results in the correct baseband signal components. For this purpose, we require the trigonometric identities

$$\sin(\theta) \cos(\theta) = \frac{\sin(2\theta)}{2}, \quad (\text{B.15})$$

$$\sin^2(\theta) = \frac{1 - \cos(2\theta)}{2}, \quad (\text{B.16})$$

$$\cos^2(\theta) = \frac{1 + \cos(2\theta)}{2}. \quad (\text{B.17})$$

By evaluating the multiplication of $x(t)$ and $\cos(\omega t)$, we get

$$i_{\text{rx}}(t) = \left(i(t) \cos(\omega t) - q(t) \sin(\omega t) \right) \left(+ \cos(\omega t) \right) \quad (\text{B.18})$$

$$= i(t) \cos^2(\omega t) - q(t) \sin(\omega t) \cos(\omega t) \quad (\text{B.19})$$

$$= \frac{1}{2} i(t) + \frac{1}{2} \left(i(t) \cos(2\omega t) - q(t) \sin(2\omega t) \right), \quad (\text{B.20})$$

and by evaluating the multiplication of $x(t)$ and $-\sin(\omega t)$, we get

$$q_{\text{rx}}(t) = \left(i(t) \cos(\omega t) - q(t) \sin(\omega t) \right) \left(- \sin(\omega t) \right) \quad (\text{B.21})$$

$$= -i(t) \sin(\omega t) \cos(\omega t) + q(t) \sin^2(\omega t) \quad (\text{B.22})$$

$$= \frac{1}{2} q(t) - \frac{1}{2} \left(i(t) \sin(2\omega t) + q(t) \cos(2\omega t) \right). \quad (\text{B.23})$$

The complex baseband signal $\tilde{x}(t)$ can be regenerated from $i_{\text{rx}}(t)$ and $q_{\text{rx}}(t)$ by

$$\tilde{x}(t) = 2 \text{LP}\{i_{\text{rx}}(t)\} + j 2 \text{LP}\{q_{\text{rx}}(t)\}, \quad (\text{B.24})$$

where $\text{LP}\{\cdot\}$ is an ideal lowpass filter with a cutoff frequency anywhere between $B/2$ and $2f_c - B/2$. In a real receiver, the lowpass filter also acts as an anti-aliasing filter before the analog-to-digital converter. In this case, its cutoff frequency can be specified as $B/2$ times the oversampling ratio of the analog-to-digital converter.



Wirtinger Calculus

Behavioral modeling and digital predistortion of RF-PAs requires complex-valued signal processing. Below, I present an introduction to Wirtinger calculus, which is an important foundation for complex optimization and the derivations in Appendices D and E.

Motivation and Background

Behavioral models of RF-PAs and digital predistorters are typically formulated in the discrete time and complex baseband domain. If we assume time-invariance and finite-memory, such models can be described by a multivariate complex function, given by

$$f(\mathbf{x}) : \mathbb{C}^m \mapsto \mathbb{C}, \quad (\text{C.1})$$

where $\mathbf{x} = (x[n], x[n-1], \dots, x[n-(m-1)])^T$ is a vector containing the current input sample and $m-1$ past input samples and $y[n] = f(\mathbf{x})$ is a scalar containing the current output sample. The models are typically parameterized by a vector of several complex coefficients \mathbf{c} . To optimize the coefficients, one has to formulate an objective function

$$J(\mathbf{c}) : \mathbb{C}^n \mapsto \mathbb{R}, \quad (\text{C.2})$$

which maps the complex coefficient vector \mathbf{c} to a real-valued loss or reward parameter. Many optimization techniques for minimizing the loss or maximizing the reward require the computation of first and second-order derivatives of $J(\mathbf{c})$ with respect to \mathbf{c} . This requires the definition of complex derivatives of real-valued functions, which is known as Wirtinger calculus. Below, I summarize basic concepts of Wirtinger calculus.

The material below mainly follows [29, 33, 129]. Wirtinger derivatives were introduced in 1927 by Wilhelm Wirtinger [234]. In the engineering literature, they first appear in papers by Brandwood [33] and Van Den Bos [29]. Recent applications of Wirtinger calculus include the direct learning of digital predistorters [138, 139], complex-valued optimization [3, 209], and complex-valued statistics [4]. Introductions to Wirtinger calculus can be found in [90, App. A], [149, Ch. 5], [197, App. B], [2, Ch. 1], and [214].

Definition of Complex Derivatives

The derivative of a complex function $f(z) : \mathbb{C} \mapsto \mathbb{C}$ is defined by

$$f'(z_0) = \left. \frac{df}{dz} \right|_{z_0} = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}, \quad (\text{C.3})$$

where $z_0 \in \mathbb{C}$ is the point where the derivative is evaluated. The limit in (C.3) only exists, if the result of the limiting process is independent of the direction from which z_0 is approached. If z is real, there are only two possible directions for the limiting process: the limit from above and the limit from below. However, if z is complex, there are infinitely many possible directions, since we can approach z_0 from every possible angle in the complex plane. A consequence of this is that complex differentiability is much more restrictive than real differentiability. If we express the complex function by

$$f(z) = u(x, y) + jv(x, y), \quad (\text{C.4})$$

where $x = \text{Re}\{z\}$, $y = \text{Im}\{z\}$, it can be shown [149, pp. 56-57] that for the limit in (C.3) to exist, the functions $u(x, y)$ and $v(x, y)$ must fulfill the Cauchy-Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \quad (\text{C.5})$$

Non-constant, real-valued functions $f(z) = u(x, y)$ do not fulfill (C.5) and are therefore not differentiable based on (C.3), even if $u(x, y)$ is real-differentiable. For practical purposes, we therefore require a less restrictive definition of complex differentiation. Such a definition is given by the Wirtinger derivatives, which are defined by [234]

$$\frac{\partial f}{\partial z} = \frac{1}{2} \left(\frac{\partial f}{\partial x} - j \frac{\partial f}{\partial y} \right), \quad (\text{C.6})$$

$$\frac{\partial f}{\partial z^*} = \frac{1}{2} \left(\frac{\partial f}{\partial x} + j \frac{\partial f}{\partial y} \right). \quad (\text{C.7})$$

A useful consequence of (C.6) and (C.7) is that Wirtinger derivatives can be computed by conventional differentiation rules, considering z and z^* as independent variables. For functions that fulfill (C.5), the Wirtinger derivatives are equivalent to (C.3).

From (C.6) and (C.7), we immediately see that

$$\left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right) = (0 \quad 0) \Leftrightarrow \frac{\partial f}{\partial z} = 0 \Leftrightarrow \frac{\partial f}{\partial z^*} = 0 \quad (\text{C.8})$$

which means that the necessary and sufficient condition for a stationary point of $f(z)$ can be expressed equivalently in terms of $\left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right)$, $\frac{\partial f}{\partial z}$, and $\frac{\partial f}{\partial z^*}$ [33, Theorem 2].

Taylor Series of a Real Function of Several Complex Variables

For locating the minimum of a real function $f(\mathbf{z})$ of several complex variables $\mathbf{z} \in \mathbb{C}^n$, we can solve any of the equations in (C.8) for each variable in \mathbf{z} to find the stationary points of $f(\mathbf{z})$. If (C.8) has no analytic solution, we have to use an iterative algorithm like a gradient descent or a Newton algorithm to find the stationary points. To derive such an algorithm and to check whether the stationary points are minima, maxima, or saddle points, we require a Taylor series approximation of the function $f(\mathbf{z})$. Below, I present three equivalent forms of second-order Taylor series approximations of $f(\mathbf{z})$.

A point $\mathbf{z} \in \mathbb{C}^n$ can be expressed by $\mathbf{r} \in \mathbb{R}^{2n}$ or $\mathbf{c} \in \mathbb{C}^{2n}$ with

$$\mathbf{r} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} \mathbf{z} \\ \mathbf{z}^* \end{pmatrix}. \quad (\text{C.9})$$

The second-order Taylor series using the real parameter vector \mathbf{r} is given by

$$f_2(\mathbf{r} + \Delta\mathbf{r}) = f(\mathbf{r}) + \mathbf{g}_r^T(\mathbf{r})\Delta\mathbf{r} + \frac{1}{2}\Delta\mathbf{r}^T \mathbf{H}_{rr}(\mathbf{r})\Delta\mathbf{r} \quad (\text{C.10})$$

and an equivalent form using the complex parameter vector \mathbf{c} is given by [29]

$$f_2(\mathbf{c} + \Delta\mathbf{c}) = f(\mathbf{c}) + \mathbf{g}_c^H(\mathbf{c})\Delta\mathbf{c} + \frac{1}{2}\Delta\mathbf{c}^H \mathbf{H}_{cc}(\mathbf{c})\Delta\mathbf{c} \quad (\text{C.11})$$

where the gradient vector is defined by

$$\mathbf{g}_z(\mathbf{z}_0) = \left. \frac{\partial f}{\partial \mathbf{z}^*} \right|_{\mathbf{z}_0} = \left(\left. \frac{\partial f}{\partial z_1^*} \quad \cdots \quad \frac{\partial f}{\partial z_n^*} \right)^T \right|_{\mathbf{z}_0} \quad (\text{C.12})$$

and the Hessian matrix is defined by

$$\mathbf{H}_{xy}(\mathbf{z}_0) = \left. \frac{\partial^2 f}{\partial \mathbf{x}^* \partial \mathbf{y}^T} \right|_{\mathbf{z}_0} = \left(\left. \begin{array}{ccc} \frac{\partial^2 f}{\partial x_1^* \partial y_1} & \cdots & \frac{\partial^2 f}{\partial x_1^* \partial y_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n^* \partial y_1} & \cdots & \frac{\partial^2 f}{\partial x_n^* \partial y_n} \end{array} \right) \right|_{\mathbf{z}_0}. \quad (\text{C.13})$$

For defining derivatives with respect to vectors as in (C.12), (C.13), and for similar operations in matrix calculus, there are two possible layout conventions: the numerator layout and the denominator layout. In the present work, I use the denominator layout as in [2, 29, 90, 149, 209, 214]. Examples for the numerator layout are [129, 197].

By expanding (C.11), we can write the Taylor series also as [129]

$$f_2(\mathbf{z} + \Delta\mathbf{z}) = f(\mathbf{z}) + 2 \operatorname{Re} \left\{ \mathbf{g}_{\mathbf{z}}^H(\mathbf{z}) \Delta\mathbf{z} \right\} + \operatorname{Re} \left\{ \Delta\mathbf{z}^H \mathbf{H}_{zz}(\mathbf{z}) \Delta\mathbf{z} + \Delta\mathbf{z}^H \mathbf{H}_{zz^*}(\mathbf{z}) \Delta\mathbf{z}^* \right\}. \quad (\text{C.14})$$

The representation in (C.14) is advantageous with respect to implementation complexity since it uses vectors of size n , whereas (C.10) and (C.11) use vectors of size $2n$. On the other hand, (C.10) and (C.11) are better suited for the derivation and analysis of optimization algorithms [129, 138, 139, 209] and for checking whether a stationary point is a minimum, a maximum, or a saddle point, using the following criteria:

- A stationary point is a local minimum, if the Hessian matrix at this point is positive definite, i.e., all eigenvalues are larger than zero.
- A stationary point is a local maximum, if the Hessian matrix at this point is negative definite, i.e., all eigenvalues are smaller than zero.
- A stationary point is a saddle point, if the Hessian matrix at this point is indefinite, i.e., it has eigenvalues of mixed sign and none of them is zero.
- A stationary point cannot be characterized by second derivatives, if the Hessian matrix at this point is singular, i.e., it has at least one eigenvalue that is zero. In this case, higher order derivatives are required, as described in [28].

It was shown in [29] that the eigenvalues of \mathbf{H}_{cc} are two times the eigenvalues of \mathbf{H}_{rr} . Therefore, either \mathbf{H}_{rr} or \mathbf{H}_{cc} can be used to characterize a stationary point.

Example for a Complex Optimization Problem

Let's consider a real function of two complex variables, given by

$$f(\mathbf{r}) = x_1 e^{-(x_1^2 + y_1^2 + x_2^2 + y_2^2)} \Leftrightarrow f(\mathbf{c}) = \frac{1}{2} (z_1 + z_1^*) e^{-(z_1^* z_1 + z_2^* z_2)}. \quad (\text{C.15})$$

The gradient of $f(\mathbf{r})$ is given by

$$\mathbf{g}_{\mathbf{r}} = \frac{\partial f}{\partial \mathbf{r}^*} = \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \frac{\partial f}{\partial y_1} \quad \frac{\partial f}{\partial y_2} \right)^T = e^{-(x_1^2 + y_1^2 + x_2^2 + y_2^2)} \begin{pmatrix} 1 - 2x_1^2 \\ -2x_1x_2 \\ -2x_1y_1 \\ -2x_1y_2 \end{pmatrix}, \quad (\text{C.16})$$

and the gradient of $f(\mathbf{c})$ is given by

$$\mathbf{g}_{\mathbf{c}} = \frac{\partial f}{\partial \mathbf{c}^*} = \left(\frac{\partial f}{\partial z_1^*} \quad \frac{\partial f}{\partial z_2^*} \quad \frac{\partial f}{\partial z_1} \quad \frac{\partial f}{\partial z_2} \right)^T = \frac{1}{2} e^{-(z_1^* z_1 + z_2^* z_2)} \begin{pmatrix} 1 - z_1(z_1 + z_1^*) \\ -z_2(z_1 + z_1^*) \\ 1 - z_1^*(z_1 + z_1^*) \\ -z_2^*(z_1 + z_1^*) \end{pmatrix}. \quad (\text{C.17})$$

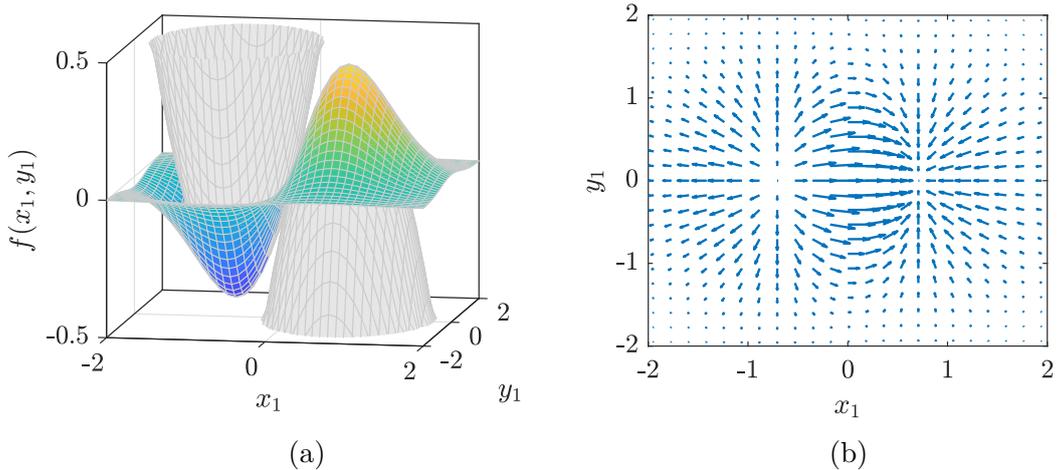


Figure C.1: Plot of a real-valued function in the complex plane and its gradient.

To find the stationary points of (C.15), we can set all elements of (C.16), or the first two or last two elements of (C.17) to zero. Solving any of these equations, we get two stationary points, located at $\mathbf{z} = (-1/\sqrt{2}, 0)^T$ and $\mathbf{z} = (+1/\sqrt{2}, 0)^T$.

To visualize the stationary points, I evaluated (C.15) on a grid spanned by the real and imaginary part of z_1 , with the second complex variable fixed at $z_2 = 0$. This resulted in the colored surface shown in Figure C.1 (a). From this plot, we clearly see that $\mathbf{z} = (-1/\sqrt{2}, 0)^T$ is a local minimum and $\mathbf{z} = (+1/\sqrt{2}, 0)^T$ is a local maximum.

To verify the Taylor series equations in (C.10), (C.11), (C.14), I computed the gradient vectors and the Hessian matrices at the two stationary points, and evaluated the second-order Taylor series approximation, resulting in the gray surfaces in Figure C.1 (a).

In the present example, the Hessian matrices \mathbf{H}_{rr} and \mathbf{H}_{cc} at both stationary points have an eigenvalue at zero, i.e., they are singular. The analytic characterization of the stationary points therefore requires higher order derivatives.

If we want to find the stationary points by an iterative algorithm, we may use the gradient vector $\mathbf{g}_z = \frac{\partial f}{\partial z^*}$, which is visualized in Figure C.1 (b). The gradient vector always points in the direction of steepest ascent. A gradient descent algorithm therefore moves along the negative gradient vector. The expression $\frac{\partial f}{\partial z}$ with $\mathbf{z} \in \mathbb{C}^n$ is not useful as a gradient, since it does not point in the direction of steepest ascent [33, 139].

D

Linear Least Squares

The method of linear least squares is an important foundation for behavioral modeling and digital predistortion of RF-PAs. In this appendix, I derive the well-known formula for linear least squares model fitting and summarize related methods for reducing the requirements of feedback receivers in linear least squares based DPD systems.

