

MINISTRY OF EDUCATION & TRAINING MINISTRY OF NATIONAL DEFENSE
MILITARY TECHNICAL ACADEMY

NGUYEN THANH

**NONLINEAR DISTORTIONS AND
COUNTERMEASURES FOR PERFORMANCE
IMPROVEMENTS IN CONTEMPORARY
RADIO COMMUNICATION SYSTEMS**

A thesis for the degree of Doctor of Philosophy

HA NOI - 2019

MINISTRY OF EDUCATION & TRAINING MINISTRY OF NATIONAL DEFENSE
MILITARY TECHNICAL ACADEMY

NGUYEN THANH

**NONLINEAR DISTORTIONS AND
COUNTERMEASURES FOR PERFORMANCE
IMPROVEMENTS IN CONTEMPORARY
RADIO COMMUNICATION SYSTEMS**

A thesis for the degree of Doctor of Philosophy

Specialization : Electronic Engineering
Specialization code : 9 52 02 03

Supervisor:
Assoc. Prof. NGUYEN QUOC BINH

HA NOI - 2019

THESIS DECLARATION

I hereby declare that all data and results shown in this thesis are my own original work created under the guidance from my supervisor. These data and results are honestly presented and are not yet published in any previous works. I also declare that, as required by academic rules and ethical conduct, I have fully cited and referenced all materials and results that are not original to this work.

Ha Noi, November 2019

Nguyen Thanh

ACKNOWLEDGMENTS

At the very first words, it takes a lot of good karma to have Assoc. Prof. Nguyen Quoc Binh as a mentor. His insightful thinking, thoughtful enthusiasm and unbounded kindness have always helped change his students' lives for the better, and I am no exception to this rule. I will always be indebted to him for igniting my passion for the profession when I was an undergraduate and then for guiding me through the most memorable years of my life doing this thesis.

My heartfelt thanks also go to respected senior colleague from Department of Communications, Faculty of Radio-Electronic Engineering, Le Quy Don Technical University, and also to other lecturers, professors and authorities for their valuable ideas, comments and reviews that actually make this work much better.

I would like to thank the staff from Office of Postgraduate Academic Affairs, Le Quy Don Technical University for their devoted help in making administrative procedures extremely convenient.

I am grateful to all my friends here at Le Quy Don Technical University and elsewhere. Each one of them, in his or her own unique way, has left on me a lasting impression that can not be described in words.

Finally, I really would like to thank my dear parents and my small family for sharing the simple yet great joy of life in every moment.

Table of Contents

Table of Contents	
List of Acronyms	v
List of Figures	ix
List of Tables	xii
List of Mathematical Notations	xiii
Foreword	1
Chapter 1. Introduction to Nonlinear Distortions and Practical MIMO-STBC Systems	14
1.1. Main causes of nonlinear distortions in radio communication systems	14
1.2. Nonlinear HPA model classification	16
1.3. Nonlinear HPA distortion impacts in SISO systems	24
1.4. Multiple-input multiple-output systems	27
1.5. MIMO in satellite communication systems	35
1.6. Nonlinear HPA distortion impacts in MIMO systems	39
1.7. Summary of chapter 1	42

Chapter 2. Nonlinear HPA Modeling and Proposed Polysine Model.....	43
2.1. Introduction.....	43
2.2. Instantaneous nonlinear models.....	45
2.2.1. Cann original model.....	45
2.2.2. Cann new model.....	47
2.3. Envelope nonlinear models.....	50
2.3.1. Envelope representation of bandpass signals.....	50
2.3.2. Saleh model.....	52
2.3.3. Rapp model.....	54
2.3.4. Cann envelope model.....	56
2.3.5. Polynomial model.....	57
2.3.6. Proposed polysine model.....	59
2.3.7. Other conventional HPA models.....	61
2.4. Applications of HPA models in communication simulation.....	63
2.4.1. Representation of envelope models.....	63
2.4.2. Simulation with two-tone testing signal.....	65
2.4.3. Simulation with continuous-spectrum testing signal.....	67
2.5. Summary of chapter 2.....	69
Chapter 3. Predistortion Methods for Nonlinear Distortions due to HPAs in MIMO-STBC Systems.....	71
3.1. Overview.....	72

3.2. Nonlinear distortion effects in MIMO-STBC systems	74
3.2.1. MIMO-STBC $2 \times n_R$ system model	74
3.2.2. Nonlinear distortion effects incurred by HPAs	77
3.3. Predistortion schemes	82
3.3.1. Ideal inverse Saleh predistortion	84
3.3.2. Adaptive secant predistortion	85
3.3.3. Adaptive Newton predistortion	87
3.3.4. Adaptive LMS polynomial-approximated predistortion	89
3.4. Performance evaluation for predistorted MIMO-STBC systems ...	90
3.4.1. System parameters and performance measures	90
3.4.2. Receive signal constellations with predistortion	91
3.4.3. Error vector module	93
3.4.4. Modulation error ratio	95
3.4.5. Bit error ratio	97
3.5. Summary of chapter 3	97
Chapter 4. Automatic Phase Estimation and Compensation for	
Nonlinear Distortions due to HPAs in MIMO-STBC Systems	99
4.1. Overview	99
4.2. Phase rotation impact due to nonlinear HPAs for the MIMO-	
STBC signals	101
4.2.1. Nonlinear MIMO-STBC system model with phase estimation	
and compensation at the receiver	101
4.2.2. Phase rotation impact due to nonlinear HPAs	103

4.3. Phase estimation problem	107
4.3.1. Gaussian approximation for the nonlinear model	107
4.3.2. Optimal blind feedforward phase estimation	108
4.3.3. Harmonic approximation	111
4.3.4. Biharmonic approximation	112
4.4. Performance evaluation of the phase estimation and phase compensation scheme	113
4.4.1. Performance of the phase estimator	114
4.4.2. Optimum proximity of the estimated phases	115
4.4.3. Total degradation	116
4.4.4. Bit error ratio	118
4.5. Summary of chapter 4	119
Final Conclusions	121
List of Publications	125
Bibliography	127

List of Acronyms

2/3D	2-/3-Dimensional
2/3/4/5G	Second/Third/Fourth/Fifth Generation
3GPP	Third Generation Partnership Project
AC	Alternative Current
ADC	Analog-to-Digital Converter
AM-AM	Amplitude Modulation-to-Amplitude Modulation
AM-PM	Amplitude Modulation-to-Phase Modulation
APSK	Amplitude and Phase-Shift Keying
ASK	Amplitude-Shift Keying
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLAST	Bell-Labs Layered Space-Time (Architecture)
BO	Back-Off
BS	Base Station
CCI	Co-Channel Interference
DAC	Digital-to-Analog Converter
dB	Decibel
dBr	dB relative to reference level
DC	Direct Current
<i>dd</i>	<i>distance degradation</i>
DVB	Digital Video Broadcasting

DVB-S2	DVB - Satellite - Second Generation
DVB-S2X	DVB-S2 Extension
DVB-SH	DVB - Satellite services to Handhelds
DVB-T	DVB - Terrestrial
EPC	Electronic Power Conditioner
ETSI	European Telecommunications Standards Institute
EVM	Error Vector Module/Magnitude
FS	Fixed Satellite
FST	Fixed Satellite Terminal
FSK	Frequency-Shift Keying
GSO	GeoStationary Orbit
HPA	High Power Amplifier
IBO	Input Back-Off
IEEE	Institute of Electrical and Electronics Engineers
IMD	Inter-Modulation Distortion
IMP	Inter-Modulation Product
IMP3/5	Third-/Fifth-order IMP
ISI	Inter-Symbol Interference
LDMOS	Laterally-Diffused Metal Oxide Semiconductor
LHCP	Left-Hand Circular Polarization
LMS	Least Mean Square
LMSat	Land Mobile Satellite
LTE	Long Term Evolution (3.9G)
LTE-A	LTE-Advanced (4G)

LOS	Line-Of-Sight
MER	Modulation Error Ratio
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MLD	Maximum-Likelihood Detection
MMSE	Minimum Mean Square Error
MRC	Maximum-Ratio Combining
MS	Mobile Satellite
MSB	Mobile Satellite Broadcasting
MST	Mobile Satellite Terminal
MU	Multi-User
NGSO	Non-GeoStationary Orbit
NLOS	Non LOS
OAPS	Optimum Additional Phase Shifting
OBO	Output Back-Off
OrbD	Orbital Diversity
OFDM	Orthogonal Frequency-Division Multiplexing
OSTBC	Orthogonal Space-Time Block Coding
PD	PreDistortion
PSK	Phase-Shift Keying
PTC	Polarization-Time Coding
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying

RF	Radio Frequency
RHCP	Right-Hand Circular Polarization
SatCom	Satellite Communications
SatD	Satellite Diversity
SD	Spatial Diversity
SEL	Soft Envelope Limiter
SER	Symbol Error Ratio
SF	Space-Frequency
SIMO	Single-Input Multiple-Output
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input Single-Output
SM	Spatial Multiplexing
SNR	Signal-to-Noise Ratio
SRRC	Square-Root Raised Cosine
SSPA	Solid-State Power Amplifier
ST	Space-Time
STBC	Space-Time Block Coding
STF	Space-Time-Frequency
STTC	Space-Time Trellis Coding
TD	Total Degradation
TR-STBC	Time-Reversal STBC
TWT	Travelling-Wave Tube
TWTA	TWT Amplifier
V-BLAST	Vertical-BLAST

List of Figures

1.1	Simplified block diagram of a typical radio transmitter.	15
1.2	The IEEE 802.11a spectrum mask for the 20 MHz bandwidth signal [5].	16
1.3	HPA modeling classification.	18
1.4	Typical amplitude and phase distortion characteristics of an HPA ^(*) .	23
1.5	Spectrum regrowth due to nonlinear HPA ^(*)	25
1.6	Constellation warping due to nonlinear HPA.	26
1.7	Nonlinear ISI due to nonlinear HPA.	26
1.8	Simplified MIMO system diagram.	28
1.9	MIMO technique classification ^(*) [68].	29
1.10	Dual-polarized MIMO land mobile satellite system model.	38
1.11	Simplified MIMO system with nonlinear HPA.	39
2.1	Characteristic functions of the Cann new model.	47
2.2	Characteristic functions of the Rapp/Cann original model (2.1) compared to that of the Cann new model (2.2).	48
2.3	Third order (a) and fifth order (b) IMPs created by the Cann new model (2.2).	49
2.4	AM-AM functions of the Cann envelope model corresponding to the instantaneous model (2.2).	52
2.5	AM-AM (a) and AM-PM (b) functions of typical envelope models.	53

2.6	AM-AM functions of the Rapp model with different sharpnesses.	55
2.7	AM-AM functions of the Cann, Rapp, polynomial, odd-order polynomial and polysine models fitted to the measured data. . .	57
2.8	Two-tone waveform, $f_1 = 7$ [Hz], $f_2 = 10$ [Hz].	63
2.9	Polar envelope model block diagram [52].	64
2.10	Third order (a) and fifth order (b) IMPs of five models in Figure 2.7.	65
2.11	Amplitude histograms of two-tone (a) and 1+7-APSK (b) testing signals.	66
2.12	Receive constellations (a) and spectra (b) created from 1+7-APSK testing signal with different nonlinear models.	68
3.1	MIMO-STBC system model with transmit/receive filters and nonlinear HPAs.	74
3.2	Receive signals after MRC for the system models with (a) and without (b) transmit/receive filters.	79
3.3	Receive signals after MRC with HPA model used in [81] for the systems models with (a) and without (b) transmit/receive filters.	80
3.4	MIMO-STBC system model with predistorters.	83
3.5	Power amplifier linearization using baseband digital predistorter.	83
3.6	Baseband digital predistorter diagram.	83
3.7	Receive signal constellations with predistortion: a) LUT; b) Secant; c) Newton; d) Polynomial.	92

3.8	EVM versus IBO of the MIMO-STBC system with different predistorters.	94
3.9	MER versus IBO of the MIMO-STBC system with different predistorters.	96
3.10	BER versus E_b/N_0 of the MIMO-STBC system with $IBO = 6$ dB and different predistorters.	97
4.1	Proposed MIMO-STBC system model with phase estimation and compensation.	101
4.2	AM-AM (a) and AM-PM (b) characteristics of considered HPA models.	102
4.3	Receive signal constellations after matched filtering: a) Fully characterized (4.3); b) Approximated (4.5).	105
4.4	BER versus compensated phase angle: a) Saleh and modified Ghorbani models; b) Modified Saleh and modified Rapp models.	116
4.5	TD versus IBO of systems with and without phase compensation at $BER = 10^{-3}$	117
4.6	BER versus E_b/N_0 of systems with and without phase compensation.	118

List of Tables

1.0	Commercialized wireless standards using MIMO.	33
2.1	Coefficients of the polynomial models (2.12), (2.13).	58
2.2	Coefficients of the polysine model (2.14).	60
2.3	Approximation performance of five models (SES σ_e^2).	61
4.1	Estimated phase values and their variances for different non-linear models.	114

List of Mathematical Notations

Notation	Meaning
a	a is a variable
\mathbf{a}	\mathbf{a} is a column vector
\mathbf{A}	\mathbf{A} is a matrix
a_{ij}	The element at i -th row, j -th column of matrix \mathbf{A}
a^*	Complex conjugate of a
\mathbf{A}^T	Transpose of matrix \mathbf{A}
\mathbf{A}^H	Conjugate (Hermitian) transpose of matrix \mathbf{A}
$(k)!$	Factorial of k
$P\{x\}$	Probability of x
$\text{Re}(x)$	Real part of x
$\text{sgn}(x)$	Sign of x
$\mathbb{E}[x]$	Expectation of x
$\mathcal{CN}(\mathbf{0}, \mathbf{N}_0)$	The circularly symmetric complex Gaussian random vector with mean vector $\mathbf{0}$ and covariance matrix \mathbf{N}_0
$\mathcal{O}(f(x))$	Order of function $f(x)$: if there exists a positive real number $M > 0$ such that, when x is sufficiently close to x_0 , $ g(z) \leq M f(z) $, then $g(x) = \mathcal{O}(f(x))$

Foreword

1. Posing problems

The online Oxford English dictionary by Oxford University Press¹ defines *linearity* as involving or exhibiting directly proportional change in two related quantities; *nonlinearity* as involving a lack of linearity between two related qualities; and *distortion* as change in the form of an electrical signal or sound wave during processing. So, the nonlinearity concept focuses on modeling and formulating, while the distortion concept concentrates on describing the phenomenon. However, it can be seen that distortion and nonlinearity have a close relation, examining the phenomenon in different points of view, with different criteria and purposes. These are basic concepts and will be the main topics discussed throughout the thesis.

Otherwise, practical parameters of an amplifying device (vacuum tube, traveling wave tube, transistor,...) in a general amplifier and especially, a high power amplifier (HPA), such as mutual conductance (or transconductance), capacitance,... are nonlinear according to the input signal amplitude [9, 25, 46, 55]; then, a practical amplifier does have a nonlinear input-output characteristic and the ideal linearity does not exist. Therefore, a general amplifier and especially, an HPA does distort its output signal. For a base-band (or low-frequency) HPA, the existence of nonlinearity, which introduces nonlinear distortion, could significantly degrade the performance of ampli-

¹<https://en.oxforddictionaries.com>

fied signals. As an intuitive example, nonlinear distortions existing in audio-frequency HPAs cause a lot of discomfort for enjoying sound, especially high fidelity (Hi-Fi) audio, and have been studied to master for the time almost parallel with the development history of electrical amplifiers [9]. For a radio-frequency (RF) HPA, with the presence of nonlinear distortions, besides the waveform deformation of baseband modulating signal, there are several serious problems that should be overcome or solved thoroughly. These might be power efficiency, spectrum efficiency, in-band interference, out-of-band interference, spectrum regrowth, error-vector magnitude (EVM),...

Therefore, modeling and simulating nonlinear HPA transfer functions, and specifically, investigating detrimental impacts of these characteristics on modern digital communication systems are still timely topics widely studied in different aspects and extents. These are subjects of many published books, papers and seminars that, for researches at a more intensive level, often lead to the conclusion of requiring more discoveries even before questions seem to be simply answered. Originally, one of the very first nonlinear HPA models is the instantaneous nonlinearity model proposed by Cann in 1980 [17]. With the transfer function which can vary its curvature, this model is quite suitable for analytical analysis as well as simulation. However, the irrationality of results created from this model was only discovered after a long time, in 1996, when Litva analyzed inter-modulation products (IMPs) generated from the two-tone test simulation [62]. Four years later, Loyka [65] showed the reason for this problem: non-analyticity of the model. Based on the Loyka's finding, Cann recently proposed an improved model [18], allowing to completely overcome the above problem with minimal complexity involved. Further, besides

working well with instantaneous signals, this new model could conveniently be used with envelope signals. However, the model's capability of approximating its characteristic to measurement data is not so good, and inferior to the Rapp classic model [18].

On the other hand, flourishing achievements in studying multi-antenna or multiple-input multiple-output (MIMO) transmission techniques in the last two decades for terrestrial digital radio communication systems have been realized through the integration of this technology in commercial standards, such as IEEE 802.11n, 802.16e, 802.16m, 802.20, 802.22, DVB-T2, 3GPP version 7, 8 (LTE, or 3.9G), 3GPP version 10 (LTE-A, or 4G) and recently, 3GPP version 15 for the 5G networks. For maintaining competition with the terrestrial counterpart, satellite communications (SatCom) is trying to pursue and also benefit from important research achievements in each area of MIMO technologies for terrestrial communications. However, MIMO is a fairly general term, including many techniques spread across various categories (such as single-user (SU) MIMO, multi-user (MU) MIMO, or distributed/virtual MIMO). Therefore, a leading question to be answered is which specific MIMO techniques can be applied to SatCom, because SatCom itself has so many different variations, each with completely different characteristics compared to terrestrial systems.

It is then very challenging to study the applicabilities of MIMO in SatCom according to the diversity mentioned above. However, for simplification, SatCom could be divided into two broad classes based on the development motivation for commercial services [75]: a) Fixed satellite (FS) systems working in geostationary orbit (GSO) at frequency bands higher than 10 GHz (such as

Ku, Ka bands) supporting fixed satellite terminals (FST) in the line-of-sight (LOS) transmission environments; b) Mobile satellite (MS) systems working in GSO or non-geostationary orbit (NGSO) at frequency bands significantly lower than 10 GHz (such as L, S, C bands) supporting mobile satellite terminals (MST) in the general non line-of-sight (NLOS) transmission environments. These two broad classes constitute the main application areas with the very successful deployment of recently developed satellite standards by the European telecommunications standards institute (ETSI), namely DVB-S2 [29] and DVB-S2X [31] for FS systems and DVB-SH [30] for MS systems. These achievements come from thorough researches over the last decade on applicabilities of MIMO SatCom with some typical works, which can be listed as follows.

Through the creation of a channel model, which allows for generating simulation results in accordance with practical measurements for land mobile satellite (LMSat) channels in the urban or highway environments, Peter R. King et al. investigated the diversity gain of satellite-MIMO space-time coding (STC) systems [58, 59]. The Alamouti STC scheme [7] was employed with improvements on synchronization and equalization adapted for the two-satellite transmit diversity (satellite diversity - SatD) and also for the channel delay and propagation characteristics. Simulation results showed that the satellite-MST configuration in the form of a 2×1 multiple-input single-output (MISO) system results in diversity order two for both shadowing (shadow correlation dependent) and multipath environments. Meanwhile, for the 2×2 MIMO configuration, shadowing diversity gain is still of order two but multipath diversity gain is now of order four.

This research group further modeled 3-dimensional (3D) polarization MIMO satellite channels [49, 50]. Differences and arising problems of the satellite channels were analyzed against the terrestrial counterparts; moreover, a statistical-physical model for satellite-to-indoor/mobile dual-polarized channels was proposed and validated through measurements. Based on these results, the authors recommended Alamouti-typed polarization-time coding (PTC) for the deployment of a dual polarization antenna (of virtually collocated radiators) in a single satellite system (a simple satellite-MIMO realization). They also verified that, compared to the single-input single-output (SISO) case, the capacity is increased by a factor of two only, though satellite MIMO channel employs 2×2 PTC. This is in contrast to the terrestrial channel case, where the capacity is increased by factor of four. Capacity also does not increase significantly when replacing the 2×2 configuration by 2×3 one. However, it is the most feasible solution considering the channel as well as structural features of satellite communication systems.

At a higher generality, this research group continued to consider applicabilities of MIMO and user cooperation (or cooperative/user diversity) to satellite communications [37]. System aspects of two up-to-date versions of the diversity concept namely, polarization diversity and satellite diversity, were investigated with preference given to the former. Main conclusions of this work could be summarized as follows: Due to the environmental differences from terrestrial, diversity in satellite communications is somewhat less efficient, the main reason of that is the poorer scattering along the propagation path; Among the possibilities available for diversity in satellite communications, PTC seems to be the most appropriate for attaining diversity

advantages yielded by MIMO; The advantages of 3D polarization can be achieved in the ground terminal only, thus maximum order of diversity is 6 for the downlink and 4 for the uplink, this could be significant in the case of low SNR and is worthwhile to applying 3D polarization; Satellite diversity produces a problem not present in PTC, i.e. inter-symbol interference (ISI) or asynchronism, making terminal hardware more complex; User cooperation plays an important role in the future satellite communications and is suitable for terrestrial gap-filling.

For another application, [8] also proposed the use of a dual polarization per beam (DPPB) MIMO scheme replacing the conventional one of single polarization per beam (SPPB) in digital video broadcast - satellite services to handheld devices (DVB-SH) systems. MIMO schemes were deployed in the forms of Alamouti STC [7], spatial multiplexing (SM) [35], or Golden codes [10] with maximum likelihood detection (MLD). The proposed dual polarization MIMO system was compared to the single polarization SISO, dual polarization non MIMO ($2 \times$ SISO) and MIMO configurations based on the DVB-SH state-of-the-art mobile satellite standard [30]. It was shown that the shift from SPPB paradigm to DPPB configuration could double the throughput of the next generation mobile satellite broadcasting (MSB) systems with practical deployment conditions of user terminal (UT) receiver, payload and antenna.

Though deployed in UTs or satellite terminals, then HPAs are still a non-negligible nonlinear elements and have been mentioned briefly in several publications such as [8, 14, 75, 77]. However, besides studies considering general nonlinear MIMO systems like [34, 74, 81, 94], there have been not so many

thorough discussions on the impacts of nonlinear distortions on MIMO Sat-Com systems and there are even fewer suggestions for overcoming these adverse effects to improve system performance [3].

2. Research motivations

The above analyses expose that the efficient representation and modeling HPA is still a widely studied topic. On the other hand, modern complex wideband modulation schemes and memory effects further complicate the matters that are inherently elaborate. However, the need to accurately assess newly proposed signals with increasingly complicated structures in practical working conditions subjected to nonlinear HPA distortions is an objective, stringent requirement directly supporting the system design, standardization, deployment,... Thoughtful understanding the causes of errors in simulating intermodulation products for conventional models such as Saleh [84], Rapp [82], polynomial,... and then improving and overcoming these defects by constructing a suitable HPA model are therefore, really strong but challenging research motivations.

MIMO technology has outstanding advantages of supporting larger data rates and higher quality of service (QoS) requirements for next generation radio communication systems. On the other hand, nonlinear distortion caused by HPA is one of the main detrimental factors significantly degrading the system performance. Therefore, studying and overcoming adverse effects of nonlinear distortion caused by HPA is an essential and urgent topic. Using the distance degradation (dd) parameter for investigation, studies in [1, 4, 11, 13] thoroughly resolved some nonlinear HPA-related problems such as evaluating separate effects of nonlinear distortion incurred by HPA, applying optimum

additional phase shifting (OAPS) solution to reduce the impacts of nonlinear distortion or evaluating concurrent effects of nonlinear and linear distortions. Recently, [3] extended these results to MIMO-STBC systems accenting on satellite communications. However, there are several topics which were not rigorously discussed and also were not extended to new directions. Therefore, this work focuses on the following purposes, tasks, scopes and methodologies.

Research purposes

- Construct an HPA model which better approximates to measurement data while more exactly simulates the spectral regrowth for newly proposed signals with complicated structures;
- Study countermeasures to limit the adverse impacts of nonlinear distortions in MIMO-STBC systems using predistortion at the transmitter;
- Propose phase estimation and compensation schemes that overcome the adverse impacts of nonlinear distortions at the receiver.

Research tasks

- Study nonlinear HPA models considering to aspects of structure, analyticity, and applicabilities in simulating inter-modulation products and spectrum regrowth;
- Study MIMO techniques with practical applications in SatCom;
- Study impacts of nonlinear distortion due to HPA in MIMO-STBC systems;
- Study predistortion methods applying to nonlinear MIMO-STBC;

- Study phase estimation method and phase compensation scheme for non-linear MIMO-STBC using M -QAM modulation.

Research scope

- Radio communication channels, modern satellite communication systems;
- Nonlinear HPA modeling and nonlinear distortions incurred by HPA;
- MIMO-STBC systems, Alamouti coding;
- Predistortion methods;
- Phase estimation and compensation methods.

Research methodology

The research method used here is a combination of modeling, analytical analysis and Monte-Carlo simulation [52]. In particular, modeling and analytical analysis methods are used to analyze HPA models and to establish system equations, combining and detecting signals at the receiver as well as developing predistortion and phase estimation algorithms. Monte-Carlo simulation is used to quantify performance of the proposed algorithms through measures such as square error, variance,... of algorithm results and to quantify performance of the proposed system through measures such as error vector module, modulation error ratio, total degradation, bit error rate,...

On the other hand, due to the complicated nature of the research problem and the limited ability of Ph.D candidate, this thesis only focuses on the main object of nonlinear distortions caused by HPAs in Alamouti MIMO-STBC systems; while other conditions such as radio transmission channels, system hardware (other than HPAs),... are temporarily assumed to be ideal.

More realistic conditions for these subjects as well as other MIMO systems will be addressed in future expansion studies.

3. Thesis contributions

Major contributions of the thesis are of scientific and practical values that can be summarized as follows.

1. Thoroughly analyzing the modeling problem for typical HPA models and proposing a nonlinear HPA model which better approximates to measurement data while ensures analyticity for all values of input signal amplitude, thus creates IMP simulation results more accurate than the Cann new model [18] does, not only with two-tone testing signals, but more importantly with newly proposed signals [28] for modern satellite communications. In design and analysis of communication systems, the precise determination of the spectrum regrowth of IMP products, especially for complex input signals, plays an important role, ensuring that the signal is in compliance with the required spectrum mask.
2. Proposing the use of transmitter-side distortion compensations (pre-compensation/predistortion) for MIMO-STBC Alamouti schemes with nonlinear distortions caused by HPAs using system model fully equipped with transmit/receive filters. Thus, all nonlinear impacts including those incurred by memory effect from the transmit/receive filters are fully analyzed so that the proposed scheme approaches the practice significantly better than previous studies [73, 74, 81, 94], also limitations and shortcomings of these works are clearly pointed out. The predistortion schemes are thoroughly analyzed in several aspects including algorithm, complex-

ity, order of convergence, performance, practical applicability and are verified by numerical simulations with measures of error vector module, modulation error ratio, bit error rate.

3. Reasonably approximating nonlinear phase distortion by a linear model, then proposing an efficient feedforward non-parametric phase estimation method and a nonlinear phase distortion compensation scheme on the receiver side (post-compensation/post-distortion) for MIMO-STBC Alamouti schemes using M -QAM signaling. The OAPS method was initially proposed by Nguyen Quoc Binh in [11], then further developed in works of [1, 4] for SISO systems and most recently extended for MIMO-STBC systems [3]. This is a manual method performing phase compensation at the receiver based on determining the optimal phase compensation angle (for the phase-shift effect of signals caused by nonlinear HPAs) using a nonlinear measure of distance degradation dd . However, the nonlinear phase rotation effect due to HPAs for these systems has not been fully analyzed to clearly determine the acting mechanism. This thesis performs thorough analyzes to explicitly show factors affecting the phase-shifting effect of the combined signal set. Then this nonlinear phase rotation is rationally approximated by a linear phase-shifting model, facilitating the phase estimation. This new proposal allows for automatic phase estimation and compensation, making a qualitative leap for the post-compensation solution that is inherently simple yet extremely effective.

4. Thesis organization

This thesis is organized as follows.

- Chapter 1: Introduction to Nonlinear Distortions and Practical MIMO-STBC Systems

This chapter discusses (a) the basic knowledge regarding to HPAs and their typical nonlinear characteristics, (b) the impacts of nonlinear HPAs on single-carrier SISO systems, (c) general introduction to MIMO systems with three basic techniques, (d) updated studies for mobile satellite MIMO communication systems and (e) nonlinear HPA impacts on MIMO-STBC systems.

- Chapter 2: Nonlinear HPA Modeling and Proposed Polysine Model

Focusing on the analyticity and data approximation capability, this chapter analyzes envelope models which have been widely used such as Saleh, Rapp, Cann, polynomial models, etc. Based on detailed assessments for the causes creating advantages and disadvantages of these models, the thesis proposes polysine model, which not only satisfies analyticity but also matches to the measured data significantly better than previous models do. Next, these models are applied to simulate nonlinear distortion impact on intermodulation products with two typical testing signals, whose results all demonstrate the preeminence of the proposed polysine model. The contents of this chapter are associated with publication 5 in the List of publications.

- Chapter 3: Predistortion Methods for Nonlinear Distortions due to HPAs in MIMO-STBC Systems

This chapter performs the analyses of nonlinear distortion impacts in the MIMO-STBC Alamouti schemes with transmit/receive filters intro-

duced in the system model. This one, therefore, is closer to the practice than others used in previous publications [73, 74, 81, 94], revealing extra detrimental nonlinear distortion effects that have not been investigated before. Then, the thesis proposes the use of transmitter-side distortion compensation (predistortion) for the system. Four typical predistortion algorithms are investigated, allowing to make comparisons between performance improvement capabilities (when applying predistortion) and complexities as well as practical applicabilities. The results of this chapter relate to publication 3 in the List of publications.

- Chapter 4: Automatic Phase Estimation and Compensation for Nonlinear Distortions due to HPAs in MIMO-STBC Systems

On the basis of thorough analysis for the nonlinear phase rotation effects in the Alamouti MIMO-STBC schemes, the thesis proposes a feedforward non-parametric phase estimator and nonlinear phase distortion compensator at the receiver. Phase estimation and compensation performance are verified with various nonlinear HPA models, representing both SSPA and TWTA technologies, confirming the rationale of theoretical analyses and the effectiveness of proposed phase estimation algorithms and phase compensation method. These results have been shown in publications 1, 2 and 4 in the List of publications.

Chapter 1

Introduction to Nonlinear Distortions and Practical MIMO-STBC Systems

The first part of this chapter presents a preliminary classification of nonlinear HPA models and considers impacts of nonlinear HPA in SISO systems. Next, MIMO systems are discussed focusing on three techniques: spatial diversity, spatial multiplexing, and smart antenna. Then, land mobile satellite MIMO systems with their particular characteristics are thoroughly analyzed. On that basis, effects of nonlinear HPAs on performance of these MIMO systems are roughly evaluated. This initiation plays the role of starting point for thorough analyses and specified proposals presented in the following chapters.

1.1. Main causes of nonlinear distortions in radio communication systems

In practice, radio transmitters often have structure consisting of several typical stages such as baseband signal processing, digital-to-analog conversion (DAC), modulation, frequency up-conversion, filtering, amplifications, matching, and antenna, as illustrated in Figure 1.1. Among these parts, RF HPA is one of the most power-consuming components. As an example, for the 2G or 3G terrestrial mobile networks, percentage of power used by base stations (BS) is the largest portion (more than 55%), of which 50%-80% is reserved for HPA [43]. The very large power consumption of HPA comes from two main reasons: limitation of maximum power efficiency that HPA

can achieve; and finite extent of dynamic range that HPA can perform linear amplification. Then, non-ideal effects in radio communication systems also arise primarily from the transmitting part [9].