Derivation of Linear Least Squares

Discrete-time models that are linear in the coefficients can be written as

$$y[n] = \sum_{q=1}^Q c_q \mathcal{F}_q\{x[n]\}, \quad (\text{D.1})$$

where $x[n]$ and $y[n]$ are complex input and output signals¹, c_q are complex coefficients, and $\mathcal{F}_q\{\cdot\}$ ² are basis functionals. For $0 \leq n \leq N - 1$, we can write (D.1) as³

$$\mathbf{y} = \mathbf{X}\mathbf{c}, \quad \text{where} \quad \begin{cases} \mathbf{x} = (x[0], \dots, x[N-1])^T, & \mathbf{X} = (F_1(\mathbf{x}), \dots, F_Q(\mathbf{x})), \\ \mathbf{y} = (y[0], \dots, y[N-1])^T, & \mathbf{c} = (c_1, \dots, c_Q)^T. \end{cases} \quad (\text{D.2})$$

The objective of least squares fitting is to find the coefficient vector \mathbf{c} that minimizes the sum of squared errors between the actual model output \mathbf{y} and a desired model output \mathbf{y}_d , with a fixed model input \mathbf{x} . The objective function is given by

$$J(\mathbf{c}) = (\mathbf{y}_d - \mathbf{y})^H (\mathbf{y}_d - \mathbf{y}) \quad (\text{D.3})$$

$$= (\mathbf{y}_d - \mathbf{X}\mathbf{c})^H (\mathbf{y}_d - \mathbf{X}\mathbf{c}) \quad (\text{D.4})$$

$$= \mathbf{y}_d^H \mathbf{y}_d - \mathbf{y}_d^H \mathbf{X}\mathbf{c} - (\mathbf{X}\mathbf{c})^H \mathbf{y}_d + (\mathbf{X}\mathbf{c})^H \mathbf{X}\mathbf{c} \quad (\text{D.5})$$

$$= \mathbf{y}_d^H \mathbf{y}_d - \mathbf{y}_d^H \mathbf{X}\mathbf{c} - \mathbf{c}^H \mathbf{X}^H \mathbf{y}_d + \mathbf{c}^H \mathbf{X}^H \mathbf{X}\mathbf{c}. \quad (\text{D.6})$$

¹For simplicity, I omit the symbol \sim on top of complex baseband variables.

²The curly braces indicate dependence on $x[n] : \mathbb{Z} \mapsto \mathbb{C}$, rather than on $x[n] \in \mathbb{C}$.

³In (D.2), we represent $\mathcal{F}_q\{x[n]\}$ by multivariate functions $F_q(\mathbf{x}) : \mathbb{C}^N \mapsto \mathbb{C}^N$.

The objective function $J(\mathbf{c})$ is a real-valued function of the complex coefficient vector \mathbf{c} . In Appendix C, we have seen that the stationary point of such a function can be found by setting the Wirtinger derivative with respect to \mathbf{c} or \mathbf{c}^* equal to zero. Wirtinger derivatives are computed using conventional rules for differentiation with the additional rule that \mathbf{c} and \mathbf{c}^* are treated as independent variables. Differentiating (D.6), we get⁴

$$\frac{\partial J(\mathbf{c})}{\partial \mathbf{c}} = -(\mathbf{y}_d^H \mathbf{X})^T + (\mathbf{c}^H \mathbf{X}^H \mathbf{X})^T. \quad (\text{D.7})$$

Setting (D.7) equal to zero, we get an equation for the stationary point $\hat{\mathbf{c}}$, given by

$$(\hat{\mathbf{c}}^H \mathbf{X}^H \mathbf{X})^T = (\mathbf{y}_d^H \mathbf{X})^T. \quad (\text{D.8})$$

Applying complex conjugation to both sides of (D.8) and re-arranging terms, we get

$$(\mathbf{X}^H \mathbf{X}) \hat{\mathbf{c}} = \mathbf{X}^H \mathbf{y}_d, \quad (\text{D.9})$$

which is called the normal equation. The expression $(\mathbf{X}^H \mathbf{X})$ on the left side of (D.9) is a Hermitian-symmetric square matrix of size $Q \times Q$. Multiplying (D.9) from the left by $(\mathbf{X}^H \mathbf{X})^{-1}$, we arrive at the solution for the stationary point, given by

$$\hat{\mathbf{c}} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{y}_d. \quad (\text{D.10})$$

To check whether this stationary point is a minimum, a maximum, or a saddle point, we compute the Hessian matrix by differentiating (D.7) a second time, resulting in

$$\frac{\partial^2 J(\mathbf{c})}{\partial \mathbf{c}^* \partial \mathbf{c}} = \mathbf{X}^H \mathbf{X}. \quad (\text{D.11})$$

We found a minimum, if the Hessian matrix is positive definite, which is defined by

$$\mathbf{z}^H \mathbf{X}^H \mathbf{X} \mathbf{z} > 0 \quad \forall \mathbf{z} \neq \mathbf{0} \quad \Leftrightarrow \quad \mathbf{X}^H \mathbf{X} \text{ is positive definite.} \quad (\text{D.12})$$

To check whether (D.12) is the case, we may write the left term as

$$\mathbf{z}^H \mathbf{X}^H \mathbf{X} \mathbf{z} = (\mathbf{X} \mathbf{z})^H \mathbf{X} \mathbf{z} = \mathbf{v}^H \mathbf{v} \geq 0, \quad (\text{D.13})$$

proving that $\mathbf{X}^H \mathbf{X}$ is positive semi-definite. It is positive definite, if additionally

$$\mathbf{z}^H \mathbf{X}^H \mathbf{X} \mathbf{z} \neq 0 \quad \forall \mathbf{z} \neq \mathbf{0}. \quad (\text{D.14})$$

If the matrix $\mathbf{X}^H \mathbf{X}$ is regular, which means that the inverse matrix $(\mathbf{X}^H \mathbf{X})^{-1}$ exists, the only solution of $\mathbf{X}^H \mathbf{X} \mathbf{z} = 0$ is the zero vector $\mathbf{z} = \mathbf{0}$. Since the zero vector is excluded in (D.14), we conclude that if $\mathbf{X}^H \mathbf{X}$ is regular, it is positive definite, and therefore the stationary point in (D.10) represents a minimum of the sum of squared errors.

⁴Like in Appendix C, we use the denominator layout for derivatives with respect to vectors.

Separation of the Real and Imaginary Part

We can write the complex system in (D.2) also as a real system, given by

$$\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_q \end{bmatrix} = \begin{bmatrix} \mathbf{X}_i & -\mathbf{X}_q \\ \mathbf{X}_q & \mathbf{X}_i \end{bmatrix} \begin{bmatrix} \mathbf{c}_i \\ \mathbf{c}_q \end{bmatrix}, \quad \text{where} \quad \begin{cases} \mathbf{y}_i = \text{Re}(\mathbf{y}), & \mathbf{y}_q = \text{Im}(\mathbf{y}), \\ \mathbf{X}_i = \text{Re}(\mathbf{X}), & \mathbf{X}_q = \text{Im}(\mathbf{X}), \\ \mathbf{c}_i = \text{Re}(\mathbf{c}), & \mathbf{c}_q = \text{Im}(\mathbf{c}). \end{cases} \quad (\text{D.15})$$

To identify the coefficients, we may apply the linear least squares solution to the complete equation system in (D.15). However, as presented in [39], we do not need all equations of (D.15), but we can identify \mathbf{c} by solving only one of the systems, given by

$$\mathbf{y}_i = [\mathbf{X}_i \quad -\mathbf{X}_q] \begin{bmatrix} \mathbf{c}_i \\ \mathbf{c}_q \end{bmatrix}, \quad \text{and} \quad \mathbf{y}_q = [\mathbf{X}_q \quad \mathbf{X}_i] \begin{bmatrix} \mathbf{c}_i \\ \mathbf{c}_q \end{bmatrix}. \quad (\text{D.16})$$

Using only one of the equations in (D.16) reduces the complexity of direct conversion feedback receivers for behavioral modeling and direct learning DPD, since only one ADC is required. For indirect learning, this method cannot be used, since in this case the feedback signal is used to construct \mathbf{X} , which must be fully available.

Reduction of the Sampling Rate

Shannon's sampling theorem [146] states that a bandlimited signal $x(t)$ can be perfectly reconstructed from its discrete-time signal $x[n]$, if the latter is uniformly sampled with a rate of at least two times the highest frequency in $x(t)$. If we use the discrete-time input and output signals $x[n]$ and $y[n]$ for system identification, it seems appropriate to use a sampling rate which allows perfect reconstruction of $x(t)$ and $y(t)$. However, for identifying memoryless nonlinear systems [249] and Volterra systems [93], it was shown that it is actually sufficient to fulfill Shannon's sampling theorem regarding the input signal. More recently, direct learning DPD systems have been investigated in which the RF-PA feedback signal can be sampled at an arbitrary low rate [225, 226]. In such systems, the DPD training is based on an undersampled equation system, given by

$$\begin{bmatrix} y[n] \\ y[n+M] \\ \vdots \\ y[n+kM] \end{bmatrix} = \begin{bmatrix} \mathcal{F}_1\{x[n]\} & \dots & \mathcal{F}_Q\{x[n]\} \\ \mathcal{F}_1\{x[n+M]\} & \dots & \mathcal{F}_Q\{x[n+M]\} \\ \vdots & & \vdots \\ \mathcal{F}_1\{x[n+kM]\} & \dots & \mathcal{F}_Q\{x[n+kM]\} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_Q \end{bmatrix}, \quad (\text{D.17})$$

where $x[n]$ is the bandlimited input signal, $y[n]$ is the broadband output signal and M is the undersampling factor. It is important to note that although the sampling rate of the feedback receiver can theoretically be reduced to an arbitrary low value, its frequency response must be flat over the complete modeling or linearization bandwidth.

E

Synchronization

Synchronization is an important preprocessing step for the behavioral modeling and DPD of RF-PAs. This appendix summarizes basic theory and techniques regarding the compensation of the time delay and complex gain between two complex baseband signals. In a stand-alone receiver, also the carrier and sampling frequency must be synchronized, but in a feedback receiver these frequencies are typically derived from the same local oscillator, which avoids the need for frequency synchronization.

Problem Statement

Let's consider two finite power, complex signals $x(t)$ and $y(t)$ ^{1,2}. To synchronize them, we have to estimate the time delay and complex gain between these signals and apply the respective operations either on $x(t)$ as shown in Figure E.1 (a) or on $y(t)$ as shown in Figure E.1 (b). Note that within these block diagrams, the error signals $e_x(t)$ and $e_y(t)$ are consequences of the input and output signals and synchronization parameters.

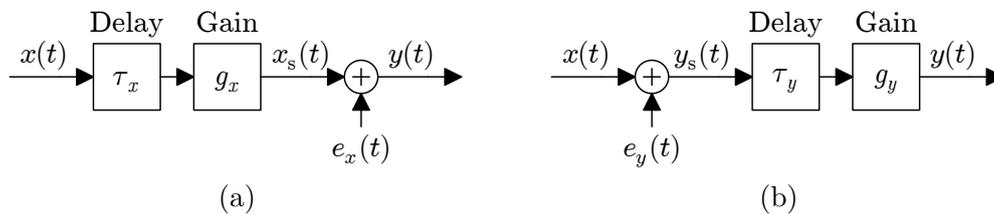


Figure E.1: System models for the delay and the complex gain correction.

In the case of Figure E.1 (a), we say that $x_s(t)$ is synchronized to $y(t)$, if the parameters τ_x and g_x are chosen such that the mean square magnitude of $e_x(t)$ is minimized.

In the case of Figure E.1 (b), we say that $y_s(t)$ is synchronized to $x(t)$, if the parameters τ_y and g_y are chosen such that the mean square magnitude of $e_y(t)$ is minimized.

¹Finite power and finite energy signals are discussed in [170, Ch. 12]

²For simplicity, I omit the symbol \sim on top of complex baseband variables.

Definitions

The mean square error (MSE) regarding an error signal $e(t)$ is defined by

$$\text{MSE} = \text{mean } |e(t)|^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} e^*(t) e(t) dt. \quad (\text{E.1})$$

The normalized MSEs (NMSEs) in dB of the setups in Figure E.1 are defined by

$$\text{NMSE}_x = 10 \log_{10} \left(\frac{\text{mean } |e_x(t)|^2}{\text{mean } |y(t)|^2} \right), \quad \text{NMSE}_y = 10 \log_{10} \left(\frac{\text{mean } |e_y(t)|^2}{\text{mean } |x(t)|^2} \right). \quad (\text{E.2})$$

The complex cross-correlation for finite power signals is defined by [170, p. 253]

$$r_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} x^*(t - \tau) y(t) dt. \quad (\text{E.3})$$

The correlation coefficient is a normalized version of (E.3), defined by^{3,4}

$$\rho_{xy}(\tau) = \frac{r_{xy}(\tau)}{\sqrt{r_{xx}(0) r_{yy}(0)}}. \quad (\text{E.4})$$

Analytical Results

The MSE optimal delay is identical for the two setups and given by

$$\tau_{xy} = \tau_x = \tau_y = \arg \max_{\tau} |r_{xy}(\tau)|. \quad (\text{E.5})$$

The MSE optimal complex gain is in general different for the two setups and given by

$$g_x = \frac{r_{xy}(\tau_{xy})}{r_{xx}(0)}, \quad g_y = \frac{r_{yy}(0)}{r_{xy}^*(\tau_{xy})}. \quad (\text{E.6})$$

The MSE optimal phase rotation is identical for the two setups and given by⁵

$$\varphi_{xy} = \angle g_x = \angle g_y = \angle r_{xy}(\tau_{xy}). \quad (\text{E.7})$$

The NMSE of the synchronized signals is identical for the two setups and given by

$$\text{NMSE} = \text{NMSE}_x = \text{NMSE}_y = 10 \log_{10} (1 - |\rho_{xy}(\tau_{xy})|^2). \quad (\text{E.8})$$

³The Pearson correlation coefficient [190] uses the covariance instead of the correlation.

⁴Due to the Cauchy-Schwarz inequality, the magnitude of (E.4) is upper bounded by one.

⁵This follows immediately from (E.6) by noting that $r_{xx}(0)$ and $r_{yy}(0)$ are real.

Proofs of Analytical Results

In the following, I present proofs of the statements in (E.5) to (E.8).

MSE optimal delay τ_x

To prove (E.5) with respect to τ_x , we express the MSE of the setup in Figure E.1 (a) as a function of the delay τ_x . The MSE of the setup in Figure E.1 (a) is defined by

$$\text{MSE}_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} e_x^*(t) e_x(t) dt. \quad (\text{E.9})$$

Inserting an expression for $e_x(t)$ into (E.9), we get

$$\text{MSE}_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} \left(y(t) - g_x x(t - \tau_x) \right)^* \left(y(t) - g_x x(t - \tau_x) \right) dt. \quad (\text{E.10})$$

Expanding the product in (E.10) results in

$$\text{MSE}_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} |y(t)|^2 - 2 \operatorname{Re} \left(g_x^* x^*(t - \tau_x) y(t) \right) + |g_x|^2 |x(t - \tau_x)|^2 dt. \quad (\text{E.11})$$

By using the definition of the cross-correlation in (E.3), we can write (E.11) as

$$\text{MSE}_x = r_{yy}(0) - 2 \operatorname{Re} \left(g_x^* r_{xy}(\tau_x) \right) + |g_x|^2 r_{xx}(0). \quad (\text{E.12})$$

Applying the identity $\operatorname{Re}(x) = |x| \cos(\angle x)$ on the middle term in (E.12), we get

$$\text{MSE}_x = r_{yy}(0) - 2 |g_x| |r_{xy}(\tau_x)| \cos \left(\angle \left(g_x^* r_{xy}(\tau_x) \right) \right) + |g_x|^2 r_{xx}(0). \quad (\text{E.13})$$

Since the angle of the MSE optimal complex gain is $\angle g_x = \angle r_{xy}(\tau_x)$, the argument of the cosine in (E.13) is zero, independent of the delay τ_x . The remaining dependency between MSE_x and τ_x expressed by (E.13) proves (E.5) with respect to τ_x .

MSE optimal complex gain g_x

To find the MSE optimal complex gain g_x , we expand (E.10) as

$$\begin{aligned} \text{MSE}_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} & y^*(t) y(t) - g_x^* x^*(t - \tau_x) y(t) \\ & - g_x x(t - \tau_x) y^*(t) + g_x^* g_x x^*(t - \tau_x) x(t - \tau_x) dt. \end{aligned} \quad (\text{E.14})$$

By computing the partial derivative of MSE_x with respect to g_x^* , we get

$$\begin{aligned} \frac{\partial \text{MSE}_x}{\partial g_x^*} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} -x^*(t - \tau_x) y(t) + g_x x^*(t - \tau_x) x(t - \tau_x) dt \\ &= -r_{xy}(\tau_x) + g_x r_{xx}(0). \end{aligned} \quad (\text{E.15})$$

Setting (E.15) to zero and solving for g_x , we obtain the left formula in (E.6).

The following proofs for the MSE optimal delay τ_y and complex gain g_y are very similar to the proofs for the MSE optimal delay τ_x and complex gain g_x presented above.

MSE optimal delay τ_y

To prove (E.5) with respect to τ_y , we express the MSE of the setup in Figure E.1 (b) as a function of the delay τ_y . The MSE of the setup in Figure E.1 (b) is defined by

$$\text{MSE}_y = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} e_y^*(t) e_y(t) dt. \quad (\text{E.16})$$

Insert an expression for $e_y(t)$ into (E.16), we get

$$\text{MSE}_y = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} \left(g_y^{-1} y(t + \tau_y) - x(t) \right)^* \left(g_y^{-1} y(t + \tau_y) - x(t) \right) dt. \quad (\text{E.17})$$

Expanding the product in (E.17) results in

$$\text{MSE}_y = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} |g_y|^{-2} |y(t + \tau_y)|^2 - 2 \text{Re} \left(g_y^{-1} x^*(t) y(t + \tau_y) \right) + |x(t)|^2 dt. \quad (\text{E.18})$$

By using the definition of the cross-correlation in (E.3), we can write (E.18) as

$$\text{MSE}_y = |g_y|^{-2} r_{yy}(0) - 2 \text{Re} \left(g_y^{-1} r_{xy}(\tau_y) \right) + r_{xx}(0). \quad (\text{E.19})$$

Applying the identity $\text{Re}(x) = |x| \cos(\angle x)$ on the middle term in (E.19), we get

$$\text{MSE}_y = |g_y|^{-2} r_{yy}(0) - 2 |g_y|^{-1} |r_{xy}(\tau_y)| \cos \left(\angle \left(g_y^{-1} r_{xy}(\tau_y) \right) \right) + r_{xx}(0). \quad (\text{E.20})$$

Since the angle of the MSE optimal complex gain is $\angle g_y = \angle r_{xy}(\tau_y)$, the argument of the cosine in (E.20) is zero, independent of the delay τ_y . The remaining dependency between MSE_y and τ_y expressed by (E.20) proves (E.5) with respect to τ_y .

MSE optimal complex gain g_y

To find the MSE optimal complex gain g_y , we expand (E.17) as

$$\text{MSE}_y = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} \frac{y^*(t)y(t)}{g_y^* g_y} - \frac{y^*(t + \tau_y)x(t)}{g_y^*} - \frac{y(t + \tau_y)x^*(t)}{g_y} + x^*(t)x(t) dt. \quad (\text{E.21})$$

By computing the partial derivative of MSE_y with respect to g_y^* , we get

$$\begin{aligned} \frac{\partial \text{MSE}_y}{\partial g_y^*} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{+T} -\frac{y^*(t)y(t)}{g_y^* g_y^* g_y} + \frac{y^*(t + \tau_y)x(t)}{g_y^* g_y^*} dt \\ &= \frac{1}{g_y^* g_y^*} \left(-\frac{r_{yy}(0)}{g_y} + r_{xy}^*(\tau_x) \right). \end{aligned} \quad (\text{E.22})$$

By setting (E.22) to zero and solving for g_y , we obtain the right formula in (E.6).

Relation between NMSE_x and ρ_{xy}

To prove (E.8) with respect to NMSE_x , we insert the expression for the MSE optimal complex gain g_x from (E.6) into (E.12), resulting in

$$\begin{aligned} \text{MSE}_x &= r_{yy}(0) - 2 \operatorname{Re} \left(\frac{r_{xy}^*(\tau_x)}{r_{xx}(0)} r_{xy}(\tau_x) \right) + \frac{|r_{xy}(\tau_x)|^2}{r_{xx}^2(0)} r_{xx}(0) \\ &= r_{yy}(0) - 2 \frac{|r_{xy}(\tau_x)|^2}{r_{xx}(0)} + \frac{|r_{xy}(\tau_x)|^2}{r_{xx}(0)} \\ &= r_{yy}(0) - \frac{|r_{xy}(\tau_x)|^2}{r_{xx}(0)}. \end{aligned} \quad (\text{E.23})$$

The definition of NMSE_x in (E.2) is equivalent to

$$\text{NMSE}_x = 10 \log_{10} \left(\frac{\text{MSE}_x}{r_{yy}(0)} \right). \quad (\text{E.24})$$

Inserting the last line of (E.23) into (E.24), we get

$$\text{NMSE}_x = 10 \log_{10} \left(1 - \frac{|r_{xy}(\tau_x)|^2}{r_{xx}(0) r_{yy}(0)} \right), \quad (\text{E.25})$$

which proves (E.8) with respect to NMSE_x .

Relation between NMSE_y and ρ_{xy}

To prove (E.8) with respect to NMSE_y , we insert the expression for the MSE optimal complex gain g_y from (E.6) into (E.19), resulting in

$$\begin{aligned} \text{MSE}_y &= \frac{|r_{xy}(\tau_x)|^2}{r_{yy}^2(0)} r_{yy}(0) - 2 \operatorname{Re} \left(\frac{r_{xy}^*(\tau_x)}{r_{yy}(0)} r_{xy}(\tau_x) \right) + r_{xx}(0) \\ &= \frac{|r_{xy}(\tau_x)|^2}{r_{yy}(0)} - 2 \frac{|r_{xy}(\tau_x)|^2}{r_{yy}(0)} + r_{xx}(0) \\ &= r_{xx}(0) - \frac{|r_{xy}(\tau_x)|^2}{r_{yy}(0)}. \end{aligned} \quad (\text{E.26})$$

The definition of NMSE_y in (E.2) is equivalent to

$$\text{NMSE}_y = 10 \log_{10} \left(\frac{\text{MSE}_y}{r_{xx}(0)} \right). \quad (\text{E.27})$$

Inserting the last line of (E.26) into (E.27), we get

$$\text{NMSE}_y = 10 \log_{10} \left(1 - \frac{|r_{xy}(\tau_x)|^2}{r_{xx}(0) r_{yy}(0)} \right), \quad (\text{E.28})$$

which proves (E.8) with respect to NMSE_y .

Discrete-Time Complex Cross-Correlation

The discrete-time version of the complex cross-correlation defined in (E.3) is given by

$$r_{xy}[m] = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{n=-N}^{+N} x^*[n-m]y[n]. \quad (\text{E.29})$$

In practice, the discrete-time signals $x[n]$ and $y[n]$ are only defined for a finite number of samples with indices $0 \leq n < N < \infty$. Therefore, we estimate (E.29) by

$$r_{xy}[m] = \frac{1}{N} \sum_{n=0}^{N-1} x^*[n-m]y[n]. \quad (\text{E.30})$$

To compute (E.30) for $m > 0$, we require additional samples of $x[n]$ for $n < 0$. Likewise, for $m < 0$, we require additional samples of $x[n]$ for $n \geq N$. If we assume that $x[n]$ is N -periodic in n , i.e., $x[n+kN] = x[n]$ for all $k \in \mathbb{Z}$, we call the result of (E.30) the circular cross-correlation. On the other hand, if we assume that $x[n] = 0$ for $n < 0$ or $n \geq N$, we call the result of (E.30) the biased linear cross-correlation. The unbiased linear cross-correlation is obtained, if we replace the normalization factor $\frac{1}{N}$ in (E.30) by $\frac{1}{N-|m|}$. This, however, increases the variance of the estimate.

The circular cross-correlation is N -periodic in the delay index m and is therefore uniquely defined by N consecutive samples of $r_{xy}[m]$. The linear cross-correlation has a limited support of $2N - 1$ samples given by the delay indices $-(N - 1) \leq m \leq +(N - 1)$.

If we require $r_{xy}[m]$ only for a few delay indices m , we may implement (E.30) directly. However, if we require $r_{xy}[m]$ for all delay indices m , it is typically more efficient to implement the computation of in the frequency domain. To compute the circular cross-correlation in the frequency domain, we define the signal vectors

$$\mathbf{x} = \begin{pmatrix} x[0] & x[1] & \dots & x[N-1] \end{pmatrix}^T, \quad (\text{E.31})$$

$$\mathbf{y} = \begin{pmatrix} y[0] & y[1] & \dots & y[N-1] \end{pmatrix}^T, \quad (\text{E.32})$$

$$\mathbf{r}_{xy} = \begin{pmatrix} r_{xy}[0] & r_{xy}[1] & \dots & r_{xy}[N-1] \end{pmatrix}^T, \quad (\text{E.33})$$

and compute \mathbf{r}_{xy} from \mathbf{x} and \mathbf{y} by [30]

$$\mathbf{r}_{xy} = \frac{1}{N} \text{IFFT} \left(\left(\text{FFT}(\mathbf{x}) \right)^* \circ \text{FFT}(\mathbf{y}) \right), \quad (\text{E.34})$$

where the symbol \circ is used for element-wise multiplication. To compute the linear cross-correlation, we may also use (E.34), if we append $N - 1$ zeros to the vectors \mathbf{x} and \mathbf{y} and the variables $r_{xy}[-(N - 1)]$ to $r_{xy}[-1]$ to the vector \mathbf{r}_{xy} . The equivalence of (E.30) and (E.34) results from two properties of the Fourier transform: Multiplication in the frequency domain corresponds to convolution in the time domain and conjugation in the frequency domain corresponds to conjugation and time-reversal in the time domain.

Discrete-Time Delay Estimation and Correction

In the following, we consider two discrete-time signals $x[n]$ and $y[n]$, defined within the range of sample indices $0 \leq n < N < \infty$. For the delay estimation and correction, we must decide if we assume N -periodic signals, or if we do not assume any knowledge of the signals outside their range of definition. Assuming N -periodic signals has the advantage that the delay-corrected signals have the same length as the original signals and that the accuracy of the delay estimate is independent of the delay. Without the assumption of N -periodic signals, the length of the delay corrected signals is $N - |d|$, where d is the estimated delay and the accuracy of the delay estimate reduces with $|d|$.

The assumption of N -periodic signals is justified if the output signal $y[n]$ represent the steady state response to an N -periodic input signal $x[n]$ and their sampling frequencies are synchronized. In a simulation or measurement setup this is typically easy to produce. In a practical DPD system within a wireless transmitter, however, the input signal $x[n]$ is typically not periodic. In this case, the signal generation and data acquisition may be synchronized by hardware such that $|d| \ll N$. This ensures that the delay-corrected signals have sufficient length and the delay estimate is sufficiently accurate.

To estimate the delay between two N -periodic signals, we compute the circular cross-correlation for all delay indices $0 \leq m < N$, using the efficient frequency domain method described above. For signals without periodicity, but assuming that the delay is limited by $|d| \leq d_{\max}$, we may compute the biased or unbiased linear cross-correlation for the limited number of delay indices $-d_{\max} \leq m \leq d_{\max}$ directly in the time domain.

After computing the cross-correlation, we find the integer part of the delay by

$$d_{\text{int}} = \arg \max_m |r_{xy}[m]|. \quad (\text{E.35})$$

The fractional part of the delay can be estimated efficiently by parabolic interpolation of $|r_{xy}[m]|$ around $m = d_{\text{int}}$ [30]. The interpolating parabola is given by

$$r(\Delta d) = c_0 + c_1 \Delta d + c_2 \Delta d^2 \quad (\text{E.36})$$

where c_0, c_1, c_2 are coefficients and Δd is the relative delay with respect to $m = d_{\text{int}}$. To uniquely determine the coefficients, we require three points of $r(\Delta d)$, given by

$$r(\Delta d) = |r_{xy}[d_{\text{int}} + \Delta d]| \quad \text{for} \quad \Delta d \in \{-1, 0, +1\}. \quad (\text{E.37})$$

To simplify the notation, we abbreviate the three points by

$$r_{\text{left}} = |r_{xy}[d_{\text{int}} - 1]|, \quad r_{\text{mid}} = |r_{xy}[d_{\text{int}}]|, \quad r_{\text{right}} = |r_{xy}[d_{\text{int}} + 1]|. \quad (\text{E.38})$$

Inserting (E.36) and (E.38) into (E.37) and solving for the coefficients, we obtain

$$c_0 = r_{\text{mid}}, \quad c_1 = \frac{r_{\text{right}} - r_{\text{left}}}{2}, \quad c_2 = -\frac{2r_{\text{mid}} - r_{\text{right}} - r_{\text{left}}}{2}. \quad (\text{E.39})$$

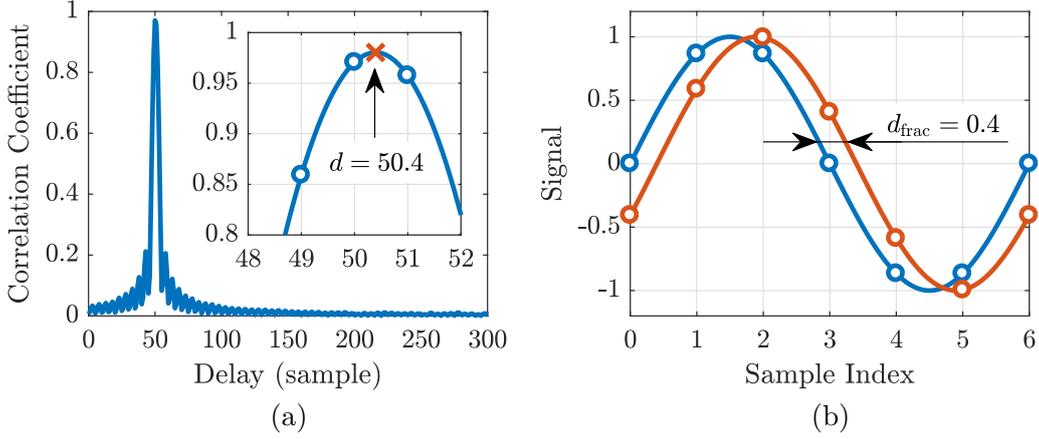


Figure E.2: Example of the discrete-time delay estimation and correction.

To locate the maximum of the parabola, we differentiate (E.36) with respect to Δd , set the result to zero, and solve for Δd , resulting in the fractional delay estimate

$$d_{\text{frac}} = \frac{r_{\text{right}} - r_{\text{left}}}{2(2r_{\text{mid}} - r_{\text{right}} - r_{\text{left}})}. \quad (\text{E.40})$$

The overall delay estimate is given by $d = d_{\text{int}} + d_{\text{frac}}$. An example for delay estimation based on this method is shown in Figure E.2 (a). To generate this plot, I used a noise-like complex baseband signal $x[n]$ with a bandwidth of $f_s/5$, where f_s is the sampling frequency. The signal $y[n]$ is derived from $x[n]$ by applying a circular delay of 50.4 samples, adding white Gaussian noise corresponding to a signal to noise ratio of 20 dB and applying an arbitrary complex gain. The correlation coefficient in Figure E.2 (a) shows a clear peak at an integer delay of $d_{\text{int}} = 50$. The fractional delay is estimated from the correlation coefficient at the delay indices 49, 50, 51, resulting in $d_{\text{frac}} = 0.4$.

To correct the estimated delay, a delay of d samples can be applied on $x[n]$, or a delay of $-d$ samples can be applied on $y[n]$. The integer part of this delay can be easily corrected by circular shifting the elements in the respective signal vector. The fractional part can be corrected by applying a phase rotation on the spectrum of \mathbf{x} , given by

$$\mathbf{x}_{\text{delay}} = \text{IFFT}\left(\text{FFT}(\mathbf{x}) \circ e^{-j\omega d}\right), \quad (\text{E.41})$$

where ω is the vector of discrete-time angular frequencies given by

$$\omega = \frac{2\pi}{N} \left(0 \quad \dots \quad \left\lfloor \frac{N}{2} \right\rfloor - 1 \quad - \left\lfloor \frac{N}{2} \right\rfloor \quad \dots \quad -1 \right)^T. \quad (\text{E.42})$$

Figure E.2 (b) shows a sinusoidal signal and its fractionally delayed version, computed by (E.41) with $d_{\text{frac}} = 0.4$. In a practical DPD system the fractional delay correction is often implemented with lower complexity based on the Farrow structure [84, 134].

Discrete-Time Complex Gain Estimation and Correction

After the delay correction, we can estimate and correct the complex gain. The discrete-time version of the MSE optimal complex gain from (E.6) is given by

$$g_x = \frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}}, \quad g_y = \frac{\mathbf{y}^H \mathbf{y}}{\mathbf{y}^H \mathbf{x}}. \quad (\text{E.43})$$

To correct the complex gain, we can either multiply $x[n]$ by g_x , or we can multiply $y[n]$ by g_y^{-1} . As shown above, both options give the same NMSE. If we synchronize the signals for fitting a behavioral model, we may want to keep the magnitude information and only correct the phase rotation using the MSE optimal phase estimate given by

$$\varphi_{xy} = \angle g_x = \angle g_y = \angle(\mathbf{x}^H \mathbf{y}). \quad (\text{E.44})$$

If we synchronize the signals for the iterative training of a DPD, we may want to ensure that the output power with the DPD converges to the same output power as without the DPD. In this case, we estimate the root mean square (RMS) gain

$$g_{\text{rms}} = \sqrt{\frac{\mathbf{y}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}}} \quad (\text{E.45})$$

from the signals without the DPD and use this real gain factor for normalizing the output signal magnitude during each iteration of DPD training.