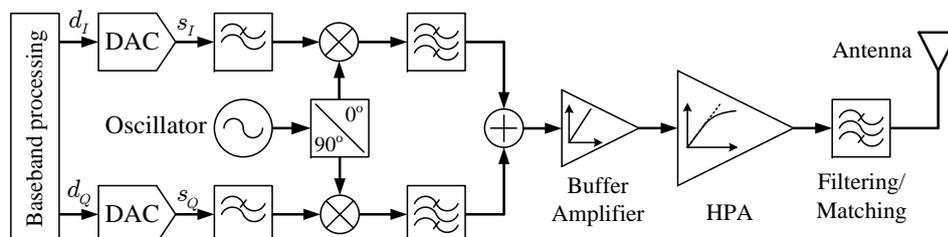


Figure 1.1: Simplified block diagram of a typical radio transmitter.

For a radio transmitter, signal distortions might come from different causes such as non-ideal amplitude/phase frequency response (e.g. in piezoelectric quartz-based devices), harmonic distortion, group delay distortion, direct current (DC) offset, I-Q imbalance,... of which, main distortions are caused by the nonlinearity of RF parts, especially by HPAs [9, 22, 23, 55].

Practically, HPA characteristic is nonlinear and ideal linearity does not exist. HPA nonlinear characteristic results in harmonic distortions (HD) and inter-modulation distortions (IMD). Here, harmonic distortions are unintentionally generated at harmonic frequencies, being positive integer multiples of the input fundamental (first harmonic) frequency; while IMDs are created from any combinations of input fundamental frequencies. These derivative products locate both inside and outside the working passband, and degrade performance of the amplified signal itself. They may also be potential interfering sources for other users/systems. Therefore, these distortions should be minimized such that all users/systems can operate normally. For higher power transmitters such as in satellite transponders or (2G/3G) base station trans-

mitters, this requirement should be strictly adhered to, since the spurious emissions, though having much smaller power than that of the desired signal, also become too large in absolute value and would cause heavy interference to other channels/systems. Another case also being worth to consider is radio systems operating in licence-free frequency bands (or industrial, scientific and medical (ISM) radio bands), where if there is no consensus of transmit powers both inside and outside the working channel/band, it will inevitably causes strong interference such that all devices/systems cannot normally operate.

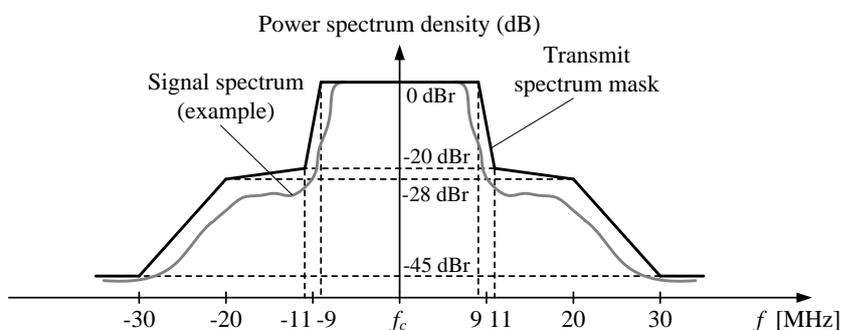


Figure 1.2: The IEEE 802.11a spectrum mask for the 20 MHz bandwidth signal [5].

Therefore, a radio system must maintain its transmit spectrum in a specified mask and is not allowed to emit its energy outside the limits of this mask. A typical example of the transmit spectrum regulation in regard to the ISM band is the spectrum mask for wireless local area networks (WLAN) according to the IEEE 802.11 standard [5], as illustrated in Figure 1.2 for the case of 20 MHz bandwidth signal.

1.2. Nonlinear HPA model classification

Mathematically, although HD and IMD are concisely defined, in practice, adjacent channel power ratio (ACPR) and error vector module/magnitude

(EVM)¹ are more commonly used to determine the level of nonlinear distortion in radio transmitters using complicated digitally-modulated signals with high linearity requirements [9, 23, 32]. ACPR is the ratio between total power of adjacent channels (inter-modulation signal) to the power of desired channel (useful signal), used for determining out-of-band distortion. Meanwhile, EVM measures how far the signal points are from the ideal positions in the constellation, defined in dB or percentage, for quantifying in-band distortion.

Since the higher the ACPR or EVM values are, the lower the quality of signal detection at the receiver is, or equivalently, the larger the degradation of energy efficiency (EE) and spectrum efficiency (SE) is; then, it is really necessary to have a suitable model that precisely represents the HPA nonlinear characteristics for designing effective EE and/or SE systems. While the HPA model at component level (vacuum tube, traveling wave tube, transistors,...) provides high accuracy but has difficulty in analyses, the system-level HPA model, including some main parameters acquired from measurements, significantly facilitates analyses with reasonable accuracy; therefore, the latter is widely used to model HPA for designing or analyzing communication system performances. This type of HPA model is further classified into two sub-categories, memory and memoryless models. Figure 1.3 describes the tree-structured classification of HPA models and summarizes main features relating to each model. In this figure, the models marked with gray background will be studied in detail throughout the thesis.

System-level memory HPA model: Due to the existence of capaci-

¹Sometimes also called relative constellation error or RCE.

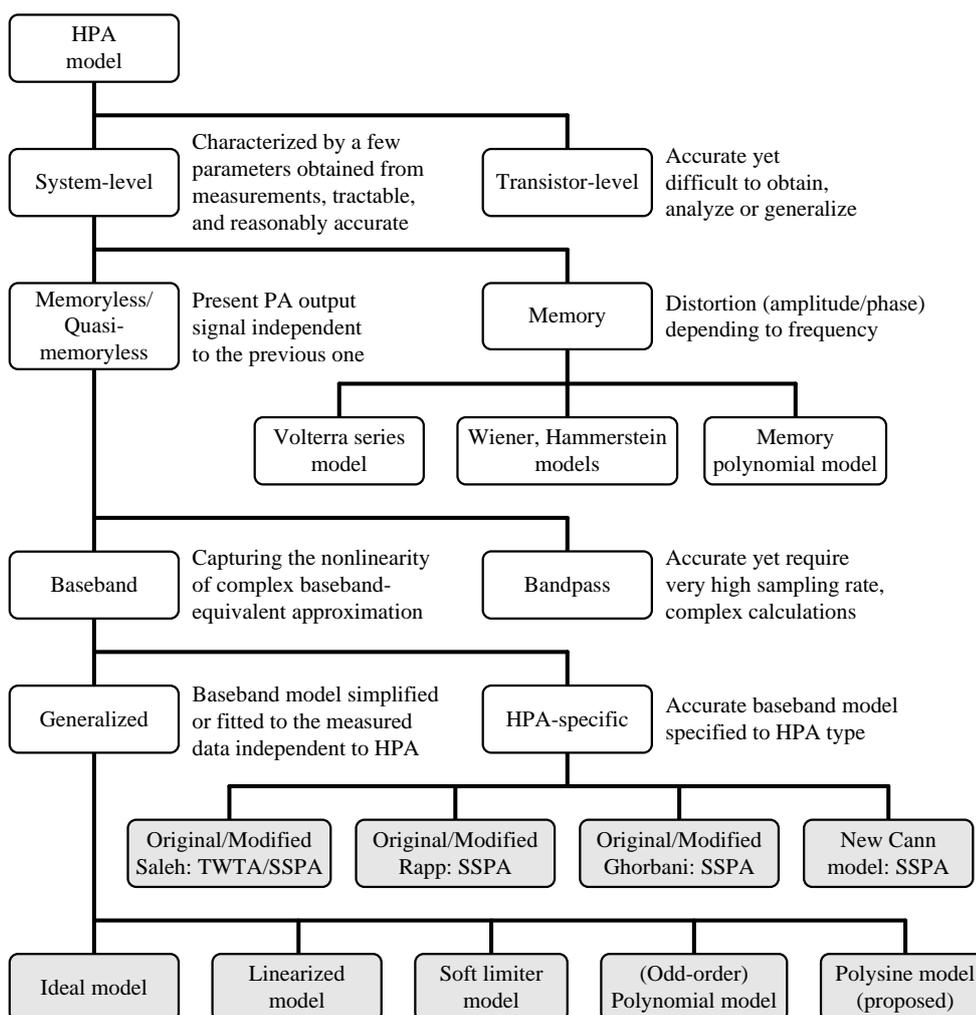


Figure 1.3: HPA modeling classification.

tance, inductance in the circuit as well as temperature variation (when the current and/or voltage changes), frequency-domain memory variation (frequency selective) appears in the HPA transfer function when its bandwidth is large enough [9, 23]. In this case, memory length (or delay) is comparable to the symbol duration or inverse of the signal bandwidth. This manifestation is quite close to the frequency-selective phenomenon of multipath radio channels [54]. As an example, for laterally diffused metal oxide semiconductor (LDMOS) operating in the 2 GHz band, memory effects could be ignored

when system bandwidth is about 1 MHz to 5 MHz; however, electric memory effects will become serious when the signal bandwidth is larger than 5 MHz [53]. Some commonly-used memory HPA models are:

- *Volterra series model* uses multi-variable polynomial (finite order, finite support) for representing HPA output as a function of the input, memory length, and kernel (polynomial) order. The model's computation complexity increases exponentially with numbers of its input parameters, and will be soon impossible for common currently-used processors [52]. Therefore, this model is only suitable for weak nonlinearity and non-realtime applications; otherwise, the approximation by truncating this series returns a model with inferior performance of accuracy.
- *Wiener, Hammerstein, and Wiener-Hammerstein models* includes two parts [32]: The linear filter (memory time-invariant system) A and memoryless nonlinearity B. Wiener, Hammerstein and Wiener-Hammerstein correspond to A-B, B-A, and A-B-A compositions. For example, Wiener model A-B is the structure that creates the output by passing the input signal through A then B blocks. These models give relatively high modeling accuracy with reasonable number of parameters and much lower complexity compared to the Volterra series model.
- *Memory polynomial model* assumes that all signal phases (passing through the model) are independent for reducing the number of Volterra model coefficients approximating to the memory length. As a result, the model's computation complexity scales linearly with the number of samples used to estimate the polynomial coefficients, and this model can be used for

real-time applications with reasonable accuracy.

Therefore, as a simple way, a memory nonlinear HPA model could be constructed from a memoryless HPA model by supplementing memory part in terms of finite impulse response (FIR) filter(s) posited before and/or after memoryless nonlinearity part. The following discussions in this thesis will focus on memoryless nonlinear HPA models. However, it should be emphasized that, practically, an HPA is normally posited before and after respectively, by a transmit pulse-shaping filter and a receive matched filter, most commonly in the form of FIR square-root raised cosine (SSRC) filters; therefore, the overall system nonlinearity characteristic is indispensably memory. Despite of this, for maximal simplification, most nonlinear HPA-related researches often do not include these filters into the system model; then, some important nonlinear impacts that practically and significantly degrade the systems performance are attentively/inattentively ignored. This will be evident through analytical analyses and simulation results in sections 3.2.2 and 4.2.2 of the thesis.

System-level memoryless HPA model: this model basically assumes that the HPA's output at the previous instants do not affect the present one. Amplitude modulation-to-amplitude modulation (AM-AM) distortion function and amplitude modulation-to-phase modulation (AM-PM) distortion function^{2,3} are used for this model. In practice, phase modulation-to-amplitude modulation (PM-AM) and phase modulation-to-phase modulation

²These distortion functions will be briefly defined just below and thoroughly discussed in section 2.3.

³Rigorously, pure memoryless model only has AM-AM characteristic. This HPA type practically causes constant phase shift (delay) regardless input magnitude level when this signal has a bandwidth small enough; then for simplification, it is reasonable to assume that the phase conversion is zero. If HPA's memory (causing delay variation for the output signal) occurs with its effectiveness within the signal's symbol period, then this HPA does cause nonlinear phase distortion. In this case, it could be further introduced an AM-PM function into the model, making it becomes quasi-memoryless.

(PM-PM) conversions are often ignored except when having significant values, for example in the case of a quadrature modulator with predistortion [53].

It is often not easy to filter out distorted components that are quite close to the carrier frequency; then, they could be thoroughly and precisely represented by using *bandpass model* with sampling rate satisfying the Nyquist sampling criterion [52, 54]. In this case, computation complexity could be really huge and redundant. Thus, for easy calculation and simulation, *baseband model* expressing the nonlinearity of complex baseband frequency approximation (also called equivalent lowpass/complex envelope) is more widely used than the passband model [52]. Most commonly used baseband models can be divided into two categories: *generalized baseband model* and *HPA-specific baseband model*. The former is either mostly simplified HPA models (ideal model, linearized model, and soft limiter model) or models which are fitted to data independent of a specific HPA type (polynomial model, polysine model,...); they allow analytical analyses that are not dependent on a particular HPA. However, simplified models may be (extremely) inaccurate. Meanwhile, HPA-specific models often use more parameters to achieve higher accuracy for the HPA nonlinear characteristics, and therefore, produce more reliable simulation results.

In Chapter 2, rigorous analyses are carried out for the following models: polynomial, polysine, Cann, Rapp, Saleh, Ghorbani, and their corresponding modified versions. Therefore, to avoid duplication, this section only mentions: ideal, linearized, and soft limiter model. Here, the concept of equivalent low-pass signal is latently used and will be discussed thoroughly in subsection

2.3.1 for better coherence. Let $r(t)$ and $\phi(t)$ respectively be the amplitude modulation (AM) component and phase modulation (PM) component of the input $x(t)$ with time variable t , $x(t) = r(t)e^{j\phi(t)}$, $j = \sqrt{-1}$. Generally, for representing nonlinear relation $F(\cdot)$ between the input $x(t)$ and output $y(t)$, an *input amplitude-dependent gain function* (or AM-AM function), $F_a(r)$, and an *input amplitude-dependent phase shift function* (or AM-PM function), $F_p(r)$, are used to present the output signal as

$$y(t) = F(x(t)) = F_a(r(t))e^{j(\phi+F_p(r(t)))}. \quad (1.1)$$

- *Ideal model* is the perfectly linearized model for HPA. Specifically, HPA's input magnitude is linearly amplified over the whole dynamic range,

$$y = gx, \quad (1.2)$$

where x and y are the short forms of $x(t)$ and $y(t)$, $g > 0$ is the linear gain. In this case, g is a positive real number, meaning that there is no phase conversion for the output signal y . Hence, $F_a(r) = gr$ and $F_p(r) = 0$.

- *Linearized model* is the simplest HPA model without considering output magnitude clipping for nonlinear characteristic

$$y = gx + n, \quad (1.3)$$

where g has the same meaning as in (1.2), n is a nonlinear distortion uncorrelated to the input x and is approximated by a Gaussian noise based on the Bussgang theorem [16, 69]⁴. This model does not show the

⁴Bussgang introduced the original theorem [16] which states that the cross-correlation function between the real-valued, Gaussian-distributed input and the output of a memoryless nonlinear-amplitude device is proportional to the input autocorrelation function. Minkoff [69] extended this result for complex-valued signals, taking into account both amplitude and phase nonlinearities.

HPA's output saturation when the input amplitude exceeds a specified threshold; therefore, it can only be applied to systems operating at a large input back-off (IBO) power [23].

- *Soft limiter* is the simplest HPA model considering output magnitude clipping [52] as

$$F_a(r) = \begin{cases} r, & |r| < A_{is} \\ A_{is}, & |r| \geq A_{is}, \end{cases} \quad (1.4)$$

$$F_p(r) = 0, \quad (1.5)$$

where A_{is} is the input saturation level (voltage) which, in this case with unity gain $g = 1$, is also the output saturation level, i.e. $A_{os} = A_{is}$. The soft limiter model represents output amplitude clipping at large power regime and is applicable if the nonlinearity in the low power regime has been mitigated by applying linearization techniques [23, 32, 55].

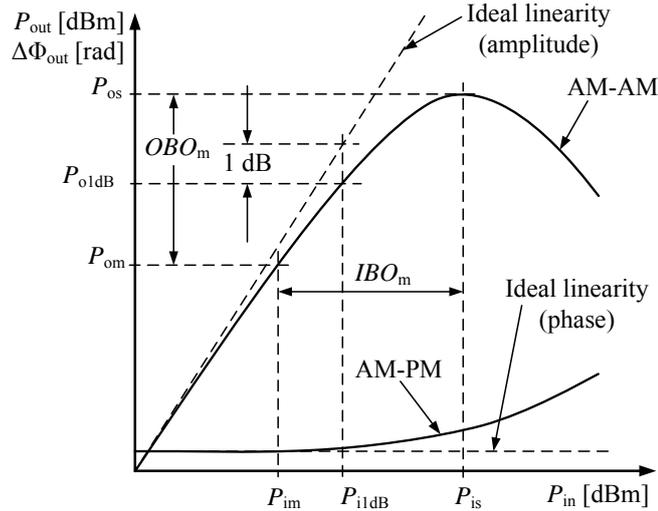


Figure 1.4: Typical amplitude and phase distortion characteristics of an HPA^(*).
^(*) Notations used for the inputs: P_{is} - Saturated input power; P_{1dB} - Input power at 1dB output power compression point; P_{im} - Average input power; IBO_m - Average input power backoff. Counterparts for the outputs: P_{os} , P_{0dB} , P_{om} , OBO_m .

In general, as shown in formula (1.1), HPA characteristics express the magnitude compressing-expanding effect, or AM-AM conversion $F_a(r)$ and phase shift effect, or AM-PM conversion $F_p(r)$. For the input argument of these functions, in some cases, especially for datasheets, input magnitude $|r|$ or input power, $P_{\text{in}} = |r|^2$, could be used instead of input envelope r . Also is the output power P_{out} . Figure 1.4 illustrates these relations for typical input/output powers, $P_{\text{out}} = F_a(P_{\text{in}})$, and phase shift $\Delta\Phi_{\text{out}} = F_p(P_{\text{in}})$. The quantities represented in this figure will be widely used in quantitative analyses in later chapters. In practice, HPA's AM-AM and AM-PM conversions are often given by manufacturers in terms of output power-dependent power gain characteristic ($\Delta G(P_{\text{out}})$) and output power-dependent phase shift characteristic ($\Delta\Phi(P_{\text{out}})$).

HPA-specific models, polynomial model and polysine model are rigorously discussed in Chapter 2 in a specific topic for HPA models and are widely used in Chapters 3 and 4 for analysing system performance.

1.3. Nonlinear HPA distortion impacts in SISO systems

Based on typical nonlinear characteristics illustrated in Figure 1.4, it can be seen that, nonlinearities have different impacts on amplitude-, frequency- and phase-modulated signals. For frequency- and phase-modulated signals, which have a constant magnitude, HPA nonlinearities do not create harmful AM-AM and AM-PM distortions. In contrast, amplitude- or amplitude-phase-modulated signals like M -ASK, M -QAM, or M -APSK are strongly affected by HPA nonlinearities, especially for high modulation order M . In fact, for single-carrier SISO systems, under the influence of HPA nonlinear

characteristics, several complex-interrelated detrimental impacts will be generated when inputting signal with non-constant envelope. However, in order for understanding the acting mechanisms in the analysis, it can be isolated into separate effects as follows [2].

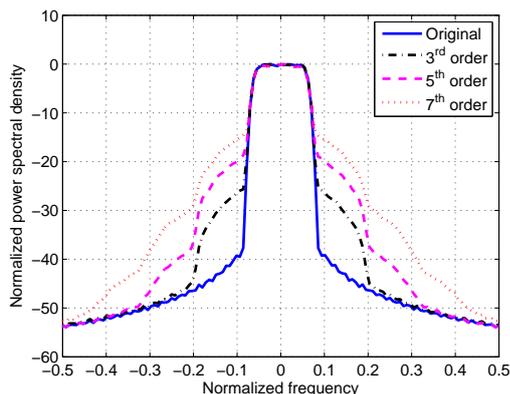


Figure 1.5: Spectrum regrowth due to nonlinear HPA^(*).

^(*) Here, different AM-AM nonlinearities are approximated by polynomials with their corresponding orders annotated.

- Creating spectrum regrowth and nonlinear noise:** If approximately representing the AM-AM characteristic of an HPA by a polynomial, some of the intermodulation products generated from the odd orders of the model can be considered as a nonlinear noise if these products fall into the input signal bandwidth. Along with the thermal noise, this nonlinear noise will contribute to reducing the signal to interference plus noise ratio (SINR), and increasing the system bit error ratio (BER). In addition, the intermodulation products may also fall outside the bandwidth, causing spectrum regrowth as illustrated in Figure 1.5. This spectrum regrowth could result in adjacent channel interference (ACI). To limit this interference, at the HPA output, there is usually a zonal filter that is responsible for keeping the transmit spectrum within the limits of a

standard spectral mask (as demonstrated in Figure 1.2).

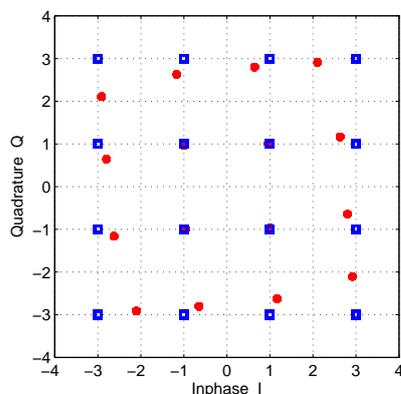


Figure 1.6: Constellation warping due to nonlinear HPA.

- **Warping constellation:** AM-AM and AM-PM characteristics make signals with different magnitudes be amplified at different amplification coefficients and be rotated at different phase rotation angles [69]. Therefore, the average positions of the receive constellation could be shifted away from the ideal ones. Signals with larger magnitudes (powers) are further shifted away, as shown in Figure 1.6. Signal points that are shifted closer to the decision boundary, significantly increase the system BER.

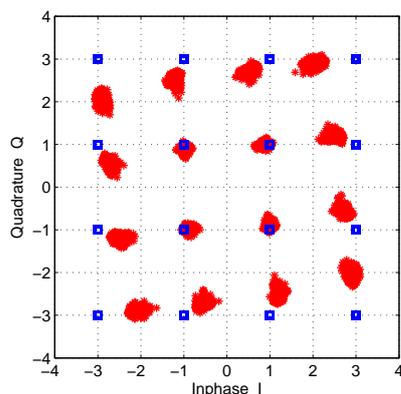


Figure 1.7: Nonlinear ISI due to nonlinear HPA.

- **Creating nonlinear ISI:** Practically, an HPA, which is sandwiched

between transmit (pulse-shaping) and receive (matched) filters⁵, causes nonlinear ISI whenever AM-AM and/or AM-PM distortions exist. The reason is that, when having an HPA with nonlinearities, the overall (linear) system transfer function initially designed to satisfy the Nyquist's zero ISI transmission condition [54] will no longer be fulfilled. The resulted nonlinear ISI could seriously increase the system BER. Under the impact of nonlinear HPA, instead of being located in the right designed positions, received signals will appear as a cluster of points in the constellation. The larger the signal magnitude, the wider the cluster of points (stronger ISI), as illustrated in Figure 1.7. More importantly, it is observed that these clusters are obviously non-Gaussian distributed. This makes analytical analyses really difficult to keep track.

1.4. Multiple-input multiple-output systems

The concept of multiple-input multiple-output schemes began to appear in the mid-1950s in circuit theory and signal filtering theory, where it was used to describe circuits/diagrams with multiple-input ports/terminals and multiple-output ports/terminals [27]. However, in the 1990s, this concept was put on a new look, for a completely different signal processing technique [26]. The concept of multiple-input in this case is used to index signals from different transmit antennas “entering” into the radio medium/channel; while the concept of multiple-output is for indexing signals “exiting” from this medium. A simplified block diagram of a communication system having a MIMO channel with n_T inputs and n_R outputs is illustrated in Figure 1.8.

⁵Here, the system model is briefly introduced to preliminarily explain the mechanism creating nonlinear ISI. This problem will be thoroughly discussed in section 3.2.

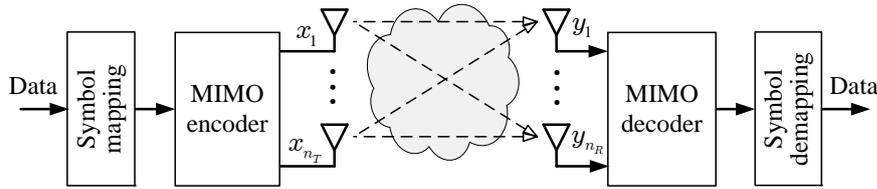


Figure 1.8: Simplified MIMO system diagram.

Based on different point of views on the effects of fading on radio signals, three new multi-antenna techniques have been developed in almost parallel: *spatial diversity* (SD), *spatial multiplexing* (SM), and *smart antenna* (SA). Each technique tries to most effectively exploit fading effects in a different way.

1. **Spatial diversity** uses multiple transmit antennas and/or multiple receive antennas to acquire transmit/receive diversity gains based on statistical independence of fading on sub-channels between each transmit-receive antenna pair⁶. In this case, similar to the channel codes, the antennas are used to improve the system error performance by transmitting and/or receiving residual versions representing the same sequence/block of information signals. The gain acquired in the form of improvement in SNR to achieve the same BER performance compared to the single-antenna case is called *diversity gain*.

Historically, the development order may be different, however, the main motivation for diversity techniques derived from transmit diversity, contrasting to receive diversity, which has been developed and almost played

⁶If the spacing between antennas in the transmitter and in the receiver is sufficiently large, separate transmission paths between transmit-receive antenna pairs become statistically independent. As a consequence, in case of occurring fading, the probability that all transmission paths are simultaneously in a bad state (deep fade) will be much smaller (according to multiplication rule) compared to the case when system having only one transmission path (SISO).

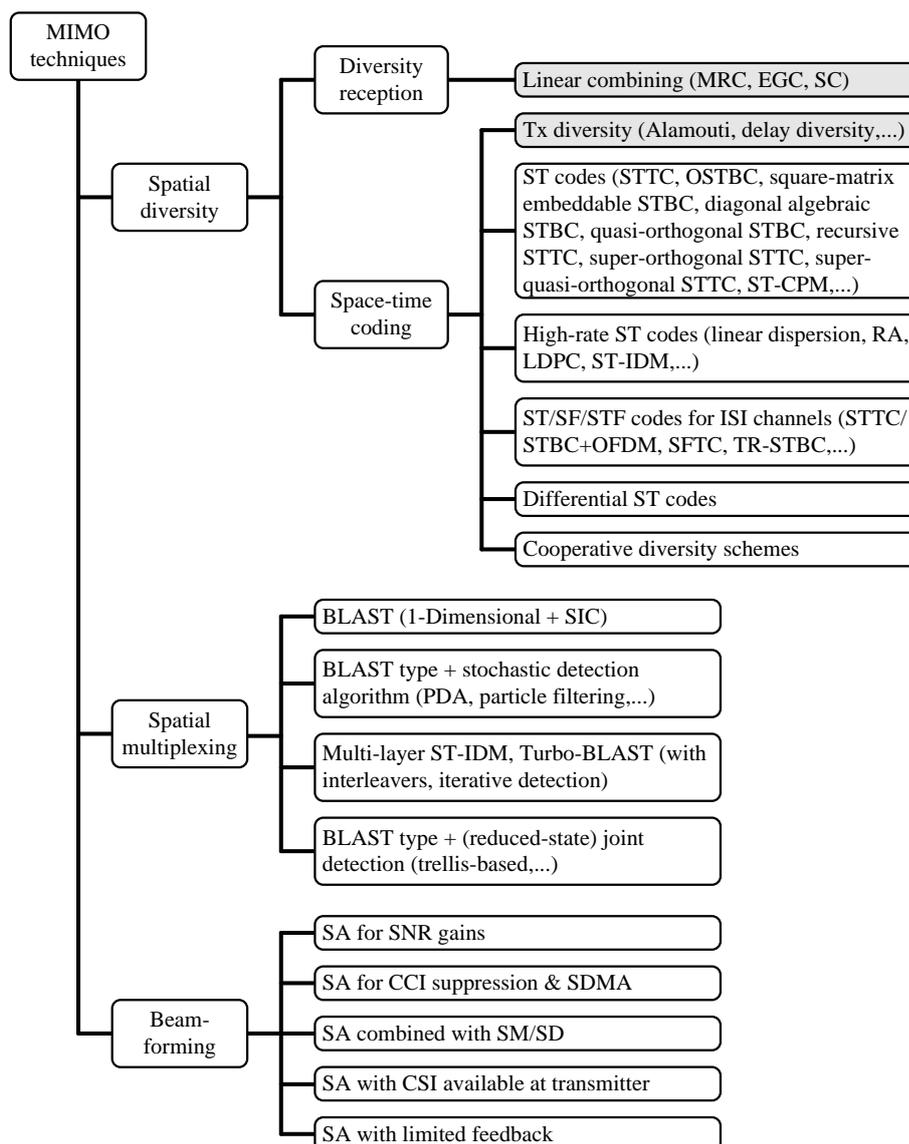


Figure 1.9: MIMO technique classification^(*) [68].

^(*) Some terminologies and acronyms used in this figure: MRC - Maximum Ratio Combining; EGC - Equal-Gain Combining; SA - Smart Antenna; SC - Selection Combining; ST-CPM - Space-Time Continuous-Phase Modulation; RA - Repeat-Accumulate (code); LDPC - Low-Density Parity-Check (code); ST-IDM - Space-Time Interleave-Division Multiplexing; TR-STBC - Time-Reversal STBC; SIC - Successive Interference Cancellation; PDA - Probabilistic Data Association; CCI - Co-Channel Interference; SDMA - Space-Division Multiple-Access; CSI - Channel-State Information.

the monopoly long ago, since the late 1950s [15]. Transmit diversity really needs for the downlink in cellular networks, especially in the 5G massive MIMO systems, and in some other scenarios: user terminals need to be as simple and compact as possible and therefore need to avoid deploying multiple antennas and multiple receivers here. Moreover, in order to

effectively differentiate diversity paths in the receiver, the transmitter should actively deploy appropriate coding methods. The first works on the spatial diversity technique might belong to Witneben A. published in 1991 [102] and to Seshadri N. and Winters J.H. [85] in 1993, applied to terrestrial mobile communication systems. Thereafter, the most prominence could be the work of Alamouti S. [7] proposing a space-time coding method that enables transmit diversity with only two antennas in a fairly simple way, while does not require too high processing complexity on the receiver side. This diagram achieves maximum transmission efficiency with minimum complexity. Since then, extended versions of the Alamouti's principle have become one of the most popular schemes for practical MIMO systems widely used until now.

In general, the combination of space dimension with other dimension(s) (time, frequency,...) depends mainly on the specific characteristics of the radio channel considered in the system design problem [47]. For frequency-flat channels, through two-dimensional (2D) coding over time and space, and therefore often called *space-time* (ST) encoding, the information sequence is spread on transmit antennas. There are two ways to implement ST encoding: *Space-time trellis codes* (STTC) allocate a trellis (convolutional) code on antennas and time slots, and thus provide coding gain and diversity gain [89]. Meanwhile, *space-time block codes* (STBC) handles each data block (in a similar way as block codes execute) for allocating on antennas and time slots, and thus also create diversity gain but do not result in coding gain [90]. ST Alamouti code [7]

is a particularly simple case of orthogonal STBC (OSTBC) in this STBC class. OSTBC codes achieve full diversity with low decoding complexity due to the orthogonality achieved from the transmit encoding [42, 105]. For frequency-selective fading channels, the solution that has been successfully applied for SISO paradigm using orthogonal frequency division multiplexing (OFDM), could be directly expanded for MIMO configuration, resulting in MIMO-OFDM structures in the forms of STTC+OFDM [70] or STBC+OFDM [63]. The OFDM component serves to convert a time-domain multipath channel into (frequency-domain) frequency-flat sub-channels (subcarriers). Then beside the space and time dimensions, MIMO-OFDM systems generate a third dimension, i.e. the frequency dimension. Therefore, more generally, in addition to space-time coding, MIMO-OFDM systems can also employ coding on (2D) *space-frequency* (SF) (on subcarriers) or on (3D) *space-time-frequency* (STF). *Space-frequency trellis codes* (SFTC) were proved to have significantly better quality than STTC codes on frequency-selective channels [47]. In addition, compared to ST codes, SF codes sometimes have more advantages in deployment, such as the relatively simple receiver structure, flexible processing and resource allocation, or the availability of existing OFDM technologies,...