Note that we can relate both the RMS gain g_{rms} and the magnitude of the correlation coefficient $|\rho_{xy}(\tau_{xy})|$ with the MSE optimal complex gain estimates g_x and g_y by

$$g_{\text{rms}} = \sqrt{g_x^* g_y}, \quad |\rho_{xy}(\tau_{xy})| = \sqrt{\frac{g_x}{g_y}}, \quad (\text{E.46})$$

which shows that if $|\rho_{xy}(\tau_{xy})| = 1$, the magnitude of all gain estimates is identical.

Discrete-Time Normalized Mean Square Error

To compute the NMSE, we may use one of the formulas, given by

$$\text{NMSE} = 10 \log_{10} \left(\frac{\mathbf{e}_x^H \mathbf{e}_x}{\mathbf{y}^H \mathbf{y}} \right) = 10 \log_{10} \left(\frac{\mathbf{e}_y^H \mathbf{e}_y}{\mathbf{x}^H \mathbf{x}} \right) = 10 \log_{10} \left(1 - \frac{|\mathbf{x}^H \mathbf{y}|^2}{(\mathbf{x}^H \mathbf{x})(\mathbf{y}^H \mathbf{y})} \right), \quad (\text{E.47})$$

where \mathbf{x} and \mathbf{y} are the delay-corrected signal vectors, $\mathbf{e}_x = \mathbf{y} - g_x \mathbf{x}$ is the error signal vector of Figure E.1 (a) and $\mathbf{e}_y = g_y^{-1} \mathbf{y} - \mathbf{x}$ is the error signal vector of Figure E.1 (b).



Combinatorics

In the context of Volterra models, we encounter several combinatoric problems. In this appendix, I present a short introduction to the twelvefold way, which is a systematic method to classify such problems, followed by an overview of the four most common combinatoric problems and their relation to Volterra models.

The Twelvefold Way and Volterra Models

Let \mathcal{K} and \mathcal{N} be finite sets with cardinalities $k = |\mathcal{K}|$ and $n = |\mathcal{N}|$. Basic combinatoric problems are to count the equivalence classes of functions $f : \mathcal{K} \mapsto \mathcal{N}$.

With respect to the type of function, we may demand

- no restriction on f , or
- that f is injective¹, or
- that f is surjective².

With respect to the type of equivalence, we may demand

- equality, or
- equality, up to a permutation of \mathcal{K} , or
- equality, up to a permutation of \mathcal{N} , or
- equality, up to a permutation of \mathcal{K} and \mathcal{N} .

The constraints listed above can be combined in twelve different ways, resulting in twelve combinatoric problems, which are discussed in [210, Ch. 1.9]. Below, we focus on the four most common problems, obtained by combining the first two constraints of each list. For this purpose, we interpret the functions f as number sequences, where the elements in \mathcal{K} are the positions within the sequences and the elements in \mathcal{N} are the possible values at each position. An overview of formulas for computing the number of sequences is given in Table F.1 and an example for listing and counting of sequences is given in Table F.2.

¹A function $f : \mathcal{K} \mapsto \mathcal{N}$ is injective, if every element in \mathcal{N} is the image of at most one element in \mathcal{K} .

²A function $f : \mathcal{K} \mapsto \mathcal{N}$ is surjective, if every element in \mathcal{N} is the image of at least one element in \mathcal{K} .

	Repetition is allowed. (no restriction on f)	Repetition is not allowed. (f must be injective)
Order is relevant. (demand equality)	n^k	$\frac{n!}{(n-k)!}$
Order is irrelevant. (allow permutation of \mathcal{K})	$\binom{\binom{n}{k}}{k} = \binom{n+k-1}{k}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Table F.1: Summary of basic formulas from enumerative combinatorics.

Listing and counting of sequences. Take k elements out of n elements.		$\mathcal{N} = \{1, 2, 3\} \quad k = 2$									
<table border="1" style="display: inline-table; vertical-align: middle;"> <tr><td>11</td></tr> <tr><td>22</td></tr> <tr><td>33</td></tr> <tr><td>12</td></tr> <tr><td>21</td></tr> <tr><td>23</td></tr> <tr><td>32</td></tr> <tr><td>13</td></tr> <tr><td>31</td></tr> </table>	11	22	33	12	21	23	32	13	31	} Repetition } Permutation } Permutation } Permutation	Sequences: $3^2 = 9$ Permutations: $\frac{3!}{(3-2)!} = 6$ Multicombinations: $\binom{3+2-1}{2} = \frac{4!}{2! \cdot 2!} = 6$ Combinations: $\binom{3}{2} = \frac{3!}{1! \cdot 2!} = 3$
11											
22											
33											
12											
21											
23											
32											
13											
31											

Table F.2: Examples of basic problems from enumerative combinatorics.

In the context of Volterra models, we encounter the following combinatoric problems:

Sequences The causal p -th order Volterra kernel with memory depth M contains all sample-lag vectors $\mathbf{m} = [m_1, \dots, m_p]^T$ with $0 \leq m_i \leq M$. The number of such vectors is the number of p -fold sequences taken from $M + 1$ values, i.e. $(M + 1)^p$.

Permutations Due to the commutativity of multiplication, all permutations of the sample-lags within a given sample-lag vector correspond to the same input product. Therefore, there are $p!$ symmetric regions in the p -th order Volterra kernel.

Multicombinations Since the main diagonal is included in each symmetric region, the number of sample-lag vectors in each symmetric region is bigger than $(M + 1)^p/p!$. It is given by the number of p -multicombinations out of $M + 1$ values.

Combinations In the derivation of the baseband Volterra series, we combine all p -fold input products with q not-conjugated and $p - q$ conjugated terms. This corresponds to the selection of q not-conjugated terms, where the order of selection is irrelevant, but repeated selection is not allowed. Therefore the number of such terms is $\binom{p}{q}$.



Indefinite Integrals

In Chapter 2, we derive equations (2.14) and (2.16), which describe the harmonic drain current amplitudes of the RF-PA model based on either the piecewise-linear or the piecewise-cubic transistor model. They contain integrals of the form

$$F_{n,k} = \int \cos^n(\theta) \cos(k\theta) d\theta, \quad (\text{G.1})$$

where n and k are non-negative integers. To evaluate (G.1), we can derive closed-form expressions for specific values of n and k using the reduction formula [105, Eq. 2.538]

$$F_{n,k} = \frac{1}{n+k} \left(\cos^n(\theta) \sin(k\theta) + n F_{n-1,k-1} \right) \quad (\text{G.2})$$

with the base cases $F_{0,0} = \theta$ and $F_{0,1} = F_{1,0} = \sin(\theta)$. In this appendix, I present such expressions for the values of n and k that are required for evaluating (2.14) and (2.16).

Formulas for the DC and RF Component

For evaluating the linearity and efficiency of the RF-PA model, we only require the DC and RF components of the drain current. Consequently, we need to evaluate (G.1) only for $k \in \{0, 1\}$. In both cases, (G.1) effectively reduces to

$$\int \cos^n(\theta) d\theta. \quad (\text{G.3})$$

The required solutions for solving (2.14) and (2.16) for $k \in \{0, 1\}$ are given by

$$\int \cos^n(\theta) d\theta = \begin{cases} \theta & n = 0, \\ \sin(\theta) & n = 1, \\ \frac{2\theta + \sin(2\theta)}{4} & n = 2, \\ \frac{9\sin(\theta) + \sin(3\theta)}{12} & n = 3, \\ \frac{12\theta + 8\sin(2\theta) + \sin(4\theta)}{32} & n = 4. \end{cases} \quad (\text{G.4})$$

Formulas for the Higher Harmonic Components

To evaluate (2.14) and (2.16) for higher harmonic drain current components, we must solve (G.1) for $n \in \{0, 1, 2, 3\}$ and arbitrary k . Such solutions can be derived with (G.2) using the additional base case $F_{0,k} = -k \sin(k\theta)$. For specific values of k , however, the solutions contain indeterminate expressions with a denominator equal to zero. In these cases, the rule of L'Hospital can be applied to obtain determinate expressions.

The solution for $n = 0$ and arbitrary k is given by

$$\int \cos(k\theta) d\theta = \begin{cases} \theta & k = 0, \\ \frac{\sin(k\theta)}{k} & \text{otherwise.} \end{cases} \quad (\text{G.5})$$

The solution for $n = 1$ and arbitrary k is given by

$$\int \cos(\theta) \cos(k\theta) d\theta = \begin{cases} \frac{2\theta + \sin(2\theta)}{4} & |k| = 1, \\ \frac{k \cos(\theta) \sin(k\theta) - \sin(\theta) \cos(k\theta)}{k^2 - 1} & \text{otherwise.} \end{cases} \quad (\text{G.6})$$

The solution for $n = 2$ and arbitrary k is given by

$$\int \cos^2(\theta) \cos(k\theta) d\theta = \begin{cases} \frac{2\theta + \sin(2\theta)}{4} & k = 0, \\ \frac{4\theta + 4 \sin(2\theta) + \sin(4\theta)}{16} & |k| = 2, \\ \frac{\sin(k\theta)}{2k} + \frac{\sin((k-2)\theta)}{4(k-2)} + \frac{\sin((k+2)\theta)}{4(k+2)} & \text{otherwise.} \end{cases} \quad (\text{G.7})$$

The solution for $n = 3$ and arbitrary k is given by

$$\int \cos^3(\theta) \cos(k\theta) d\theta = \begin{cases} \frac{12\theta + 8 \sin(2\theta) \sin(4\theta)}{32} & |k| = 1, \\ \frac{12\theta + 18 \sin(2\theta) + 9 \sin(4\theta) + 2 \sin(6\theta)}{96} & |k| = 3, \\ \frac{3 \sin((k+1)\theta)}{8(k+1)} + \frac{3 \sin((k-1)\theta)}{8(k-1)} & \\ + \frac{\sin((k-3)\theta)}{8(k-3)} + \frac{\sin((k+3)\theta)}{8(k+3)} & \text{otherwise.} \end{cases} \quad (\text{G.8})$$



Definite Integrals

In Chapter 3, we derive passband-baseband transformation pairs of memoryless models using the Chebyshev transform. In this context, we encounter several definite integrals. In this appendix, I show how these integrals can be solved.

Definite Integral from Equation (3.12)

In (3.12) on page 36, we encounter the expression

$$F_p = \frac{2}{\pi} \int_0^{\pi} \cos^{p+1}(\theta) d\theta, \quad (\text{H.1})$$

where p is a positive integer. Due to symmetry, we have

$$F_p = \frac{2}{\pi} \int_{-\pi}^0 \cos^{p+1}(\theta) d\theta. \quad (\text{H.2})$$

Adding (H.1) and (H.2) and dividing by two, we get

$$F_p = \frac{1}{\pi} \int_{-\pi}^{+\pi} \cos^{p+1}(\theta) d\theta. \quad (\text{H.3})$$

Applying the substitution $\cos(\theta) = \frac{1}{2}(e^{j\theta} + e^{-j\theta})$, we get

$$F_p = \frac{1}{\pi} \frac{1}{2^{p+1}} \int_{-\pi}^{+\pi} (e^{j\theta} + e^{-j\theta})^{p+1} d\theta. \quad (\text{H.4})$$

To further evaluate (H.4), we use the binomial theorem given by

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}, \quad (\text{H.5})$$

where the binomial coefficient is given by

$$\binom{n}{k} = \frac{n!}{(n-k)! k!} \quad (\text{H.6})$$

and the factorial is defined by $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$.

Applying (H.5) to (H.4), we get

$$F_p = \frac{1}{\pi} \frac{1}{2^{p+1}} \int_{-\pi}^{+\pi} \sum_{q=0}^{p+1} \binom{p+1}{q} e^{j\theta q} e^{-j\theta(p+1-q)} d\theta \quad (\text{H.7})$$

$$= \frac{1}{\pi} \frac{1}{2^{p+1}} \sum_{q=0}^{p+1} \binom{p+1}{q} \int_{-\pi}^{+\pi} e^{j\theta(2q-p-1)} d\theta. \quad (\text{H.8})$$

Applying the substitution $k = 2q - p - 1$, we get

$$F_p = \frac{1}{\pi} \frac{1}{2^{p+1}} \sum_{k=-(p+1)}^{+(p+1)} \binom{p+1}{\frac{p+1+k}{2}} \int_{-\pi}^{+\pi} e^{j\theta k} d\theta. \quad (\text{H.9})$$

For integer k , we have

$$\int_{\theta_0}^{\theta_0+2\pi} e^{j\theta k} d\theta = \begin{cases} 2\pi & k = 0, \\ 0 & k \neq 0. \end{cases} \quad (\text{H.10})$$

The index k in (H.9) reaches zero only for odd p . Therefore (H.9) simplifies to

$$F_p = \begin{cases} \lambda_p & p \text{ is odd,} \\ 0 & p \text{ is even,} \end{cases} \quad \text{where} \quad \lambda_p = \frac{1}{2^p} \binom{p+1}{\frac{p+1}{2}}. \quad (\text{H.11})$$

The expression for λ_p in (H.11) is similar to λ_p (3.13). To show that these expressions are actually identical, we use properties of the binomial coefficient that follow from its relation with Pascal's triangle. Using the definition of Pascal's triangle, given by

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad \text{for} \quad \left\{ \begin{array}{l} n, k \in \mathbb{N} \\ 1 \leq k \leq n-1 \end{array} \right\} \quad \text{with} \quad \binom{n}{0} = \binom{n}{n} = 1, \quad (\text{H.12})$$

we can write λ_p in (H.11) as

$$\lambda_p = \frac{1}{2^p} \left(\binom{p}{\frac{p-1}{2}} + \binom{p}{\frac{p+1}{2}} \right). \quad (\text{H.13})$$

Due to the symmetry of the binomial coefficient, we have

$$\binom{n}{k} = \binom{n}{n-k} \implies \binom{p}{\frac{p-1}{2}} = \binom{p}{\frac{p+1}{2}}. \quad (\text{H.14})$$

Therefore, we can simplify (H.13) to

$$\lambda_p = \frac{1}{2^{p-1}} \binom{p}{\frac{p}{2}}, \quad (\text{H.15})$$

which is the same result as in (3.13).

Definite Integral from Equation (3.23)

In (3.23) on page 37, we encounter the expression

$$\frac{p+1}{2} \int_0^{\pi/2} \cos^p(\theta) d\theta, \quad (\text{H.16})$$

where p is a positive integer. The integral in (H.16) and a similar integral with the sine instead of the cosine are commonly known as Wallis' integrals, named after the English mathematician John Wallis. These integrals are related to Wallis' product, given by

$$\frac{\pi}{2} = \frac{2 \cdot 2}{1 \cdot 3} \cdot \frac{4 \cdot 4}{3 \cdot 5} \cdot \frac{6 \cdot 6}{5 \cdot 7} \cdot \frac{8 \cdot 8}{7 \cdot 9} \cdot \frac{10 \cdot 10}{9 \cdot 11} \cdots \quad (\text{H.17})$$

A well-known solution of Wallis' integrals for positive integers p is given by [227]

$$W_p = \int_0^{\pi/2} \sin^p(\theta) d\theta = \int_0^{\pi/2} \cos^p(\theta) d\theta = \begin{cases} \frac{(p-1)!!}{p!!} & p \text{ is odd,} \\ \frac{(p-1)!!}{p!!} \frac{\pi}{2} & p \text{ is even,} \end{cases} \quad (\text{H.18})$$

where the double factorial is defined by

$$p!! = \begin{cases} 1 \cdot 3 \cdot 5 \cdots p & p \text{ is odd,} \\ 2 \cdot 4 \cdot 6 \cdots p & p \text{ is even.} \end{cases} \quad (\text{H.19})$$

Using integration by parts, one can furthermore show that

$$W_p = \frac{p-1}{p} W_{p-2}, \quad W_{p+1} = \frac{p}{p+1} W_{p-1}, \quad W_{p+2} = \frac{p+1}{p+2} W_p. \quad (\text{H.20})$$

In the following, I derive more general solutions of Wallis' integrals using the Euler integrals of the first and second kind, which are also known as the beta and the gamma function, respectively. After that, I show how (H.16) can be related to λ_p in (H.15).