2. **Spatial multiplexing**, instead of using multiple antennas to eliminate the “adverse” effects of fading, this technique takes advantages from the independence of those fading fluctuations to increase the system capacity. Spatial multiplexing techniques perform simultaneous transmission

of independent signal sequences, often called layers, on multiple antennas. With the use of n_T transmit antenna and accepting the penalty of hardware expansion (in terms of cost, complexity, space,...), the total bit rate, compared to the single antenna system, can therefore increase maximally n_T times without requirements of additional bandwidth or transmit power. The achieved gain in the form of bit rate increase compared to the single antenna system, is called *multiplexing gain*.

At the transmitter, since layers are (linearly) positioned on each other during the transmission, they need to be separated in the receiver by some kind of interference-suppression algorithms (often combined with receive diversity). Foschini G.J. was the first author to propose the concept of spatial multiplexing called *Bell-labs layered space-time* (BLAST) structure in 1996 [35]. Two years later, it was also this author and his colleagues who introduced the first system model to implement BLAST technique called Vertical-BLAST (V-BLAST) [103]. The BLAST diagram then is further expanded in directions such as reducing the detection complexity or applying OFDM solution for multipath frequency-selective channels,...

3. **Smart antenna** does not directly aim at increasing the overall system data rate or reducing the error rate, but focusing on improving the signal-to-noise ratio (SNR) in the receiver and reducing co-channel interference in the multi-user scenario by controlling the antenna array's directivity (radio beam). The concept of antenna array with adaptive beamforming was actually not new because of its origination in the field of radar

Table 1.0: Commercialized wireless standards using MIMO.

Wireless standard	Maximum antenna configuration
IEEE 802.11n (WiFi)	4×4
IEEE 802.16e (WiMAX)	4×4
HSPA ⁺ (Evolved HSPA)	2×2
LTE (3.9G)	4×4
LTE-Advanced (4G)	8×8
802.11ac (extension of 802.11n)	8×8

(as in the case of target tracking) and space technology (with very large phased-array antenna systems) [44]. However, in-depth research on smart antennas for radio communication systems began only in the late 1990s maybe with the first work of Godara L.C. applied for mobile communication systems [39]. By using beamforming techniques, the radiation patterns of the transmit and receive antenna arrays can be driven in the desired direction, while the undesirable directions (e.g. the directions of main noise/interfering sources from/to other users/systems) can be suppressed (nulled). This beamforming technique can be seen as a linear filtering technique on the spatial domain. The SNR gain achieved by beamforming technique is often referred to as *antenna/array gain*.

In fact, a clear distinction between these three MIMO techniques is not easy. For example, spatial multiplexing techniques can also provide diversity gain if the receiver uses maximum likelihood detection [105]. Similarly, spatial diversity systems can be used to increase the bit rate⁷ when combined with adaptive channel coding/modulation schemes [40], which is inherently available in most modern digital radio communication standards.

⁷If the system BER achieved by MIMO technique is much smaller than a minimum requirement, it is possible to switch the system to use a modulation scheme with higher order or to use a channel coding scheme with a higher coding rate. By this way, the system can achieve higher bit rates with acceptable error rates without reducing system power efficiency.

Commercial MIMO technology first appeared in 2001 by Iospan Wireless for broadband fixed wireless networks with the trademark AirBurstTM, while the first WiMAX standard using MIMO technology was introduced in 2005. Currently, MIMO technologies in terms of spatial diversity, spatial multiplexing or smart antenna technology have been widely adopted for commercial standards of wireless networks. Table 1.0 lists some common commercial wireless standards that support MIMO and the maximum antenna configuration supported in each standard. The standards listed in this table support the following multi-antenna techniques [42]:

- Alamouti space-time coding for transmit diversity;
- BLAST spatial multiplexing architectures;
- Eigen-beamforming spatial multiplexing;
- Conventional beam and null forming;
- Conventional receive diversity.

The summary classification of MIMO techniques is shown in Figure 1.9, where blocks marked with gray background (Alamouti transmit diversity and linear receive combining) are the research subjects of this thesis considering the nonlinear effect of HPA. This is a relatively classical MIMO configuration and has many practical applications, such as in terrestrial radio communication systems and especially in satellite communications. However, a thorough examination of the system considering impacts of nonlinear HPAs has not been mentioned widely yet.

1.5. MIMO in satellite communication systems

Although MIMO techniques were originally developed for terrestrial systems, they have many advantages for other applications, of which satellite communications is an example. However, compared to terrestrial counterparts, satellite systems have many clear differences from radio transmission channel characteristics to system structures, etc. Therefore, the application of MIMO techniques in satellite communications is slower carried out since more researches are needed to fill these distinct gaps.

The prerequisite conditions for single-satellite configurations to provide full advantages of spatial diversity and spatial multiplexing as theoretically determined are a sufficiently large spacing between antennas, as well as a rich scattering transmission medium so that the fadings occurring in radio links between pairs of transmit-receive antennas are independent. However, the great distances between satellites and ground stations make these radio links actually become a *keyhole* (or *pinhole*) channel⁸ [76], with almost only one connection link existing. The high correlation between MIMO subchannels caused by degeneration of multipath environments, thus causes a significant reduction in channel capacity compared to the ideal case determined by theoretical calculations [93].

Although studies have shown that a high-altitude platform (HAP), flying at a height of 20 km above the ground, can be used effectively to exploit the advantages of MIMO for terrestrial communications [58]; however, for the case

⁸Keyhole channel effect involves cases where strong scattering occurs around the transmitter and/or receiver, creating a low correlation between the propagating signals; but other wave propagation phenomena such as diffraction or waveguiding occur in the intermediate space, causing dependent channel states and thus rank deficient of the MIMO channel transfer function matrix. Some transmission environments that could create this type of channel are corridor, tunnel or satellite-ground,...

of satellites at significantly higher altitudes (from hundreds of kilometers to tens of thousands of kilometers), the deployment of multiple antennas at each individual satellite does not seem to be beneficial due to space restrictions. Specifically, the authors in [58] have shown that, for the case of geostationary satellite orbit (GSO) the spacing between antenna elements must be at least 1.5×10^5 wavelengths, or about 30 km for the operating frequency of 1.54 GHz (L band), to achieve low enough antenna correlation (less than 0.1). Therefore, the vast majority of researches relating to multi-antenna satellite systems only focuses on exploiting the following diversity configurations (or their combinations):

- *Site diversity* (for downlink transmission) consists of connecting two or more ground stations (posited far enough apart), all receive the same signal: if the signal to be received is severely faded in one region, other ground station(s), which are highly in good transmission conditions can accommodate for it;
- *Satellite diversity*, also called *angular diversity* or *orbit diversity*, through the deployment of multiple (usually two) satellites (posited far enough apart) and a ground station equipped with single-polarized antennas or diversity via hybrid satellite-terrestrial MIMO systems⁹;
- *Polarization diversity* or *pattern diversity*, in which a satellite with dual-polarization antenna communicates to a ground station also equipped with the same type of antenna. This technique are becoming increasingly popular complements - and even alternatives - to space diversity [93].

⁹This system includes a satellite and one or more auxiliary terrestrial base stations, also known as gap filler. This is one of the most promising solutions for terrestrial mobile satellite networks, ensuring coverage for several different environments: urban, suburban, rural, vehicular, open area,...

Connections in fixed satellite systems are based on high-directional antennas; for such systems, although a significant multiplexing gain acquired by site diversity has recently been demonstrated in practice [48], but currently, it could be unworthy compared to the penalty of system cost, complexity, implementation difficulties. That is why most studies of MIMO satellite communications are currently focused on land mobile satellite (LMSat) systems.

On the other hand, polarization diversity is a promising solution due to recent advances in the design of MIMO (dual polarization) antennas overcoming the onboard spatial limitations and disadvantages for multi-satellite structure, for example, problem of wasting satellite bandwidth (which is inherently very scarce) to transmit the same signal, problem of synchronizing the receive signals (becoming very complicated when incurred by large transmission delay), or problems of scheduling, ISI, high deployment costs, etc. Therefore, the general trend of MIMO LMSat studies accents to the use of polarization diversity [8, 20, 37, 50, 59]. However, maximally, polarization diversity can only double throughput, while satellite diversity can allow increasing throughput proportional to the number of satellites in the system. In addition, extreme obstruction (on/off blockage) and highly correlated environments (by rain attenuation, mainly in frequency bands greater than 10 GHz) can cause significant performance degradation for polarization diversity.

It also should be noted that terrestrial systems often use linear polarization for simplification, while satellite systems, in a certain conditions, could choose circular polarization to overcome the Faraday rotation effect in the ionosphere¹⁰. Other effects in the ionosphere and stratosphere can be omitted

¹⁰Generally, the Faraday rotation is a magneto-optical phenomenon, that is an interaction between light and a magnetic field in a medium. For satellite communications, this phenomenon results from the electron plasma of the

for the L and S band signal [20]. Therefore, only the multipath propagation by local scattering near (ground) mobile stations is mostly interested. In addition, it is possible to combine satellite diversity and polarization diversity when using multiple satellites, each deploying a dual-polarization antenna. Space-time-frequency-polarization coding can be exploited and thus increasing the system capacity. Configuration of a MIMO LMSat system using dual right/left hand circular polarization (RHCP/LHCP) is illustrated in Figure 1.10. Here, with the basic transmit-receive structure 2×2 , the following coding techniques are considered: parallel SISO system $2 \times \text{SISO}$, Alamouti transmit diversity [7], Golden code [10], spatial multiplexing [35].

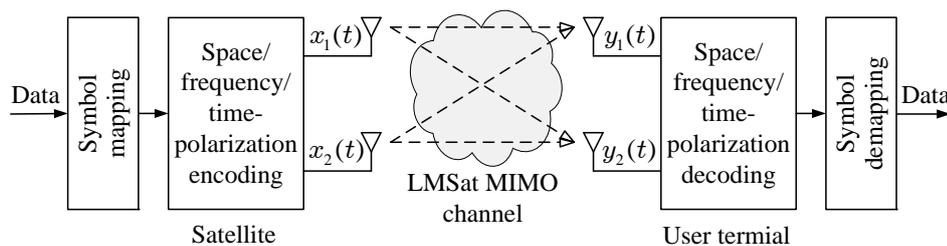


Figure 1.10: Dual-polarized MIMO land mobile satellite system model.

Analyzing the dual-polarization MIMO LMSat system performance has thoroughly been studied, basically on linear channel, [20, 37, 49, 50, 75]. The general conclusion is that it is possible to deploy a dual-polarization antenna on mobile terminals for practical communication. In addition, the expansion of using MIMO techniques for hybrid satellite-terrestrial systems has also been taken into account [20, 37, 75, 77]. However, most of these works do not mention the practical nonlinear HPA in the system, or introduce the nonlinear

ionosphere and the presence of Earth's magnetic field. As a linearly polarized radio wave passes through the ionosphere, its electric field, with orientation described by a polarization vector, will gradually rotate in the plane perpendicular to the direction of propagation. Then, this wave becomes de-polarized, causing attenuation for the desired component and undesirable reception of the opposing polarity. However, the Faraday rotation has no detrimental effect for circularly-polarized signals. Otherwise, since the rotation is proportional to the square of the carrier wavelength, the effect might be significant for carriers in low frequencies such as L, S, C band and negligible at higher ones, such as Ku Band[24].

HPA into the system model to evaluate performance by simulation but do not perform any quantitative analysis or assessment of the specific effects of nonlinear factors on system performance. In the next section, impacts of nonlinear HPAs on the MIMO system as shown in the Figure 1.10 will be analyzed briefly to get an overview of the problems that will be solved next in the thesis.

1.6. Nonlinear HPA distortion impacts in MIMO systems

In addition to incurring the similar effects as for conventional SISO systems analyzed in section 1.3, performance degradation due to nonlinear HPAs for MIMO systems using non-constant envelope signals could also come from the fact that nonlinearly-distorted transmit signal sets could introduce extra interference. More seriously, the interference might interact in an involved manner making poly-degradation. This can be easily observed through an example of a simple MIMO system using the STBC Alamouti coding [7] with nonlinear HPA as illustrated in Figure 1.11.

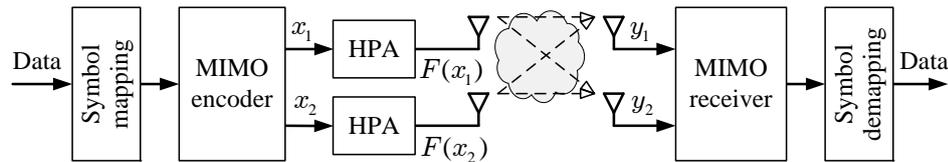


Figure 1.11: Simplified MIMO system with nonlinear HPA.

First, M -level messages m_k and m_{k+1} are fed into the symbol mapper, such as an M -QAM modulator, which maps these values to the output modulated signals s_k and s_{k+1} , correspondingly. The MIMO encoder receives this input

symbol pair and generates the output encoding matrix \mathbf{X} in the form of

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} x_{1,k} & x_{1,k+1} \\ x_{2,k} & x_{2,k+1} \end{bmatrix} = \begin{bmatrix} s_k & -s_{k+1}^* \\ s_{k+1} & s_k^* \end{bmatrix}, \quad (1.6)$$

where the columns of \mathbf{X} correspond to time dimension of slots k and $k + 1$, while the rows correspond to the spatial dimension of the first and second transmit antennas¹¹, $(\cdot)^*$ is the complex conjugate operator.

It is not difficult to realize that this is an orthogonal design, namely

$$\mathbf{x}_1 \mathbf{x}_2^H = [s_k \quad -s_{k+1}^*] \begin{bmatrix} s_{k+1}^* \\ s_k \end{bmatrix} = 0, \quad (1.7)$$

where $(\cdot)^H$ is the complex conjugate (Hermitian) transpose operator. In other words, signals on each transmit antenna do not interfere with signals on the other antenna; and it is the orthogonality of the STBC Alamouti code that gives full transmit diversity order while allows maximum simplification for decoding in the receiver through linear combination processing without the need for any other interfering cancellation measures [7, 42].

The above property of the STBC Alamouti code is only valid conditioned on the linearity of HPA. In a general case with nonlinear HPA, its output signal will relate to the input one according to a nonlinear function $F(\cdot)$ as shown in (1.1), representing both AM-AM and AM-PM distortions, that means changing not only the amplitude but also the phase of s_k and s_{k+1} in a nonlinear way. Under such conditions, it can be seen that the orthogonality

¹¹For conventional MIMO systems, these antennas are usually linearly polarized and placed far enough apart to minimize the correlation. However, in the case of using dual-polarized antennas (for polarization diversity) as, for example, in MIMO LMSat systems discussed above, these (two) antennas can be posited almost in the same place, corresponding to two symmetrical right- and left-handed circular polarizations. These orthogonal polarization antennas may provide low levels of correlation with minimum or no antenna spacing while making the communication link robust to polarization rotations in the channel [93].

of the STBC Alamouti code (1.7) will be broken:

$$\mathbf{x}_1 \mathbf{x}_2^H = [F(s_k) \quad F(-s_{k+1}^*)] \begin{bmatrix} F(s_{k+1}^*) \\ F(s_k) \end{bmatrix} \neq 0. \quad (1.8)$$

Therefore, in this case, the linear combination decoder used at the receiver will no longer be optimal, and the system performance (including transmit diversity order) will degrade compared to the ideal linear case.

Thus, it is the HPA nonlinearity that eliminates the orthogonality of the Alamouti code and the diversity gain is deteriorated under the appearance of non-orthogonal components due to nonlinear distortions. The problem will become even more complicated when further considering the memory effects of transmit/receive filters. For example, as analyzed in section 1.3, the nonlinear ISI, generated from each individual transmit branch continues to affect orthogonality in a manner similar to what useful signals influenced as shown above, or the nonlinear inter-antenna interference (non-orthogonality component) continues to deteriorate receive signal in each antenna branch under the memory effect of the matched filter. Thus, the system performance is poly-degraded in an involved manner.

This simple example reveals that nonlinear HPA distortion effects in MIMO systems are rather more complicated than those in SISO counterparts, especially when further considering the memory in the system model. Detailed analyses will be presented in the Chapter 3 and 4, to clearly show the acting mechanisms. These are the bases for simple yet efficient proposed treatments that correctly resolve the detrimental impacts of distortion.

1.7. Summary of chapter 1

The background knowledge directly related to the research objects has been discussed in this chapter. They include the nonlinear HPA model, MIMO techniques with specific implementations to the LMSat systems, and the impacts of nonlinear HPAs in MIMO communication systems. These analyses have clearly shown urgent issues and updated research directions that the thesis can pursue. This is the foundation for specific studies presented in the next chapters.

Chapter 2

Nonlinear HPA Modeling and Proposed Polysine Model

The first part of this chapter summarizes modern HPA technologies and nonlinear HPA modeling problems. Next, analyses are devoted to the Cann original model for instantaneous signals and its defect encountered in creating intermodulation products when performing simulation experiments with the two-tone testing signals. Thereafter, the Cann new model is examined on the aspect of overcoming this defect. The problem is then extended to envelope nonlinear models that are now widely used such as Saleh, Rapp, polynomial models, and our proposed polysine model. The latter not only surmounts defects in analyticity compared to the previous ones but also outperforms in data approximating compared to the Cann new model. Finally, simulation experiments with two different types of testing signals are presented, illustrating the pros and cons of the considered models. The content of this chapter relates to the publication 5 in the List of Publications.

2.1. Introduction

Besides a relatively small portion using other technologies for specific applications, HPAs are currently implemented by the two most popular technologies, namely the traveling wave tube (TWT) and semiconductor [100]. A traveling wave tube amplifier (TWTA) consists of an amplification element being TWT and an electronic power conditioner (EPC), which contains DC-

to-DC, DC-to-AC adapters, sensors, protection and (remote) control circuits. Similarly, a solid-state power amplifier (SSPA) consists of a transistor and an EPC (yet having quite different parameters compared to TWTA's EPC).

For satellite communications (SatCom), historically, the TWTA was first used on satellite transponders with features that can be listed as [100]: high power, wide operating frequency range (and can be very high), high power efficiency, large bandwidth, high reliability, good thermal durability; but it also is a highly nonlinear element, bulky in size and needs a very high voltage source (up to kilo volts between cathode and collector). With the advancements of semiconductor technologies, SSPA gradually replaces TWTA in low frequency bands such as L, S, C, X and lower portion of Ku¹. The main features of SSPA in comparison with TWTA are lighter in weight, lower in cost, better in linearity, safer in operating with low DC voltage (but very high DC current), lower in operating frequency (currently, commercial products are popular up to Ku band), inferior in power efficiency (and decrease by operating frequency), smaller in output power, lower in reliability, and poorer in heat durability (can be more cumbersome than TWTA when attached to the heat sink). In addition to being deployed on satellite transponders, SSPAs have been traditionally and commonly used in terrestrial communications equipment.

In fact, HPA's nonlinear characteristics (input-output relations) are often measured at discrete points based on testing signals in the forms of one or two unmodulated carrier(s) (then called single-/two-tone test); therefore it needs to perform interpolation or extrapolation from these data to obtain other

¹Corresponding to frequency bands 1-2, 2-4, 4-8, 8-12, 12-18 GHz.

necessary value(s) for the system analytical analysis or numerical simulation. With reasonable accuracy, a model approximated by closed-form mathematical functions is thus a very convenient tool to replace.

Initially following this direction, in 1980, Cann [17] introduced an instantaneous nonlinear model for HPAs with variable knee sharpness, relatively convenient for analytical analysis as well as numerical simulation. However, it must be a long time later, Litva [62] discovered that this model produces erroneous results for intermodulation products in the two-tone test. Other studies further showed that this problem occurred particularly for the two-tone testing signal and did not occur with other practically-used signals [18].

Obviously, there has been an objective, urgent and practical requirement for the applicability of nonlinear HPA models satisfying not only good approximation to actual measurement data but also accurate simulation of nonlinear phenomena occurring with different testing signals. On the other hand, in the theoretical and systematic aspects, this requirement needs proper attention right from the beginning of a research on communication systems with nonlinear HPA distortions. Therefore, the remaining sections in this chapter will in turn proceed detailed analyses of arising problems as well as corresponding solutions for the HPA modeling complication.

2.2. Instantaneous nonlinear models

2.2.1. Cann original model

To represent a signal passing through a nonlinear semiconductor amplifier, in 1980, Cann [17] proposed an instantaneous nonlinear model with variable knee sharpness, particularly suitable for analyzing and simulating compres-

sors, expanders, and detectors

$$y = \frac{A_{\text{os}} \cdot \text{sgn}(x)}{\left[1 + \left(\frac{A_{\text{os}}}{g|x|}\right)^s\right]^{1/s}} = \frac{gx}{\left[1 + \left(\frac{g|x|}{A_{\text{os}}}\right)^s\right]^{1/s}}, \quad (2.1)$$

where x, y , in the full form of $x(t), y(t)$, are respectively the instantaneous input and output voltages (hence the name of this model) with the time variable t ; $\text{sgn}(\cdot)$ is the sign operator; g is the small-signal (linear) gain; A_{os} is the output saturation level; and s is the sharpness (smoothness) parameter. This model is usually regarded as one of the earliest nonlinear models for representing HPA [52].

However, until 1996, Litva [62] found that this model gave incorrect results for the third and higher odd order IMPs in the two-tone test. Four years later, Loyka [65] discovered that the reason is the use of modulus ($|\cdot|$) function in (2.1), some of whose derivatives at zero do not exist, are undefined, or are infinite. In other words, the function is non-analytic², despite the deceptively smooth appearance of the plotted curves.

Incidentally, in 1991, Rapp [82] introduced a complex envelope model³ for SSPAs that almost resembles to the Cann instantaneous model (2.1) except for the modulus operator in the denominator and the exponent of $2s$ instead of s . However, to the best of our knowledge, there are not so many discussions on the defect of this frequently used model. Detailed analysis on this topic will be given in subsection 2.3.3.

²A complex function is said to be analytic (or, equivalently, holomorphic) on a region R if it is complex differentiable at every point in R ; then this function is infinitely differentiable in R . This condition is more strict than the ordinary differentiable property: a differentiable function may not be analytic; however, the converse is always true. <http://mathworld.wolfram.com/AnalyticFunction.html>

³Hitherto, this is one of the most common models proposed in studies relating to nonlinearity due to SSPA.

2.2.2. Cann new model

Based on the magnitude Bode plot of a simple phase-lead network's transfer function in the form of $(1 + j\omega)/(a + j\omega)$ (with a constant a and an angular frequency variable ω), that is analytic and symmetric regarding to ω , Cann [18] suggested a new nonlinear instantaneous model in the scaled normalized form as

$$y = \frac{A_{os}}{s} \ln \frac{1 + e^{s(gx/A_{os}+1)}}{1 + e^{s(gx/A_{os}-1)}} - A_{os}, \quad (2.2)$$

with variables x, y and parameters g, A_{os}, s have the same meanings as what they are in the original model (2.1). It is not difficult to show that the derivatives of the new model (2.2) exist and well behave, even with a fractional s .

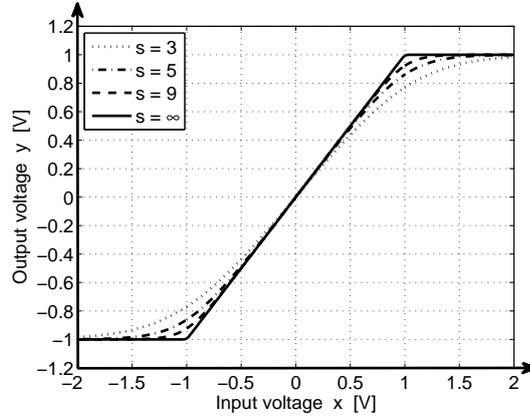


Figure 2.1: Characteristic functions of the Cann new model.

Figure 2.1 illustrates characteristics of this model with different sharpnesses s while other parameters are normalized, $g = 1$ and $A_{os} = 1$ [V]. Obviously, increasing s leads to a sharper knee, that effectively concentrates the nonlinearity to smaller region while further broadens the saturation range. Extremely, the case of infinity sharpness, $s = \infty$, usually regarded as *ideal*

limiter, has pointwise nonlinearity and is represented in the following form

$$y = \begin{cases} -1, & x < -1 \\ x, & -1 \leq x \leq 1 \\ 1, & x > 1. \end{cases} \quad (2.3)$$

This is the normalized soft limiter (1.4) with $A_{\text{is}} = 1$ for instantaneous signals.

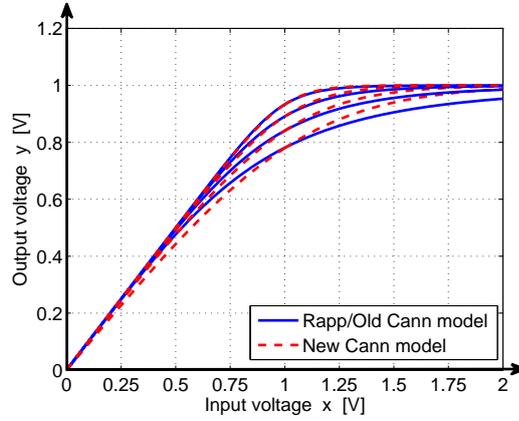


Figure 2.2: Characteristic functions of the Rapp/Cann original model (2.1) compared to that of the Cann new model (2.2).

The new model (2.2) is graphically compared to the old one (2.1) in Figure 2.2. Here, the sharpnesses of each model are chosen such that their corresponding curves intersect at the input value of 1 [V], which in some sense, is the center of these curves. Then, from bottom up, the curve sharpnesses respectively are, for Cann old model (2.1): $s = 2.8, 4, 6, 10$, corresponding to Rapp model (2.11): $s = 1.4, 2, 3, 5$, and for Cann new model (2.2): $s = 3.14, 4.35, 6.36, 10.35$. Surely, these models could fulfill any sharpnesses. Further, they almost coincide when s is large enough. However, though alike in the graphical forms, they behave quite differently in simulating IMPs. These are the non-analyticities yet having deceptive smooth appearances of

the Rapp and Cann old models that lead to the surprised fault of IMPs simulations. This problem will be demonstrated in section 2.4.

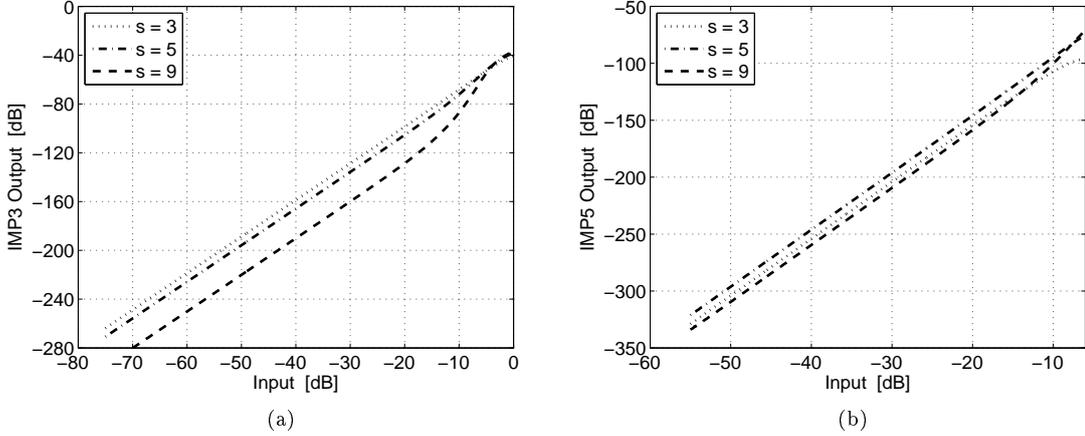


Figure 2.3: Third order (a) and fifth order (b) IMPs created by the Cann new model (2.2).

The reasonableness of third- and fifth-order IMPs for the two-tone test simulation using the Cann new model (2.2) is illustrated in Figures 2.3(a) and 2.3(b). Here, the sharpnesses s vary in a quite large range revealing the model's effectiveness. It is observed that these lines have the expected slopes as what happening in a real-world experiment [38, 64]: 3 dB/dB for third order (in sub-figure 2.3(a)) and 5 dB/dB for fifth order (in sub-figure 2.3(b)), respectively. Moreover, the IMPs' slopes remain the same for all sharpnesses. This confirms the suitability of Cann new model (2.2), yielding simulation results conforming to what happening in reality. Therefore, model (2.2) eliminates the shortcomings of previous one (2.1). This is the analyticity and symmetry of the original phase-lead network's transfer function to resolve the problem.

2.3. Envelope nonlinear models

2.3.1. Envelope representation of bandpass signals

Practically, to comply with spectral regulations (reconsider Figure 1.2, for example), a transmitter with nonlinear HPA often has a bandpass zonal filter that restricts the output to the first spectral zone, suppressing all harmonics and even-order IMPs. Such a system is referred as narrowband or bandpass, meaning that the bandwidth is considerably smaller than the center frequency⁴. This attribute allows huge saving of computation since the required sampling rate is then determined not by the highest frequency of the signal but by its bandwidth, of course plus a suitable redundance for significant IMPs representation and anti aliasing [52, 54]. The resulting model is the lowpass equivalent representation of the bandpass system and is regarded as *envelope model*.

A narrowband radio-frequency (RF) signal can be represented as

$$v(t) = r(t) \cos[\omega_0 t + \phi(t)] = \text{Re}[r(t)e^{j[\omega_0 t + \phi(t)]}], \quad (2.4)$$

where $r(t)$ is the amplitude modulation component, and $\phi(t)$ is the phase modulation component, both varying slowly regarding to the carrier with angular frequency ω_0 . When being observed in a reference plane rotating at this frequency, the resulting signal is *complex envelope* or *lowpass equivalent*⁵

$$x(t) = r(t)e^{j\phi(t)} = r(t) \cos \phi(t) + jr(t) \sin \phi(t). \quad (2.5)$$

It is noteworthy that in (2.5), the carrier ω_0 disappears but all modulating

⁴In practice, e.g., satellite communications, typical values of the RF signal bandwidth and its carrier frequency respectively are 40 MHz (36 MHz for channel bandwidth and 4 MHz for guard band) and 4 GHz (downlink C band). Hence, referring to baseband, the IMP spectrum will be negligible for $|f| > 100$ MHz, which satisfies this requirement by a large margin.

⁵For a lowpass signal the complex envelope of the signal is the signal itself. But for bandpass signal, the complex envelope representation allows us to easily separate out the carrier; hence is the name.

information (carried in both amplitude and phase) still exists. On the other hand, the complex amplitude signal $x(t)$ in (2.5) is represented by both polar and quadrature forms. All are useful for envelope models, of which the latter leads to the quadrature model, being especially suitable when the HPA nonlinearity characteristic has a phase shift component.