Solution of Wallis' Integrals in terms of the Beta and Gamma Functions

The beta function is defined by

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt, \quad (\text{H.21})$$

where $x, y \in \mathbb{C}$, $\text{Re}(x) > 0$, $\text{Re}(y) > 0$. Substituting $t = \sin^2(\theta)$ in (H.21), we get

$$B(x, y) = 2 \int_0^{\pi/2} \sin^{2x-1}(\theta) \cos^{2y-1}(\theta) d\theta. \quad (\text{H.22})$$

Setting $(x, y) = (\frac{p}{2}, \frac{1}{2})$ in (H.22), we get after some rearrangement

$$W_p = \frac{1}{2} B\left(\frac{p+1}{2}, \frac{1}{2}\right). \quad (\text{H.23})$$

The gamma function is defined by

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt, \quad (\text{H.24})$$

where $x \in \mathbb{C}$, $\text{Re}(x) > 0$. It is related to the beta function by

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \quad (\text{H.25})$$

Substituting (H.25) into (H.23) and noting that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, we get

$$W_p = \frac{\sqrt{\pi} \Gamma\left(\frac{p+1}{2}\right)}{2 \Gamma\left(\frac{p+2}{2}\right)}. \quad (\text{H.26})$$

Applying $\Gamma(x+1) = x\Gamma(x)$ to the denominator of (H.26), we get

$$W_p = \frac{\sqrt{\pi} \Gamma\left(\frac{p+1}{2}\right)}{2 \frac{p}{2} \Gamma\left(\frac{p}{2}\right)} = \frac{\pi}{2p} \left(\frac{2 \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{p}{2}\right)} \right). \quad (\text{H.27})$$

The right side of (H.27) contains a factor equal to $\frac{1}{W_{p-1}}$, so we have

$$W_p = \frac{\pi}{2p} \frac{1}{W_{p-1}} \quad \text{and} \quad W_{p+1} = \frac{\pi}{2(p+1)} \frac{1}{W_p}. \quad (\text{H.28})$$

Relation of Wallis' integrals with the Factor λ_p

In (3.24) on page 37, it is stated that (H.16) is equal to $\frac{1}{\lambda_p}$, where λ_p is given by (H.15), which was derived from (H.1). To prove this statement, we demonstrate that (H.1) can also be expressed in terms of Wallis' integrals. The expression (H.1) is given by

$$F_p = \frac{2}{\pi} \int_0^{\pi} \cos^{p+1}(\theta) d\theta. \quad (\text{H.29})$$

Separating the integration range of (H.29) into two equal parts, we get

$$F_p = \frac{2}{\pi} \left(\int_0^{\pi/2} \cos^{p+1}(\theta) d\theta + \int_{\pi/2}^{\pi} \cos^{p+1}(\theta) d\theta \right), \quad (\text{H.30})$$

which can equivalently be written as

$$F_p = \frac{2}{\pi} \left(\int_0^{\pi/2} \cos^{p+1}(\theta) d\theta + \int_0^{\pi/2} \cos^{p+1}(\theta + \pi/2) d\theta \right). \quad (\text{H.31})$$

Since $\cos(\theta + \pi/2) = -\sin(\theta)$, we can write (H.31) as

$$F_p = \frac{2}{\pi} \left(\int_0^{\pi/2} \cos^{p+1}(\theta) d\theta + (-1)^{p+1} \int_0^{\pi/2} \sin^{p+1}(\theta) d\theta \right). \quad (\text{H.32})$$

Since Wallis' sine and cosine integrals give the same result, it follows from (H.32) that

$$F_p = \begin{cases} \lambda_p & p \text{ is odd,} \\ 0 & p \text{ is even,} \end{cases} \quad (\text{H.33})$$

where the factor λ_p for odd p is given by

$$\lambda_p = \frac{4}{\pi} W_{p+1}. \quad (\text{H.34})$$

Applying the reduction formula for W_{p+1} in (H.28) to W_{p+1} in (H.34), we get

$$\lambda_p = \frac{2}{p+1} \frac{1}{W_p} \quad \text{and therefore} \quad \frac{1}{\lambda_p} = \frac{p+1}{2} W_p, \quad (\text{H.35})$$

which proves that (H.16) evaluates to the inverse of λ_p in (H.15) for odd p .

Computation of the Factor λ_p

We may compute λ_p not only for integers p , but in general for non-negative real numbers $p \in \mathbb{R}_{\geq 0}$. We can derive a formula for this by inserting (H.26) into (H.34), resulting in

$$\lambda_p = \frac{2}{\sqrt{\pi}} \frac{\Gamma\left(\frac{p+2}{2}\right)}{\Gamma\left(\frac{p+3}{2}\right)}, \quad (\text{H.36})$$

which is the result presented in (3.25). For odd values of p , (H.36) gives the same result as (H.15). For other values of p , (H.36) generalizes (H.15). This generalization can also be derived by substituting $n! = \Gamma(n+1)$ in (H.15), resulting in

$$\lambda_p = \frac{1}{2^{p-1}} \frac{\Gamma(p+1)}{\Gamma\left(\frac{p+1}{2}\right) \Gamma\left(\frac{p+3}{2}\right)}. \quad (\text{H.37})$$

Applying the duplication formula of the gamma function, given by [105, Eq. 8.335-1]

$$\Gamma(2x) = \frac{2^{2x-1}}{\sqrt{\pi}} \Gamma(x) \Gamma\left(x + \frac{1}{2}\right) \quad (\text{H.38})$$

to the numerator of (H.37), we can show that (H.36) and (H.37) are equivalent.

For positive integers p , we can insert (H.18) into (H.34), resulting in

$$\lambda_p = \begin{cases} 2 \frac{p!!}{(p+1)!!} & p \text{ is odd,} \\ \frac{4}{\pi} \frac{p!!}{(p+1)!!} & p \text{ is even.} \end{cases} \quad (\text{H.39})$$

For positive integers p , we may also use a recursive formula for λ_p , derived from the recursive formula for W_p in (H.20). The recursion is given by

$$\lambda_p = \frac{p}{p+1} \lambda_{p-2} \quad \text{with} \quad \lambda_1 = 1 \quad \text{and} \quad \lambda_2 = \frac{1}{\pi} \frac{8}{3}. \quad (\text{H.40})$$

Definite Integral from Equation (3.36)

In (3.36) on page 40, we encounter the expression

$$F_{p,k} = \frac{2}{\pi} \int_0^{\pi/2} \cos^p(x) \cos(k\theta) d\theta. \quad (\text{H.41})$$

The work of Nelson Blachman indicates that the solution of (H.41) can be written very concisely using a generalized binomial coefficient¹, given by [26, Tab. I (T4)]

$$F_{p,k} = \frac{1}{2^p} \binom{p}{\frac{p-k}{2}}, \quad \text{where} \quad \binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(x-y+1)\Gamma(y+1)}. \quad (\text{H.42})$$

Nelson Blachman presented (H.42) without a derivation.

Using the table of integrals by Gradshteyn and Ryzhik, one finds an alternative solution for the expression in (H.41)², given by [105, Eq. 3.631-9]

$$F_{p,k} = \frac{1}{2^p (p+1) B\left(\frac{p+k+2}{2}, \frac{p-k+2}{2}\right)}. \quad (\text{H.43})$$

The integrals in [105] are also presented without derivations.

We can verify that (H.42) and (H.43) are equivalent by substituting the beta function in (H.43) with its representation by the gamma function in (H.25), resulting in

$$F_{p,k} = \frac{1}{2^p (p+1)} \frac{\Gamma(p+2)}{\Gamma\left(\frac{p+k+2}{2}\right) \Gamma\left(\frac{p-k+2}{2}\right)}. \quad (\text{H.44})$$

Substituting $\Gamma(p+2) = (p+1)\Gamma(p+1)$ in the numerator of (H.44), we get

$$F_{p,k} = \frac{1}{2^p} \frac{\Gamma(p+1)}{\Gamma\left(\frac{p+k+2}{2}\right) \Gamma\left(\frac{p-k+2}{2}\right)}, \quad (\text{H.45})$$

and by noting that the binomial coefficient in (H.42) can be written as

$$\binom{p}{\frac{p-k}{2}} = \frac{\Gamma(p+1)}{\Gamma\left(\frac{p+k+2}{2}\right) \Gamma\left(\frac{p-k+2}{2}\right)}, \quad (\text{H.46})$$

we have verified the equivalence of (H.42) and (H.43).

¹An investigation of this generalized binomial coefficient can be found in [92].

²Note that the factor $2/\pi$ is included in our definition of $F_{p,k}$.

An actual derivation of a solution for the integral in (H.41) can be found in [88]. In this work, the author investigates more general integrals similar to (H.41), given by

$$\int_0^{\pi/2} \cos^p(\theta) e^{jk\theta} d\theta \quad \text{and} \quad \int_0^{\pi/2} \sin^p(\theta) e^{jk\theta} d\theta. \quad (\text{H.47})$$

By solving (H.47), the author also provides a solution for (H.41)³, given by [88, Eq. 13 a]

$$F_{p,k} = \frac{1}{\pi} \frac{1}{2^p} \sin\left(\frac{k-p}{2}\pi\right) \Gamma\left(\frac{k-p}{2}\right) \frac{\Gamma(p+1)}{\Gamma\left(\frac{k+p+2}{2}\right)}. \quad (\text{H.48})$$

In the following, we verify that (H.48) is equivalent to (H.42) and (H.43).

By multiplying (H.48) with $\Gamma(\frac{p-k+2}{2})/\Gamma(\frac{p-k+2}{2})$, we can write it as

$$F_{p,k} = \frac{1}{\pi} \frac{1}{2^p} \sin\left(\frac{k-p}{2}\pi\right) \Gamma\left(\frac{k-p}{2}\right) \Gamma\left(\frac{p-k+2}{2}\right) \binom{p}{\frac{p-k}{2}}. \quad (\text{H.49})$$

By rewriting the argument in the second gamma function of (H.49), we get

$$F_{p,k} = \frac{1}{\pi} \frac{1}{2^p} \sin\left(\frac{k-p}{2}\pi\right) \Gamma\left(\frac{k-p}{2}\right) \Gamma\left(1 - \frac{k-p}{2}\right) \binom{p}{\frac{p-k}{2}}. \quad (\text{H.50})$$

Now, we can apply Euler's reflection formula, given by [105, Eq. 8.334-3]

$$\Gamma(x) \Gamma(1-x) = \frac{\pi}{\sin(x\pi)}, \quad (\text{H.51})$$

to simplify (H.49) to

$$F_{p,k} = \frac{1}{2^p} \frac{\sin\left(\frac{k-p}{2}\pi\right)}{\sin\left(\frac{k-p}{2}\pi\right)} \binom{p}{\frac{p-k}{2}} = \frac{1}{2^p} \binom{p}{\frac{p-k}{2}}, \quad (\text{H.52})$$

which is identical to (H.42) and therefore also equivalent to (H.43).

In the special case of $p = 0$, (H.48) can be reduced to a sinc function. For this purpose, we note that $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$, so we can write

$$F_{0,k} = \frac{1}{\pi} \sin\left(\frac{k\pi}{2}\right) \Gamma\left(\frac{k}{2}\right) \frac{1}{\Gamma\left(\frac{k}{2}+1\right)} = \frac{\sin\left(\frac{k\pi}{2}\right)}{\frac{k\pi}{2}}. \quad (\text{H.53})$$

³Note that the factor $2/\pi$ is included in our definition of $F_{p,k}$.

Bibliography

- [1] 3GPP TSG RAN WG1 37, *Comparison of PAR and cubic metric for power de-rating*, Tdoc R1-040522, Motorola, Montreal, Canada, May 2004 (cited on page 20).
- [2] T. Adalı and S. Haykin, *Adaptive Signal Processing: Next Generation Solutions*. Wiley-IEEE Press, 2010 (cited on pages 88, 90).
- [3] T. Adalı and P. J. Schreier, “Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation”, *IEEE Signal Processing Magazine*, volume 31, number 5, pages 112–128, Sep. 2014 (cited on page 88).
- [4] T. Adalı, P. J. Schreier, and L. L. Scharf, “Complex-valued signal processing: The proper way to deal with impropriety”, *IEEE Transactions on Signal Processing*, volume 59, number 11, pages 5101–5125, Nov. 2011 (cited on page 88).
- [5] W. T. Adams and J. Brady, “Magnitude approximations for microprocessor implementation”, *IEEE Micro*, volume 3, number 5, pages 27–31, Oct. 1983 (cited on page 84).
- [6] S. Afsardoost, T. Eriksson, and C. Fager, “Digital predistortion using a vector-switched model”, *IEEE Transactions on Microwave Theory and Techniques*, volume 60, number 4, pages 1166–1174, Apr. 2012 (cited on page 75).
- [7] W. Ahmed, “Fast orthogonal search for training radial basis function neural networks”, Master’s thesis, University of Maine, Aug. 1994 (cited on page 64).
- [8] A. Alizadeh, M. Frounchi, and A. Medi, “Dual-band design of integrated Class-J power amplifiers in GaAs pHEMT technology”, *IEEE Transactions on Microwave Theory and Techniques*, volume 65, number 8, pages 3034–3045, Aug. 2017 (cited on page 27).
- [9] C. M. Andersson, D. Gustafsson, K. Yamanaka, E. Kuwata, H. Otsuka, M. Nakayama, Y. Hirano, I. Angelov, C. Fager, and N. Rorsman, “Theory and design of Class-J power amplifiers with dynamic load modulation”, *IEEE Transactions on Microwave Theory and Techniques*, volume 60, number 12, pages 3778–3786, Dec. 2012 (cited on page 28).
- [10] L. Anttila, P. Handel, and M. Valkama, “Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters”, *IEEE Transactions on Microwave Theory and Techniques*, volume 58, number 4, pages 730–739, Apr. 2010 (cited on page 64).
- [11] L. Anttila, M. Valkama, and M. Renfors, “Frequency-selective I/Q mismatch calibration of wideband direct-conversion transmitters”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, volume 55, number 4, pages 359–363, Apr. 2008 (cited on pages 3, 64).

- [12] P. Asbeck and Z. Popovic, “ET comes of age: Envelope tracking for higher-efficiency power amplifiers”, *IEEE Microwave Magazine*, volume 17, number 3, pages 16–25, Mar. 2016 (cited on page 28).
- [13] F. M. Barradas, T. R. Cunha, P. M. Lavrador, and J. C. Pedro, “Polynomials and LUTs in PA behavioral modeling: A fair theoretical comparison”, *IEEE Transactions on Microwave Theory and Techniques*, volume 62, number 12, pages 3274–3285, Dec. 2014 (cited on page 75).
- [14] F. M. Barradas, L. C. Nunes, J. C. Pedro, T. R. Cunha, P. M. Lavrador, and P. M. Cabral, “Accurate linearization with low-complexity models using cascaded digital predistortion systems”, *IEEE Microwave Magazine*, volume 16, number 1, pages 94–103, Feb. 2015 (cited on pages 67, 79).
- [15] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd edition. Kluwer Academic Publishers, 2004 (cited on page 31).
- [16] T. Barton, “Not just a phase: Outphasing power amplifiers”, *IEEE Microwave Magazine*, volume 17, number 2, pages 18–31, Feb. 2016 (cited on page 28).
- [17] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, “2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters”, *IEEE Transactions on Microwave Theory and Techniques*, volume 59, number 10, pages 2547–2553, Oct. 2011 (cited on page 75).
- [18] E. Bedrosian and S. O. Rice, “The output properties of Volterra systems (non-linear systems with memory) driven by harmonic and Gaussian inputs”, *Proceedings of the IEEE*, volume 59, number 12, pages 1688–1707, Dec. 1971 (cited on page 31).
- [19] E. Bedrosian, “The analytic signal representation of modulated waveforms”, *Proceedings of the IRE*, volume 50, number 10, pages 2071–2076, Oct. 1962 (cited on page 85).
- [20] A. Behravan and T. Eriksson, “PAPR and other measures for OFDM systems with nonlinearity”, in *Proceedings of the 5th International Symposium on Wireless Personal Multimedia Communications*, Honolulu, Hawaii, Oct. 2002, pages 149–153 (cited on page 20).
- [21] A. Behravan and T. Eriksson, “Some statistical properties of multicarrier signals and related measures”, in *Proceedings of the 63rd IEEE Vehicular Technology Conference*, Melbourne, Australia, May 2006, pages 1854–1858 (cited on page 20).
- [22] S. Benedetto, E. Biglieri, and R. Daffara, “Modeling and performance evaluation of nonlinear satellite links – A Volterra series approach”, *IEEE Transactions on Aerospace and Electronic Systems*, volume 15, number 4, pages 494–507, Jul. 1979 (cited on pages 31, 32, 53).
- [23] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. Kluwer Academic Publishers, 1999 (cited on page 34).