Reiterating that as described by (1.1), the envelope model is characterized by its complex transfer function $F(\cdot)$, including the AM-AM transfer function $F_a(r)$, output amplitude as a function of input amplitude, and the AM-PM transfer function $F_p(r)$, phase shift as a function of input amplitude, all for a single frequency signal. In case of the Cann new model, for example, Figure 2.4 illustrates the AM-AM functions $F_a(r)$ of this envelope model corresponding to the instantaneous model (2.2). These curves are amplitudes of the fundamental (first order) harmonic component of the output after passing a single-tone sine wave through the instantaneous model (2.2). Obviously, these characteristics are not simply the right halves of instantaneous model's curves depicted in Figure 2.1, but are totally different from them. For example, the ideal limiter's characteristic, where $s = \infty$, does not have a broken point at the input magnitude of 1 [V] and then saturating at a constant level. Instead of this, the output magnitude monotonically increases up to the asymptotic value of $4/\pi$. Its analytical expression is of the following form [52, 88]

$$F_a(r) = \begin{cases} r, & 0 \leq r < 1 \\ \frac{2}{\pi} \left[r \sin^{-1} \left(\frac{1}{r} \right) + \left(1 - \frac{1}{r^2} \right)^{1/2} \right], & r \geq 1. \end{cases} \quad (2.6)$$

Actually, there does not exist any practical HPA approaching this ideal

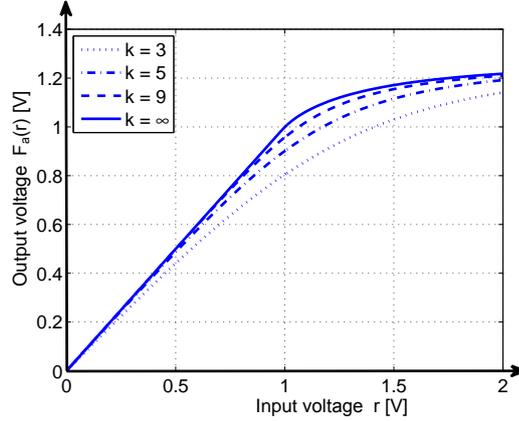


Figure 2.4: AM-AM functions of the Cann envelope model corresponding to the instantaneous model (2.2).

limiter's characteristic. However, there is an asymptotic exception: A well combination of a predistorter and an HPA allows larger linear amplification range and sharper knee before saturation [32, 100].

2.3.2. Saleh model

In 1981, Saleh, a researcher working at Bell Labs in Crawford Hill, New Jersey, United States, introduced a TWTA envelope model [84], which then has been widely used in the literature studying nonlinear distortion due to HPAs. This model includes both AM-AM and AM-PM functions, of which the former has typical turndown after saturation, in accordance with reality. These amplitude and phase conversion characteristics are given by

$$F_a(r) = \frac{\alpha_a r}{1 + \beta_a r^2}, \quad (2.7)$$

$$F_p(r) = \frac{\alpha_p r^2}{1 + \beta_p r^2}, \quad (2.8)$$

where r is the input amplitude, $F_a(r)$ is the output amplitude (voltage), $F_p(r)$ is the phase shift (degree/radian), α_a is the small-signal (linear) gain, together with $\beta_a, \alpha_p, \beta_p$ forming the specific shapes of amplitude and phase

conversion curves. They are respectively illustrated in sub-figures 2.5(a) and 2.5(b) with the linear gain and output saturation level normalized, i.e. $\alpha_a = 1$, $A_{os} = \alpha_a/2\sqrt{\beta_a} = 1$. For comparison purpose, these figures also include different AM-AM and AM-PM characteristics from other typical HPA models that are then discussed below.

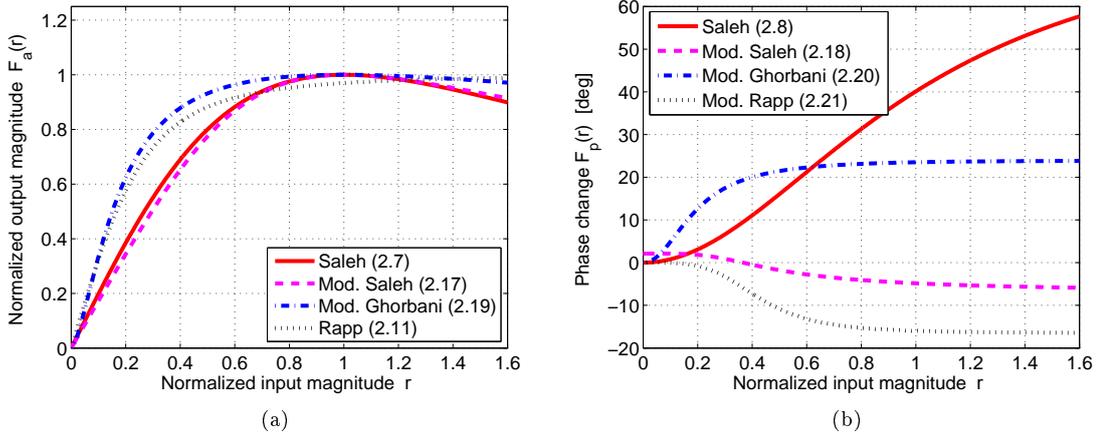


Figure 2.5: AM-AM (a) and AM-PM (b) functions of typical envelope models.

Saleh reminded that the amplitude r might be negative, thus, (2.7) must be an odd function. Moreover, this model does not support adjusting the knee sharpness of the AM-AM characteristic. Otherwise, the curvature of (2.7) is too smooth in regard to a typical SSPA amplitude characteristic, which also does not fall down after saturation⁶ [32].

Moreover, Saleh also proposed the quadrature form of this model, which is particularly suitable to characterize HPAs having additional AM-PM dis-

⁶Following the law of energy conservation and considering HPA as a closed system then it should observe the turndown region after saturation when further raising the input level since the output amplitude is constrained by limited DC supply voltage while the increased temperature (because of useful signal energy converted to heat) reduces the amplification efficiency. However, the input dynamic range and the overheating endurance capability of SSPA are much different (inferior) compared to that of TWTA [100]. Thus, it is not easy to observe the turndown region after saturation; SSPA might be destroyed before approaching saturation by overheating or input overvoltage.

tortion. This form is described by

$$I(r) = \frac{er}{1 + fr^2}, \quad (2.9)$$

$$Q(r) = \frac{gr^3}{(1 + hr^2)^2}, \quad (2.10)$$

where parameters e, f, g, h form the curve shapes. Also noting that (2.9) and (2.10) are both odd functions.

2.3.3. Rapp model

In 1991, when studying the impacts of nonlinear HPAs in digital audio broadcasting systems, Rapp proposed an envelope model with variable knee sharpness for SSPAs as follows [82]

$$F_a(r) = \frac{gr}{\left[1 + \left(\frac{gr}{A_{os}}\right)^{2s}\right]^{1/2s}}, \quad (2.11)$$

where r and $F_a(r)$ correspondingly are the input and output amplitudes, g is the small-signal (linear) gain, A_{os} is the output saturation level, and s is the curve's sharpness controlling the transition from the linear region to the limiting region. It is noteworthy that this model assumed zero AM-PM conversion and by changing the sharpness parameter s , the AM-AM characteristic could have any curvature. On the other hand, (2.11) is an odd function (Saleh condition) only for integer s .

Several examples of (2.11) with different knee sharpnesses s are illustrated in Figure 2.6 with normalized linear gain and output saturation level, i.e. $g = 1, A_{os} = 1$ [V]. In addition to this, the normalized characteristic curve of the ideal limiter (the upper most curve in Figure 2.4) is included for reference purpose. Noting again, this is an upper bound for any real-world amplifier

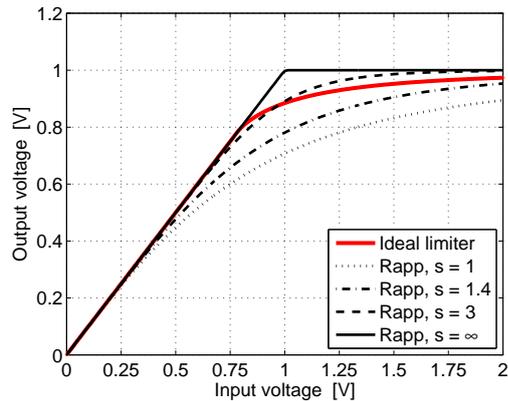


Figure 2.6: AM-AM functions of the Rapp model with different sharpnesses.

(with an asymptotic exception of ideal predistorter-amplifier combination [32, 100]).

Incidentally, this model resembles to the instantaneous model (2.1) excepting the absence of modulus operator ($|\cdot|$) in the denominator. Thus, it seems to avoid the problem of (2.1) for the suitability of IMPs resulted by simulation, but this is not the case. The Rapp’s model has been widely used for roughly a quarter of century without any considerations for its reasonableness and also its suspicious results until the publication of Cann [18].

Thorough investigation leads to the conclusion that the problem of (2.1) in simulating IMPs only manifests with signals whose magnitude distributions concentrating around zero, such as the signal used in the two-tone test (that will be adequately investigated in subsection 2.4.2). For real-world signals like M -FSK, M -PSK, M -QAM, M -APSK, OFDM, . . . the Rapp’s model behaves almost perfectly well. Therefore, resembling to the case of instantaneous models, all envelope AM-AM models should ideally be odd and analytic⁷ over the

⁷The analyticity requirement for both instantaneous and envelope models could be intuitively explained as follows. To yield precise IMP results from time-domain (instantaneous or envelope) signals being passed through a nonlinear model, then information related to all frequency components (physically, being rates of change over time) of these waveform signals should be uninterruptedly transferred through the model. This means the model’s transfer function should be “smooth” for all frequency components, or mathematically equivalent to being infinitely differentiable. When this condition is not satisfied, some frequency components will not be transferred through the model, directly affecting

expected amplitude range. An envelope model, which is asymmetric and is not analytic at zero, should be used with caution and only for signal waveforms that are sufficiently complex to have a widespread amplitude distribution with small possibility of zero amplitude. Thus, this problem, once having not been widely understood though being relatively important, is now clear. It is then clearer when considering the mutual effects of three factors: model's analyticity, data-approximation accuracy and signal's amplitude distribution in simulating IMPs. This is particularly discussed in section 2.4.

2.3.4. *Cann envelope model*

Although originally developed as an instantaneous model, (2.2) can be used equally as an envelope model. In this case, it is not the right halves of curves in Figure 2.4 but those in Figure 2.1. This model should find broad applications, particularly for SSPAs since, like Rapp model, it has adjustable knee sharpness and does not turn down after saturation. But, unlike the Rapp model, it is analytic everywhere and therefore valid for any signal waveform. Moreover, if the phase conversion is significant, an AM-PM function, like Saleh's (2.8), could be included.

Similar to the Rapp model (2.11), envelope model (2.2) could support any curvatures, especially in the region above $s = 2.5$, suitable for AM-AM characteristics of most SSPAs [32]. The approximations of the Cann new model (2.2) and Rapp model (2.11) to the real-world data are verified by curve fitting of these functions to the measured data from the L band Quasonix 10W amplifier [86]. The results are, for Rapp model: $g = 29.4$, $A_{os} = 30$ [V], $s = 4.15$, Squared Error Sum (SES) $\sigma_e^2 = 0.963$; for Cann

the IMP values estimated at the output.

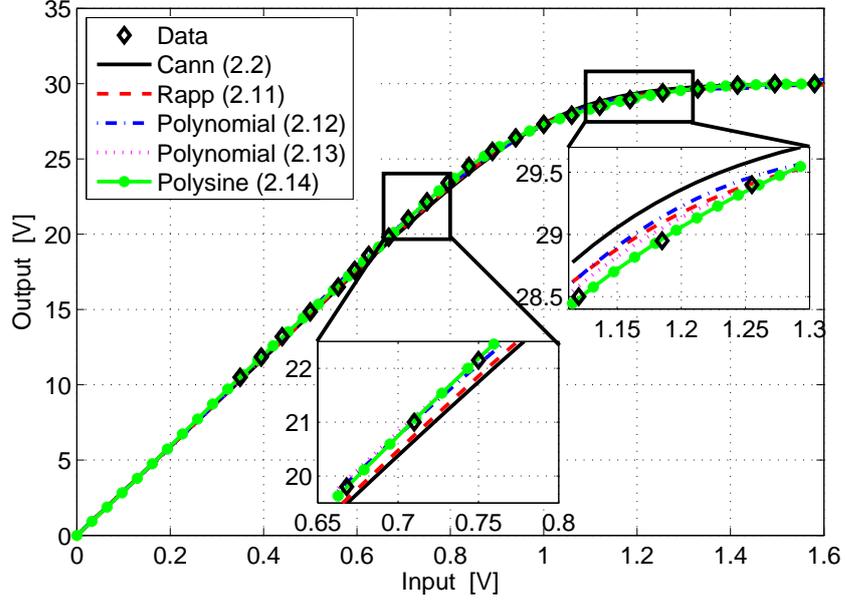


Figure 2.7: AM-AM functions of the Cann, Rapp, polynomial, odd-order polynomial and polysine models fitted to the measured data.

new model: $g = 29.4$, $A_{os} = 30$ [V], $s = 8.9$ and $\sigma_e^2 = 1.786$. For this particular HPA, Rapp model is little better fitted than Cann model. Figure 2.7 illustrates these fittings with the inclusion of other approximated curves discussed next.

2.3.5. Polynomial model

Considering the measured data in Figure 2.7, it is not difficult to recognize that there is a simple yet efficient method approaching the closed-form AM-AM function by approximation using polynomials. In this case, the complex envelope nonlinearity, $F(\cdot)$, can be represented by a complex polynomial power series of a finite order N such that [32]:

$$y = F(x) = \sum_{n=1}^N a_n |x|^{n-1} x = \sum_{n=1}^N a_n \Psi_n^P[x], \quad (2.12)$$

Table 2.1: Coefficients of the polynomial models (2.12), (2.13).

Model	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
(2.12)	30.02	-8.665	33.68	-40.19	12.39	0	0	0	0
(2.13)	28.60	0	8.310	0	-15.06	0	6.257	0	-0.872

where $\Psi_n^P[x] = |x|^{n-1}x$ are the basis functions of the polynomial model, and a_n are the model's complex coefficients of order n , $n = 1, 2, \dots, N$.

Obviously, model (2.12) is not analytic at $r = |x| = 0$ by the existence of modulus operators ($|\cdot|$). However, if even order coefficients a_{2n} vanish, then for real-valued $x(t)$, (2.12) turns into the odd-order polynomial model of the form

$$y = \sum_{n=1}^N a_{2n-1} |x|^{2(n-1)} x = \sum_{n=1}^N a_{2n-1} x^{2n-1}. \quad (2.13)$$

Model (2.13) is clearly analytic at $r = 0$ and is used as a counter example to model (2.12), showing that though having almost similar structure, they give quite different results. This is discussed in the section 2.4 below. The measured data of the L band Quasonix 10W amplifier [86] is then used to fit the polynomial models (2.12) and (2.13) with the same number of coefficients $N = 5$. Figure 2.7 depicts the approximated characteristics with their corresponding parameters listed in Table 2.1.

It is not difficult to show that at large enough order, polynomial models are better fitted to the real-world data than Rapp model (2.11) and Cannon model (2.2). Moreover, with the same N , odd-order polynomial of (2.13) is obviously smoother than full order one of (2.12) resulting in better fitting performance for the former. However, it is also worthy of noting that using too high order for polynomial fitting leads to the Runge phenomenon [101],

which also does increase the approximating errors.

2.3.6. Proposed polysine model

It can be seen that the sine/cosine functions are distinctly better than polynomial ones in terms of both analyticity and smoothness. Thus, while remaining to be analytic, the former are better fitted to the real-world data than the latter. Based on this fact, we propose a nonlinear model of the form

$$y = \sum_{n=1}^N a_n \sin(b_n x), \quad (2.14)$$

where a_n and b_n are respectively amplitude and phase coefficients.

Intuitively, the introduction of b_n allows (2.14) better adapting to the data variations, thus further improves the approximation performance. On the other hand, mathematically, it is not too difficult to recognize that polysine model is equivalent to odd-order polynomial model. Definitely, applying the Taylor expansion at $x = 0$, function $\sin(x)$ could be decomposed as [41]

$$\sin(x) = \sum_{m=1}^M (-1)^{m-1} \frac{x^{2m-1}}{(2m-1)!} + \mathcal{O}(x^{2M-1}), \quad (2.15)$$

where $(k)!$ is the factorial of order k , $\mathcal{O}(x^{2M-1})$ is the order of approximation error. Then, (2.14) is recast as

$$\begin{aligned} y &= \sum_{n=1}^N a_n \sin(b_n x) \\ &= \sum_{n=1}^N a_n \left(\sum_{m=1}^M (-1)^{m-1} \frac{(b_n x)^{2m-1}}{(2m-1)!} + \mathcal{O}(x^{2M-1}) \right) \\ &= \sum_{m=1}^M \left(\sum_{n=1}^N (-1)^{m-1} \frac{a_n (b_n)^{2m-1}}{(2m-1)!} \right) x^{2m-1} + \mathcal{O}(x^{2M-1}) \\ &= \sum_{m=1}^M a'_m x^{2m-1} + \mathcal{O}(x^{2M-1}). \end{aligned} \quad (2.16)$$

This is of the form (2.13) plus approximation error polynomial of order $2M - 1$, $\mathcal{O}(x^{2M-1})$. Thus, mathematically, odd-order polynomial model is equivalent to polysine model. However, technically, it is the quantity of $\mathcal{O}(x^{2M-1})$ that allows polysine model being always better fitted to data compared to odd-order polynomial model (2.13) and thus also being superior to full order polynomial (2.12) as analyzed above. Besides the analyticity, the requirement of small approximation error in modeling becomes truly critical when the model's output is used to evaluate high-order harmonics and intermodulation products, which commonly are quantities much smaller than the fundamental harmonics/components. Obviously, model with lower approximation error yields more precise results. Demonstrating examples in section 2.4 could clarify this fact.

Table 2.2: Coefficients of the polysine model (2.14).

n	1	2	3	4	5
a_n	30.73	-0.6586	-0.1061	0.00955	0.1859
b_n	1.045	5.312	12.91	18.61	8.107

Using the Matlab curve fitting tool, (2.14) is fixed to the AM-AM characteristic of the L band Quasonix 10W amplifier data in Figure 2.7 resulting in parameters listed in Table 2.2. The fitting performances of these five models are quantified using Squared-Error Sum (SES) measure and are compared in Table 2.3. The odd-order polynomial model (2.13) and the polysine model (2.14) are both analytic and much better fitted to the real data than the Cann model (2.2). This is illustrated in Figure 2.7 with sub-figures focusing on segments with significant differences where the data are rather harder to

Table 2.3: Approximation performance of five models (SES σ_e^2).

Model	Cann (2.2)	Rapp (2.11)	Polynomial (2.12)	Polynomial (2.13)	Polysine (2.14)
SES	1.786	0.963	0.533	0.346	0.032

fit. The better fitting performance is the closer to the data these curves approach. With almost one order of magnitude better in SES than the rest, the polysine model's curve always coincide with all data points. The fitting performance of these models will be reflected in the nonlinearity simulation results discussed in section 2.4 below.

2.3.7. Other conventional HPA models

Besides the AM-AM characteristic, updated envelope models for SSPAs at higher frequencies (of order GHz) and larger bandwidth (of order tens MHz) all consider the AM-PM conversion and generally better fit to the measured data than the previous models. However, it is not difficult to see that all models discussed below are not analytic or symmetric at $r = 0$ for most of the parameter sets and thus problem of (2.11) still exists. The AM-AM and AM-PM characteristics of these models are graphically illustrated in sub-figures 2.5(a) and 2.5(b) for comparison purpose.

- *Modified Saleh model:* [72] is proposed for popular laterally diffused metal oxide semiconductor (LDMOS) HPAs, that are very common for the base station (BS) amplifiers of 2G, 3G and 4G mobile networks (in the L, S,

C bands). The AM-AM and AM-PM conversion functions are

$$F_a(r) = \frac{\alpha_a r}{\sqrt{1 + \beta_a r^3}}, \quad (2.17)$$

$$F_p(r) = \frac{\alpha_p}{\sqrt[3]{1 + r^4}} - \varepsilon_p, \quad (2.18)$$

where $\alpha_a = 1.0536$, $\beta_a = 0.086$, $\alpha_p = 0.161$, and $\varepsilon_p = 0.124$ is a typical parameter set.

- *Modified Ghorbani model*: [6] is suited for GaAs pHEMT (Gallium arsenide pseudomorphic High-electron-mobility transistor) HPAs that operate at frequencies up to 28 GHz (K band, largely proposed for the 5G millimeter wave cells [96]) and are dominant in terms of production technologies and market shares compared to other power semiconductor technologies. This model assumes the following characteristics

$$F_a(r) = \frac{\alpha_1 r^{\alpha_2} + \alpha_3 r^{\alpha_2+1}}{1 + \alpha_4 r^{\alpha_2}}, \quad (2.19)$$

$$F_p(r) = \frac{\beta_1 r^{\beta_2} + \beta_3 r^{\beta_2+1}}{1 + \beta_4 r^{\beta_2}}, \quad (2.20)$$

where the model parameters are given by $\alpha_1 = 7.851$, $\alpha_2 = 1.5388$, $\alpha_3 = -0.4511$, $\alpha_4 = 6.3531$, $\beta_1 = 4.6388$, $\beta_2 = 2.0949$, $\beta_3 = -0.0325$, $\beta_4 = 10.8217$.

- *Modified Rapp model*: [21] is introduced for GaAs pHEMT/CMOS (Complementary metal-oxide-semiconductor) HPA model operating at 60 GHz band, the new band for communication industry, with AM-AM function of (2.11) and AM-PM described as

$$F_p(r) = \frac{\alpha r^{q_1}}{\left(1 + \left(\frac{r}{\beta}\right)^{q_2}\right)}, \quad (2.21)$$

where the parameter set are $g = 16$, $A_{\text{os}} = 1.9$, $s = 1.1$, $\alpha = -345$, $\beta = 0.17$, $q_1 = q_2 = 4$.

2.4. Applications of HPA models in communication simulation

This section describes the applications of envelope models investigated above for representing nonlinear HPAs in communication systems and analyses typical experiments with testing signals having discrete and continuous spectra to reveal the applicabilities and reasonableness of these models.

2.4.1. Representation of envelope models

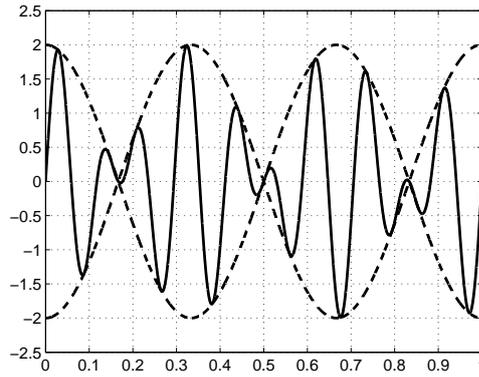


Figure 2.8: Two-tone waveform, $f_1 = 7$ [Hz], $f_2 = 10$ [Hz].

Consider the finding of IMPs in a two-tone test with a signal consisting of two equal-amplitude ($A_0/2$) unmodulated sinusoid waveforms at frequencies f_1 and $f_2 > f_1$. This (instantaneous) testing signal could be equivalently regarded as a double-sideband suppressed carrier AM of the form

$$\begin{aligned} x_{\text{inst}}(t) &= \frac{1}{2}A_0[\sin(2\pi f_1 t) + \sin(2\pi f_2 t)] \\ &= A_0 \cos(2\pi f_m t) \sin(2\pi f_c t), \end{aligned} \quad (2.22)$$

where $f_m = \frac{1}{2}(f_2 - f_1)$ is the modulating frequency, and $f_c = \frac{1}{2}(f_2 + f_1)$ is the (center) carrier frequency. Waveform (2.22) with $f_1 = 7$ [Hz], $f_2 = 10$ [Hz] is

illustrated in Figure 2.8. It is observed that the carrier f_c (solid line) manifests inside the envelope and is the average of f_1 and f_2 , while the envelope (dashed line) is the modulating signal at frequency f_m .

For representation convenience, with the 90° phase shifting of the modulating envelope component, (2.22) could be recast in the form of

$$x_{\text{inst}}(t) = A_0 \sin(2\pi f_m t) \sin(2\pi f_c t). \quad (2.23)$$

Therefore, its envelope form is

$$x_{\text{env}}(t) = A_0 \sin(2\pi f_m t). \quad (2.24)$$

Since the envelope model requires non-negative input, thus, the sinusoid waveform of (2.24) is decomposed into the polar form as

$$x_{\text{env}}(t) = r(t)e^{j\phi(t)} = A_0 |\sin(2\pi f_m t)| e^{j\phi(t)}, \quad (2.25)$$

where

$$r(t) = A_0 |\sin(2\pi f_m t)|,$$

$$e^{j\phi(t)} = \begin{cases} 0, & \sin(2\pi f_m t) \geq 0 \\ \pi, & \sin(2\pi f_m t) < 0. \end{cases}$$

In other words, the amplitude component $r(t)$ is the full-wave-rectified sinusoid, and the phase component $\phi(t)$ is the 180° square wave.

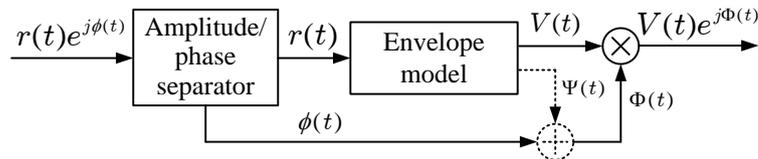


Figure 2.9: Polar envelope model block diagram [52].

For simulating with envelope model in a general case of arbitrary $r(t)$ and $\phi(t)$, the amplitude component $r(t)$ is input to the model, while the

phase component $\phi(t)$ is bypassed as depicted in Figure 2.9. The distorted amplitude output $V(t)$ is then combined with the phase part, resulting the output waveform for analysis. If AM-PM conversion is included, then the distorted phase $\Psi(t)$ is added up to the input phase $\phi(t)$ resulting in output phase $\Phi(t)$, that finally is combined with $V(t)$ in the polar form.

2.4.2. Simulation with two-tone testing signal

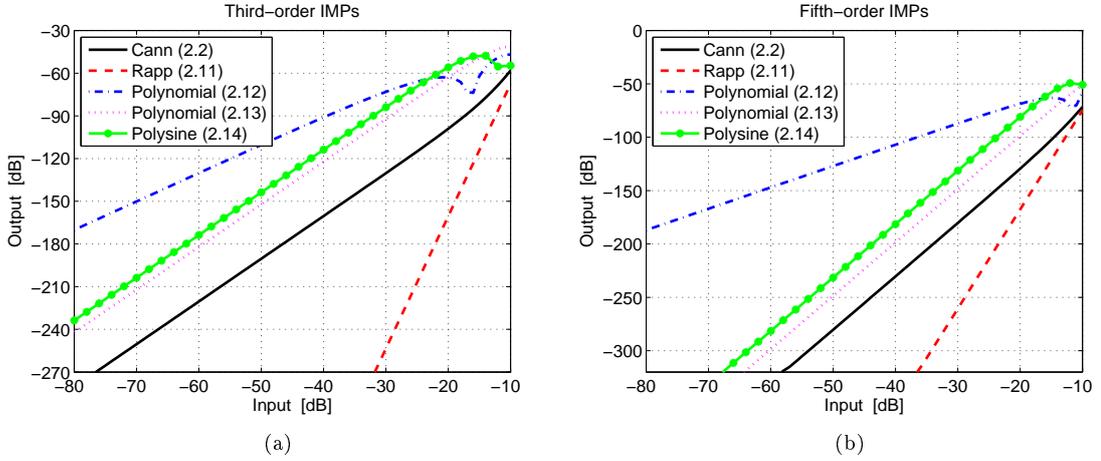


Figure 2.10: Third order (a) and fifth order (b) IMPs of five models in Figure 2.7.

Simulation procedure is depicted in Figure 2.9 with the following parameters: simulation time 1 [s], sampling rate 1000 [Hz], input signal waveform as in Figure 2.8, five models with their corresponding characteristics depicted in Figure 2.7 are considered. Output signals will be used for IMPs analysis.

The third- and fifth-order IMPs (IMP3/5) are respectively shown in sub-figures 2.10(a) and 2.10(b). As observed, Cann new model (2.2), odd order polynomial model (2.13) and polysine model (2.14) result in the required slope of 3 [dB/dB] and 5 [dB/dB] correspondingly for IMP3 and IMP5 as expected [38, 64]. With almost the same structure as (2.13), however, the full order polynomial model (2.12) fails in simulating the odd IMPs, revealing

the problem as found by Litva in [62] for Cann old model (2.1). So does the Rapp model. In these cases, full order polynomial model results in slope of 2 [dB/dB] for both IMP3 and IMP5, while this quantity is 9 [dB/dB] for Rapp model. They are obviously irrelevant and quite different. Surely, those are the differences in non-analyticities of these models that create the results.

It can be seen that IMPs are much smaller in magnitude than the fundamental component (input) signal. Further, for three analytic models, there are different constant gaps between IMPs created. Obviously, smaller error in fitting approximation should result in better performance of simulation, particularly for small-value components. Thus, Cann new model (2.2) produces less confident results than what created by odd-order polynomial model (2.13) and especially by polysine model (2.14).

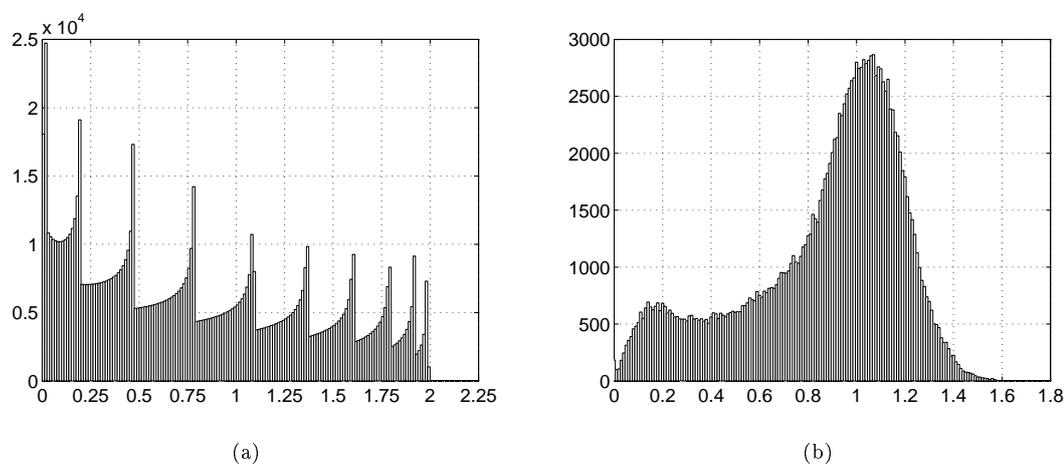


Figure 2.11: Amplitude histograms of two-tone (a) and 1+7-APSK (b) testing signals.

Reconsider the signal processing depicted in Figure 2.9, it is recognized that the amplitude/phase separator indirectly yields the modulus operation, causing the former problem. Thus, to receive reasonable results for the two-tone test, the envelope model should be analytic at $r = 0$, as the same as

what found by Loyka [65] for the instantaneous model.

To clearly see the defect of Rapp model (2.11) and polynomial model (2.12) under the effect of the signal amplitude distribution to the IMPs created, consider the amplitude histogram of this two-tone testing signal illustrated in sub-figure 2.11(a). It is undoubtedly inferred that, for non-analytic models, the very high concentration of signal amplitude around $r = 0$ results in irrelevant model outputs, of which some or many frequency components disappeared. This finally yields the failure in calculating IMPs as illustrated in sub-figures 2.10(a) and 2.10(b).

2.4.3. Simulation with continuous-spectrum testing signal

Consider a real-world updated signal as the input for such models investigated above. Amplitude-phase shift keying (APSK) is commonly used for communication systems with considerations in spectral and power efficiencies [3, 22]. 1 + 7 APSK has recently been introduced as an efficient modulation scheme for satellite communications [28]. The signal constellation of this modulation scheme includes one signal point at the origin ($r = 0$) and seven others evenly distributed in a circle. Deceptively according to the above argument about the effects of model's analyticity to the simulated IMPs, the simulation test with this 1 + 7 APSK input signal could result in the evident fail of models (2.11) and (2.12). However, the outcome is more complicated.

With the inclusion of transmit shaping filter and receive matched filter⁸, the testing signal now has a continuous spectrum with the amplitude distribution in the form of histogram depicted in sub-figure 2.11(b). It is clearly

⁸Both are SRRC filters with roll-off factor $\alpha = 0.2$, input sampling rate $F_d = 1$, output sampling rate $F_s = 16$, for reasonably representing harmonics up to 7-th order.

seen that there is so less concentration at the origin coordinates $r = 0$, totally different to the counterpart of two-tone waveform amplitude distribution in sub-figure 2.11(a). This somehow relieves the defect of non-analytic models (2.11)) and (2.12) investigated in the previous section.

Applying this testing signal to five systems with their corresponding HPA models used in the previous section, the output signals are then analysed showing the spectrum regrowth. Sub-figure 2.12(a) illustrates the receive constellations resulted from the use of Cann (2.2) model and odd order polynomial model (2.13), manifesting a relatively strong effect of nonlinearity created by HPAs. Sub-figure 2.12(b) depicts the spectral regrowths corresponding to all five models.

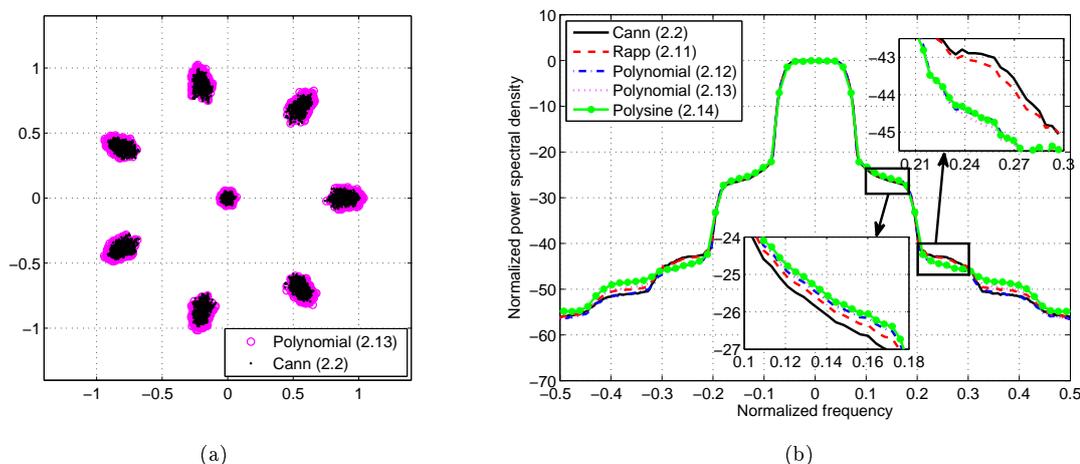


Figure 2.12: Receive constellations (a) and spectra (b) created from 1+7-APSK testing signal with different nonlinear models.

Obviously, at high magnitude levels of spectra in the main lobe, there is almost no difference in results from all models, both analytic and non-analytic ones. However, for lower magnitude levels in adjacent sidelobes (spectral regrowth zones), there are divergences for the third- and fifth-order spectrum regrowths. This gap is up about 0.5 dB between the curve yielded by Cann

model and that of polysine model in the first sidelobe (IMP3) and is up about 2 dB in the second sidelobe (IMP5). The differences further increase for higher-order IMPs. Insignificant discrepancy between results by odd-order polynomial model and by polysine model reveals defects of new Cann and Rapp models: either non-analyticity or larger data-approximation error could result in less accurate or even irrelevant IMPs.

Reconsider the amplitude histogram of testing-signal in sub-figure 2.11(b), then for non-analytic models, a relatively small possibility of signal with zero amplitude would still result in significant errors for lower level components (high-order IMPs) of the output. This will be important for the case of verifying the conformity to spectral mask standard as discussed in Chapter 1.

2.5. Summary of chapter 2

Through analyses and verification with numerical and visual simulation results, several important problems related to HPA modeling have been elucidated. It is now clear that nonlinear HPA models satisfying analyticity and well-matching to measurements can ensure reliable simulation results, especially for newly proposed signals with complex structures.

The results achieved in this chapter:

- Analyzing the causes of defects in simulating inter-modulation products for commonly used HPA envelope models including polynomial, Saleh, Rapp, modified Saleh, modified Ghorbani, and modified Rapp;
- Verifying the validity and showing the limits due to poor approximation of the new Cann model (2.2) by simulation experiment using both discrete-spectrum and continuous-spectrum testing signals;

- Showing that the odd-order polynomial model (2.13) can replace the new Cann model with the advantages of simpler structure and higher accuracy in approximating to measurement data;
- Proposing polysine model with higher approximation accuracy, well suitable for evaluating the spectral regrowth for testing signals with complex structure.

Chapter 3

Predistortion Methods for Nonlinear Distortions due to HPAs in MIMO-STBC Systems

Based on the detailed investigations of nonlinear HPA models in the previous part, this chapter applies them to the MIMO-STBC system. Previous publications such as [73, 74, 81, 94] have mentioned the predistortion method for nonlinear MIMO-STBC systems. However, these works all use discrete models, without considering transmit/receive filters; and therefore, as preliminarily analyzed in section 1.6, some nonlinear impacts have not been taken into account. This creates a gap between the analysed results and practice. Therefore, in this chapter, we investigate in detail the impacts of nonlinear HPAs on MIMO-STBC system with a model including typical transmit/receive filters and typical nonlinearities having both AM-AM and AM-PM characteristics. Thereby, limitations in the approaches of previous works like [81] with the assumptions of a discrete system model and simplified HPA (having only AM-AM distortion) are indicated. Also through the detailed analyses of the nonlinear HPA effects on the system, four predistortion schemes are applied, allowing to efficiently minimize these adverse effects. The content of this study relates to publication 3 in the List of Publications.

3.1. Overview

With the actual nonlinearity characteristics that cannot be ignored when high power efficiency is required, HPA causes many adverse effects and is still a problem not easily solved, particularly for satellite communication systems [22]. It is the diversity of HPA nonlinear characteristics as well as the way these nonlinearities affect the system that cause many difficulties for modeling and analyzing their impacts on the system. Therefore, specific studies often try to greatly simplify nonlinear characteristics when introduced in the system and/or to maximally simplify the system modeling with nonlinear distortion(s) to facilitate analytical analysis. This creates a (large) gap between the proposed model and practice and leads to less reliable results.

In fact, on the transmitter side, the pulse-shaping filter, in the most common form of square root raised cosine (SRRC) filter, is installed before the HPA and thus creates memory in the HPA input. As mentioned in Chapter 1, for single carrier SISO systems using non-constant envelope signals, the (memory) HPA nonlinearities cause constellation warping, spectrum regrowth and nonlinear ISI. Furthermore, MIMO systems, in general, when operating with nonlinear HPA are subjected to additional nonlinear inter-channel interference (ICI) [34].

To limit the effects of nonlinear distortion, one of the effective measures to be mentioned is the predistortion (PD), which has been applied quite intensively in SISO systems initially for single-carrier [22] and then for multi-carrier [97]. However, research and development of this technique for MIMO-STBC systems has not been thoroughly discussed. In [94], predistortion using

adaptive polynomial was applied while [81] analyzed performance of MIMO-STBC system under the effect of simplified nonlinear distortion (including only AM-AM characteristic) and proposed a distortion compensation method at the receiver by determining the optimal decision regions with assumption of knowing the HPA model parameters. More recently, Oussama et al in [73, 74] qualitatively analyzed (capacity, symbol error rate, error vector module) of the MIMO-STBC system with distortion compensation by nonlinear equalization at the receiver side designed using a separate, reliable link to inform HPA parameters “learned” by a neural network installed in the transmitter.

However, all considered system models do not include transmit/receive filters, and therefore only memoryless nonlinear effects are considered and resolved. As preliminarily discussed earlier in Chapter 1 and will be more clearly seen in detailed analyses next, memory (appears under the effects of transmit/receive filtering) when combined with HPA nonlinear distortion will cause serious adverse effects to the system performance and therefore should be given proper attention.

Driven by the aforementioned factors, in this chapter, the thesis will carry out a thorough analysis for the MIMO-STBC system incurred memory nonlinear distortions. Standard transmit/receive filters and typical HPAs are introduced into the system model, allowing to further approach to practice than previously mentioned works. On the other hand, with the introduction of typical nonlinear characteristics including both AM-AM and AM-PM conversions, nonlinear distortion effects are fully expressed, from which limitations in previous publications are figured out. Then, four predistortion schemes will be analyzed and applied to the system. Simulation results for error vec-

tor module (EVM), modulation error ratio (MER), bit error rate (BER) will be used to verify effectiveness of the proposed schemes as well as to assess the system performance with predistortion applied.

3.2. Nonlinear distortion effects in MIMO-STBC systems

3.2.1. MIMO-STBC $2 \times n_R$ system model

Figure 3.1 describes the baseband-equivalent MIMO-STBC system model using the Alamouti's coding [7], $n_T = 2$ transmit antennas, and n_R receive antennas.

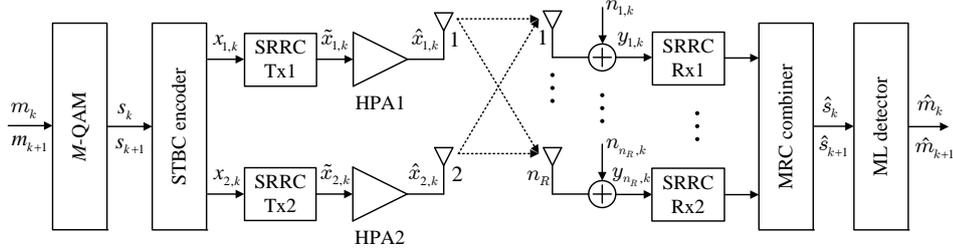


Figure 3.1: MIMO-STBC system model with transmit/receive filters and nonlinear HPAs.

Assume that the M -QAM modulator inputs M -level messages m_k, m_{k+1} and outputs complex signals s_k, s_{k+1} . These signals are then processed by the STBC coder resulting in the coded matrix \mathbf{X}_k of the form

$$\mathbf{X}_k = \begin{bmatrix} x_{1,k} & x_{1,k+1} \\ x_{2,k} & x_{2,k+1} \end{bmatrix} = \begin{bmatrix} s_k & -s_{k+1}^* \\ s_{k+1} & s_k^* \end{bmatrix}, \quad (3.1)$$

where columns of \mathbf{X}_k correspond to transmit time slots $k, k+1$ and rows of \mathbf{X}_k correspond to the first and second transmit antennas. Also assume that all SRRC pulse-shaping filters at transmitter have the same continuous-time

impulse response $h(t)$ with roll-off factor $\alpha \in [0,1]$, given by [33]

$$h(t) = \frac{1}{\sqrt{T}} \frac{\sin \left[\pi \frac{t}{T} (1 - \alpha) \right] + 4\alpha \frac{t}{T} \cos \left[\pi \frac{t}{T} (1 + \alpha) \right]}{\pi \frac{t}{T} \left[1 - \left(4\alpha \frac{t}{T} \right)^2 \right]}, \quad (3.2)$$

where T is the period of input signal x and $t \in (-\infty, \infty)$. In digital format, these filters are usually configured as symmetric finite impulse response (FIR) filters with odd length N , input sampling frequency $F_d = 1/T$, and output sampling frequency $F_s = M_{\text{ovs}} F_d$, M_{ovs} is the oversampling factor, which is an integer greater than 1. The non-causal discrete-time response of (3.2) can be drawn after replacing¹

$$h[n] = h(t)|_{t=n\frac{T}{M}}, \quad -\frac{N-1}{2} \leq n \leq \frac{N-1}{2}. \quad (3.3)$$

At the k -th time slot, the output of the i -th filter, $i = 1, 2$, thus is of the form

$$\tilde{x}_{i,k} = x_{i,k} h[0] + \sum_{l \neq k, -\frac{N-1}{2} \leq l \leq \frac{N-1}{2}} x_{i,l} h[k-l] = x_{i,k} h[0] + n_{i,k}^{\text{ISI}}. \quad (3.4)$$

Therefore, apart from the desired component $x_{i,k} h[0]$, signals inputted to the HPA also contain the ISI term $n_{i,k}^{\text{ISI}}$ resulted by SRRC filter's memory. For linear system, this ISI will be totally canceled out at the receive matched filter. However, in system with nonlinearity as being considered, both desired and ISI components will be nonlinearly amplified when passing through HPA and then resulting in nonlinear ISI at the output.

As discussed in Chapter 2, high power amplifiers could be fully described by input-output AM-AM and AM-PM conversions, that respectively represent the input-output amplitude and phase distortions. However, for the purposed of comparison with previous publications, without considering the

¹Practically, the causal discrete-time form of (3.2) is given by $h[n] = h(t)|_{t=(n-\frac{N-1}{2})\frac{T}{M}}$, $0 \leq n \leq N-1$.

spectral regrowth effect as analysed in Chapter 2, here we use conventional HPA models with typical nonlinearity characteristics. These include TWTA with Saleh model and SSPA with Rapp model. The former one has the corresponding AM-AM and AM-PM functions of (2.7), and (2.8) where, α_a , β_a and α_p , β_p are model parameter set acquired by curve fitting from real-world data measurements. The Rapp model only includes AM-AM conversion (2.11), where, g is the small-signal (linear) gain, A_{os} is the output saturation amplitude and $s > 0$ is the sharpness factor.

As being observed from Figure 2.6, for s large enough (2.11) approximates the soft envelope limiter (1.4). SEL could be used to model HPA with ideal PD. The overall transfer function of the PD-HPA system, then has linear form for the input range before saturation, and output distortion (constant amplitude) occurs when the input level is higher than this level. So, regarding to the performance, SEL could be considered as a reference to compare with other distortion algorithms, that are discussed further in the following sections.

For Saleh model (2.7), it is not difficult to determine that the input saturation level is $A_{is} = 1/\sqrt{\beta_a}$ and the output saturation level is $A_{os} = \alpha_a/2\sqrt{\beta_a}$. In practice, the HPA's operating point is set below this (input/output) saturation level and is regarded as (input/output) back-off. The input backoff (IBO) is defined as

$$IBO = 10\log_{10} \frac{P_{is}}{P_{im}}, \quad (3.5)$$

where $P_{is} = A_{is}^2$ is the input saturation power, and P_{im} is the input mean power. These quantities are illustrated in Figure 1.4

3.2.2. Nonlinear distortion effects incurred by HPAs

Assume that the fading is quasi-static (channel state unchanges over each transmit frame), and is known perfectly at receiver. Moreover, for simplifying representation without loss of generality, the time slot index k will be omitted in the following formulas. Then, the received vector \mathbf{y} at this time slot is of the form

$$\mathbf{y} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{n}, \quad (3.6)$$

where $\hat{\mathbf{x}} = [\hat{x}_{1,k}, \hat{x}_{2,k}]^T \in \mathbb{C}^{2 \times 1}$ is the transmit signal vector (HPA output), $\mathbf{y} = [y_{1,k}, y_{2,k}, \dots, y_{n_R,k}]^T \in \mathbb{C}^{n_R \times 1}$ is the receive signal vector, $\mathbf{H} = [h_{m,l}]^{n_R, 2} \in \mathbb{C}^{n_R \times 2}$ represents channel matrix $n_R \times 2$ with entry $h_{m,l}$ being the channel coefficient between the l -th transmit antenna và the m -th receive antenna, \mathbf{n} is the proper complex additive white Gaussian noise (AWGN) vector $\mathbf{n} = [n_{1,k}, n_{2,k}, \dots, n_{n_R,k}]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{N}_0)$, with mean vector $\mathbf{0}$ and covariance matrix \mathbf{N}_0 .

Signal on each receive branch is then passed through the SRRC matched filter (similar to the transmit one). To focus on the effects of nonlinear distortion existing in this signal, the input-output relation of the HPA on the first transmit branch, between \tilde{s}_k and \hat{s}_k , in time slot k (similarly for the second branch) is recast using the linearized model (1.3) as

$$\hat{s}_k = \gamma_k \tilde{s}_k + d_k, \quad (3.7)$$

where γ_k is a constant determined by the HPA's AM-AM characteristic and the input power, $\gamma_k = \mathbb{E}[\tilde{s}_k^* \hat{s}_k] / \mathbb{E}[\tilde{s}_k^* \tilde{s}_k]$, d_k represents uncorrelated (to the input \tilde{s}_k) nonlinear noise with zero mean and variance given by $\sigma_{d_k}^2 = \mathbb{E}[\hat{s}_k^* \hat{s}_k] - \gamma_k^2 \mathbb{E}[\tilde{s}_k^* \tilde{s}_k]$.

Receive signals on the l -th branch, $1 \leq l \leq n_R$, i.e. the l -th equation of (3.6) in the time slots k -th and $(k+1)$ -th are given by

$$y_{l,k} = h_{l,1}\hat{s}_k + h_{l,2}\hat{s}_{k+1} + n_{l,k}, \quad (3.8)$$

$$y_{l,k+1} = -h_{l,1}\hat{s}_{k+1}^* + h_{l,2}\hat{s}_k^* + n_{l,k+1}. \quad (3.9)$$

Substitute (3.7) into (3.8) and (3.9) then reconsider (3.4), it is not difficult to realized the appearance of nonlinear noise (d_k, d_{k+1}) and nonlinear ISI $(F(n_k^{\text{ISI}}), F(n_{k+1}^{\text{ISI}}))$ existing in the receive signals at time slots k -th and $(k+1)$ -th.

Applying maximum ratio combining (MRC) [7], in the l -th branch, receive signals are processed as

$$\bar{s}_{l,k} = h_{l,1}^*y_{l,k} + h_{l,2}y_{l,k+1}^*, \quad (3.10)$$

$$\bar{s}_{l,k+1} = h_{l,2}^*y_{l,k} - h_{l,1}y_{l,k+1}^*. \quad (3.11)$$

Apply the same analysis as for (3.8) and (3.9), it can be seen that the signals after MRC combining are affected by both nonlinear noise, nonlinear ISI and nonlinear ICI since the STBC code has lost its orthogonality after passing through HPAs as described by (1.8)². These effects are illustrated in sub-figure 3.2(a), which is the simulation result with basic parameters: 16-QAM modulation; transmit/receive filters with rolloff factor $\alpha = 0.2$, input sampling rate F_d equals to the 16-QAM symbol rate, output sampling rate $F_s = 16F_d$ (for representing harmonics up to 7-th order), group delay $Dl = 10$ (for representing ISI statistics); HPA represented by Saleh model with $\alpha_a = 2$, $\beta_a = 1$, $\alpha_p = \pi/3$, $\beta_p = 1$, $IBO = 10$ dB (moderate nonlin-

²Nonlinear ICI manifests if considering the orthogonality of Alamouti code results in two independent subchannels, over which two antennas transmit two coded symbols in parallel in each transmission block of two time slots. This convention has been applied in [81] and will be further clarified below.

earity).

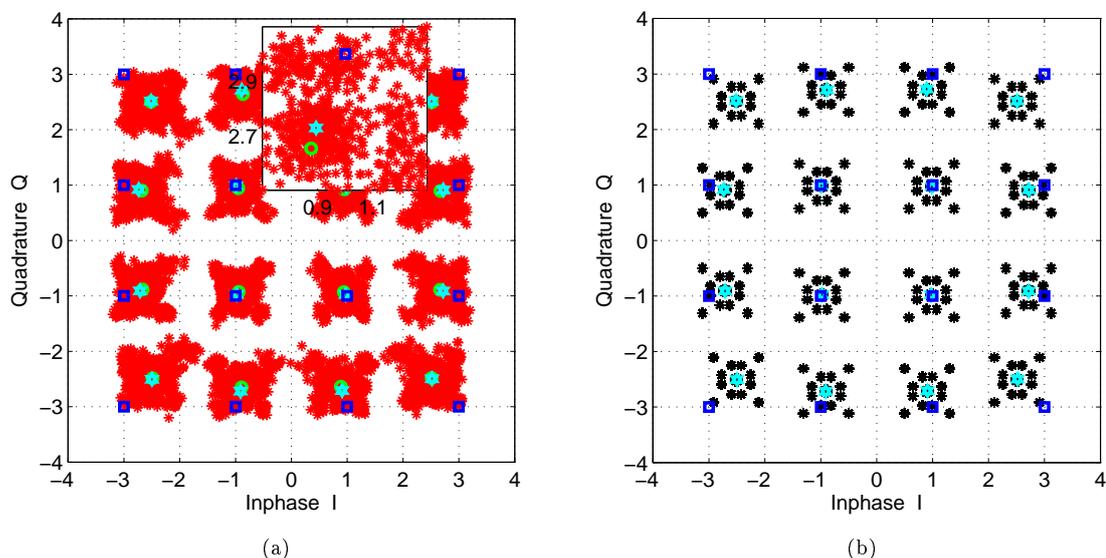


Figure 3.2: Receive signals after MRC for the system models with (a) and without (b) transmit/receive filters.

In sub-figure 3.2(a), square markers denote ideal signal points, circle markers are geometrical centers of the receive signal points (in asterisk markers) created by system with transmit/receive filtering, and hexagonal star markers, which are included for comparison purpose, are geometrical centers of the receive signal clusters created by system without transmit/receive filtering; zoomed-in sub-figure corresponds to transmit signal points with coordinate (1,3).

Under the (memory) influence of filters, nonlinear ISIs appear, changing the statistics of receive signals (asymmetric signal clusters and distinct differences between the geometric centers of these points for systems with and without transmit/receive filtering. On the other hand, though incurred by moderate nonlinearity ($IBO = 10$ dB), signal after MRC combining deteriorates significantly, manifested by the nonlinear expansion of receive signal

clusters, although there is no impact of thermal noise yet.

For comparison purpose, in case of a counterpart model with the same nonlinearity but without transmit/receive filtering, the signal received after the MRC combining is shown in sub-figure 3.2(b) (in asterisk markers). It is easy to see that models without transmit/receive filtering such as those in [73, 74, 81, 94] do not fully represent HPA's nonlinear impacts and therefore give (much) more optimistic results than reality. The receive signal quality in this case is much better than in the case of system model with filtering illustrated in sub-figure 3.2(a). More importantly, the change of receive signal statistics, that actually reflects the nonlinear ISI and ICI effects, is not observed in sub-figure 3.2(b), making deceptive assumption for analytical analyses.

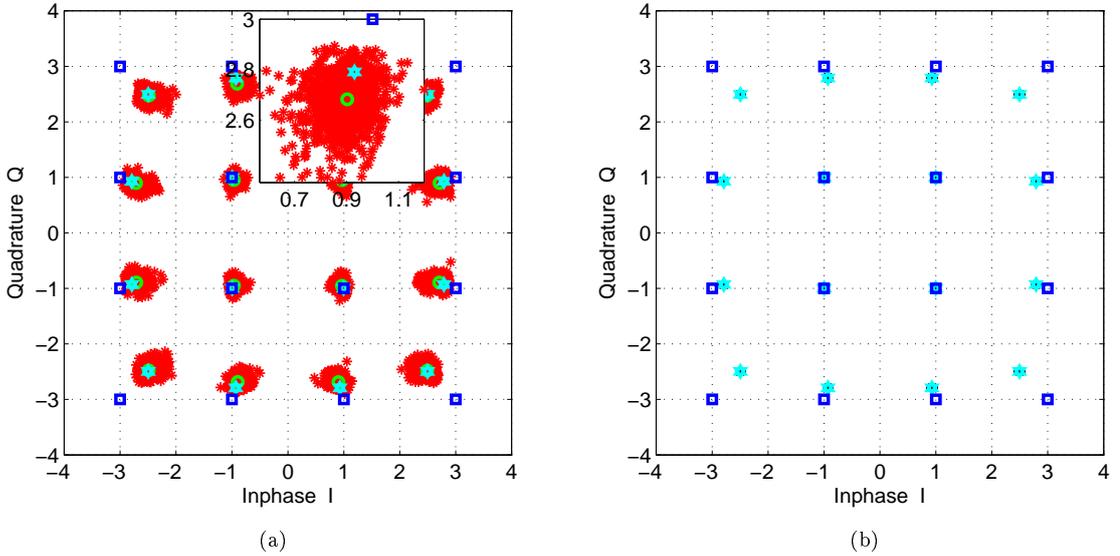


Figure 3.3: Receive signals after MRC with HPA model used in [81] for the systems models with (a) and without (b) transmit/receive filters.

On the other hand, from sub-figure 3.2(b) it can be observed that in case without filtering, ICI still exists under the effects of nonlinear HPA since the STBC code transmits two symbols s_k and s_{k+1} in the same time slot k

but these two symbols lose their orthogonality after passing through HPA as described by (1.8). Therefore, the assumption of using orthogonality of the Alamouti STBC code to decompose the MIMO system into SISO equivalent ones as equation (8) in [81] is an over simplification creating a large gap from reality. The theoretical and simulation results in [81] agree with each other because the nonlinearity used in the simulation are relatively weak³. In this case, the ICI on sub-figure 3.2(b) will disappear, since there is no phase rotation effect; the cluster of points shrinks into a single point and the system will incur only one effect of amplitude compression with a moderate nonlinearity. The impacts of HPAs in the MIMO-STBC system will therefore be no difference from what being in the classic SISO system. This fact is demonstrated by the simulation results shown in sub-figure 3.3(b) with the HPA parameters similar to those used in [81]. Therefore, the simplification/approximation in [81] is obviously less convincing if comparing to the results in sub-figure 3.2(b) when further considering the HPA's AM-PM phase rotation effect and especially to the results in sub-figure 3.2(a) when the system model is fully introduced with transmit/receive filters. Even when using the HPA as shown in [81], if the system model having transmit/receive filters as it practically is, the signal quality is also significantly degraded as illustrated in sub-figure 3.3(a). Obviously, the receive signals are non-Gaussian, making analytical analyses hardly tractable.

Therefore, it can be seen that, although the theoretical results in [81] are consistent with the numerical ones using the Gaussian model, but in essence, these analytical results are Gaussian approximations, that are relatively in-

³This work uses SSPA following Rapp model (2.11) without AM-PM distortion; moreover, the knee sharpness of $s = 2$ makes the nonlinear AM-AM close to the soft limiter (1.4); please reconsider sub-figure 2.6 for reference.

adequate and far away from reality. Also using a Gaussian approximation with thorough analysis and validity demonstration, Chapter 4 of the thesis presents a proposal approximating the phase rotation effect incurred by nonlinear HPAs in MIMO-STBC system as well as an effective solution to overcome this harmful impact.

All signals in n_R branches are combined in the same manner as (3.10) and (3.11) in the time slot k and $k + 1$, and then are input to the maximum likelihood (ML) detector [7], returning results \hat{m}_k and \hat{m}_{k+1} corresponding to the m_k and m_{k+1} messages.

So, in this section, the effects of nonlinear HPAs, with typical characteristics including both amplitude and phase distortions, on the MIMO-STBC system have been analyzed in detail. Thereby, the limitations and shortcomings of previous publications have been clearly figured out. The adequate consideration of these nonlinear adverse effects is the basis for implementing compensation schemes that are effective and closer to reality. This is the content presented in the next section.

3.3. Predistortion schemes

In this section, four predistortion techniques belonging to the signal predistortion methods [22, 32] will be discussed in detail. The proposed system model is illustrated in Figure 3.4. This is a very brief baseband-equivalent model that represents the MIMO-STBC system with digital predistortion. A more specific diagram describing the predistortion part of this type will have blocks representing relationships between the HPA output, inherently being an RF signal, with digital baseband or intermediate frequency PD

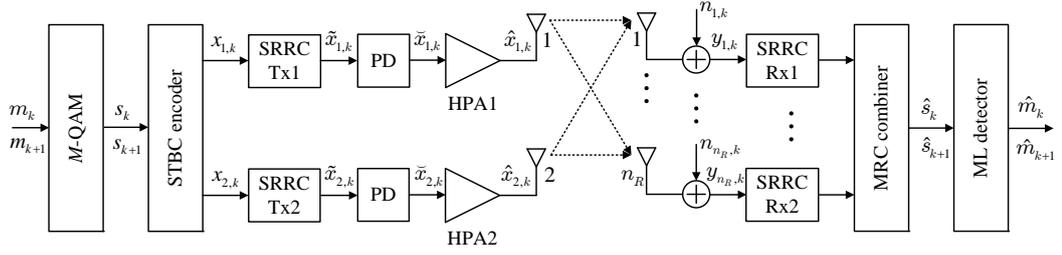


Figure 3.4: MIMO-STBC system model with predistorters.

part (namely frequency conversion, filtering and ADC/DAC parts) as well as functional blocks and their connections to control, adapt, update parameters according to a certain predistortion algorithms (control parts) as described in further detail in Figure 3.5. These diagrams are thoroughly discussed in, for example, [22, 23, 32, 91], with the baseband predistorter illustrated in Figure 3.6. Assume that these blocks function ideally, namely there are no errors, delays, phase differences, I/Q imbalances,... when processing signals across.

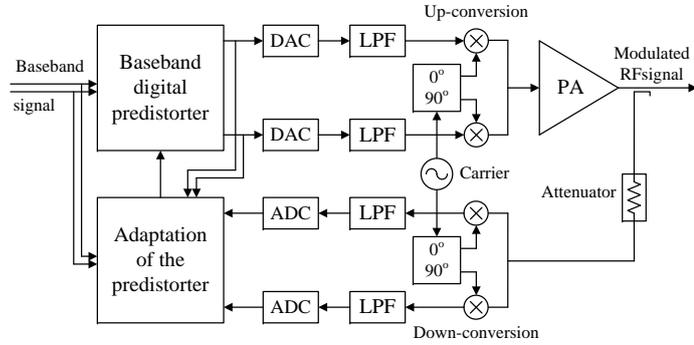


Figure 3.5: Power amplifier linearization using baseband digital predistorter.

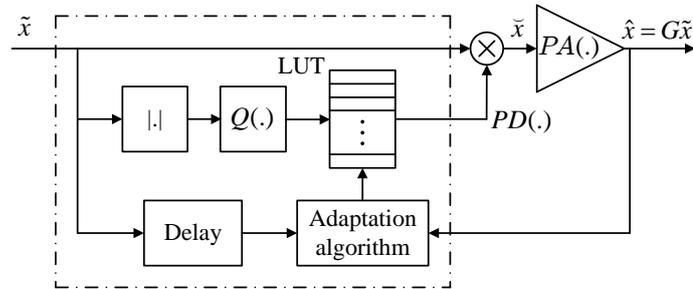


Figure 3.6: Baseband digital predistorter diagram.

The predistortion algorithms below are developed on the basis of consulting ideas from published works with recalibration by careful calculations or from original references.

3.3.1. Ideal inverse Saleh predistortion

Considering HPA Saleh model (2.7), and (2.8). It is not difficult to realize that the AM-AM function (2.7) is monotonic on the input range before saturation point, $0 \leq r \leq A_{is}$. Therefore, the amplitude predistortion function $P_a(r)$ is uniquely defined (inverse injective) only in this range; beyond the saturation point, the predistorter output is then normalized. Generally, the AM-PM function (2.8) is monotonic over all input domain r ; however, by using amplitude r as the input variable, the phase predistortion function $P_p(r)$ has the same domain as $P_a(r)$. These amplitude and phase predistortion functions are

$$P_a(r) = \begin{cases} \frac{\alpha_a - \sqrt{\alpha_a^2 - 4\beta_a r^2}}{2\beta_a r}, & r \leq A_{os} \\ 1/\sqrt{\beta_a}, & r > A_{os}, \end{cases} \quad (3.12)$$

$$P_p(r) = \begin{cases} \frac{-\alpha_p(\alpha_a - \sqrt{\alpha_a^2 - 4\beta_a r^2})^2}{\beta_p(\alpha_a - \sqrt{\alpha_a^2 - 4\beta_a r^2})^2 + 4\beta_a^2 r^2}, & r \leq A_{os} \\ \frac{-\alpha_p}{\beta_a + \beta_p}, & r > A_{os}. \end{cases} \quad (3.13)$$

Thus, in order to realize the ideal Saleh predistortion, it is necessary to know the parameter set of HPA model. There have been many different proposals for estimating this set of parameters, such as neural networks [73, 74] or particle swarm optimization (PSO) [87], etc. Ideal PD can be accomplished using lookup table (LUT) with a quantizer $Q(\cdot)$ operating on digital numbers represented by a certain number of quantized bits (reconsider Figure

3.6). The size of LUT determines the number of discrete points on which the predistortion function is determined. This may be a disadvantage of the predistortion method: Random access memory (RAM) must be large enough to store complex PD values (consisting of two I/Q components) if high resolution is required, about 2 Kword for 10 bit representation of input signal, for example. Moreover, these points might not be uniformly distributed (quantized) over the dynamic range of input signal amplitude. In some applications, it is necessary to arrange the predistortion function points over the dynamic range of input signal amplitude using squared or logarithmic laws to create denser density of coefficients at higher input power levels, where HPA more clearly demonstrates the strong nonlinearities [57].