- [24] A. Birafane, M. El-Asmar, A. B. Kouki, M. Helaoui, and F. M. Ghannouchi, “Analyzing LINC systems”, *IEEE Microwave Magazine*, volume 11, number 5, pages 59–71, Aug. 2010 (cited on page 28).
- [25] N. Blachman, “Band-pass nonlinearities”, *IEEE Transactions on Information Theory*, volume 10, number 2, pages 162–164, Apr. 1964 (cited on pages 32, 34, 65).
- [26] N. Blachman, “Detectors, bandpass nonlinearities, and their optimization: Inversion of the Chebyshev transform”, *IEEE Transactions on Information Theory*, volume 17, number 4, pages 398–404, Jul. 1971 (cited on pages 32, 33, 35, 40, 65, 114).
- [27] H. S. Black, “Inventing the negative feedback amplifier”, *IEEE Spectrum*, volume 14, number 12, pages 55–60, Dec. 1977 (cited on page 29).
- [28] T. S. Bolis, “Degenerate critical points”, *Mathematics Magazine*, volume 53, number 5, pages 294–299, Nov. 1980 (cited on page 91).
- [29] A. van den Bos, “Complex gradient and Hessian”, *IEE Proceedings - Vision, Image and Signal Processing*, volume 141, number 6, pages 380–382, 1994 (cited on pages 88, 90, 91).
- [30] R. Boucher and J. Hassab, “Analysis of discrete implementation of generalized cross correlator”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, volume 29, number 3, pages 609–611, Jun. 1981 (cited on pages 101, 102).
- [31] S. Boyd and L. Chua, “Fading memory and the problem of approximating nonlinear operators with Volterra series”, *IEEE Transactions on Circuits and Systems*, volume 32, number 11, pages 1150–1161, Nov. 1985 (cited on page 31).
- [32] R. N. Braithwaite, “A comparison of indirect learning and closed loop estimators used in digital predistortion of power amplifiers”, in Proceedings of the IEEE International Microwave Symposium (IMS), Phoenix, Arizona, May 2015, pages 1–4 (cited on page 77).
- [33] D. H. Brandwood, “A complex gradient operator and its application in adaptive array theory”, *IEE Proceedings F: Communications, Radar and Signal Processing*, volume 130, number 1, pages 11–16, Feb. 1983 (cited on pages 88, 89, 92).
- [34] P. M. Cabral, J. C. Pedro, and N. B. Carvalho, “Dynamic AM-AM and AM-PM behavior in microwave PA circuits”, in Proceedings of the Asia-Pacific Microwave Conference (APMC), Suzhou, China, Dec. 2005, pages 1–4 (cited on page 30).
- [35] V. Camarchia, M. Pirola, R. Quaglia, S. Jee, Y. Cho, and B. Kim, “The Doherty power amplifier: Review of recent solutions and trends”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 2, pages 559–571, Feb. 2015 (cited on page 28).

- [36] H. Cao, H. M. Nemati, A. S. Tehrani, T. Eriksson, and C. Fager, “Digital predistortion for high efficiency power amplifier architectures using a dual-input modeling approach”, *IEEE Transactions on Microwave Theory and Techniques*, volume 60, number 2, pages 361–369, Feb. 2012 (cited on page 28).
- [37] H. Cao, A. S. Tehrani, C. Fager, T. Eriksson, and H. Zirath, “I/Q imbalance compensation using a nonlinear modeling approach”, *IEEE Transactions on Microwave Theory and Techniques*, volume 57, number 3, pages 513–518, Mar. 2009 (cited on page 64).
- [38] R. Caverly, F. Raab, and J. Staudinger, “High-efficiency power amplifiers”, *IEEE Microwave Magazine*, volume 13, number 7, pages S22-S32, Nov. 2012 (cited on pages 6, 26).
- [39] J. Chani-Cahuana, M. Özen, C. Fager, and T. Eriksson, “Digital predistortion parameter identification for RF power amplifiers using real-valued output data”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, volume 64, number 10, pages 1227–1231, Oct. 2017 (cited on page 95).
- [40] W. Chen, S. Zhang, Y.-J. Liu, F. M. Ghannouchi, Z. Feng, and Y. Liu, “Efficient pruning technique of memory polynomial models suitable for PA behavioral modeling and digital predistortion”, *IEEE Transactions on Microwave Theory and Techniques*, volume 62, number 10, pages 2290–2299, Oct. 2014 (cited on page 64).
- [41] X. Chen, S. Zhang, and W. Chen, “Two-dimensional crest factor reduction for performance improvement of concurrent dual-band power amplifiers”, *Electronics Letters*, volume 49, number 18, pages 1163–1165, Aug. 2013 (cited on page 71).
- [42] S. Chi, P. Singerl, and C. Vogel, “Efficiency optimization for burst-mode multi-level radio frequency transmitters”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 60, number 7, pages 1901–1914, Jul. 2013 (cited on page 28).
- [43] H. Chireix, “High power outphasing modulation”, *Proceedings of the IRE*, volume 23, number 11, pages 1370–1392, Nov. 1935 (cited on page 28).
- [44] L. O. Chua and C. Y. Ng, “Frequency domain analysis of nonlinear systems: General theory”, *IEE Journal on Electronic Circuits and Systems*, volume 3, number 4, pages 165–185, Jul. 1979 (cited on page 31).
- [45] L. O. Chua and C. Y. Ng, “Frequency-domain analysis of nonlinear systems: Formulation of transfer functions”, *IEE Journal on Electronic Circuits and Systems*, volume 3, number 6, pages 257–269, Nov. 1979 (cited on page 31).
- [46] P. Colantonio, F. Giannini, and E. Limiti, *High Efficiency RF and Microwave Solid State Power Amplifiers*. Wiley, 2009 (cited on page 26).
- [47] D. Cox, “Linear amplification with nonlinear components”, *IEEE Transactions on Communications*, volume 22, number 12, pages 1942–1945, Dec. 1974 (cited on page 28).

- [48] C. Crespo-Cadenas, M. J. Madero-Ayora, J. Reina-Tosina, and J. A. Becerra-González, “A widely nonlinear approach to compensate impairments in I/Q modulators”, in Proceedings of the European Microwave Conference (EuMC), Paris, France, Sep. 2015, pages 506–509 (cited on page 64).
- [49] C. Crespo-Cadenas, M. J. Madero-Ayora, J. Reina-Tosina, and J. Becerra-González, “Formal deduction of a Volterra series model for complex-valued systems”, *Signal Processing*, volume 131, pages 245–248, 2017 (cited on page 64).
- [50] C. Crespo-Cadenas, J. Reina-Tosina, and M. J. Madero-Ayora, “Volterra behavioral model for wideband RF amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 55, number 3, pages 449–457, Mar. 2007 (cited on page 64).
- [51] C. Crespo-Cadenas, J. Reina-Tosina, M. J. Madero-Ayora, and J. Munoz-Cruzado, “A new approach to pruning Volterra models for power amplifiers”, *IEEE Transactions on Signal Processing*, volume 58, number 4, pages 2113–2120, Apr. 2010 (cited on page 64).
- [52] S. C. Cripps, *Advanced Techniques in RF Power Amplifier Design*, 1st edition. Artech House, 2002 (cited on page 26).
- [53] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd edition. Artech House, 2006 (cited on pages 6–8, 12–15, 26, 27).
- [54] S. C. Cripps, “The intercept point deception”, *IEEE Microwave Magazine*, volume 8, number 1, pages 44–50, Feb. 2007 (cited on page 32).
- [55] S. C. Cripps, “Intermodulation squared response”, *IEEE Microwave Magazine*, volume 16, number 5, pages 102–106, Jun. 2015 (cited on page 32).
- [56] S. C. Cripps, P. J. Tasker, A. L. Clarke, J. Lees, and J. Benedikt, “On the continuity of high efficiency modes in linear RF power amplifiers”, *IEEE Microwave and Wireless Components Letters*, volume 19, number 10, pages 665–667, Oct. 2009 (cited on pages 3, 27).
- [57] T. R. Cunha, E. G. Lima, and J. C. Pedro, “A polar-oriented Volterra model for power amplifier characterization”, in Proceedings of the IEEE International Microwave Symposium (IMS), Anaheim, California, May 2010, pages 556–559 (cited on page 32).
- [58] T. R. Cunha, E. G. Lima, and J. C. Pedro, “Validation and physical interpretation of the power-amplifier polar Volterra model”, *IEEE Transactions on Microwave Theory and Techniques*, volume 58, number 12, pages 4012–4021, Dec. 2010 (cited on page 32).
- [59] R. Darraji, P. Mousavi, and F. M. Ghannouchi, “Doherty goes digital: Digitally enhanced Doherty power amplifiers”, *IEEE Microwave Magazine*, volume 17, number 8, pages 41–51, Aug. 2016 (cited on page 28).

- [60] M. Deumal, A. Behravan, T. Eriksson, and J. L. Pijoan, “Evaluation of performance improvement capabilities of PAPR-reducing methods”, *Wireless Personal Communications*, volume 47, number 1, pages 137–147, Oct. 2008 (cited on page 20).
- [61] M. Deumal, A. Behravan, and J. L. Pijoan, “On cubic metric reduction in OFDM systems by tone reservation”, *IEEE Transactions on Communications*, volume 59, number 6, pages 1612–1620, Jun. 2011 (cited on page 20).
- [62] L. Ding and G. T. Zhou, “Effects of even-order nonlinear terms on power amplifier modeling and predistortion linearization”, *IEEE Transactions on Vehicular Technology*, volume 53, number 1, pages 156–162, Jan. 2004 (cited on pages 5, 31).
- [63] L. Ding, “Digital predistortion of power amplifiers for wireless applications”, Doctoral thesis, Georgia Institute of Technology, Mar. 2004 (cited on pages 5, 31, 81).
- [64] L. Ding, Z. Ma, D. R. Morgan, M. Zierdt, and J. Pastalan, “A least-squares/Newton method for digital predistortion of wideband signals”, *IEEE Transactions on Communications*, volume 54, number 5, pages 833–840, May 2006 (cited on page 5).
- [65] L. Ding, Z. Ma, D. R. Morgan, M. Zierdt, and G. T. Zhou, “Compensation of frequency-dependent gain/phase imbalance in predistortion linearization systems”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 55, number 1, pages 390–397, Feb. 2008 (cited on pages 5, 64, 81).
- [66] L. Ding, F. Mujica, and Z. Yang, “Digital predistortion using direct learning with reduced bandwidth feedback”, in Proceedings of the IEEE International Microwave Symposium (IMS), Seattle, Washington, Jun. 2013, pages 1–3 (cited on page 77).
- [67] L. Ding, H. Qian, N. Chen, and G. T. Zhou, “A memory polynomial predistorter implemented using TMS320C67XX”, in Proceedings of the Texas Instruments Developer Conference, Houston, Texas, 2004, pages 1–7 (cited on page 5).
- [68] L. Ding, R. Raich, and G. T. Zhou, “A Hammerstein predistortion linearization design based on the indirect learning architecture”, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, Florida, May 2002, pages 2689–2692 (cited on page 5).
- [69] L. Ding and G. T. Zhou, “Effects of even-order nonlinear terms on predistortion linearization”, in Proceedings of the IEEE 10th Digital Signal Processing Workshop and 2nd Signal Processing Education Workshop, Pine Mountain, Georgia, Oct. 2002, pages 1–6 (cited on page 31).
- [70] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, “A robust digital baseband predistorter constructed using memory polynomials”, *IEEE Transactions on Communications*, volume 52, number 1, pages 159–165, Jan. 2004 (cited on pages 5, 81).

- [71] W. H. Doherty, “A new high efficiency power amplifier for modulated waves”, *Proceedings of the IRE*, volume 24, number 9, pages 1163–1182, Sep. 1936 (cited on page 28).
- [72] M. M. Ebrahimi, M. Helaloui, and F. M. Ghannouchi, “Delta-sigma-based transmitters: Advantages and disadvantages”, *IEEE Microwave Magazine*, volume 14, number 1, pages 68–78, Jan. 2013 (cited on page 28).
- [73] H. Enzinger, *bibget – The easiest way to get BibTeX entries from IEEE Xplore*, MATLAB source code, available online at: <http://mathworks.com/matlabcentral/fileexchange/53412>, Oct. 2015 (cited on page 4).
- [74] H. Enzinger, *Joint linearity-efficiency model*, MATLAB source code, available online at: <http://mathworks.com/matlabcentral/fileexchange/53413>, Oct. 2015 (cited on pages 4, 6, 12, 14, 15).
- [75] H. Enzinger, *Fast time-domain Volterra filtering implemented in C*, C source code, available online at: <http://spsc.tugraz.at/tools/fast-time-domain-volterra-filtering>, Oct. 2016 (cited on pages 4, 54).
- [76] H. Enzinger, K. Freiberger, G. Kubin, and C. Vogel, “A survey of delay and gain correction methods for the indirect learning of digital predistorters”, in Proceedings of the IEEE International Conference on Electronics, Circuits and Systems (ICECS), Monte Carlo, Monaco, Dec. 2016, pages 285–288 (cited on pages 4, 81).
- [77] H. Enzinger, K. Freiberger, G. Kubin, and C. Vogel, “Baseband Volterra filters with even-order terms: Theoretical foundation and practical implications”, in Proceedings of the 50th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, Nov. 2016, pages 220–224 (cited on pages 4, 30, 64, 81).
- [78] H. Enzinger, K. Freiberger, G. Kubin, and C. Vogel, “Fast time-domain Volterra filtering”, in Proceedings of the 50th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, Nov. 2016, pages 225–228 (cited on pages 4, 31, 54, 57, 81).
- [79] H. Enzinger, K. Freiberger, and C. Vogel, “Analysis of even-order terms in memoryless and quasi-memoryless polynomial baseband models”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, May 2015, pages 1714–1717 (cited on pages 4, 30, 81).
- [80] H. Enzinger, K. Freiberger, and C. Vogel, “A joint linearity-efficiency model of radio frequency power amplifiers”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Montreal, Canada, May 2016, pages 281–284 (cited on pages 4, 6, 14, 81).
- [81] H. Enzinger, K. Freiberger, and C. Vogel, “Competitive linearity for envelope tracking: Dual-band crest factor reduction and 2D-vector-switched digital predistortion”, *IEEE Microwave Magazine*, volume 19, number 1, pages 69–77, 2018 (cited on pages 4, 66, 67, 79, 81).

- [82] H. Enzinger and C. Vogel, “Analytical description of multilevel carrier-based PWM of arbitrary bounded input signals”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Melbourne, Australia, Jun. 2014, pages 1030–1033 (cited on pages 4, 28, 34, 65, 81).
- [83] M. Eron, B. Kim, F. Raab, R. Caverly, and J. Staudinger, “The head of the class”, *IEEE Microwave Magazine*, volume 12, number 7, pages S16-S33, Dec. 2011 (cited on page 12).
- [84] L. Erup, F. M. Gardner, and R. A. Harris, “Interpolation in digital modems – Part II: Implementation and performance”, *IEEE Transactions on Communications*, volume 41, number 6, pages 998–1008, Jun. 1993 (cited on page 103).
- [85] L. Feldman, “Class G high efficiency Hi-Fi amplifier”, *Radio Electronics*, volume 47, number 8, pages 47–49, Aug. 1976, <http://americanradiohistory.com> (cited on page 27).
- [86] L. Feldman, “Class H variproportional amplifier”, *Radio Electronics*, volume 48, number 10, pages 53–55, Oct. 1977, <http://americanradiohistory.com> (cited on page 27).
- [87] X. Feng, P. Zhang, and T. Cui, “Effects of even-order nonlinear terms and over-sampling rate on digital baseband predistortion linearization”, in Proceedings of the International Conference on Communication, Electronics and Automation Engineering, Xi’an, China, Aug. 2012, pages 53–59 (cited on page 32).
- [88] H. E. Fettis, “More trigonometric integrals”, *Mathematics of Computation*, volume 43, number 168, pages 557–564, 1984 (cited on page 115).
- [89] A. Filip, “Linear approximations to $\sqrt{x^2 + y^2}$ having equiripple error characteristics”, *IEEE Transactions on Audio and Electroacoustics*, volume 21, number 6, pages 554–556, Dec. 1973 (cited on page 84).
- [90] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Wiley IEEE Press, 2002 (cited on pages 88, 90).
- [91] P. O. Fisher and S. F. Al-Sarawi, “An optimized segmented quasi-memoryless nonlinear behavioral modeling approach for RF power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 66, number 1, pages 294–305, Jan. 2017 (cited on page 50).
- [92] D. Fowler, “The binomial coefficient function”, *The American Mathematical Monthly*, volume 103, number 1, pages 1–17, Jan. 1996 (cited on page 114).
- [93] W. A. Frank, “Sampling requirements for Volterra system identification”, *IEEE Signal Processing Letters*, volume 3, number 9, pages 266–268, Sep. 1996 (cited on page 95).
- [94] K. Freiberger, “Measurement methods for estimating the error vector magnitude of OFDM transceivers”, Doctoral thesis, Graz University of Technology, Dec. 2017 (cited on page 22).