Simulation results show that, with a relatively small number of quantization bits (6 bits, corresponding to about 130 words of memory) and a uniform quantizer, then linearity improvement is very positive. Therefore, this method is often considered first when dealing with HPA nonlinear distortion. However, it should be reemphasized that this method requires prior knowledge of HPA nonlinear characteristics and is incapable of adapting to the changes of these characteristics.

3.3.2. Adaptive secant predistortion

The concept of predistortion by secant method was proposed by Cavers J.K. [19] when he found an adaptive solution for PD by referring to the problem of solving algebraic equations using the secant method. Omitting the antenna branch index i and the time index k , assuming PD has a complex transfer function (including both AM-AM and AM-PM distortions) of $PD(\cdot)$,

and HPA has complex transfer function of $PA(\cdot)$, then

$$\tilde{x} = \tilde{x} \cdot PD(|\tilde{x}|^2), \quad (3.14)$$

$$\hat{x} = \tilde{x} \cdot PA(|\tilde{x}|^2). \quad (3.15)$$

An integrated system including the predistorter and the power amplifier, PD-HPA, can be represented by the following relationship

$$\tilde{x} \cdot PD(|\tilde{x}|^2) \cdot PA(|\tilde{x}|^2) = \hat{x} = K \cdot \tilde{x}, \quad (3.16)$$

where, K is the overall system gain, often determined slightly below the gain at the middle point of the HPA dynamic range⁴. This equation could be recast using single input variable \tilde{x} as

$$\tilde{x} \cdot PD(|\tilde{x}|^2) \cdot PA(|\tilde{x}|^2 \cdot |PD(|\tilde{x}|^2)|^2) = K \cdot \tilde{x}. \quad (3.17)$$

If considering $PD(\tilde{x})$ as a variable of equation (3.17), then the system error will be determined by

$$e(PD(\tilde{x})) = \tilde{x} \cdot PD(|\tilde{x}|^2) \cdot PA(|\tilde{x}|^2 \cdot |PD(|\tilde{x}|^2)|^2) - K \cdot \tilde{x}. \quad (3.18)$$

The determination of predistorter, thus becomes a problem of finding root $PD_m(\tilde{x})$ for equation $e(PD(\tilde{x})) = 0$. This equation could be solved using secant method [98, 101] by the recurrence relation

$$PD_{m+1}(\tilde{x}) = \frac{PD_{m-1}(\tilde{x}) \cdot e(PD_m(\tilde{x})) - PD_m(\tilde{x}) \cdot e(PD_{m-1}(\tilde{x}))}{e(PD_m(\tilde{x})) - e(PD_{m-1}(\tilde{x}))}, \quad (3.19)$$

starting with initial values $PD_0(\tilde{x}_k)$ and $PD_1(\tilde{x}_k)$. The iterates (3.19) of the secant method converge to a root if the initial values are sufficiently close to this root. The order of convergence is approximate to 1.6 (precisely being

⁴Reconsidering equation (3.12) and noting that $A_{os} = \alpha_a/2\sqrt{\beta_a}$, where, α_a is the linear gain of the model.

$(1 + \sqrt{5})/2$, or golden ratio) [98, 101], and therefore resulting in relatively fast convergence.

The secant algorithm allows to adaptively determine the predistorter according to the changes of HPA characteristics depending on the operating conditions such as temperature, power supply, load variations, or components aging,... On the other hand, the complex characteristic function of PD (3.14) has a variable in the form of input power, that is to work with real numbers instead of complex ones; thus it can be simplified the LUT (see Figure 3.6) since the table indexes are determined by the real-valued quantized input power levels.

3.3.3. Adaptive Newton predistortion

To reduce computational complexity, [104] proposed the root finding $PD_m(\tilde{x})$ in (3.17) by adaptive Newton method. Recursive formula for finding root by the Newton method [99, 101] is given as:

$$PD_{m+1}(\tilde{x}) = PD_m(\tilde{x}) - \left\{ \frac{\partial PA(|\tilde{x}|^2 PD^2(|\tilde{x}|^2))}{\partial PD(|\tilde{x}|^2)} \Big|_{PD_m(\tilde{x})} \right\}^{-1} \{PA(|\tilde{x}|^2 PD^2(|\tilde{x}|^2)) - K\tilde{x}\}. \quad (3.20)$$

It can be decomposed the complex equation (3.18) into amplitude and phase components as

$$\begin{cases} P_a(|\tilde{x}|^2) F_a(|\tilde{x}|^2 |P_a(|\tilde{x}|^2)|^2) - K = 0, \\ F_p(|\tilde{x}|^2 |P_a(|\tilde{x}|^2)|^2) + P_p(|\tilde{x}|^2) = 0, \end{cases} \quad (3.21)$$

where, $P_a(\cdot)$, $P_p(\cdot)$, $F_a(\cdot)$, $F_p(\cdot)$ correspondingly are amplitude and phase functions of PD and HPA introduced in the previous equations. For equation system (3.21), $P_a(\cdot)$ and $P_p(\cdot)$ are respectively desired roots determining the

predistorter. This equation system could be recast as

$$\mathbf{h} = \begin{cases} h_1(P_a, P_p) = 0, \\ h_2(P_a, P_p) = 0. \end{cases} \quad (3.22)$$

For determining roots, Jacobian matrix (first-order partial derivatives) is defined as

$$\mathbf{J}_0 = \begin{pmatrix} \frac{\partial h_1}{\partial P_a} & \frac{\partial h_1}{\partial P_p} \\ \frac{\partial h_2}{\partial P_a} & \frac{\partial h_2}{\partial P_p} \end{pmatrix}. \quad (3.23)$$

Since it is impossible to analytically calculate these derivatives, an approximation by differential entropy is used instead [104]:

$$\mathbf{J} = \begin{pmatrix} \frac{h_1(P_a^m, P_p^m) - h_1(P_a^{m-1}, P_p^m)}{P_a^m - P_a^{m-1}} & \frac{h_1(P_a^m, P_p^m) - h_1(P_a^m, P_p^{m-1})}{P_p^m - P_p^{m-1}} \\ \frac{h_2(P_a^m, P_p^m) - h_2(P_a^{m-1}, P_p^m)}{P_a^m - P_a^{m-1}} & \frac{h_2(P_a^m, P_p^m) - h_2(P_a^m, P_p^{m-1})}{P_p^m - P_p^{m-1}} \end{pmatrix}. \quad (3.24)$$

If \mathbf{J} is nonsingular, the matrix inverse can be defined as follows

$$\mathbf{J}^{-1} = \frac{1}{J_{11}J_{22} - J_{12}J_{21}} \begin{pmatrix} J_{22} & -J_{12} \\ -J_{21} & J_{11} \end{pmatrix}. \quad (3.25)$$

Then, recursive formula (3.20) of the Newton method becomes

$$\mathbf{x}_{m+1} = \mathbf{x}_m - \mathbf{J}^{-1} \cdot \mathbf{h}, \quad (3.26)$$

where $\mathbf{x}_m = [P_a^m, P_p^m]^T$ is the root of m -th iteration with initial value \mathbf{x}_0 .

The replacement of exact derivative (3.23) by approximation (3.24) reduces the order of convergence of the algorithm. On the other hand, the initial value \mathbf{x}_0 has a strong influence on the accuracy of iterated solution. These factors together degrade the algorithm performance. However, the separation of complex equation (3.20) into the real-valued equation system (3.21) allows a significant reduction in computational complexity and makes ease of deployment on hardware.

3.3.4. Adaptive LMS polynomial-approximated predistortion

Amplitude and phase predistortion functions, $P_a(r)$ and $P_p(r)$, could be approximated by the following polynomials [94]

$$P_a(r) = a_1r + a_2r^2 + a_3r^3 + \dots + a_lr^l = \mathbf{a}^T \mathbf{r}_a, \quad (3.27)$$

$$P_p(r) = p_0 + p_1r + p_2r^2 + \dots + a_mr^m = \mathbf{p}^T \mathbf{r}_p, \quad (3.28)$$

where $r = |\tilde{x}|$ is the PD input amplitude, $\mathbf{r}_a = [r, r^2, r^3, \dots, r^l]^T$ is the amplitude power polynomial vector of order l , $\mathbf{r}_p = [1, r, r^2, \dots, r^m]^T$ is the phase power polynomial vector of order m , $\mathbf{a} = [a_1, a_2, a_3, \dots, a_l]^T$ is the amplitude power polynomial coefficient vector, $\mathbf{p} = [p_0, p_1, p_2, \dots, p_m]^T$ is the phase power polynomial coefficient vector. The least mean square (LMS) algorithm [45] is employed for calculating amplitude and phase coefficients as

$$\mathbf{a}(m+1) = \mathbf{a}(m) + \mu_a(m)\mathbf{r}_a(m)e_a(m), \quad (3.29)$$

$$\mathbf{p}(m+1) = \mathbf{p}(m) + \mu_p(m)\mathbf{r}_p(m)e_p(m), \quad (3.30)$$

where $\mu_a(m)$ and $\mu_p(m)$ correspondingly are (adaptive) amplitude and phase step sizes for updating, and amplitude error $e_a(m)$ and phase error $e_p(m)$ are given by

$$e_a(m) = |\tilde{x}(m)| - |\hat{x}(m)|, \quad (3.31)$$

$$e_p(m) = \angle\tilde{x}(m) - \angle\hat{x}(m), \quad (3.32)$$

where $|\cdot|$ and $\angle(\cdot)$ respectively are modulus and argument operators.

It is easy to see that, with the introduction of polynomials representing amplitudes and phases, the complexity of this algorithm increases significantly. On the other hand, polynomial approximation does not well perform

for strong nonlinearities. Therefore, this algorithm gives inferior performance while requires much higher complexity compared to previous algorithms.

In summary, this section mentions four different distortion algorithms in terms of many aspects such as algorithm structure, computational complexity, convergence order, and applicability. These properties will be quantified specifically through the measures and simulation results presented in the next section.

3.4. Performance evaluation for predistored MIMO-STBC systems

3.4.1. System parameters and performance measures

This section presents and analyzes numerical simulation results for four predistortion algorithms applied in the MIMO-STBC system illustrated in Figure 3.4. Since the receive branches process independently, the 2×1 MISO Alamouti scheme can be used without changing the main properties of simulation results. The system parameters are remained the same as in the case simulating Figure 3.2(a), HPA represented by Saleh model with $\alpha_a = 2$, $\beta_a = 1$, $\alpha_p = \pi/3$, $\beta_p = 1$.

Basic parameters of predistortion algorithms: Ideal inverse Saleh using LUT with 10 bits indexing; adaptive secant method initialized by $PD_0 = 0$, $PD_1 = 1$, absolute recursive error $|PD_m - PD_{m+1}| < 10^{-6}$; adaptive Newton method initialized by $(P_a, P_p)_0 = (0, 0)$, $(P_a, P_p)_1 = (1, 1)$, absolute recursive error $|(P_a, P_p)_m - (P_a, P_p)_{m+1}| \leq 10^{-2}$; LMS polynomial approximation method using amplitude polynomial of order 5, phase polynomial of order 4, amplitude error $\varepsilon_a \leq 10^{-5}$, phase error $\varepsilon_p \leq 10^{-6}$, adaptive

step sizes determined by their corresponding errors, $\mu_a(m) = |e_a(m)|/2$, $\mu_p(m) = |e_p(m)|/2$. All adaptive algorithms have same maximum iterations $N_{\text{ite}} \leq 50$, avoiding the case of unsatisfying error conditions.

First, receive signal constellations of predistortion methods are depicted for qualitative comparison of performance improvement capabilities to the case without predistortion and also to each other. Then, non-ideal effects incurred by HPA and performance improvement capabilities of predistorters are quantitatively evaluated using measures of EVM, MER. Finally, the general performance of predistorters are estimated using BER versus E_b/N_0 ,

3.4.2. Receive signal constellations with predistortion

Receive signal constellations after MRC combining of four predistorters are illustrated in Figure 3.7: a) Ideal inverse Saleh method (LUT); b) Adaptive secant method (Secant); c) Adaptive Newton method (Discrete Newton); d) Adaptive LMS polynomial approximation method (Polynomial). Simulation results are generated using relative small input power backoff (strong nonlinearities), $IBO = 6$ dB, and large enough $E_b/N_0 = 30$ dB, such that thermal noise effect can be ignored. Obviously, this IBO value is significantly lower (thus stronger nonlinear distortions) than the case of $IBO = 10$ dB used for generating results in Figure 3.2(a), however, receive signals have indistinguishably better performance compared to the case without predistortion.

The failure in finding roots PD_m or $P_{a,m}$ and $P_{p,m}$ results in drifts of signal points (in asterisk markers) apart from ideal ones (in square markers); though having different patterns but generally, these shifted points are sparse and relatively random. Moreover, the geometrical centers of receive signal clusters

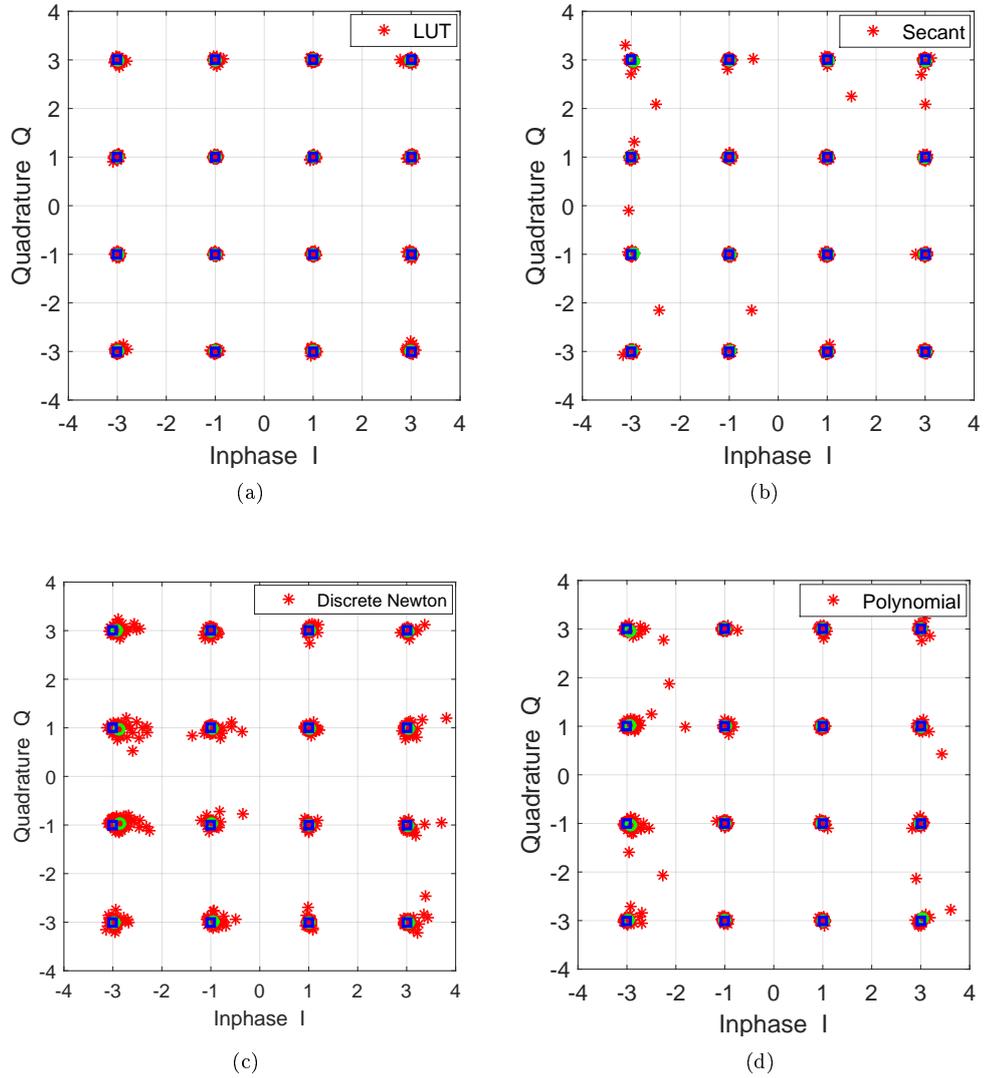


Figure 3.7: Receive signal constellations with predistortion: a) LUT; b) Secant; c) Newton; d) Polynomial.

(in circle markers) almost coincide with the ideal signal points. That expresses capabilities of predistorters effectively compensating nonlinear effects (including impacts both from AM-AM and AM-PM distortions as analyzed in section 3.2.2), which inherently appear in the form of non-Gaussian distributions for clusters in Figure 3.2(a) (though in this case, the distortions are significantly stronger). Preliminarily, it can be seen that, performance improvement capabilities of these four predistorters decrease in the following

order: Ideal inverse Saleh method, adaptive secant method, adaptive Newton method and adaptive LMS polynomial approximation method, conforming to the above analyses.

3.4.3. Error vector module

Error vector module is defined as the difference between reference signal and measurement one at a certain point in system [32]:

$$\begin{aligned} EVM_{\text{RMS}}(\%) &= \sqrt{\frac{P_{\text{error}}}{P_{\text{reference}}}} \times 100\% \\ &= \sqrt{\frac{\frac{1}{N} \sum_{k=1}^N (I_k - \tilde{I}_k)^2 + (Q_k - \tilde{Q}_k)^2}{\frac{1}{N} \sum_{k=1}^N (I_k^2 + Q_k^2)}} \times 100\%, \quad (3.33) \end{aligned}$$

where I_k and Q_k respectively are the reference (ideal) values in the inphase and quadrature axes, \tilde{I}_k and \tilde{Q}_k are the measurement values.

EVM gives a measure for aggregate effect caused by non-ideal working conditions to the transmit symbols. These conditions include both deterministic factors and non-deterministic ones. Completely deterministic and static factors simply shift the signals from the ideal positions a fixed amount (e.g., the amplitude compression and phase shift effects of nonlinear HPA); but random factors such as ISI, ICI and noise, can cause signals to fluctuate randomly around the ideal points, forming a cluster of error points.

Figure 3.8 depicts EVMs versus IBO of systems: ideal inverse Saleh (LUT), soft envelope (SEL), adaptive secant method (Secant), adaptive Newton method (Newton), adaptive LMS polynomial approximation (Polynomial), and none predistortion (HPA only). The ability to improve the signal quality of the predistortion algorithms also decreases in this order.

The above results show that, in general, the considered predistortion algo-

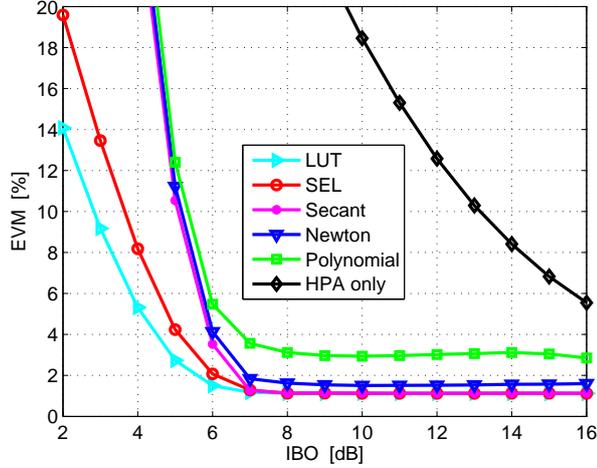


Figure 3.8: EVM versus IBO of the MIMO-STBC system with different predistorters.

gorithms can perform effectively to a certain threshold of IBO (around $6 \sim 7$ dB, for the simulation parameter set and the HPA model introduced). Revising the results from Figure 3.7 (with $IBO = 6$ dB) reveals that, above this IBO threshold, the nonlinear distortions will almost be eliminated, the receive signals are approximately Gaussian distributed, becoming ideal for the optimal receiver. Therefore, when using predistortion, at least the HPA operating point can reliably be pushed to this threshold. Beyond this value, nonlinear distortions are too large, and cannot be compensated even with the ideal predistortion scheme (inverse Saleh), so the quality of the system decreases very quickly.

On the other hand, it can be seen that LUT outperforms SEL especially at small IBO , the reason is SEL (1.5) does not consider the phase distortion as LUT does in (3.13). Thus, once again, it is clearer that the phase distortion greatly deteriorates signal quality. This was clarified in [69] that AM-PM conversion can only degrade, never improve, the intermodulation-noise performance of memoryless nonlinear systems. The ignoring of phase distortion

as in [81] causes the nonlinear distortion impacts become negligible and such a system can be analytically analyzed as previously discussed.

Real systems generally accept EVM at a certain threshold [60] [Table 1]; therefore, the results from Figure 3.8 allows determining the *IBO* level (or HPA operating point) according to the specified predistortion algorithm.

3.4.4. Modulation error ratio

The modulation error ratio [32] and EVM basically measure the same quantity and can be easily converted to each other when the signal set is known:

$$\begin{aligned} MER(dB) &= 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{error}}}\right) \\ &= 10\log_{10}\left(\frac{\frac{1}{N}\sum_{k=1}^N(I_k^2 + Q_k^2)}{\frac{1}{N}\sum_{k=1}^N(I_k - \tilde{I}_k)^2 + (Q_k - \tilde{Q}_k)^2}\right). \end{aligned} \quad (3.34)$$

In fact, MER is preferred because: 1) Values and the unit of MER are very similar to the previous familiar measurements such as carrier-to-noise ratio C/N, signal-to-noise ratio SNR; 2) MER can be considered as a form of SNR, which accurately characterizes the receiver's ability to successfully demodulate signals, since it not only includes Gaussian noise but also includes all errors unable to be compensated of the receiver's signal. If only Gaussian noise plays the role of signal distortion, then MER will be equivalent to SNR.

Figure 3.9 represents MER measures of six systems considered in Figure 3.8. Similar to the conclusions drawn from the sooner, but Figure 3.9 gives a clearer view for the performance of predistortion algorithms considered. Adaptive LMS polynomial approximation algorithm and adaptive Newton

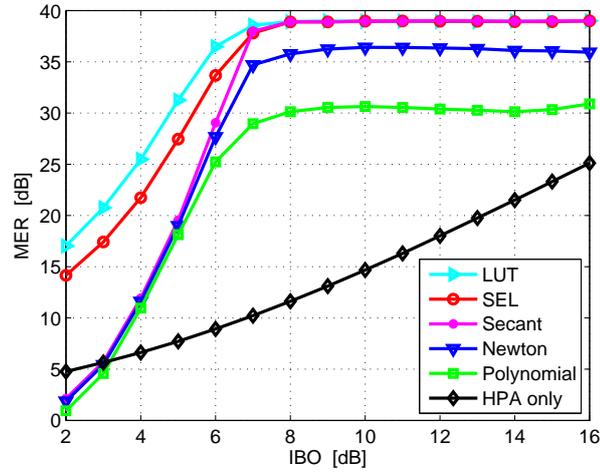


Figure 3.9: MER versus IBO of the MIMO-STBC system with different predistorters.

algorithm are less capable of operating in stronger nonlinear region (small IBO), but in weaker nonlinear region (large IBO) they also do not well perform. The reason is the lack of capability of finding correct predistortion roots in the system equations examined. On the other hand, these are also the most complex algorithms among the previous distortion algorithms assessed. The simulation run time of these two algorithms is dozens of times larger (for the Newton method) to hundreds of times larger (for the polynomial approximation method) than that of the reference method of ideal inverse Saleh algorithm. However, it should be repeated that these algorithms are completely independent of HPA and their hardware deployment is also simpler, so they can still be realistic choices in certain cases, for example, when the system has to work with different HPAs, or with the same HPA, but has different operating modes while the system hardware has sufficient computation capability.

3.4.5. Bit error ratio

The aggregate performance of the MIMO-STBC system with different pre-distorters is expressed in terms of bit error rate by E_b/N_0 as shown in Figure 3.10. According to the Figure 3.9 when not to mention the effect of thermal noise, the setting of IBO at 6 dB in this case will generate a MER of approximately 37 dB for ideal Saleh distortion, 29 dB for adaptive secant method, 27.5 dB for Newton method, and 25.5 dB for polynomial approximation method. These results are clearly consistent with what observed from Figure 3.10. At this IBO level, ideal Saleh predistortion (LUT) is well approximated to the perfect linear system (dashed curve with notation “Linear”).

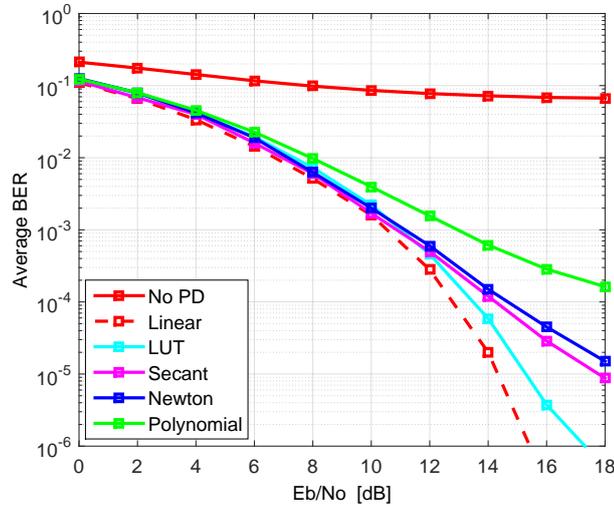


Figure 3.10: BER versus E_b/N_0 of the MIMO-STBC system with $IBO = 6$ dB and different pre-distorters.

3.5. Summary of chapter 3

This chapter aims to fully investigate the impacts of nonlinear distortions on MIMO-STBC system that have not been mentioned in previous publi-

cations. The analyses show that the transmit/receive filters significantly increase the detrimental effects of nonlinear distortions in the system, the distribution of receive signals becomes non-Gaussian and so it is not easy to perform the analytical analysis. Thereby, limitations in the previous works are shown. Based on these analyses, four predistortion schemes are proposed to apply to the system. These diagrams are analyzed in detail and their performance are compared through specific measures including EVM, MER and BER.

The results achieved in this chapter:

- Introducing transmit/receive filters into the MIMO-STBC system model with typical HPAs, that include both amplitude and phase distortions; then carrying out detailed and complete analyses of nonlinear distortion impacts on the system; and pointing out limitations and shortcomings of previous publications;
- Revising predistortion schemes: Ideal inverse Saleh, adaptive secant, adaptive Newton, adaptive polynomial approximation. The analyses mentioned include: algorithm structure, convergence, complexity, applicability;
- Evaluating the performance of predistortion diagrams by measures EVM, MER, BER. Analysis of advantages and disadvantages of each diagram, relating to the reality.

Chapter 4

Automatic Phase Estimation and Compensation for Nonlinear Distortions due to HPAs in MIMO-STBC Systems

Although the optimal phase compensation method for the MIMO-STBC with nonlinear distortions incurred by HPAs has recently been mentioned in [71], there has been no detailed discussion on the mechanism of phase-shifting effects of nonlinear HPA for the MIMO-STBC signals. Therefore, in this chapter, the thesis specifically analyzes the effects of nonlinear phase rotation for the above system. An interesting result acquired is that, it is possible to well approximate this nonlinear effect to a linear phase rotation model. This is the basis for the effective estimation of the rotated phase angle value of the receive signals and for the approximately optimal automatic phase compensation proposed. The contents of this chapter are related to published works 1, 2, and 4 in the List of Publications.

4.1. Overview

As introduced in the first part of Chapter 3, some solutions to overcome the adverse effects of nonlinear HPA on MIMO-STBC systems have recently been proposed in [73, 74, 81, 94]. However, all models considered ignore transmit/receive filters and therefore only memoryless nonlinear distortion effects are concerned with. In more detail, it can be seen that the HPA nonlinearities, when combined with the memory generated by transmit/receive filters

cause serious detrimental effects on the receive signals. These effects, therefore, need to be adequately attended to when studying MIMO-STBC systems with HPA nonlinear distortions.

On the other hand, for the non-ideal characteristics of HPA, AM-AM distortion, or the signal amplitude compression effect can be compensated relatively effectively at the receiver by automatic gain control (AGC) schemes [22]. However, the AM-PM distortion, inherently existing in practical HPAs for both TWTAs [84] and SSPAs [6, 21, 32, 72] causes phase rotation for transmit signals, thus destroys the orthogonality of STBC codes as represented in (1.8), and consequently, seriously reduces signal quality. Although being a strongly harmful factor, until now, this phase-shifting effect has not been fully discussed, especially in MIMO-STBC systems.

For phase-shifting effects in linear SISO M -QAM systems with fixed phase rotation unknown at the receiver, optimal feedforward phase estimation methods with harmonic and biharmonic approximations were proposed by Sergienko A. [78, 79]. This is an efficient estimator [56], approaching the Cramer-Rao lower bound [79, 83]. However, there is no similar proposal for the phase estimator of nonlinear MIMO-STBC signal.

Driven by the effectiveness of Sergienko's method and based on detailed analysis of the phase rotation effects of the nonlinear MIMO-STBC signal, this chapter presents a proposal of phase estimation and phase compensation for this nonlinear MIMO-STBC system. Some typical HPA models, representing both TWTA and SSPA technologies, with very different nonlinear characteristics are included in the analyses and simulation to assess the effectiveness and reliability of the proposed phase estimation algorithms and the

performance of the nonlinear system with phase compensation.

4.2. Phase rotation impact due to nonlinear HPAs for the MIMO-STBC signals

4.2.1. Nonlinear MIMO-STBC system model with phase estimation and compensation at the receiver

Figure 4.1 describes the proposed model, which is the supplementation of Figure 3.1 with phase estimation and phase compensation blocks succeeding SRRC receive filters.

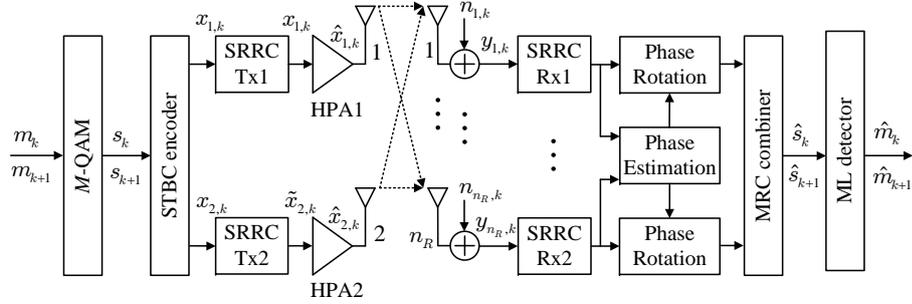


Figure 4.1: Proposed MIMO-STBC system model with phase estimation and compensation.

The signal processing in the transmit part has already been analysed in sub-section 3.2.1 with equations representing from Alamouti encoding, (3.1), to the output response of SRRC transmit shaping filter, (3.4). However, since the most important feature of the phase estimation proposal is the Gaussian approximation for the non-Gaussian model [67], then one of the most concerns is the approximating suitability, on which the phase estimation performance strongly depend. For that reason, several typical HPA nonlinearities, including both AM-AM and especially, AM-PM characteristics investigated in Chapter 2 will be used to generate diversified nonlinearities for the system in consideration. The following HPAs are focused in.

For TWTAs, being essentially common in satellite communication systems, particularly for transponders, the AM-AM and AM-PM are represented by Saleh model (2.7) and (2.8) with parameters $\alpha_a = 2.1587$, $\beta_a = 1.1517$, $\alpha_p = 4.0033$, $\beta_p = 9.1040$ [84], that is a typical parameter set often referred in up-to-date publications.

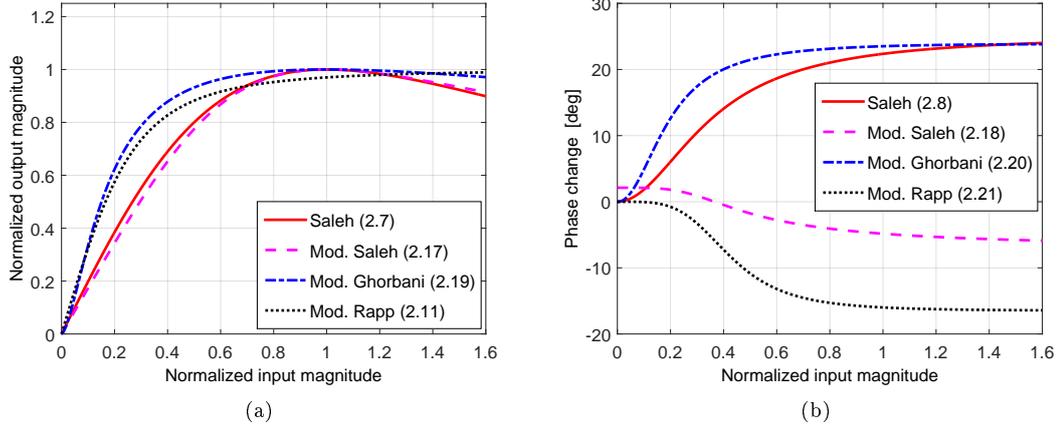


Figure 4.2: AM-AM (a) and AM-PM (b) characteristics of considered HPA models.

For SSPAs, published works frequently use Rapp model [82] with AM-PM conversion ignored. However, as discussed in Chapter 2, updated models for SSPA technologies all concern with this distortion. Among them, modified Saleh model is proposed for LDMOS (Laterally diffused metal oxide semiconductor) power amplifiers, popularly installed as HPAs for base stations in 2G, 3G or 4G mobile networks, working at the L, S, C bands with the AM-AM function (2.17) and AM-PM function (2.18), where, $\alpha_a = 1.0536$, $\beta_a = 0.086$, $\alpha_p = 0.161$, và $\epsilon_p = 0.124$ [6].

Modified Ghorbani model is appropriate to GaAs pHEMT FETs (Gallium-arsenide pseudomorphic high-electron-mobility transistor field-effect transistor), having operating frequency band up to 26 GHz (K band) and increasingly having advantages in terms of manufacturing technology and also mar-

ket share compared to other power semiconductor technologies, with amplitude and phase distortions given by (2.19) and (2.20), where, the model parameters are $x_1 = 7.851$, $x_2 = 1.5388$, $x_3 = -0.4511$, $x_4 = 6.3531$, $y_1 = 4.6388$, $y_2 = 2.0949$, $y_3 = -0.0325$, $y_4 = 10.8217$ [72].

Modified Rapp model is proposed for GaAs pHEMT/CMOS (Complementary metal-oxide-semiconductor) at frequency band 60 GHz with AM-AM characteristic (2.11) and AM-PM conversion (2.21), where, the model parameter set is $g = 16$, $A_{os} = 1.9$, $s = 1.1$, $\alpha = -345$, $\beta = 0.17$, $q_1 = q_2 = 4$ [21].

Figure 4.2 illustrates the amplitude and phase characteristics of all four model above with normalized input and output magnitudes to their corresponding input/output saturation levels for nonlinearity comparison purpose. Obviously, these characteristics are quite different in terms of amplitude and especially of phase distortions. These nonlinearity dissimilarities could affect signal passed through in very different extents and amounts; then, is the proposed phase estimator affected. Details are further discussed in the following sections.

4.2.2. Phase rotation impact due to nonlinear HPAs

Assume that the fading channel is quasi-static, and is perfectly known at receiver. For simplifying representation without loss of generality, the time slot index k is omitted in the following expressions. Then, the receive vector \mathbf{y} at this time slot is of the form

$$\mathbf{y} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{n}, \quad (4.1)$$

where $\hat{\mathbf{x}} = [\hat{x}_{1,k}, \hat{x}_{2,k}]^T \in C^{2 \times 1}$ is the transmit vector at HPA output, $\mathbf{y} = [y_{1,k}, y_{2,k}, \dots, y_{n_R,k}]^T \in C^{n_R \times 1}$ is the receive vector, $\mathbf{H} = [h_{m,l}]^{n_R, 2} \in C^{n_R \times 2}$ represents $n_R \times 2$ channel matrix with entry $h_{m,l}$ being the channel coefficient between the l -th transmit antenna and the m -th receive antenna, \mathbf{n} is the proper AWGN vector $\mathbf{n} = [n_{1,k}, n_{2,k}, \dots, n_{n_R,k}]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{N}_0)$ with mean vector $\mathbf{0}$ and covariance matrix \mathbf{N}_0 .

The signal on each receive branch is then passed through the SRRC matched filter (similar to the transmit one). To focus on the effects of nonlinear distortion existing in this signal, the relation of HPA input \tilde{s}_k and output \hat{s}_k for time slot k described in (4.1), is recast using the linearized model (1.3) as

$$\hat{s}_k = \gamma_k \tilde{s}_k + d_k, \quad (4.2)$$

where γ_k is a constant determined by the AM-AM function and the input power, $\gamma_k = \mathbb{E}[\tilde{s}_k^* \hat{s}_k] / \mathbb{E}[\tilde{s}_k^* \tilde{s}_k]$, and d_k represents nonlinear noise uncorrelated to the input \tilde{s}_k . This noise has zero mean and variance given by $\sigma_{d_k}^2 = \mathbb{E}[\hat{s}_k^* \hat{s}_k] - \gamma_k^2 \mathbb{E}[\tilde{s}_k^* \tilde{s}_k]$.

Signal on the l , $1 \leq l \leq n_R$, receive branch, namely the l -th equation of (4.1) in time slot k -th and $k+1$ -th has the following form

$$\begin{aligned} y_{l,k} &= h_{l,1} \hat{s}_k + h_{l,2} \hat{s}_{k+1} + n_{l,k}, \\ y_{l,k+1} &= -h_{l,1} \hat{s}_{k+1}^* + h_{l,2} \hat{s}_k^* + n_{l,k+1}. \end{aligned} \quad (4.3)$$

Replacing (4.2) into (4.3) and reconsider (3.4), it is not difficult to realize the existence of nonlinear noise (d_k, d_{k+1}) and nonlinear ISI ($F(n_k^{\text{ISI}})$, $F(n_{k+1}^{\text{ISI}})$, generated by nonlinear function $F(\cdot)$ as described in (1.1)) in the received signal at time slot k -th and $(k+1)$ -th. This is illustrated in sub-figure 4.3(a), resulted from simulation with parameters: 16-QAM modulation;

SRRC transmit/receive filters with roll-off factor $\alpha = 0.2$, input sampling rate F_d equals to 16-QAM symbol rate, output sampling rate $F_s = 16F_d$, group delay $Dl = 10$; HPA follows modified Ghorbani model (2.19) and (2.20) with the amplitude and phase characteristics plotted in Figure 4.2, $IBO = 14$ dB; $E_b/N_0 = 20$ dB, automatic gain control used at receiving part.

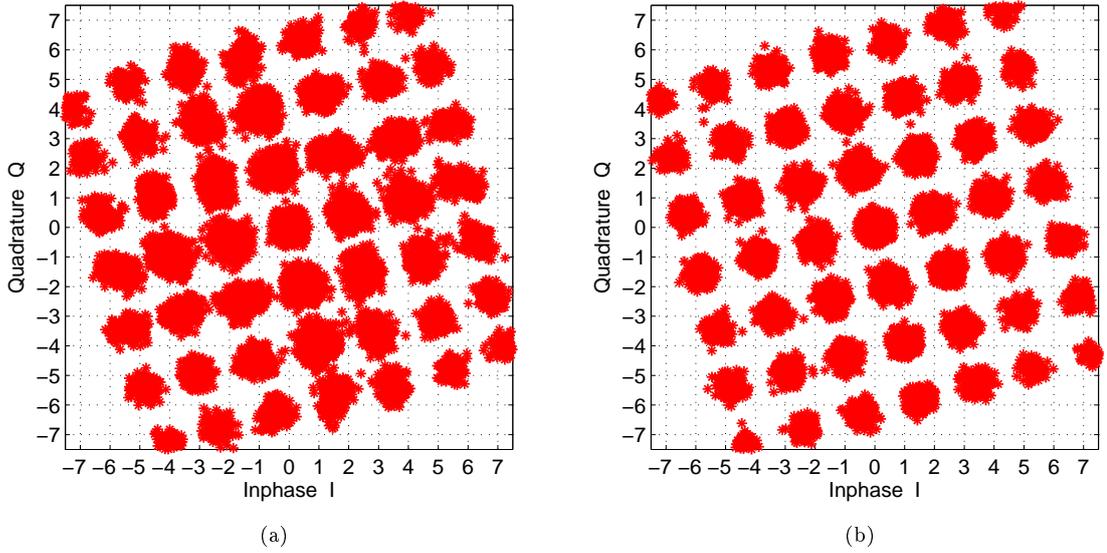


Figure 4.3: Receive signal constellations after matched filtering: a) Fully characterized (4.3); b) Approximated (4.5).

Noting that the nonlinear impacts in this MIMO case are different to what occurring in typical SISO scenario [80]. The receive signal as being a complex conjugate combination of distorted transmit signals \hat{s}_k and \hat{s}_{k+1} has different distribution to what has normally been observed in the nonlinear SISO case. Larger magnitude points (corresponding to I, Q values of ± 6 or ± 4 in the linear case) seem to less occur compared to smaller magnitude points (with I, Q values of ± 2 or 0). Nonlinear distortion appears to be stronger in these smaller magnitude points. However, these points are actually created by (distorted) components of \hat{s}_k and \hat{s}_{k+1} with maybe large magnitudes, and

thus incurring stronger nonlinear distortions, but cancelling themselves in the combination (4.3).

More importantly, it is further observed that, under the HPA's phase conversion effects, receive signal clusters tend to be almost rotated by the same angle, approximating to the phase conversion for the largest magnitude component signal. The reason is that for every combined signal $y_{l,k}$ by (4.3), there are always components with the largest magnitudes, the main factor causing phase rotation for $y_{l,k}$ ¹. Therefore, it is reasonable to have a good approximation of this nonlinear system to the linear one affected by a fixed phase rotation and an additive Gaussian-equivalent noise. This model will be discussed in more detail in the following section with graphical illustration depicted in Figure 4.3(b).

Noting further that models without transmit/receive filtering as introduced in [73, 74, 81, 94] will not fully represent nonlinear impacts caused by HPAs into the system (nonlinear ISI and nonlinear ICI), and thus (at the same nonlinearity) yield much more optimistic results than what happening in reality. This has been discussed thoroughly in section 3.2.2.

Finally, it could be seen that the phase compensation solution for these MIMO-STBC systems gains larger performance improvements for the receive signals than what would get when carrying out this solution for the conventional SISO systems. The reason is that, for MIMO-STBC systems, phase distortion could be fully compensated (for the largest phase rotation corresponding to the component signal with the largest magnitude). Obviously, it could not approach to such compensation for the conventional SISO sys-

¹This attribute is totally different to what actually happens in SISO system, where signal points (without the combinations) are phase rotated at different levels depending on their magnitudes; reconsidering Figure 1.6.

tems, where separate signals with different magnitudes incur different phase rotations. In [71], numerical results showed these large gains when manually carrying out the phase compensation solution. However, this work still owes a clear explanation for the results acquired.

Applying maximum ratio combining [7], in the l -th branch, the received signal is processed as

$$\begin{aligned}\tilde{s}_{l,k} &= h_{l,1}^* y_{l,k} + h_{l,2} y_{l,k+1}^*, \\ \tilde{s}_{l,k+1} &= h_{l,2}^* y_{l,k} - h_{l,1} y_{l,k+1}^*.\end{aligned}\tag{4.4}$$

4.3. Phase estimation problem

4.3.1. Gaussian approximation for the nonlinear model

In this work, fading channel effects is temporarily ignored and will be considered in future studies; then the channel coefficients in (4.3) could all be set to unity. Further, by the analysis discussed in previous section, it is reasonable to approximate the signal in time slot k and $k + 1$ as

$$\begin{aligned}y_k &= (\bar{s}_k + \bar{s}_{k+1})e^{j\phi_0} + n_k^{equ}, \\ y_{k+1} &= (\bar{s}_k^* - \bar{s}_{k+1}^*)e^{j\phi_0} + n_{k+1}^{equ},\end{aligned}\tag{4.5}$$

where $k = 1, 3, 5, \dots, 2K - 1$, \bar{s}_k and \bar{s}_{k+1} are modulated signals randomly carved from the M -QAM constellation, ϕ_0 is a phase rotation incurred by the combined effects investigated above and n_k^{equ} is the equivalent complex Gaussian noise, including the AWGN noise and nonlinear ISI, with variances of both real and imaginary components equal to σ^2 , $2K$ is the signal sequence (frame) length for phase estimation. Noting that approximation (4.5) insists on the phase rotation while neglecting the amplitude compression of nonlinear effects. Sub-figure 4.3(b) illustrates this approximation with phase rotation

$\phi_0 = 16.2^\circ$, which is the phase conversion of signal point (3, 3) in the 16-QAM constellation under the same nonlinearity generating sub-figure 4.3(a). Regardless of the almost indistinguishable amplitude compression (for large magnitude combined signals) in sub-figure 4.3(a), then there is a close similarity of models (4.3) and (4.5). This underlines for the efficient estimation of phase rotation caused by HPA's discussed next.

Moreover, (4.5) is a Gaussian approximation for the fully characterized nonlinear system (4.3). Then, different nonlinearities, incurred by both amplitude and phase conversions of the HPA models, result in different nonlinear ISI, nonlinear ICI, and nonlinear noise; thus the approximation accuracy could be affected in different ways. Therefore, four typical nonlinearity models introduced in section 4.2.1 could be used to verify the efficiency of proposed phase estimation algorithm and also phase compensation scheme.

4.3.2. Optimal blind feedforward phase estimation

The phase estimation problem could be formulated as follow. The maximum likelihood (ML) estimation of rotated phase ϕ_0 in (4.5) is determined by maximizing, with respect to ϕ_0 , the log-likelihood function (LLF)

$$\hat{\phi}_0 = \arg \max_{\phi_0} LLF(\phi_0 | \{y_k\}), \quad (4.6)$$

here $LLF(\cdot)$ is given by

$$LLF(\phi_0 | \{y_k\}) = \sum_{k=1}^{2K} F_\phi(\phi_0 | y_k), \quad (4.7)$$

where $F_\phi(\phi_0 | y)$ is the probability density function of sample $y = re^{j\phi}$

$$F_\phi(\phi_0 | y) = \log \left(\frac{1}{2\pi\sigma^2 M^2} \sum_{\substack{m=1, \\ n=1}}^M e^{\left(-\frac{|re^{j(\phi-\phi_0)} - s_m - s_n|^2}{2\sigma^2} \right)} \right), \quad (4.8)$$

with s_m and s_n are signal points carved from the standard M -QAM constellation. Since regarding to ϕ , $F_\phi(\cdot)$ in (4.8) is a $(\pi/2)$ periodic function, it is possible to be recast in the form of circular harmonic expansion [51] as

$$LLF(\phi_0|re^{j\phi}) = \frac{A_0(r)}{2} + \sum_{n=1}^{\infty} A_n(r) \cos(n\phi - n\phi_0 + \theta_n(r)), \quad (4.9)$$

where, $A_n(r)$ and $\theta_n(r)$ (depending on the magnitude r of received sample y) are correspondingly the amplitude and phase of the n -th harmonic of the Fourier series.

$$A_n(r)e^{j\theta_n(r)} = \frac{1}{\pi} \int_0^{2\pi} LLF(0|re^{j\phi})e^{-jn\phi} d\phi. \quad (4.10)$$

Without losing generality, it is possible to assume that $\theta_n(r) = 0$ or π so that $A_n(r)$ could be regarded as (real) *weighting function* of the n -th harmonic.

Therefore, the maximization in (4.6) becomes maximization of polyharmonic function (4.9). Then, the decrease of the harmonic's magnitudes by their order n allows approximating of $LLF(\cdot)$ by truncating the Fourier series at a certain order. Further, the I - Q symmetry of standard M -QAM constellation for components signal \bar{s}_k leads to the I - Q symmetry of the composite signal set y_k in (4.5), forcing to zeros all harmonics whose orders are not divisible to 4. Besides, the zero-order harmonic $A_0(r)$ has no meaning in the angular estimation. Thus, after truncating (4.9), the target function is of the form

$$\begin{aligned} \hat{\phi}_0 &= \arg \max_{\phi_0} \operatorname{Re} \left(\sum_{n=1}^N F_{4n}(\{y_k\}) e^{-j4n\phi_0} \right) \\ &= \arg \max_{\phi_0} f(\phi_0), \end{aligned} \quad (4.11)$$

where function $\operatorname{Re}(x)$ returns the real part of x , N is the number of harmonics

used and

$$F_{4n}(\{y_k\}) = \sum_{k=1}^{2K} A_{4n}(r_k) \exp(j4n\phi_k) \quad (4.12)$$

is a nonlinear transform of receive sequence $\{y_k = r_k e^{j\phi_k}\}$.

The use of truncated Fourier series for the phase estimation is the global approximation minimizing integral squared error. To improve the accuracy, it is better to use local approximation of $LLF(\cdot)$ in the vicinity of its maximum. This approximation is obtained by optimizing the weighting functions $A_{4n}(r)$ so that minimizing the estimation's variance of harmonic approximation, $N = 1$, or biharmonic approximation, $N = 2$, of the $LLF(\cdot)$ in (4.11).

The estimation variance is determined by approximating the target function $f(\phi_0)$ in (4.11) to the second order Taylor series in the vicinity of ϕ_0 , assumed to be zero,

$$f(\phi_0) \approx f(0) + \phi_0 f'(0) + \phi_0^2 f''(0)/2 \rightarrow \max_{\phi_0}. \quad (4.13)$$

Simply, the maximum of this approximation is

$$\hat{\phi}_0 = -f'(0)/f''(0). \quad (4.14)$$

Estimation variance of ϕ_0 is calculated based on moments of ϕ_0 in (4.14) and has the following form [79]

$$D\{\hat{\phi}_0\} = m_2\{\hat{\phi}_0\} - m_1^2\{\hat{\phi}_0\} = \frac{K-1}{K} \frac{\overline{X_1^2}}{(\overline{X_2})^2} + \frac{1}{K^2} \frac{\overline{X_1^2 X_2^2}}{(\overline{X_2})^4}, \quad (4.15)$$

where $\overline{X_1}$ and $\overline{X_2}$ respectively are means of X_1 and X_2 , that are nonlinear functions of the receive signal samples, and will be explained below. The summations of these functions over the signal sequence (frame) correspondingly

are the first and second derivatives of the target function $f(\phi_0)$

$$f'(0) = \sum_{k=1}^{2K} X_{1k}, \quad f''(0) = \sum_{k=1}^{2K} X_{2k}. \quad (4.16)$$

The average is calculated over the probability density distribution of sample y , represented in the polar coordinate as

$$p(r, \phi) = \frac{r}{2\pi\sigma^2 M^2} \sum_{\substack{m=1, \\ n=1}}^M \exp\left(\frac{-1}{2\sigma^2} (C_1 - C_2 - C_3 + C_4)\right) \quad (4.17)$$

$$C_1 = r^2 + r_m^2 + r_n^2$$

$$C_2 = 2rr_m \cos(\phi - \phi_m)$$

$$C_3 = 2rr_n \cos(\phi - \phi_n)$$

$$C_4 = 2r_m r_n \cos(\phi_m - \phi_n),$$

where $r_m e^{j\phi_m}$ is the polar form of the m -th signal point in the standard M -QAM constellation.

Simulation results show that when the signal to noise ratio (SNR) is high enough ($SNR > 10$ dB, as a normal operating condition for the 16-QAM), then the second term in the right hand side of (4.15) will be small enough for neglecting. Thus, the optimum weighting functions $\{A_{4n}(r)\}$ could be determined by minimizing the first term in the right hand side of (4.15)

$$\overline{X_1^2} / (\overline{X_2})^2 \rightarrow \min_{A_{4n}(r)}. \quad (4.18)$$

4.3.3. Harmonic approximation

For this approximation, only the 4-th order harmonic is considered ($N = 1$). From (4.11) and (4.12), X_1 and X_2 are drawn as

$$X_1 = 4A_4(r) \sin(4\phi), \quad X_2 = -16A_4(r) \cos(4\phi). \quad (4.19)$$

When there is no a priori phase knowledge, it seems reasonable to assume that this variable is uniformly distributed [61], then for the averaging in (4.18), it is only necessary to focus on the amplitude variable

$$\int_0^{\infty} A_4^2(r) N_4(r) dr \bigg/ \left(\int_0^{\infty} A_4(r) D_4(r) dr \right)^2 \rightarrow \min_{A_{4n}(r)}, \quad (4.20)$$

where

$$N_4(r) = 16 \int_0^{2\pi} \sin^2(4\phi) p(r, \phi) d\phi, \quad (4.21)$$

$$D_4(r) = -16 \int_0^{2\pi} \cos(4\phi) p(r, \phi) d\phi. \quad (4.22)$$

(4.20) could be solved using the Lagrange multiplier method with solution [79]

$$A_4(r) = -\lambda D_4(r) / (2N_4(r)), \quad (4.23)$$

where λ is a Lagrange multiplier, having no effect to the estimation result.

4.3.4. Biharmonic approximation

In this case, both the 4-th and 8-th order harmonics are used ($N = 2$).

Similar calculations result in X_1 and X_2 as

$$X_1 = 4A_4(r) \sin(4\phi) + 8A_8(r) \sin(8\phi), \quad (4.24)$$

$$X_2 = -16A_4(r) \cos(4\phi) - 64A_8(r) \cos(8\phi). \quad (4.25)$$

Also using the Lagrange solver, the weighting functions then are figured as

$$A_4(r) = \frac{\lambda(N_{48}(r)D_8(r) - 2N_8(r)D_4(r))}{4N_4(r)N_8(r) - N_{48}^2(r)}, \quad (4.26)$$

$$A_8(r) = \frac{\lambda(N_{48}(r)D_4(r) - 2N_4(r)D_8(r))}{4N_4(r)N_8(r) - N_{48}^2(r)}, \quad (4.27)$$

where besides $N_4(r)$ and $D_4(r)$ which are correspondingly given in (4.21) and (4.22), the remains are of the forms

$$N_{48}(r) = 64 \int_0^{2\pi} \sin(4\phi) \sin(8\phi) p(r, \phi) d\phi, \quad (4.28)$$

$$N_8(r) = 64 \int_0^{2\pi} \sin^2(8\phi) p(r, \phi) d\phi, \quad (4.29)$$

$$D_8(r) = -64 \int_0^{2\pi} \cos(8\phi) p(r, \phi) d\phi. \quad (4.30)$$

Thus, it is then possible to determine the optimum weighting functions for the phase estimation by both the harmonic and biharmonic methods.

4.4. Performance evaluation of the phase estimation and phase compensation scheme

This section presents and analyzes numerical simulation results of the proposed phase estimation algorithm and phase compensation scheme. The system parameters used for simulation resemble to what generating sub-figure 4.3(a). The MIMO-STBC signals are transmitted in frames each of size $2K = 2000$ symbols. The received signal's phase is then estimated by (4.14) using biharmonic approximation. Then, the frame-based estimated phase values will be averaged over multiframe of size 100 frames.

Four HPA models introduced in section 4.2.1 are included in the simulations. They are typical and updated models covering both TWTA and SSPA technologies. Each model has distinguished amplitude and phase nonlinearities, resulting in different nonlinear effects into the system considered. Because the proposed phase estimation method is based on the Gaussian approximation (4.5), it is important to verify the performance and efficiency of the proposed phase estimation algorithm and phase compensation scheme by

Table 4.1: Estimated phase values and their variances for different nonlinear models.

Saleh				M. Saleh				M. Ghorbani				M. Rapp			
IBO	ϕ_{33}	$\hat{\phi}_0$	$\text{var}(\hat{\phi}_0)$	IBO	ϕ_{33}	$\hat{\phi}_0$	$\text{var}(\hat{\phi}_0)$	IBO	ϕ_{33}	$\hat{\phi}_0$	$\text{var}(\hat{\phi}_0)$	IBO	ϕ_{33}	$\hat{\phi}_0$	$\text{var}(\hat{\phi}_0)$
[dB]	[deg]	[deg]	[deg] ²	[dB]	[deg]	[deg]	[deg] ²	[dB]	[deg]	[deg]	[deg] ²	[dB]	[deg]	[deg]	[deg] ²
6	19.7	17.5	0.085	4	-4.3	-3.1	0.137	9	21.2	18.6	0.334	10	-8.2	-5.4	0.183
7	18.6	16.3	0.025	5	-3.8	-2.7	0.042	10	20.5	19.7	0.198	11	-6.4	-4.3	0.121
8	17.5	15.2	0.019	6	-3.3	-2.3	0.019	11	19.6	17.1	0.085	12	-4.7	-3.3	0.042
9	16.2	13.9	0.015	7	-2.8	-1.8	0.008	12	18.6	16.1	0.032	13	-3.3	-2.4	0.031
10	14.8	12.6	0.008	8	-2.1	-1.2	0.008	13	17.5	15.0	0.025	14	-2.2	-1.7	0.020
11	13.4	11.2	0.007	9	-1.5	-0.7	0.007	14	16.2	13.9	0.019	15	-1.5	-1.1	0.010

considering such a comprehensive set of amplifier types.

4.4.1. Performance of the phase estimator

Estimated phase values, $\hat{\phi}_0$, and their variances, $\text{var}(\hat{\phi}_0)$, when applying Saleh, modified Saleh, modified Ghorbani and modified Rapp models into the simulation system, are respectively listed in Table 4.1. The IBO values used for each model are chosen such that the resulting system bit error rates should be in practically operational range at E_b/N_0 level of 20 dB. Included for referring purpose are the phase rotations ϕ_{33} (under the same nonlinearity conditions) corresponding to signal point s_{33} of the standard 16-QAM constellation.

Reconsidering Figure 4.2 and analyses in section 4.2.2, it is recognized that phase rotation of receive constellation strongly depends on the HPA's AM-PM characteristic but is also weakly controlled by the AM-AM function; and generally, the estimated rotation phases are close to ϕ_{33} . This is reflected in the difference (estimation bias) between the estimated value $\hat{\phi}_0$ and ϕ_{33} and in estimation variance $\text{var}(\hat{\phi}_0)$ in each case. Modified Saleh and modified Rapp models though create (relatively) small phase rotations but their estimation

biases and variances are considerable (compared to that of other cases with larger phase rotations such as Saleh model, for example). The reason is either a) phase rotation is really small yet complicatedly varying²; or b) both AM-AM and AM-PM are significantly nonlinear (for the case of modified Rapp model).

However, the estimation quality, in terms of estimation variance, $\text{var}(\hat{\phi}_0)$, is reliable. With frame and multiframe sizes as set, for all cases, the standard deviation is always smaller than 0.6° , which is a relatively small value for the phase estimation problem, even for terrestrial digital microwave or satellite applications [22, 61]. Moreover, in small phase rotation cases (modified Saleh or modified Rapp models at larger IBOs), the standard deviation is always about one tenth of the estimated value. Therefore, it is not necessary to increase the frame and multiframe sizes to improve the estimation reliability.

In summary, it is reasonable to affirm that though under complicated effects of both AM-AM, and especially, AM-PM conversions in different nonlinearities as discussed above, the accuracy and reliability of the proposed phase estimator are quite satisfactory.

4.4.2. *Optimum proximity of the estimated phases*

The phase compensation as illustrated in Figure 4.1, is carried out after averaging the estimated phases over a multiframe. The optimity of phase compensation is verified by determining the relationship between system's BER versus compensated phase rotations at a given value of IBO. Results are depicted in Figure 4.4, where each curve is noted with a solid square

²For the case of modified Saleh model, though the AM-AM conversion is relatively linear, the AM-PM conversion is strongly nonlinear overall with a positive-to-negative transition in small input range.

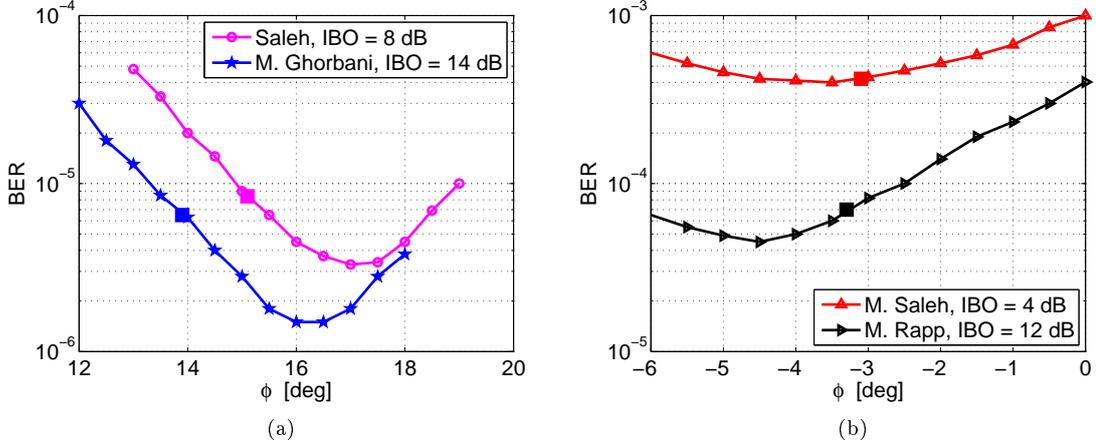


Figure 4.4: BER versus compensated phase angle: a) Saleh and modified Ghorbani models; b) Modified Saleh and modified Rapp models.

marker corresponding to the compensation using estimated phase.

It is observed that though being incurred by different nonlinearities depending on the HPA models, the optimal compensating phases (corresponding to the lowest position in each curve) always approximate to ϕ_{33} as analysed. In general, the proposed phase compensations are suboptimal but performance gains in terms of BER improvements are promising, especially for cases with larger phase rotations (such as for Saleh or modified Ghorbani models). These performance improvements are discussed further in the following section.

4.4.3. Total degradation

The benefits of phase compensations are measured using an important performance metric, the total degradation (TD), defined as [22]

$$TD = E_b/N_0^{NL} - E_b/N_0^L + IBO, \quad (4.31)$$

where E_b/N_0^{NL} is the E_b/N_0 value required to achieve a given target BER (in this case is 10^{-3}) on the nonlinear channel, E_b/N_0^L is the same quantity considering on the linear AWGN channel, and IBO is the input power back-

off; all quantities are in dB. The total degradation results in a convex function of the IBO , taking the minimum value at the optimum IBO . This function can be obtained by following the quasi-analytical procedure described in [80].

Figure 4.5 presents the variation of TD versus IBO for both systems without (HPA only) and with the proposed phase compensation at different nonlinearities generated using all four models. Clearly, huge TD gains could be achieved when applying the phase compensations especially for nonlinearities with strong phase conversions such as for Saleh or modified Ghorbani models.

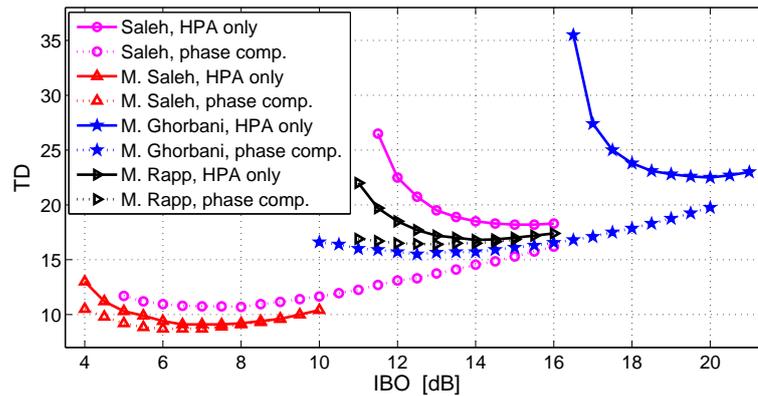


Figure 4.5: TD versus IBO of systems with and without phase compensation at $BER = 10^{-3}$.