- [95] K. Freiberger, H. Enzinger, and C. Vogel, “A noise power ratio measurement method for accurate estimation of the error vector magnitude”, *IEEE Transactions on Microwave Theory and Techniques*, volume 65, number 5, pages 1632–1645, May 2017 (cited on pages 4, 22, 81).
- [96] K. Freiberger, H. Enzinger, and C. Vogel, “SLIC EVM – Error vector magnitude without demodulation”, in Proceedings of the 89th ARFTG Microwave Measurement Conference, Honolulu, Hawaii, Jun. 2017, pages 1–4 (cited on pages 4, 22, 81).
- [97] K. Freiberger, H. Enzinger, and C. Vogel, “The error power ratio estimates EVM for a wide class of impairments: Monte Carlo simulations”, in Proceedings of the International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INMMiC), Graz, Austria, Apr. 2017, pages 1–3 (cited on pages 4, 22, 81).
- [98] K. Freiberger, M. Wolkerstorfer, H. Enzinger, and C. Vogel, “Digital predistorter identification based on constrained multi-objective optimization of WLAN standard performance metrics”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, May 2015, pages 862–865 (cited on pages 4, 81).
- [99] S. Gao, “High efficiency class-F RF/microwave power amplifiers”, *IEEE Microwave Magazine*, volume 7, number 1, pages 40–48, Feb. 2006 (cited on page 27).
- [100] F. M. Ghannouchi and O. Hammi, “Behavioral modeling and predistortion”, *IEEE Microwave Magazine*, volume 10, number 7, pages 52–64, Dec. 2009 (cited on pages 3, 64).
- [101] F. M. Ghannouchi, O. Hammi, and M. Helaoui, *Behavioral Modeling and Predistortion of Wideband Wireless Transmitters*. Wiley, 2015 (cited on pages 3, 29, 64).
- [102] P. L. Gilabert, “Multi look-up table digital predistortion for RF power amplifier linearization”, Doctoral thesis, Polytechnic University of Catalonia, Dec. 2007 (cited on page 61).
- [103] P. L. Gilabert and G. Montoro, “3-D distributed memory polynomial behavioral model for concurrent dual-band envelope tracking power amplifier linearization”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 2, pages 638–648, Feb. 2015 (cited on pages 70, 78).
- [104] P. L. Gilabert, G. Montoro, D. López, N. Bartzoudis, E. Bertran, M. Payaró, and A. Hourtane, “Order reduction of wideband digital predistorters using principal component analysis”, in Proceedings of the IEEE International Microwave Symposium (IMS), Seattle, Washington, Jun. 2013, pages 1–7 (cited on page 64).
- [105] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 8th edition. Academic Press, 2014 (cited on pages 107, 113–115).

- [106] A. Grebennikov, “Early history of switching-mode Class-E techniques for high-efficiency power amplification”, in Proceedings of the IEEE International Microwave Symposium (IMS), Honolulu, Hawaii, Jun. 2017, pages 1315–1318 (cited on page 26).
- [107] A. Grebennikov, N. O. Sokal, and M. J. Franco, *Switchmode RF and Microwave Power Amplifiers*, 2nd edition. Academic Press, 2012 (cited on pages 26–28).
- [108] L. Guan and A. Zhu, “Dual-loop model extraction for digital predistortion of wideband RF power amplifiers”, *IEEE Microwave and Wireless Components Letters*, volume 21, number 9, pages 501–503, Sep. 2011 (cited on page 77).
- [109] L. Guan and A. Zhu, “Simplified dynamic deviation reduction-based Volterra model for Doherty power amplifiers”, in Proceedings of the International Workshop on Integrated Nonlinear Microwave and Millimetre-wave Circuits (INM-MiC), Vienna, Austria, Apr. 2011, pages 1–4 (cited on pages 29, 31).
- [110] L. Guan and A. Zhu, “Green communications: Digital predistortion for wideband RF power amplifiers”, *IEEE Microwave Magazine*, volume 15, number 7, pages 84–99, Nov. 2014 (cited on pages 1, 3).
- [111] S. H. Han and J. H. Lee, “An overview of peak-to-average power ratio reduction techniques for multicarrier transmission”, *IEEE Wireless Communications*, volume 12, number 2, pages 56–65, Apr. 2005 (cited on page 71).
- [112] J. Harmon and S. G. Wilson, “Iterative approach to the indirect learning architecture for baseband digital predistortion”, in Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Miami, Florida, Dec. 2010, pages 1–5 (cited on page 31).
- [113] J. Harmon and S. G. Wilson, “Complex nonlinear adaptive predistortion”, in Proceedings of the 46th Annual Conference on Information Sciences and Systems (CISS), Princeton, New Jersey, Mar. 2012, pages 1–6 (cited on page 77).
- [114] K. Hausmair, S. Chi, P. Singerl, and C. Vogel, “Aliasing-free digital pulse-width modulation for burst-mode RF transmitters”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 60, number 2, pages 415–427, Feb. 2013 (cited on pages 28, 65).
- [115] K. Hausmair, S. Chi, and C. Vogel, “How to reach 100% coding efficiency in multilevel burst-mode RF transmitters”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, May 2013, pages 2255–2258 (cited on page 28).
- [116] M. Isaksson, D. Wisell, and D. Ronnow, “A comparative analysis of behavioral models for RF power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 54, number 1, pages 348–359, Jan. 2006 (cited on page 31).
- [117] X. Jiang, “Fundamentals of audio class D amplifier design: A review of schemes and architectures”, *IEEE Solid-State Circuits Magazine*, volume 9, number 3, pages 14–25, 2017 (cited on pages 26, 27).

- [118] T. Johnson and S. P. Stapleton, “RF class-D amplification with bandpass sigma-delta modulator drive signals”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 53, number 12, pages 2507–2520, Dec. 2006 (cited on page 26).
- [119] L. R. Kahn, “Single-sideband transmission by envelope elimination and restoration”, *Proceedings of the IRE*, volume 40, number 7, pages 803–806, Jul. 1952 (cited on page 28).
- [120] A. Katz, R. Gray, and R. Dorval, “Truly wideband linearization”, *IEEE Microwave Magazine*, volume 10, number 7, pages 20–27, Dec. 2009 (cited on page 29).
- [121] A. Katz, J. Wood, and D. Chokola, “The evolution of PA linearization: From classic feedforward and feedback through analog and digital predistortion”, *IEEE Microwave Magazine*, volume 17, number 2, pages 32–40, Feb. 2016 (cited on pages 3, 29).
- [122] A. Kaye, D. George, and M. Eric, “Analysis and compensation of bandpass nonlinearities for communications”, *IEEE Transactions on Communications*, volume 20, number 5, pages 965–972, Oct. 1972 (cited on page 46).
- [123] N. Kelly, W. Cao, and A. Zhu, “Preparing linearity and efficiency for 5G: Digital predistortion for dual-band Doherty power amplifiers with mixed-mode carrier aggregation”, *IEEE Microwave Magazine*, volume 18, number 1, pages 76–84, Jan. 2017 (cited on pages 67, 71, 79).
- [124] P. B. Kenington, *High-Linearity RF Amplifier Design*. Artech House, 2000 (cited on pages 27, 29).
- [125] J. Kim and K. Konstantinou, “Digital predistortion of wideband signals based on power amplifier model with memory”, *Electronics Letters*, volume 37, number 23, pages 1417–1418, Nov. 2001 (cited on page 31).
- [126] W.-J. Kim, K.-J. Cho, S. P. Stapleton, and J.-H. Kim, “An efficient crest factor reduction technique for wideband applications”, *Analog Integrated Circuits and Signal Processing*, volume 51, number 1, pages 19–26, Apr. 2007 (cited on page 71).
- [127] M. J. Korenberg, “A robust orthogonal algorithm for system identification and time-series analysis”, *Biological Cybernetics*, volume 60, number 4, pages 267–276, Feb. 1989 (cited on page 64).
- [128] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*. John Wiley & Sons, 1980 (cited on page 27).
- [129] K. Kreutz-Delgado, *The complex gradient operator and the CR-calculus*, available online at: <https://arxiv.org/abs/0906.4835v1>, Jun. 2009 (cited on pages 52, 88, 90, 91).
- [130] H. Ku, M. D. McKinley, and J. S. Kenney, “Quantifying memory effects in RF power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 50, number 12, pages 2843–2849, Dec. 2002 (cited on page 81).

- [131] H. Ku and J. S. Kenney, “Behavioral modeling of nonlinear RF power amplifiers considering memory effects”, *IEEE Transactions on Microwave Theory and Techniques*, volume 51, number 12, pages 2495–2504, Dec. 2003 (cited on page 31).
- [132] L. J. Kushner, “Output performance of idealized microwave power amplifiers”, *Microwave Journal*, volume 32, number 10, pages 103–112, Oct. 1989 (cited on page 6).
- [133] A. K. Kwan, M. Younes, R. Darraji, and F. M. Ghannouchi, “On track for efficiency: Concurrent multiband envelope-tracking power amplifiers”, *IEEE Microwave Magazine*, volume 17, number 5, pages 46–59, May 2016 (cited on page 69).
- [134] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, “Splitting the unit delay”, *IEEE Signal Processing Magazine*, volume 13, number 1, pages 30–60, Jan. 1996 (cited on page 103).
- [135] P. Lancaster and K. Šalkauskas, *Curve and surface fitting: An introduction*. Academic Press, 1986 (cited on page 75).
- [136] P. N. Landin, S. Gustafsson, C. Fager, and T. Eriksson, “WebLab: A web-based setup for PA digital predistortion and characterization”, *IEEE Microwave Magazine*, volume 16, number 1, pages 138–140, Feb. 2015 (cited on page 66).
- [137] P. N. Landin and D. Rönnow, “RF PA modeling considering odd-even and odd order polynomials”, in Proceedings of the IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT), Luxembourg City, Luxembourg, Nov. 2015, pages 1–6 (cited on page 32).
- [138] N. Lashkarian, J. Shi, and M. Forbes, “Direct learning adaptation of power amplifier pre-distortion based on Wirtinger calculus”, in Proceedings of the 47th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, Nov. 2013, pages 1286–1290 (cited on pages 77, 88, 91).
- [139] N. Lashkarian, J. Shi, and M. Forbes, “A direct learning adaptive scheme for power-amplifier linearization based on Wirtinger calculus”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 61, number 12, pages 3496–3505, Dec. 2014 (cited on pages 77, 88, 91, 92).
- [140] P. M. Lavrador, T. R. Cunha, P. M. Cabral, and J. C. Pedro, “The linearity-efficiency compromise”, *IEEE Microwave Magazine*, volume 11, number 5, pages 44–58, Aug. 2010 (cited on page 6).
- [141] E. G. Lima, T. R. Cunha, and J. C. Pedro, “A physically meaningful neural network behavioral model for wireless transmitters exhibiting PM–AM/PM–PM distortions”, *IEEE Transactions on Microwave Theory and Techniques*, volume 59, number 12, pages 3512–3521, Dec. 2011 (cited on pages 32, 64).
- [142] E. G. Lima, T. R. Cunha, H. M. Teixeira, M. Pirola, and J. C. Pedro, “Base-band derived volterra series for power amplifier modeling”, in Proceedings of the IEEE International Microwave Symposium (IMS), Boston, Massachusetts, Jun. 2009, pages 1361–1364 (cited on page 32).

- [143] A. Lollo, G. Bollati, and R. Castello, “A Class-G headphone amplifier in 65 nm CMOS technology”, *IEEE Journal of Solid-State Circuits*, volume 45, number 12, pages 2530–2542, Dec. 2010 (cited on page 27).
- [144] L. F. Long, Ed., *Digital Front-End in Wireless Communications and Broadcasting*. Cambridge University Press, 2011 (cited on page 29).
- [145] 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”, *IEEE Microwave Magazine*, volume 17, number 1, pages 79–87, Jan. 2016 (cited on pages 67, 71, 79).
- [146] H. D. Lüke, “The origins of the sampling theorem”, *IEEE Communications Magazine*, volume 37, number 4, pages 106–108, Apr. 1999 (cited on page 95).
- [147] M. J. Madero-Ayora, M. Allegue-Martínez, C. Crespo-Cadenas, and J. Reina-Tosina, “On the significance of even-order nonlinear terms in quadrature modulators and power amplifiers”, in Proceedings of the 40th European Microwave Conference (EuMC), Paris, France, Sep. 2010, pages 485–488 (cited on page 32).
- [148] S. Maier, D. Wiegner, M. Zierdt, W. Kuebart, U. Seyfried, C. Haslach, A. Pascht, S. Maroldt, and R. Quay, “900 MHz pulse-width-modulated class-S power amplifier with improved linearity”, in Proceedings of the IEEE International Microwave Symposium (IMS), Baltimore, Maryland, Jun. 2011, pages 1–4 (cited on page 27).
- [149] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters*. Wiley, 2009 (cited on pages 88–90).
- [150] K. W. Martin, “Complex signal processing is not complex”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 51, number 9, pages 1823–1836, Sep. 2004 (cited on page 85).
- [151] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. Wiley, 2000 (cited on pages 31, 50).
- [152] E. McCune, “A technical foundation for RF CMOS power amplifiers: Part 1 to 6”, *IEEE Solid-State Circuits Magazine*, 2015-2016 (cited on pages 6, 23, 26, 28, 78).
- [153] E. McCune, *Practical Digital Wireless Signals*. Cambridge University Press, 2010 (cited on page 1).
- [154] E. McCune, “Envelope tracking or polar – Which is it?”, *IEEE Microwave Magazine*, volume 13, number 4, pages 34–56, May 2012 (cited on page 28).
- [155] E. McCune, “Operating modes of dynamic power supply transmitter amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 62, number 11, pages 2511–2517, Nov. 2014 (cited on page 28).
- [156] E. McCune, *Dynamic Power Supply Transmitters: Envelope Tracking, Direct Polar, and Hybrid Combinations*. Cambridge University Press, 2015 (cited on page 28).

- [157] E. McCune, “Fundamentals of switching RF power amplifiers”, *IEEE Microwave and Wireless Components Letters*, volume 25, number 12, pages 838–840, Dec. 2015 (cited on page 26).
- [158] P. K. Meher, J. Valls, T.-B. Juang, K. Sridharan, and K. Maharatna, “50 years of CORDIC: Algorithms, architectures, and applications”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 56, number 9, pages 1893–1907, Sep. 2009 (cited on page 84).
- [159] S. Mendel, C. Vogel, and N. D. Dalt, “A phase-domain all-digital phase-locked loop architecture without reference clock retiming”, *IEEE Transactions on Circuits and Systems II: Express Briefs*, volume 56, number 11, pages 860–864, Nov. 2009 (cited on page 3).
- [160] N. Messaoudi, M.-C. Fares, S. Boumaiza, and J. Wood, “Complexity reduced odd-order memory polynomial pre-distorter for 400-watt multi-carrier Doherty amplifier linearization”, in Proceedings of the IEEE International Microwave Symposium (IMS), Atlanta, Georgia, Jun. 2008, pages 419–422 (cited on page 31).
- [161] J. Moon, J. Kim, and B. Kim, “Investigation of a Class-J power amplifier with a nonlinear C_{out} for optimized operation”, *IEEE Transactions on Microwave Theory and Techniques*, volume 58, number 11, pages 2800–2811, Nov. 2010 (cited on page 27).
- [162] J. Moon, J. Lee, J. Son, J. Kim, S. Jee, S. Kim, and B. Kim, “Effects of even-order terms on behavior model of envelope tracking transmitters”, in Proceedings of the 41st European Microwave Conference (EuMC), Manchester, United Kingdom, Oct. 2011, pages 1193–1196 (cited on page 32).
- [163] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, “A generalized memory polynomial model for digital predistortion of RF power amplifiers”, *IEEE Transactions on Signal Processing*, volume 54, number 10, pages 3852–3860, Oct. 2006 (cited on pages 31, 61, 64, 73, 74).
- [164] B. Murmann, C. Vogel, and H. Koeppl, “Digitally enhanced analog circuits: System aspects”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Seattle, Washington, May 2008, pages 560–563 (cited on page 1).
- [165] T. Mury and V. F. Fusco, “Inverse Class-E amplifier with transmission-line harmonic suppression”, *IEEE Transactions on Circuits and Systems I: Regular Papers*, volume 54, number 7, pages 1555–1561, Jul. 2007 (cited on page 26).
- [166] N. Narahariseti, P. Roblin, C. Quindroit, and S. Gheitanchi, “Efficient least-squares 2-D-cubic spline for concurrent dual-band systems”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 7, pages 2199–2210, Jul. 2015 (cited on page 75).

- [167] H. M. Nemati, H. Cao, B. Almgren, T. Eriksson, and C. Fager, “Design of highly efficient load modulation transmitter for wideband cellular applications”, *IEEE Transactions on Microwave Theory and Techniques*, volume 58, number 11, pages 2820–2828, Nov. 2010 (cited on page 28).
- [168] A. Nuttall, “Some windows with very good sidelobe behavior”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, volume 29, number 1, pages 84–91, Feb. 1981 (cited on page 17).
- [169] H. Ochiai, “An analysis of band-limited communication systems from amplifier efficiency and distortion perspective”, *IEEE Transactions on Communications*, volume 61, number 4, pages 1460–1472, Apr. 2013 (cited on page 23).
- [170] A. Papoulis, *The Fourier Integral and Its Applications*. McGraw-Hill, 1962 (cited on pages 33, 96, 97).
- [171] J.-C. Park, J.-G. Yook, B. H. Park, N. Jeon, K.-S. Seo, D. Kim, W.-S. Lee, and C.-S. Yoo, “Hybrid current-mode Class-S power amplifier with GaN schottky diode using chip-on-board technique for 955 MHz LTE signal”, *IEEE Transactions on Microwave Theory and Techniques*, volume 61, number 12, pages 4168–4178, Dec. 2013 (cited on page 27).
- [172] S. Pavan, R. Schreier, and G. C. Temes, *Understanding Delta-Sigma Data Converters*, 2nd edition. Wiley - IEEE Press, 2017 (cited on page 28).
- [173] J. C. Pedro and N. B. Carvalho, *Intermodulation Distortion in Microwave and Wireless Circuits*. Artech House, 2003 (cited on page 34).
- [174] R. Pengelly, C. Fager, and M. Ozen, “Doherty’s legacy: A history of the Doherty power amplifier from 1936 to the present day”, *IEEE Microwave Magazine*, volume 17, number 2, pages 41–58, Feb. 2016 (cited on page 28).
- [175] B. Picinbono and P. Chevalier, “Widely linear estimation with complex data”, *IEEE Transactions on Signal Processing*, volume 43, number 8, pages 2030–2033, Aug. 1995 (cited on page 64).
- [176] A. Prata, R. F. Cordeiro, D. C. Dinis, A. S. R. Oliveira, J. Vieira, and N. B. Carvalho, “All-digital transceivers – Recent advances and trends”, in Proceedings of the IEEE International Conference on Electronics, Circuits and Systems (ICECS), Monte Carlo, Monaco, Dec. 2016, pages 233–236 (cited on pages 3, 28).
- [177] D. Psaltis, A. Sideris, and A. A. Yamamura, “A multilayered neural network controller”, *IEEE Control Systems Magazine*, volume 8, number 2, pages 17–21, Apr. 1988 (cited on page 76).
- [178] F. H. Raab, “Radio frequency pulsewidth modulation”, *IEEE Transactions on Communications*, volume 21, number 8, pages 958–966, Aug. 1973 (cited on pages 28, 65).
- [179] F. H. Raab, “Idealized operation of the class E tuned power amplifier”, *IEEE Transactions on Circuits and Systems*, volume 24, number 12, pages 725–735, Dec. 1977 (cited on page 26).

- [180] F. H. Raab, “Efficiency of outphasing RF power-amplifier systems”, *IEEE Transactions on Communications*, volume 33, number 10, pages 1094–1099, Oct. 1985 (cited on page 28).
- [181] F. H. Raab, “Average efficiency of Class-G power amplifiers”, *IEEE Transactions on Consumer Electronics*, volume CE-32, number 2, pages 145–150, May 1986 (cited on page 27).
- [182] F. H. Raab, “Efficiency of doherty RF power-amplifier systems”, *IEEE Transactions on Broadcasting*, volume BC-33, number 3, pages 77–83, Sep. 1987 (cited on page 28).
- [183] F. H. Raab, “Class-E, Class-C, and Class-F power amplifiers based upon a finite number of harmonics”, *IEEE Transactions on Microwave Theory and Techniques*, volume 49, number 8, pages 1462–1468, Aug. 2001 (cited on page 27).
- [184] F. H. Raab, “High-efficiency linear amplification by dynamic load modulation”, in Proceedings of the IEEE International Microwave Symposium (IMS), Philadelphia, Pennsylvania, Jun. 2003, pages 1717–1720 (cited on page 28).
- [185] F. H. Raab, P. Asbeck, S. C. Cripps, P. B. Kenington, Z. B. Popovic, N. Pothechary, J. F. Sevic, and N. O. Sokal, “RF and microwave power amplifier and transmitter technologies – Part 1 to 5”, *High Frequency Electronics*, 2003-2004 (cited on pages 26, 27, 29).
- [186] F. H. Raab, P. Asbeck, S. C. Cripps, P. B. Kenington, Z. B. Popovic, N. Pothechary, J. F. Sevic, and N. O. Sokal, “Power amplifiers and transmitters for RF and microwave”, *IEEE Transactions on Microwave Theory and Techniques*, volume 50, number 3, pages 814–826, Mar. 2002 (cited on pages 26, 27).
- [187] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. Yu, and S. Cruces, “Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 2, pages 745–753, Feb. 2015 (cited on page 64).
- [188] C. G. Rey, “Adaptive polar work-function predistortion”, *IEEE Transactions on Microwave Theory and Techniques*, volume 47, number 6, pages 722–726, Jun. 1999 (cited on page 31).
- [189] P. Roblin, C. Quindroit, N. Naraharisetti, S. Gheitanchi, and M. Fitton, “Concurrent linearization: The state of the art for modeling and linearization of multiband power amplifiers”, *IEEE Microwave Magazine*, volume 14, number 7, pages 75–91, Nov. 2013 (cited on page 68).
- [190] J. L. Rodgers and W. A. Nicewander, “Thirteen ways to look at the correlation coefficient”, *The American Statistician*, volume 42, number 1, pages 59–66, Feb. 1988 (cited on page 97).
- [191] W. J. Rugh, *Nonlinear system theory*. Johns Hopkins Univ. Press, 1981, republished as free e-book in 2002, available online at: <http://sites.google.com/site/wilsonjrugh> (cited on pages 31, 50).

- [192] N. Safari, P. Fedorenko, J. S. Kenney, and T. Roste, “Spline-based model for digital predistortion of wide-band signals for high power amplifier linearization”, in Proceedings of the IEEE International Microwave Symposium (IMS), Honolulu, Hawaii, Jun. 2007, pages 1441–1444 (cited on page 75).
- [193] A. A. M. Saleh, “Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers”, *IEEE Transactions on Communications*, volume 29, number 11, pages 1715–1720, Nov. 1981 (cited on page 81).
- [194] C. Sánchez-Pérez, M. Özen, C. M. Andersson, D. Kuylenstierna, N. Rorsman, and C. Fager, “Optimized design of a dual-band power amplifier with SiC varactor-based dynamic load modulation”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 8, pages 2579–2588, Aug. 2015 (cited on page 28).
- [195] M. Schetzen, “Nonlinear system modeling based on the Wiener theory”, *Proceedings of the IEEE*, volume 69, number 12, pages 1557–1573, Dec. 1981 (cited on page 31).
- [196] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. Krieger, 2006, reprint of the first edition from 1980 (cited on pages 31, 50).
- [197] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 2010 (cited on pages 88, 90).
- [198] C. Siegl and R. F. Fischer, “Out-of-band power reduction using selected mapping with power-amplifier-oriented metrics”, in Proceedings of the 14th International OFDM Workshop, Hamburg, Germany, Sep. 2009, pages 1–5 (cited on page 20).
- [199] P. Singerl, A. Agrawal, A. Garg, Neelabh, G. Kubin, and H. Eul, “Complex baseband predistorters for nonlinear wideband RF power amplifiers”, in Proceedings of the 49th IEEE International Midwest Symposium on Circuits and Systems, San Juan, Puerto Rico, Aug. 2006, pages 675–678 (cited on page 5).
- [200] P. Singerl and H. Koepl, “A low-rate identification method for digital predistorters based on Volterra kernel interpolation”, in Proceedings of the 48th IEEE International Midwest Symposium on Circuits and Systems, Covington, Kentucky, Aug. 2005, pages 1533–1536 (cited on page 5).
- [201] P. Singerl and H. Koepl, “Volterra kernel interpolation for system modeling and predistortion purposes”, in Proceedings of the International Symposium on Signals, Circuits, and Systems (ISSCS), Iași, Romania, Jul. 2005, pages 251–254 (cited on page 5).
- [202] P. Singerl and G. Kubin, “Chebyshev approximation of baseband Volterra series for wideband RF power amplifiers”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Kobe, Japan, May 2005, pages 2655–2658 (cited on page 5).
- [203] P. Singerl and G. Kubin, “Constructing memory-polynomial models from frequency-dependent AM/AM and AM/PM measurements”, in Proceedings of the 50th IEEE International Midwest Symposium on Circuits and Systems, Montreal, Canada, Aug. 2007, pages 321–324 (cited on pages 5, 81).

- [204] P. Singerl, “Complex baseband modeling and digital predistortion for wideband RF power amplifiers”, Doctoral thesis, Graz University of Technology, Dec. 2006 (cited on page 5).
- [205] N. O. Sokal and A. D. Sokal, “Class E – A new class of high-efficiency tuned single-ended switching power amplifiers”, *IEEE Journal of Solid-State Circuits*, volume 10, number 3, pages 168–176, Jun. 1975 (cited on page 26).
- [206] N. O. Sokal, “Class-E RF power amplifiers”, *QEX*, volume 204, pages 9–20, Jan. 2001 (cited on page 26).
- [207] J. Sombrin, “Non-analytic at the origin, behavioral models for active or passive non-linearity”, *International Journal of Microwave and Wireless Technologies*, volume 5, pages 133–140, Apr. 2013 (cited on pages 32, 40).
- [208] J. Sombrin, G. Soubercaze-Pun, and I. Albert, “Discontinuity at origin in Volterra and band-pass limited models”, in Proceedings of the IEEE International Microwave Symposium (IMS), Seattle, Washington, Jun. 2013, pages 1–4 (cited on page 32).
- [209] L. Sorber, M. V. Barel, and L. D. Lathauwer, “Unconstrained optimization of real functions in complex variables”, *SIAM Journal on Optimization*, volume 22, number 3, pages 879–898, 2012 (cited on pages 88, 90, 91).
- [210] R. P. Stanley, *Enumerative Combinatorics: Volume 1*, 2nd edition. Cambridge University Press, 2012 (cited on page 105).
- [211] A. S. Tehrani, “Behavioral modeling of wireless transmitters for distortion mitigation”, Doctoral thesis, Chalmers University of Technology, 2012 (cited on page 5).
- [212] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, “A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models”, *IEEE Transactions on Microwave Theory and Techniques*, volume 58, number 6, pages 1510–1520, Jun. 2010 (cited on pages 5, 64, 81).
- [213] A. S. Tehrani, T. Eriksson, and C. Fager, “Modeling of long term memory effects in RF power amplifiers with dynamic parameters”, in Proceedings of the IEEE International Microwave Symposium (IMS), Montreal, Canada, Jun. 2012, pages 1–3 (cited on page 5).
- [214] S. Trampitsch, “Complex-valued data estimation”, Master’s thesis, University of Klagenfurt, Apr. 2013 (cited on pages 88, 90).
- [215] L. N. Trefethen, *Approximation Theory and Approximation Practice*. Society for Industrial and Applied Mathematics (SIAM), 2013 (cited on pages 30, 33, 35).
- [216] C. Vogel and H. Johansson, “Time-interleaved analog-to-digital converters: Status and future directions”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Kos, Greece, May 2006, pages 3386–3389 (cited on page 3).

- [217] J. E. Volder, “The CORDIC trigonometric computing technique”, *IRE Transactions on Electronic Computers*, volume 8, number 3, pages 330–334, Sep. 1959 (cited on page 84).
- [218] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*. Dover, 2005, reprint of the first edition from 1930 (cited on page 31).
- [219] J. S. Walling and D. J. Allstot, “Pulse-width modulated CMOS power amplifiers”, *IEEE Microwave Magazine*, volume 12, number 1, pages 52–60, Feb. 2011 (cited on page 28).
- [220] J. S. Walling, S. S. Taylor, and D. J. Allstot, “A Class-G supply modulator and Class-E PA in 130 nm CMOS”, *IEEE Journal of Solid-State Circuits*, volume 44, number 9, pages 2339–2347, Sep. 2009 (cited on page 27).
- [221] M. P. Wand and M. C. Jones, *Kernel Smoothing*. CRC Press, 1994 (cited on page 19).
- [222] Z. Wang, *Envelope Tracking Power Amplifiers for Wireless Communications*. Artech House, 2014 (cited on page 28).
- [223] Z. Wang, “Demystifying envelope tracking: Use for high-efficiency power amplifiers for 4G and beyond”, *IEEE Microwave Magazine*, volume 16, number 3, pages 106–129, Apr. 2015 (cited on page 28).
- [224] Z. Wang, *High-Efficiency Load Modulation Power Amplifiers for Wireless Communications*. Artech House, 2017 (cited on page 28).
- [225] Z. Wang, L. Guan, and R. Farrell, “Compact undersampled digital predistortion for flexible single-chain multi-band RF transmitter”, in Proceedings of the IEEE International Microwave Symposium (IMS), Honolulu, Hawaii, Jun. 2017, pages 1542–1545 (cited on page 95).
- [226] Z. Wang, L. Guan, and R. Farrell, “Undersampling observation-based compact digital predistortion for single-chain multiband and wideband direct-to-RF transmitter”, *IEEE Transactions on Microwave Theory and Techniques*, volume 65, number 12, pages 5274–5283, Dec. 2017 (cited on page 95).
- [227] Z. Wei, Y. Luo, J. Li, and X. Zheng, “A note on Wallis’ formula”, *Journal of Advances in Applied Mathematics*, volume 1, number 2, pages 120–123, Apr. 2016 (cited on page 111).
- [228] S. B. Weinstein, “The history of orthogonal frequency-division multiplexing”, *IEEE Communications Magazine*, volume 47, number 11, pages 26–35, Nov. 2009 (cited on page 17).
- [229] E. W. Weisstein, “Polynomial order”, in *MathWorld – A Wolfram Web Resource*, <http://mathworld.wolfram.com/PolynomialOrder.html>, 2017 (cited on page 32).
- [230] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms”, *IEEE Transactions on Audio and Electroacoustics*, volume 15, number 2, pages 70–73, Jun. 1967 (cited on page 17).

- [231] A. Wentzel, F. Hühn, and W. Heinrich, “The digital power amplifier for the wireless infrastructure: Status and prospects”, in Proceedings of the IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR), Phoenix, Arizona, Jan. 2017, pages 14–17 (cited on page 27).
- [232] E. Westesson and L. Sundstrom, “A complex polynomial predistorter chip in CMOS for baseband or IF linearization of RF power amplifiers”, in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Orlando, Florida, Jul. 1999, pages 206–209 (cited on page 31).
- [233] N. Wiener, *Nonlinear Problems in Random Theory*. Martino Fine Books, 2013, reprint of the first edition from 1958 (cited on page 31).
- [234] W. Wirtinger, “Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen”, German, *Mathematische Annalen*, volume 97, number 1, pages 357–376, 1927 (cited on pages 88, 89).
- [235] D. Wisell, J. Jalden, and P. Händel, “Behavioral power amplifier modeling using the LASSO”, in Proceedings of the IEEE Instrumentation and Measurement Technology Conference, Victoria, Canada, May 2008, pages 1864–1867 (cited on page 64).
- [236] K. Witrisal, “OFDM air-interface design for multimedia communications”, Doctoral thesis, Delft University of Technology, Apr. 2002 (cited on page 17).
- [237] Y. Y. Woo, Y. Yang, and B. Kim, “Analysis and experiments for high-efficiency class-F and inverse class-F power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 54, number 5, pages 1969–1974, May 2006 (cited on page 27).
- [238] J. Wood, *Behavioral Modeling and Linearization of RF Power Amplifiers*. Artech House, 2014 (cited on pages 3, 29).
- [239] J. Wood, “System-level design considerations for digital pre-distortion of wireless base station transmitters”, *IEEE Transactions on Microwave Theory and Techniques*, volume 65, number 5, pages 1880–1890, May 2017 (cited on page 3).
- [240] P. Wright, J. Lees, J. Benedikt, P. J. Tasker, and S. C. Cripps, “A methodology for realizing high efficiency Class-J in a linear and broadband PA”, *IEEE Transactions on Microwave Theory and Techniques*, volume 57, number 12, pages 3196–3204, Dec. 2009 (cited on page 27).
- [241] S.-M. Yoo, J. S. Walling, O. Degani, B. Jann, R. Sadhwani, J. C. Rudell, and D. J. Allstot, “A Class-G switched-capacitor RF power amplifier”, *IEEE Journal of Solid-State Circuits*, volume 48, number 5, pages 1212–1224, May 2013 (cited on page 27).
- [242] G. T. Zhou, “Analysis of spectral regrowth of weakly nonlinear power amplifiers”, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Istanbul, Turkey, Jun. 2000, pages 2737–2740 (cited on page 19).

- [243] G. T. Zhou, “Analysis of spectral regrowth of weakly nonlinear power amplifiers”, *IEEE Communications Letters*, volume 4, number 11, pages 357–359, Nov. 2000 (cited on page 19).
- [244] G. T. Zhou and J. S. Kenney, “Predicting spectral regrowth of nonlinear power amplifiers”, *IEEE Transactions on Communications*, volume 50, number 5, pages 718–722, May 2002 (cited on page 19).
- [245] G. T. Zhou, H. Qian, L. Ding, and R. Raich, “On the baseband representation of a bandpass nonlinearity”, *IEEE Transactions on Signal Processing*, volume 53, number 8, pages 2953–2957, Aug. 2005 (cited on pages 30, 31).
- [246] G. T. Zhou and R. Raich, “Spectral analysis of polynomial nonlinearity with applications to RF power amplifiers”, *EURASIP Journal on Advances in Signal Processing*, number 12, pages 1687–6180, Dec. 2004 (cited on page 19).
- [247] A. Zhu, “Decomposed vector rotation-based behavioral modeling for digital pre-distortion of RF power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 63, number 2, pages 737–744, Feb. 2015 (cited on pages 64, 75).
- [248] A. Zhu, J. C. Pedro, and T. J. Brazil, “Dynamic deviation reduction-based Volterra behavioral modeling of RF power amplifiers”, *IEEE Transactions on Microwave Theory and Techniques*, volume 54, number 12, pages 4323–4332, Dec. 2006 (cited on page 64).
- [249] Y.-M. Zhu, “Generalized sampling theorem”, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, volume 39, number 8, pages 587–588, Aug. 1992 (cited on page 95).