Last but not least, by using the proposed phase compensation, it is definitely possible to push the HPA's operating point closer to the saturation point (more than 6 dB for the IBO compared to the case without phase compensation for both TWTA and SSPA) while still maintain the same system performance (in this case is $BER = 10^{-3}$, for example). In practice, this leads to other beneficial consequences in system design, deployment, and maintenance such as power usage effectiveness, space usage efficiency, service life, reliability, dependability,...

4.4.4. Bit error ratio

For calculation simplicity without loss of result meaning, the proposed system is simulated with MISO configuration of $n_T = 2$ transmit antennas and $n_R = 1$ receive antenna. Relation between system BER performances by E_b/N_0 for the cases with and without phase compensation are illustrated in Figure 4.6. Moreover, two BER curves, one for weakly nonlinearity ($IBO = 20$ dB, Saleh model) and the other for the true linearity with AWGN only, are supplemented for comparison purpose.

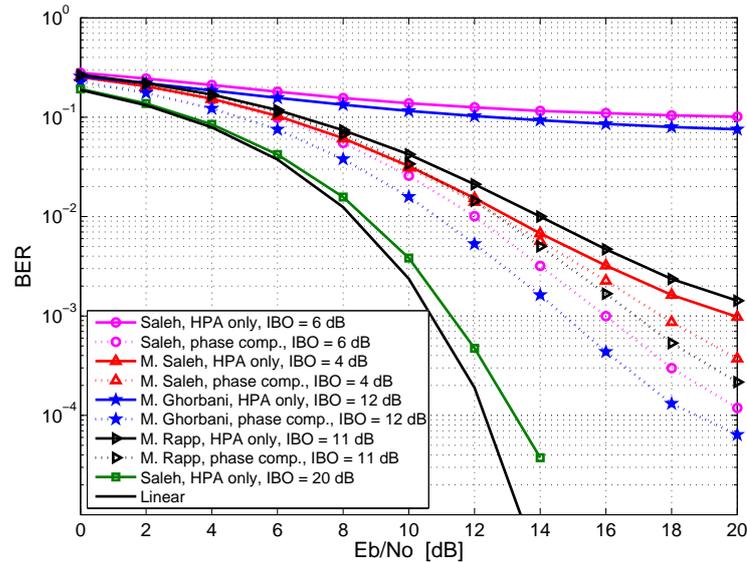


Figure 4.6: BER versus E_b/N_0 of systems with and without phase compensation.

It is clearly seen that the savings of E_b/N_0 for phase-compensated systems with nonlinearities having small phase conversions are still significant (more than 2 dB for modified Saleh model and more than 3 dB for modified Rapp model at $BER = 10^{-3}$). The gains for phase-compensated systems with strong phase conversions (with Saleh, or modified Ghorbani models) are huge.

4.5. Summary of chapter 4

In this chapter, the effects of nonlinear distortion, especially the effects of nonlinear phase distortion incurred by HPAs on the MIMO-STBC system with model including transmit/receive filters are analyzed in detail. It can be seen that the effects of nonlinear distortions for the MIMO-STBC system differ to what have been known for the corresponding effects in the SISO system. Then, limitations and omissions of previous works are pointed out. Based on these detailed analyzes, a phase estimation algorithm and phase compensation scheme are proposed, allowing a significant improvement in the receive signal quality. The effectiveness of proposed scheme is tested using a series of typical HPA models, representing both TWTA and SSPA technologies.

The results achieved in this chapter:

- Introducing transmit/receive filters into the MIMO-STBC system model, using typical HPA models for detailed analyses of the nonlinear phase distortion effects on the receive signals, also pointing out the limitations and shortcomings of previous related publications;
- Proposing the phase estimation algorithm and phase compensation scheme for the receive signals based on the reasonable approximation of nonlinear distortions to the linear model;
- Evaluating the quality of phase estimation and phase compensation scheme through the use of typical HPA models representing both TWTA and SSPA technologies;
- Demonstrating the effectiveness and reliability of the estimation algo-

rithm as well as the efficiency of the phase compensation scheme proposed under the impact of nonlinear distortions at different levels.

Final Conclusions

This thesis deals with the basic problems of nonlinear HPA modeling; MIMO-STBC communications with specific applications for terrestrial mobile satellite communication systems; effects of nonlinear distortion and corrective measures on this system. Contributions, limitations and suggestions for the future research directions are detailed below.

Summary of major findings and contributions

1. Evaluating nonlinear HPA models regarding to advantages and disadvantages in simulating intermodulation products. Proposing polysine models for precise simulation of intermodulation products, especially for modern signals with complex structures based on detailed investigations and analyses of causes giving rise to advantages and disadvantages of typical nonlinear HPA models. These are analyticity and better approximation capability of the proposed model that allow to produce more reliable results than other models which have been widely used in studies involving nonlinear distortions caused by HPA, particularly with the analysis of intermodulation products and the spectrum regrowth;
2. Proposing the use of predistortion schemes for MIMO-STBC systems based on thorough analyses of the nonlinear HPA effects on these systems with transmit/receive filters introduced in the model. Also through these analyzes, limitations and shortcomings in previous publications regarding

to the research issues are explicitly shown; The use of system model with transmit/receive filters allows the results of this proposal to be more reliable and closer to reality. The predistortion diagrams are compared and analyzed in many aspects like algorithm, complexity, performance, and applicability;

3. Reasonably approximating nonlinear phase distortion by a linear model. Based on that, proposing an automatic, asymptotically optimal phase estimation and compensation diagram for MIMO-STBC systems using M -QAM signaling. The signal phase rotation effects incurred by nonlinear HPA are rigorously analyzed, clearly demonstrating the acting mechanisms. This is the basis for a reasonable approximation of this nonlinear effect to a linear phase rotation model, allowing efficient phase estimation and compensation for the nonlinear HPA effect on the system. Typical HPAs, representing both TWTA and SSPA technologies, are included in the survey, evaluating the effectiveness and optimity of the proposed phase estimation and compensation scheme. This proposal is inherited and far developed from a simple yet particularly effective solution for phase compensation that the research group has pursued.

Limitations

- The commonly used HPA models and the proposed polysine model are just tested by simulation and with a data set of an HPA measured in a relatively low frequency band. There need additional tests with other HPA data sets at higher frequency bands, as well as with other commonly used testing signals, and also there need to control simulation results with

hardware experimental results to confirm the superiority of the proposed HPA model;

- Predistortion diagrams also need to be tested on hardware to demonstrate higher practical applicability;
- Estimation and compensation of phase distortion is being carried out with the assumption of ideal transmission channel. Obviously, the system model with fading channels needs to be investigated to ensure the generality of the solution.

Suggested extensions

Although the thesis has focused on the basic theory of nonlinear HPA models as well as proposed two solutions to limit the detrimental effects of nonlinear distortions on MIMO-STBC system, there are still some problems that should be continued to clarify:

- The simulation results have initially confirmed the advantages of the proposed polysine model as well as the pre-compensation and pos-compensation schemes for nonlinear distortions, the hardware experimental tests will solidify the achieved results and confirm the practical applicability of these proposals. These hardware experiments are being developed by the research group;
- Researches on the effects of nonlinear distortions for updated MIMO technologies and systems such as spatial modulation, multi-user MIMO, etc. are still very limited;

- Another research direction that has not been widely discussed for MIMO-STBC systems is the evaluation of system performance degradation under the simultaneous effects of nonlinear distortions and other effects such as linear distortions, or hardware impairments like local oscillator phase noise, sampling jitter, sampling frequency offset, carrier frequency offset, IQ imbalance, RF coupling, cross-talk,...
- The M -APSK modulation schemes are preferred in the new satellite communication standards since they have many advantages over M -QAM schemes. However, nonlinear distortions with the phase rotation effect are always present. The ability to apply a phase estimation and compensation solution for these M -APSK schemes is still left open;
- 5G mobile communication systems uses MIMO beamforming technologies and mm wave communications. There have not been many studies on the issues of assessing the effects of nonlinear distortions caused by HPA for these systems.

List of Publications

1. **Nguyen Thanh**, Nguyen Tat Nam, and Nguyen Quoc Binh, “Automatic phase compensation in MIMO-STBC systems with nonlinear distortion incurred by high power amplifiers,” in *Proceedings of the 2017 Advanced Technology for Communications Conference - ATC 2017*, Quy Nhon, Viet Nam, pp. 86-91, Oct. 18-20, 2017.
2. **Nguyen Thanh**, Nguyen Tat Nam, and Nguyen Quoc Binh, “Performance of a phase estimation method under different nonlinearities incurred by high power amplifiers in MIMO-STBC systems,” in *Proceedings of the Conference on Information and Computer Science - NICS 2017*, Ha Noi, Viet Nam, pp. 42-47, Nov. 24-25, 2017.
3. **Nguyễn Thành**, Nguyễn Tất Nam, Nguyễn Quốc Bình, “Ảnh hưởng của méo phi tuyến do bộ KDCS đến hệ thống MIMO-STBC trong trường hợp có sử dụng bộ méo trước và bộ lọc tạo dạng tín hiệu,” *Tạp chí Khoa học và Kỹ thuật, Học viện Kỹ thuật Quân sự*, trang 74-88, số 188, tháng 2 năm 2018.
4. **Nguyen Thanh**, Nguyen Quoc Binh, Nguyen Thi Phuong Hoa, “Phase estimation and compensation under different nonlinearities incurred by high power amplifiers in MIMO-STBC systems,” *Journal of Science and Technique - Military Technical Academy*, pp. 59-74, No. 191, Jun. 2018.

5. **Nguyen Thanh**, Nguyen Tat Nam, Nguyen Quoc Binh, “On the reasonableness of nonlinear models for high power amplifiers and their applications in communication system simulations,” *Journal of Military Science and Technology - Academy of Military Science and Technology*, pp. 86-99, No. 55, Jun. 2018.

Bibliography

Vietnamese

- [1] Nguyễn Thành Biên (2010), *Méno phi tuyến trong các hệ thống vô tuyến số M-QAM*, Luận án Tiến sĩ Kỹ thuật, Viện Khoa học và Công nghệ Quân sự.
- [2] Nguyễn Quốc Bình (2001), *Kỹ thuật truyền dẫn số*, Nhà xuất bản Quân đội nhân dân, Hà nội.
- [3] Nguyễn Tất Nam (2017), *Nghiên cứu tác động của méno phi tuyến trong hệ thống MIMO và thông tin vệ tinh*, Luận án tiến sĩ kỹ thuật, Học viện kỹ thuật quân sự.
- [4] Nguyễn Thị Hằng Nga (2003), *Nghiên cứu nâng cao chất lượng hệ thống vi ba số dung lượng cao*, Luận án Tiến sĩ Kỹ thuật, Học viện Công nghệ Bưu chính Viễn thông, Bộ Giáo dục Đào tạo.

English

- [5] 802.11 Working Group, LAN/MAN Standards Committee, IEEE Computer Society (2016), *IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012).

- [6] Aghasi A., Ghorbani A. and Amindavar H. (2006), “Polynomial based predistortion for solid state power amplifier nonlinearity compensation”, in *Proc. 2006 IEEE North-East Workshop on Circuits and Systems*, QC, Canada, pp. 181-184.
- [7] Alamouti S. (1998), “A simple transmit diversity technique for wireless communications”, *IEEE J. on Sel. Areas in Commun.*, vol. 16, no. 8, pp. 1451-1458.
- [8] Arapoglou P.-D., Burzigotti P., Bertinelli M., Alamanac A. B., and Gaudenzi R. D. (2011), “To MIMO or not to MIMO in mobile satellite broadcasting systems”, *IEEE Trans. on Wireless Commun.*, vol. 10, no. 9, pp. 2807-2811.
- [9] Behzad Razavi (2014), *Fundamentals of microelectronics*, 2nd Edition, John Wiley & Sons, New York.
- [10] Belfiore J.-C., Rekaya G., and Viterbo E. (2005), “The Golden code: A 2×2 full-rate space-time code with nonvanishing determinants”, *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432-1436.
- [11] Binh N. Q. (1995), *STM-1 Transmission over about 30 MHz microwave channels*, Dissertation for Ph.D, Hungarian Academy of Sciences.
- [12] Binh N. Q., Bérces J., and Frigyes I.(1995), “Estimation of the effect of nonlinear high power amplifier in M-QAM radio relay systems”, *Proc. The Periodica Polytechnica SER. EL. ENG.*, Hungary.
- [13] Binh N. Q., Bien N. T., and Thang N.T.(2008), “The usability of distance degradation in estimation of signal to noise ratio degradation caused by the effect of nonlinear transmit amplifiers and optimum additional phase

- shift in 256-QAM systems”, *Proc. of Int. Conf. on Advanced Technol. for Commun. - ATC 2008*, pp. 258-261, Ha Noi, Viet Nam.
- [14] Bolea Alamanac, A.; Burzigotti, P.; Cohen, M.; De Gaudenzi, R.; Liva, G.; Lipp, S.; Pulvirenti, O.; Roullet, L.; Stadali, H. (2011), “Performance validation of the DVB-SH standard for satellite/terrestrial hybrid mobile broadcasting networks”, *IEEE Trans. on Broadcasting*, vol. 57, no. 4, pp. 802-825.
- [15] Brennan D. G. (1959), “Linear diversity combining techniques”, *Proc. IRE*, vol. 47, pp. 1075-1102, Reprint: *Proc. IEEE*, vol. 91, no. 2, pp. 331-356, Feb. 2003.
- [16] Bussgang J.J. (1952), “Crosscorrelation functions of amplitude-distorted Gaussian signals”, *Tech. Rep., no. 216*, Massachusetts Institute of Technology, MA.
- [17] Cann, A. (1980), “Nonlinearity model with variable knee sharpness”, *IEEE Trans. on Aerospace and Electronic Systems*, vol. 16, no. 6, pp. 874-877.
- [18] Cann, A. (2012), “Improved nonlinearity model with variable knee sharpness”, *IEEE Trans. on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3637 - 3646.
- [19] Cavers J. K. (1990), “Amplifier linearization using a digital predistorter with fast adaptation and low memory requirements”, *IEEE Trans. Veh. Technol.*, vol. 19, no. 4, pp. 374-382.
- [20] Chatzinotas S., Ottersten B., and De Gaudenzi R. (2015), *Cooperative and Cognitive Satellite Systems*, Elsevier, London, UK.

- [21] Choi C.-S., Shoji Y., Harada H., Funada R., Kato S., Maruhashi K., Toyoda I., and Takahashi K. (2006), “RF impairment models for 60GHz-band SYS/PHY simulation”, *Tech. Rep. IEEE 802.15-06-0477-01-003c*.
- [22] Corazza, G. E. (2007), *Digital satellite communications*, Springer, Berlin.
- [23] Cripps, Steve C. (2002), *Advanced techniques in RF power amplifier design*, Artech House, Norwood, MA.
- [24] Davis Linda M.; David Haley (2015), “Geometric polarization and Faraday effects for VHF satellite communication links”, in *Proc. of the 2015 IEEE Int. Conf. on Commun., ICC '15*, London, UK, pp. 910-915.
- [25] Department of the Army Technical Manual, (1952), *Basic theory and application of electron tubes*, Departments of the Army and the Air Force.
- [26] Driessen P.F. and Foschini G.J. (1999), “On the capacity formula for multiple input-multiple output wireless channels: a geometric interpretation”, in *Proc. of the 1999 IEEE Int. Conf. on Commun., ICC '99*, Vancouver, British Columbia, Canada, pp. 1603-1607.
- [27] Elias P., Feinstein A., and Shannon C. (1956), “A note on the maximum flow through a network”, *IRE Trans. on Inform. Theory*, vol. 2, no. 4, pp. 117-119.
- [28] Eroz M. and Lee L-N. (2014), “Method and apparatus for improved high order modulation”, US Patent no. 8,674,758.
- [29] ETSI (2009), *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications*, ETSI, EN 302 307.

- [30] ETSI (2011), *Digital Video Broadcasting (DVB); System Specifications for Satellite services to Handheld devices (SH) below 3 GHz*, ETSI TS 102 585 V1.2.1.
- [31] ETSI (2014), *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications Part II: S2 - Extensions (DVB-S2X)*, ETSI, EN 302 307-2.
- [32] Fadhel M. Ghannouchi, Oualid Hammi, Mohamed Helaoui (2015), *Behavioral modeling and predistortion of wideband wireless transmitters*, John Wiley & Sons, New York.
- [33] Farhang-Boroujeny B. (2010), *Signal processing techniques for software radios*, Chapter 3, Lulu Publishing House.
- [34] Fernandes C. A. R. (2009), “Nonlinear MIMO communication systems: Channel estimation and information recovery using Volterra models”, PhD thesis, Universite de Nice Sophia Antipolis, France.
- [35] Foschini G. J. (1996), “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas”, *Bell Labs Technical J.*, vol. 1, no. 2, pp. 41-59.
- [36] Foschini G. J., Michael. J. Gans (1998). “On limits of wireless communications in a fading environment when using multiple antennas”, *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335.
- [37] Frigyes I., and Horváth P. (2008), “Application of modern diversity methods in satellite communications”, in *Proc. of the 26th Int. Commun. Satellite Systems Conf. (ICSSC)*, San Diego, CA, USA, pp. 1-13.

- [38] Giovanni G. and Marco P. (2018), *Microwave electronics*, Cambridge University Press, New York.
- [39] Godara L. C. (1997), “Application of antenna arrays to mobile communications: Part I: Performance improvement, feasibility, and system considerations; Part II: Beam-forming and direction-of-arrival considerations”, *Proc. of the IEEE*, vol. 85, no. 7/8, pp. 1031-1060, 1195-1245.
- [40] Goldsmith A. (2005), *Wireless communications*, Cambridge University Press, New York.
- [41] Gradshteyn I.S. and Ryzhik I.M. (2015), *Table of integrals, series, and products*, Academic Press, London, UK.
- [42] Hampton J. R. (2013), *Introduction to MIMO communications*, Cambridge University Press, New York.
- [43] Hasan Z., Boostanimehr H., and Bhargava V. K. (2011), “Green cellular networks - A survey, some research issues and challenges”, *IEEE Commun. Surveys & Tutorials*, vol. 13, no. 4, pp. 524-540.
- [44] Haupt L. Randy (2010), *Antenna arrays - A computational approach*, John Wiley & Sons, New Jersey.
- [45] Haykin S. (1996), *Adaptive filter theory*, 3rd Edition, Prentice-Hall, Upper Saddle River, NJ. 07458.
- [46] Herbert J. Reich, John G. Skalnik, Philip F. Ordnung, and Herbert L. Krauss (1952), *Microwave principles*, D. Van Nostrand Company Inc., New York.

- [47] Hesham El Gamal, A. Roger Hammons, Jr., Youjian (Eugene) Liu, Michael P. Fitz, and Oscar Y. Takeshita (2003), “On the design of space-time and space-frequency codes for MIMO frequency-selective fading channels”, *IEEE Trans. on Inform. Theory*, vol. 49, no.9, pp. 2277-2292.
- [48] Hofmann C., Storek K.-U., Schwarz R. T., Knopp A. (2016), “Spatial MIMO over satellite: A proof of concept”, in *Proc. of the 2016 IEEE Int. Conf. on Communication, ICC16*, Kuala Lumpur, Malaysia.
- [49] Horváth P., and Frigyes I. (2006), “Application of the 3D polarization concept in satellite MIMO systems”, in *Proc. of the 49th Annual IEEE Global Telecommun. Conf. (GLOBECOM '06)*, San Francisco, California, USA.
- [50] Horváth P., Karagiannidis K. G., King R. P., Stavrou S., and Frigyes I. (2007), “Investigations in satellite MIMO channel modelling: Accent on polarization”, *EURASIP J. on Wireless Commun. and Networking*, vol. 2007, Article ID 98942, pp. 1-10.
- [51] Jacovitti G., and Neri A. (2000), “Multiresolution circular harmonic decomposition”, *IEEE Trans. Signal Processing*, vol. 48, no. 11, pp. 3242-3247.
- [52] Jeruchim, M., Balaban, P., and Shanmugan, K. (2000), *Simulation of communication systems*, Plenum Press, Berlin.
- [53] Jingon Joung, Chin Keong Ho, Koichi Adachi, and Sumei Sun (2014), “A survey on power-amplifier-centric techniques for spectrum and energy efficient wireless communications”, *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 1, pp. 315-333.

- [54] John Proakis, and Masoud Salehi (2008), *Digital communications*, 5th Edition, McGraw-Hill Education, New York.
- [55] John L.B. Walker (2012), *Handbook of RF and microwave power amplifiers*, Cambridge University Press, New York.
- [56] Kay, S. M. (1993), *Fundamentals of statistical signal processing: Estimation theory*, Prentice Hall, New Jersey.
- [57] Kelly Mekechuk, Wan-Jong Kim, Shawn P. Stapleton, and Jong Heon Kim (2004), “Linearizing power amplifiers using digital predistortion, EDA tools and test hardware”, *High Frequency Electronics*, Summit Technical Media, LLC, http://www.summittechmedia.com/highfrequelec/Apr04/HFE0404_Stapleton.pdf.
- [58] King P.R., Evans B.G., and Stavrou S. (2005), “Physical-statistical model for the land mobile-satellite channel applied to satellite/HAP MIMO”, in *Proc. of the 11th European Wireless Conf.*, Nicosia, Cyprus, pp. 198-204.
- [59] King R.P., Horváth P., Pérez-Fontán F., Frigyes I., and Stavrou S. (2005), “Satellite channel impairment mitigation by diversity techniques”, in *Proc. 2005 IST Mobile and Wireless Communication Summit*, Dresden, Germany, pp 1-5.
- [60] Lei Guan, and Anding Zhu (2014), “Green communications: Digital predistortion for wideband RF power amplifiers”, *IEEE Microwave Mag.*, vol. 15, no. 7, pp. 84-99.
- [61] Ling Fuyun (2017), *Synchronization in digital communication systems*, Cambridge University Press, Cambridge, UK.

- [62] Litva, J. and Lo, T. K-Y (1996), *Digital beamforming in wireless communications*, Artech House, Norwood MA.
- [63] Liu Z., Giannakis G. B., Barbarossa S., and Scaglione A. (2001), "Transmit antennae space-time block coding for generalized OFDM in the presence of unknown multipath", *IEEE J. Select. Areas Commun.*, vol. 19, no. 7, pp. 1352-1364.
- [64] Loyka S.L., Mosig J.R. (2000), "New behavioral-level simulation technique for RF/microwave applications. Part I: Basic concepts", *Int. J. of RF and Microwave Computer-Aided Engineering*, vol. 10, no. 4, pp. 221-237.
- [65] Loyka S.L. (2000), "On the use of Cann' model for nonlinear behavioral-level simulation", *IEEE Trans. on Veh. Tech.*, vol. 49, no. 5, pp. 1982-1985.
- [66] Mietzner J., Hoehner P.A. , and M. Sandell (2003) "Compatible improvement of the GSM/EDGE system by means of space-time coding techniques", *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 690-702.
- [67] Middleton D. (2012), *Non-Gaussian Statistical Communication Theory*, John Wiley & Sons, Inc., Hoboken, New Jersey.
- [68] Mietzner J., et al. (2009), "Multiple-antenna techniques for wireless communications - a comprehensive literature survey", *IEEE Commun. Surveys & Tutorials*, vol. 11, no. 2, pp. 87-105.
- [69] Minkoff J. (1985), "The role of AM-to-PM conversion in memoryless nonlinear", *IEEE Trans. on Wireless Commun.*, vol. 33, no. 2, pp. 139-144.

- [70] Naguib A. F., Seshadri N., and Calderbank A. R. (2000), “Increasing data rate over wireless channels”, *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 76-92.
- [71] Nguyen Tat Nam, and Nguyen Quoc Binh (2015), “Using optimal additional phase shift to reduce the nonlinear distortion effects in MIMO STBC systems”, in *Proc. The 2015 National Conf. on Elect., Commun. and Inform. Technol. - REV-ECIT 2015*, Ho Chi Minh city, Vietnam, pp. 303-308.
- [72] O’Droma M., Meza S., and Lei Y. (2009), “New modified Saleh models for memoryless nonlinear power amplifier behavioural modelling”, *IEEE Commun. Lett.*, vol. 13, no. 6, pp 399-401.
- [73] Oussama B. Belkacem, Mohamed L. Ammari, Rafik Zayani and Ridha Bouallegue. (2013), “Capacity analysis of MIMO-STBC system in the presence of nonlinear distortion and neural network compensator”, *Proc. of The Tenth Int. Symp. on Wireless Commun. Systems (ISWCS)*, Ilmenau, Germany, pp. 1-5.
- [74] Oussama B. Belkacem, Mohamed L. Ammari, Rafik Zayani and Ridha Bouallegue (2014), “On the effect of neural network compensation on MIMO-STBC systems in the presence of HPA nonlinearity”, *Trans. on Emerging Telecommun. Technol.*, vol. 26 no. 9, pp. 1119-1130.
- [75] Pantelis-Daniel A., Konstantinos L., Massimo B., Athanasios P., Panayotis C., Riccardo D. G. (2011), “MIMO over satellite: A review”, *IEEE Communication Surveys & Tutorials*, vol. 13, no. 1, pp. 27-51.

- [76] Peter Almers, Fredrik Tufvesson, and Andreas F. Molisch (2006), “Key-hole effect in MIMO wireless channels: Measurements and theory”, in *IEEE Trans. on Wireless Commun.*, vol. 5, no. 12, pp. 3596-3604.
- [77] Perez-Neira A, et al. (2011), “MIMO channel modeling and transmission techniques for multi-satellite and hybrid satellite-terrestrial mobile networks”, *Elsevier’s Physical Commun.*, vol. 4, no. 2, pp. 127-139.
- [78] Petrov A., and Sergienko A. (2012), “Analytical evaluation of performance for harmonic and biharmonic methods of blind phase offset estimation”, in *Proc. 13th IEEE Int. Symp. on Problems of Redundancy in Inform. and Control Systems*, Saint-Petersburg, pp. 57-61.
- [79] Petrov A., and Sergienko A. (2013), “Optimal blind biharmonic feedforward phase offset estimation for QAM signals“ in *Proc. 2013 IEEE Int. Conf. on Commun. (ICC)*, Budapest Hungary, pp. 4756 - 4760.
- [80] Pupolin S., and Greenstein L.J. (1987), “Performance analysis of digital radio links with nonlinear transmit amplifiers”, *IEEE J. Sel. Areas Commun.*, vol. SAC-5, no. 3, pp. 535-546.
- [81] Qi J., and Aissa S. (2010), “Analysis and compensation of power amplifier nonlinearity in MIMO transmit diversity systems”, *IEEE Trans. on Veh. Technol.*, vol. 59, no. 6, pp. 2921 - 2931.
- [82] Rapp Christoph (1991), “Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal”, *Proc. of European Conf. on Satellite Commun.*, pp. 179-184, Liege, Belgium.

- [83] Rice F., Cowley B., Moran B., and Rice M. (2001), “Cramer-Rao lower bounds for QAM phase and frequency estimation”, *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1582-1591.
- [84] Saleh A. A. M. (1981) “Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers”, *IEEE Trans. on Commun.*, vol. 29, no. 11, pp. 1715-1720.
- [85] Seshadri N. and Winters J. H. (1993), “Two signaling schemes for improving the error performance of Frequency-Division-Duplex (FDD) transmission systems using transmitter antenna diversity”, in *Proc. of the 1993 IEEE Veh. Technol. Conf. (VTC)*, Secaucus, New Jersey, pp. 508-511.
- [86] Shaw, C. and Rice, M. (2009), “Turbo-coded APSK for aeronautical telemetry”, in *Proc. of IEEE Int. Conf. on Waveform Diversity and Design*, Orlando FL, USA, pp. 317-321.
- [87] Sheng Chen (2011), “An efficient predistorter design of compensating nonlinear memory high power amplifier”, *IEEE Trans. Broadcasting*, vol. 57, no. 4, pp. 856-865.
- [88] Siljak D. Dragoslav (1969), *Nonlinear systems: Parameter analysis and design*, John Wiley & Sons, New York.
- [89] Tarokh V., Seshadri N., and Calderbank A. R. (1998), “Space-time codes for high data rate wireless communication: Performance criterion and code construction”, *IEEE Trans. on Inform. Theory*, vol. 44, no.2, pp. 744-765.

- [90] Tarokh V., Jafarkhani H., and Calderbank A. R. (1999) “Space-time block codes from orthogonal designs”, *IEEE Trans. on Inform. Theory*, vol. 45, no. 5, pp. 1456-1467.
- [91] Teikari Ilari, *Digital predistortion linearization methods for RF power amplifiers*, PhD thesis, Helsinki University of Technol., Espoo, Finland, 2008.
- [92] Telatar E. (1999), “Capacity of multi-antenna Gaussian channels”, *European Trans. on Telecommun.*, vol. 10, no. 6, pp. 585-596.
- [93] Tulino A. M., Lozano A., and Verdu S. (2005), “Impact of antenna correlation on the capacity of multiantenna channels”, *IEEE Trans. on Inform. Theory*, vol. 51, no. 7, pp. 2491-2509.
- [94] Tuyen T. D., Tan T. D., Vu T. A., and Tue H. H. (2009), “Performance of STBC MIMO-OFDM using pilot-aided channel estimation and adaptive pre-distortion”, in *Proc. Int. Conf. on Advanced Technol. for Commun. - ATC 2009*, Hai Phong, Vietnam, pp. 104-107.
- [95] Van Nee, R. and Prasad, R. (2000), *OFDM for wireless multimedia communications*, Artech House, Norwood MA.
- [96] Vasjanov A., Barzdenas V. (2018), “A review of advanced CMOS RF power amplifier architecture trends for low power 5G wireless networks”, *Electronics* 7, no. 11: 271. <https://doi.org/10.3390/electronics7110271>.
- [97] Wasaff H.D., *Adaptive pre-distortion for nonlinear high power amplifiers in OFDM systems*, PhD thesis, The Polytechnic University of Catalonia, Catalonia, Spain, 2004.

- [98] Weisstein, Eric W. “Secant Method.” From MathWorld—A Wolfram web resource <http://mathworld.wolfram.com/SecantMethod.html>.
- [99] Weisstein, Eric W., “Newton’s Method.” From MathWorld—A Wolfram web resource <http://mathworld.wolfram.com/NewtonsMethod.html>.
- [100] Whitney Q. Lohmeyer, Raichelle J. Aniceto, and Kerri L. Cahoy (2016), “Communication satellite power amplifiers: Current and future SSPA and TWTA technologies”, *Int. J. of Satellite Comm. and Networking*, vol. 34, no. 2 pp. 95-113.
- [101] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery (2007), *Numerical recipes - The art of scientific computing*, Third edition, Cambridge University Press, New York.
- [102] Wittneben A. (1991), “Basestation modulation diversity for digital simulcast”, in *1991 Proc. of the 41st Veh. Technol. Conf. - Gateway to the Future Technol. in Motion*, pp. 848-853.
- [103] Wolniansky P. W., Foschini G. J., Golden G. D., and Valenzuela R. (1998), “V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel”, in *Proc. of the 1998 URSI Int. Symp. on Signals, Systems, and Electronics*, pp. 295-300.
- [104] Xiaochen Lin, Minglu Jin, and Aifei Liu (2007), “An improved adaptive digital predistortion using discrete Newton’s method”, in *Proc. 2007 Int. Conf. on Wireless Communication, Networking and Mobile Computing*, Shanghai, pp. 834-837.

- [105] Zheng, L. and Tse, D. (2003), “Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels”, *IEEE Trans. on Inform. Theory*, vol. 49, no. 5, pp. 1073-1096.